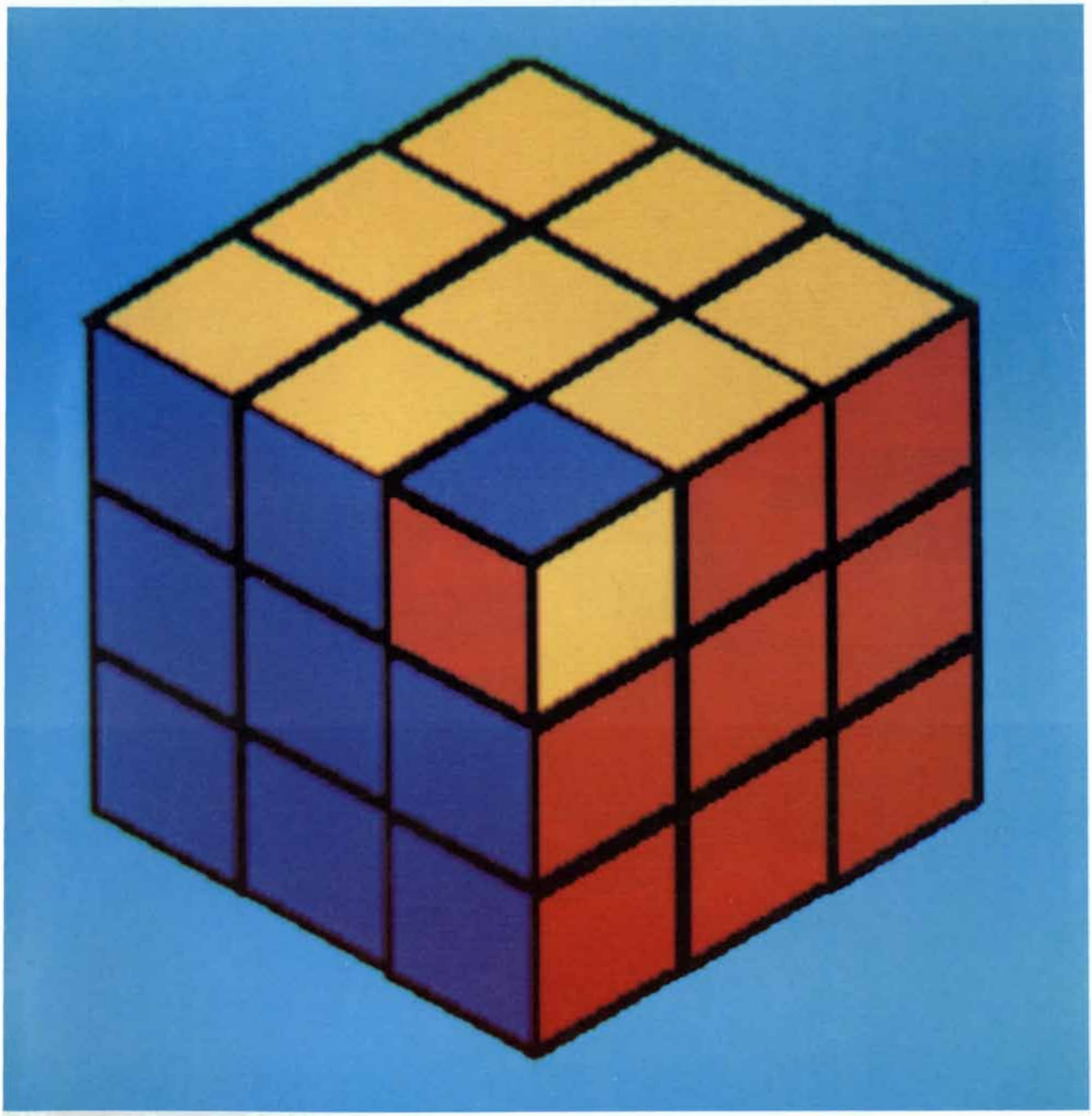
SCIENTIFIC AMERICAN



MAGIC CUBOLOGY

$2.00

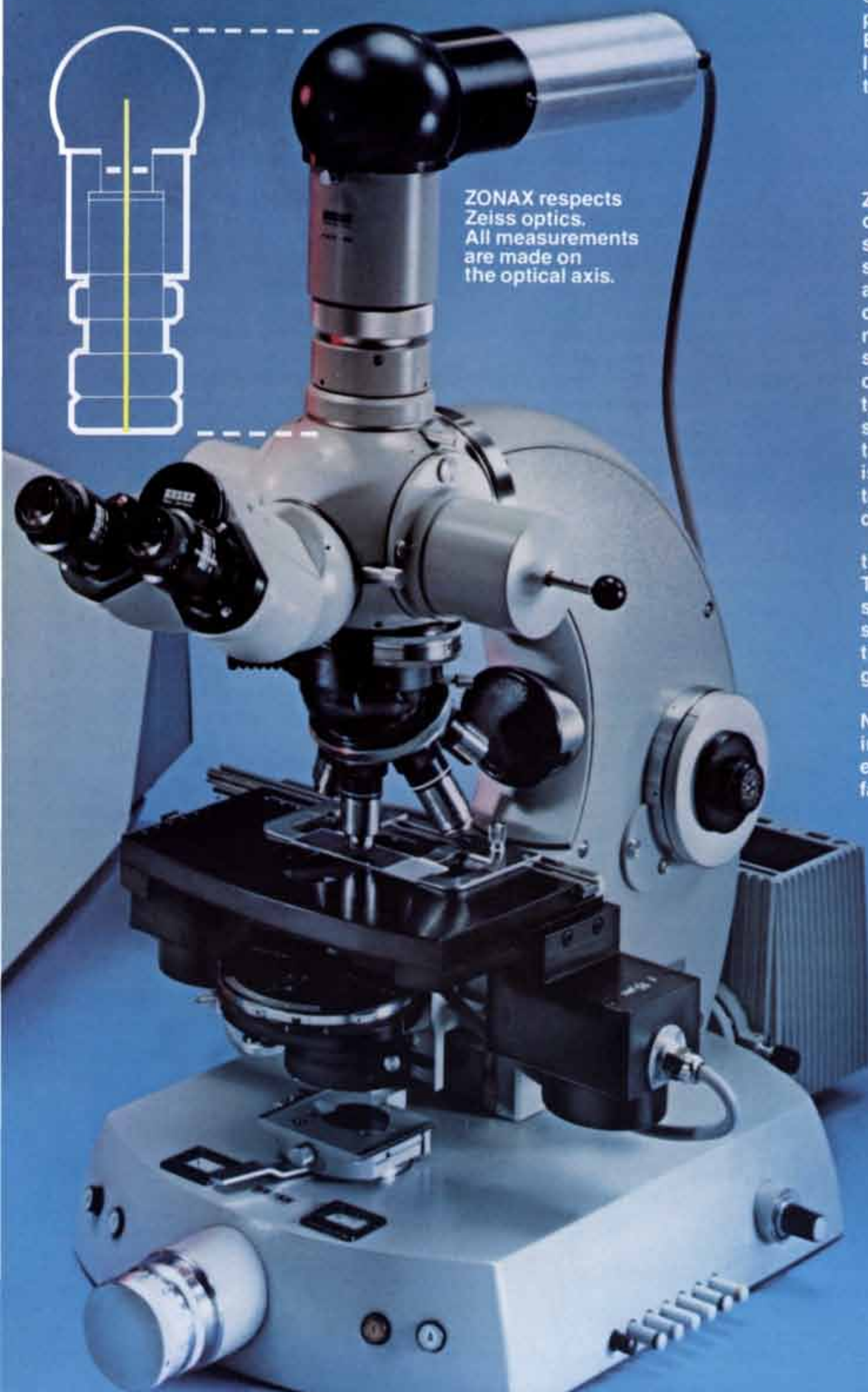
March 1981

Introducing Zonax

The next step in microscopy!

For about $9000,
ZONAX does what the eye
and the camera cannot do–
it quantifies your
microscopic image.

A quantitative microscopy system that makes full use of Zeiss optics.



Now, for about $9000, any Zeiss Microscope Photometer can become an advanced analytical system that displays results with brilliant color graphics.

Zonax so smart, it's simple!

You can use ZONAX from the moment you set it up. No programming to learn! Data analysis at your fingertips.

But, if you want to program it yourself, you can. It comes with complete software documentation for BASIC, FORTRAN or ASSEMBLER. It easily talks with larger computers through its RS232C port.

Zonax respects optics!

ZONAX, with Zeiss optics, is the only system that lets you use microscope optics to the fullest. All measurements are made on the optical axis. The specimen moves, not the optics. You take advantage of the maximum resolution of the microscope. Light goes directly from the objective lens to the photomultiplier to give the best signal quality possible in any system. Only the part of the specimen that's being measured is illuminated; there's no stray light to dilute the signal reaching the detector.

Because the optics are Zeiss, they're the best in the world. They're accompanied by the finest selection of accessories: scanning stages including a 0.125μ resolution stage, illuminators, photographic attachments.

You get everything that the Great Name in Optics can offer: engineering of the highest precision and expert service nationwide from factory-trained personnel.

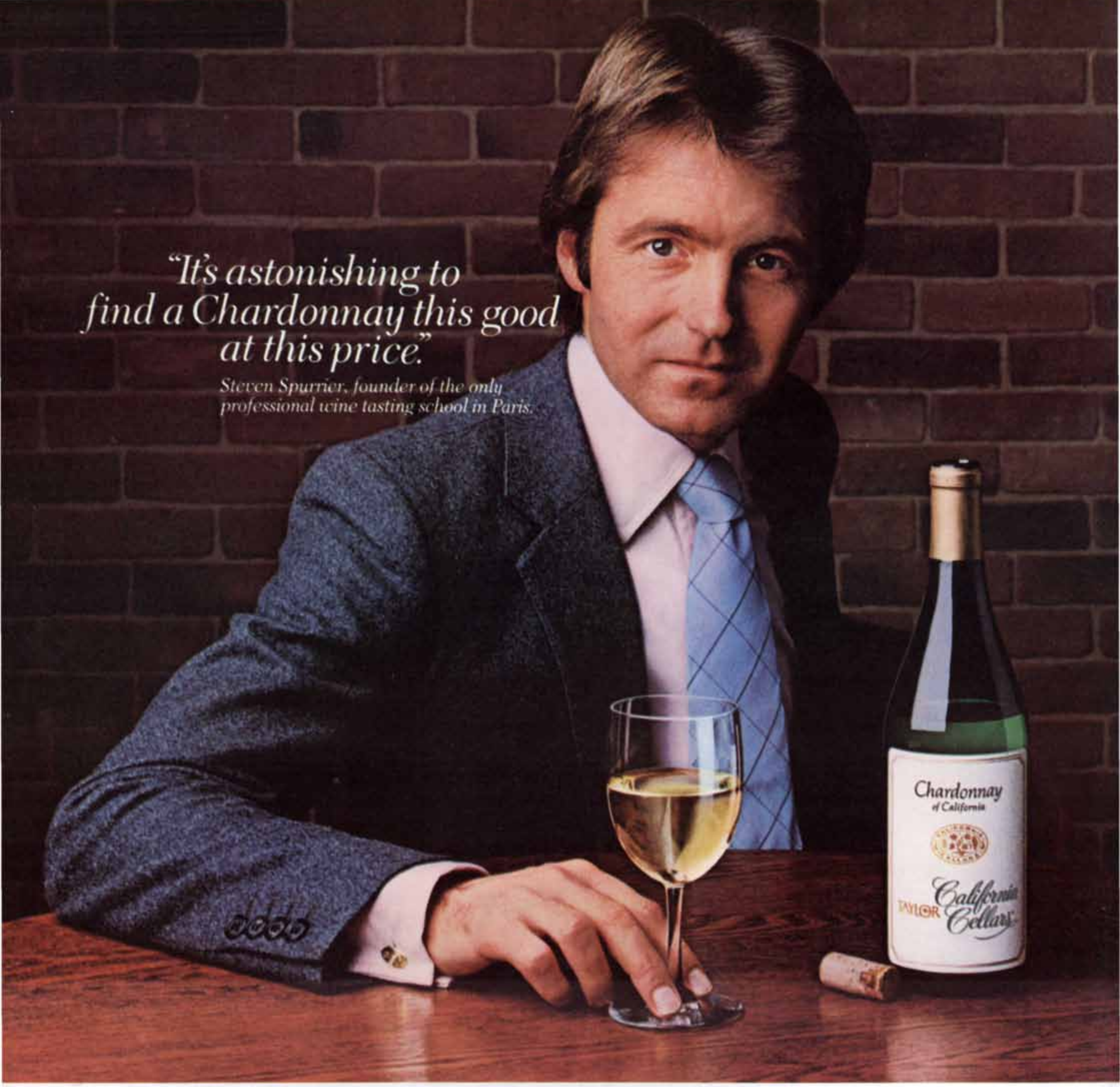
For a demonstration, call your local Zeiss dealer, or contact any Zeiss office.

The great name in optics

ZEISS

West Germany

Carl Zeiss, Inc., 444 5th Ave.
New York, N.Y. 10018
(212) 730-4400
Branches: Atlanta, Boston, Chicago, Houston, Los Angeles, San Francisco, Washington, D.C.



Selling at around $4, Taylor California Cellars varietal wines are an outstanding value. Because they're better than the leading varietals in their price range. But don't take our word for it.

Read what one of the world's leading wine experts, Steven Spurrier, has to say about our Chardonnay: "It's dry, fresh and clean-tasting, with more character than I ever expected to find in a wine at the price. I'm especially impressed by the fruit/acid balance. All in all, this wine is really quite an achievement."

Enjoy Taylor California Cellars Chardonnay, as well as our other exceptional varietals—Zinfandel, Cabernet Sauvignon, Chenin Blanc, Riesling and Sauvignon Blanc. Around $4.00.

Varietal wines from Taylor® California Cellars™.
Better wines. Judge for yourself.

Established 1845 **SCIENTIFIC AMERICAN** March 1981 Volume 244 Number 3

ARTICLES

DEPARTMENTS

SCIENTIFIC AMERICAN (ISSN 0036-8733), PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. COPYRIGHT © 1981 BY SCIENTIFIC AMERICAN, INC. ALL RIGHTS RESERVED. PRINTED IN THE U.S.A. NO PART OF THIS ISSUE MAY BE REPRODUCED BY ANY MECHANICAL, PHOTOGRAPHIC OR ELECTRONIC PROCESS, OR IN THE FORM OF A PHONOGRAPHIC RECORDING, NOR MAY IT BE STORED IN A RETRIEVAL SYSTEM, TRANSMITTED OR OTHERWISE COPIED FOR PUBLIC OR PRIVATE USE WITHOUT WRITTEN PERMISSION OF THE PUBLISHER. CONTROLLED CIRCULATION POSTAGE PAID AT NEW YORK, N.Y. (USPS 509-530), AND AT CHICAGO, ILL. (USPS 508-030). AUTHORIZED AS SECOND-CLASS MAIL BY THE POST OFFICE DEPARTMENT, OTTAWA, CANADA, AND FOR PAYMENT OF POSTAGE IN CASH. SUBSCRIPTION RATE: $21 PER YEAR, U.S., ITS POSSESSIONS AND CANADA; $27 PER YEAR, ALL OTHER COUNTRIES. POSTMASTER: SEND ADDRESS CHANGES TO SCIENTIFIC AMERICAN, 415 MADISON AVENUE, NEW YORK, N.Y 10017

executive health

the report that briefs you on what to watch

©WORLD COPYRIGHT RESERVED 1981 BY EXECUTIVE HEALTH PUBLICATIONS

Volume XVII, Number 5 • February, 1981 • Pickfair Bldg., Rancho Santa Fe, Calif. 92067 • Area 714:756-2600

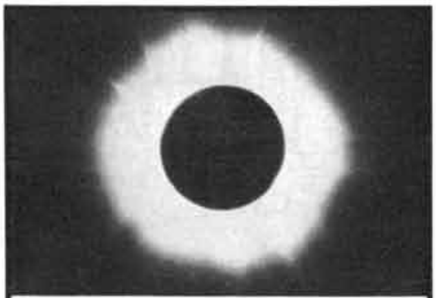
Leonard Hayflick, Ph.D.:

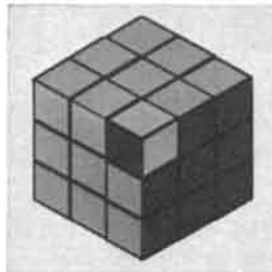
ON THE ROLE OF SLOW VIRUSES AS THE CAUSE OF NEUROLOGICAL DISEASES

Mysterious deaths in New Guinea cannibals have revolutionized our concepts about virus diseases. Several neurological conditions that occur in middle age and beyond may be caused by viruses encountered in our youth that behave like biological time bombs.

PUBLISHER'S NOTE: *Dr. Leonard Hayflick, a member of our Editorial Board, is Senior Research Cell Biologist at Children's Hospital Medical Center, Oakland, California. He is a leading authority on the biology of aging, having published over 150 research papers and books. Dr. Hayflick is the recipient of several major awards for his research in this field, and has just received the nation's most prestigious award in gerontology called The Brookdale Award from the Gerontology Society of America. He also writes on a variety of other complex biological subjects, interpreting them for non-scientific audiences.*

Most of us think of viruses as sub-microscopic organisms that may infect us and then produce disease symptoms in a matter of days. Typical of such diseases are the common cold, measles, mumps, flu and some gastro-intestinal upsets. All viruses were thought to have this quick-strike capability. Now all that has changed. New viruses have been discovered that, after infecting the human body, sit around for months, years, or even decades before disease symptoms are produced. They are true biological time bombs because they can produce diseases that are far more devastating than the flu or a cold. The devastation





THE COVER

The picture on the cover is a computer-graphical display of the Magic Cube, a best-selling puzzle that is currently maddening large numbers of puzzle addicts around the world (see "Metamagical Themas," page 20). In the cube's Start position, which is not the one on the cover, each face consists of nine "cubies" of the same color. Each plane of nine cubies can be rotated around its center, so that the colors of the 26 cubies can be quickly scrambled by a few rotations. The basic puzzle is to get the scrambled cubies back to Start. The number of possible rotations in the process of unscrambling is astronomical, so that keen analysis is called for. The picture on the cover shows the Magic Cube with a corner cubie twisted 120 degrees clockwise and all other cubies, including those on the three back faces, as they were at Start. This configuration cannot be achieved, and so it is called a quark (in analogy with the elusive entity of particle physics). If, on the other hand, there is an oppositely twisted cubie at the diagonally opposite corner of the cube (an antiquark), the configuration can be achieved. The computer-graphical display of the Magic Cube was created by David Christman of the Massachusetts Institute of Technology and Bernard Greenberg of Symbolics, Inc., of Cambridge, Mass. The display is accomplished with a "Lisp machine," one of a family of computers developed at the Artificial Intelligence Laboratory of M.I.T. for the computer language Lisp. With the display program the Magic Cube can be seen from any angle and its faces can be given any twist. Thus the program can re-create any attainable position of the cube and can test routines for transforming positions.

THE ILLUSTRATIONS

Cover photograph by Ralph Morse

Page	Source	Page	Source
21–39	Alan D. Iselin	116	Palomar Sky Survey
49–55	Albert E. Miller	118–119	John N. Bahcall and Raymond M. Soneira
56–58	Keith R. Porter	134	John Pettitt
59	Tom Prentiss	136–138	Ilil Arbel
60–67	Keith R. Porter	139–141	John Pettitt
69	U.S. Geological Survey	146–152	Jerome Kuhl
70–71	Andrew Tomko	155	James B. Jensen (*top*), William Trager (*bottom*)
73	Robert Decker	158–161	Patricia J. Wynne
74–75	National Aeronautics and Space Administration	162	Milton J. Friedman
76	Robert Decker	163–164	Patricia J. Wynne
77–80	Andrew Tomko	166	Jeremy Whitaker
92–94	Allen Beechel	168	Goldsmith Company's Library, University College London
96	United Kingdom–European Southern Observatory Sky Survey (*top*), Allen Beechel (*bottom*)	170	Historical Society of Pennsylvania
100–103	Allen Beechel	172–174	Alan D. Iselin
104	Mount Stromlo Observatory	176	University Library, Cambridge, England
106	Las Campanas Observatory	178–179	Alan D. Iselin
108	Allen Beechel	184–192	Michael Goodman
110	Cerro Tololo Inter-American Observatory		
115	Allen Beechel		

LETTERS

Sirs:

I should like to make a small addition to Reuel A. Stallones' interesting and thoughtful article "The Rise and Fall of Ischemic Heart Disease" [SCIENTIFIC AMERICAN, November, 1980]. In giving reasons why the decline in cigarette smoking might not be responsible for the decline in mortality from ischemic heart disease he stated that "a concordance between [regional variations in cigarette smoking] and the geographic pattern of the mortality from ischemic heart disease also does not seem likely." Actually, using the tobacco-tax data from 44 U.S. states as an admittedly crude index of per capita cigarette consumption, a substantial correlation between cigarette consumption and the state mortality rates for ischemic heart disease has been demonstrated (*Journal of Chronic Diseases,* Vol. 20, pages 769–779; October, 1967).

GARY D. FRIEDMAN, M.D.

Kaiser-Permanente Medical Care Program
Oakland, Calif.

Sirs:

Professor Stallones has made two important points. First, he has demonstrated that the decline of the death rate from changes in the coronary arteries differs from that caused by myocardial infarction (which may suggest different etiologies). Second, he has emphasized that such factors as serum cholesterol levels or the amount and the quality of the fat ingested cannot explain the seeming decrease in the death rate from arteriosclerotic heart disease that has taken place, particularly in the past decade.

Professor Stallones believes, however, that the most important (immediate) cause of myocardial infarction is the obstruction to coronary blood flow by the atherosclerotic atheroma. In particular, blood clots caused by the growing atheroma supposedly occlude the artery and induce sudden heart attacks.

This concept has been frequently attacked. I should like to cite papers by David Spain and A. W. Branwood (*Diseases of the Chest,* Vol. 58, page 107, 1970, and *Lipids,* Vol. 13, page 378, 1978). In these papers it was demonstrated that coronary thrombi are in the majority of cases absent in people who die immediately after the infarction. The incidence of thromboses increases the longer the patient survives the attack. Therefore it appears that the clot is a consequence rather than the cause of the attack. This evidence (and other evidence published by European workers) suggests that myocardial infarction in human beings is most often a primary disease of the heart muscle usually accompanied by coronary changes. Such a concept has more than mere academic interest; the old theory is one of the reasons thousands of coronary bypass operations are performed every year at enormous expense. While it is true that such operations may alleviate anginal pain, the life expectancy of the patients is not on the average increased. Neither is the rate of myocardial infarctions decreased after the operation. This too supports the concept that the changes leading to myocardial infarction are at least not a direct consequence of coronary changes.

HANS KAUNITZ, M.D.

Clinical Professor (retired)
Department of Pathology
College of Physicians & Surgeons of Columbia University
New York

Sirs:

I was, of course, aware of Dr. Friedman's publication of the association between 1950 and 1960 ischemic-heart-disease mortality rates and tobacco-tax revenues by state. The product moment correlation coefficients were .63 and .55 respectively, very notable given the nature of the data. As Dr. Friedman reported, these correlations are in fact buried in a matrix of surprisingly high coefficients. Gratifyingly the 1950 death rates for lung cancer were somewhat more strongly associated with the tobacco data ($r = .81$), but the proportion of the population of the states classed as urban residents was correlated with tobacco-tax data about as strongly ($r = .57$). The correlation coefficients are not helpful in selecting from among the possible inferences: (1) cigarette smoking causes ischemic heart disease, (2) ischemic heart disease causes cigarette smoking, (3) urban residence causes ischemic heart disease, (4) ischemic heart disease causes urban residence, (5) cigarette smoking causes urban residence and (6) urban residence causes cigarette smoking.

Although some of these inferences can be rejected out of hand because they do not fit into a preconceived biological rationale, giving credence only to those associations that fit our preconceptions is one way to perpetuate error. Some years ago I prepared an exercise for students in epidemiology based on the association between the cumulative total of lynchings for each state between 1882 and 1956 and the mortality rates in 1950 for arteriosclerotic heart disease. The rank-order correlation between these variables is $-.67$, which is in the same realm of respectability as the correlations reported by Dr. Friedman. I do not believe that a social proclivity for lynching provides protection against the ravages of atherosclerosis, but I do believe that the network of associations between causes of death and other characteristics of states is extremely complex and that the correlation matrix must be traversed with extreme caution. An obvious extension of Dr. Friedman's work is to perform a similar analysis for 1970 data, when the relative geographic shift of ischemic-heart-disease mortality was further advanced. The results should be interesting.

Dr. Kaunitz brings to our attention the long-standing debate about the interpretation of events immediately preceding a myocardial infarction. I am not qualified to enter the debate and simply report that I employed the concept that seems to represent the majority view. In a report issued in June, 1971, by the National Heart and Lung Institute Task Force on Arteriosclerosis, that distinguished panel said: "Atherosclerosis ultimately produces symptoms because the plaques cause marked narrowing or obstruction of affected arteries.... Very severe or total occlusion may occur rather suddenly as a result of a blood clot developing on the plaque."

REUEL A. STALLONES, M.D., M.P.H.

Dean
School of Public Health
Health Science Center at Houston
University of Texas
Houston

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 C. John Kirby, Advertising Director, 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 60 cents.

Subscription correspondence should be addressed to Subscription Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. For change of address, notify us at least four weeks in advance. Send both old and new addresses and enclose an address imprint from a recent issue. (Date of last issue on your subscription is shown at upper right-hand corner of each month's mailing label.)

Name ____________________

New Address ____________________

____

____

Old Address ____________________

____

____

Give your family the security that will never go up in price. We guarantee it.

You will grow older. Your health may decline. And prices will rise. You know it. And so do I as a New York Life Agent.

Yet, no matter what else changes, there's one thing that will always remain the same: the premiums you pay for New York Life's permanent whole life insurance.

They'll never go up. No, never. Because we spread the cost equally over your lifetime.

In addition, your policy pays dividends and builds cash value. And just look at all the things you can do with this cash reserve.

At retirement, you can use it to buy a paid-up policy for continued protection. You may choose to buy an annuity that will give you a monthly income for life. Or between now and retirement, you can borrow against the cash value to keep your policy in force if you can't afford to pay a premium. And, of course, you can borrow that cash for any kind of emergency.

Is whole life the right insurance for you? Ask me. As a New York Life Agent, I have the facts to help you decide. And if the answer is yes, we'll hold the line on cost.

We guarantee it.

"Ask me."

NEW YORK LIFE

New York Life Insurance Company, 51 Madison Ave., New York, N.Y. 10010. In Canada: 1240-444 St. Mary Avenue, Winnipeg, Manitoba R3C3T1. Life, Group and Health Insurance, Annuities, Pension Plans.

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

MARCH, 1931: "Hermann Oberth of Mediasch, Germany, proposes that we take advantage of the dual forces of gravitation and centrifugal force to send a rocket aloft that will circle the earth continuously. Rising to a height of about 500 miles, equipped with all manner of observation devices, the rocket will acquire a speed of some five miles per second parallel to the earth's surface and then, with its power shut off, will continue to circle the earth indefinitely. From this 'station,' which will circle the earth every two hours, a scientist-observer will watch cloud movements, observe the stars and note changes in the sun, with their resulting effect on terrestrial weather, radio communication and other electric phenomena. As many stations may be established as desired, each, of course, covering a circumferential segment of the terrestrial sphere. Such studies of outer space as are made possible by the station must of course precede attempts to conquer outer space and make a journey to the moon or to one of the other planets."

"Among the discoveries made in fundamental chemistry last year two stand out: the demonstration of the existence of rotating molecules in solid compounds and the solving of the chemical puzzle of the structure of the crystal of the silicates. The existence of rotating molecules in solid compounds was reported by Linus Pauling of the California Institute of Technology and Sterling B. Hendricks of the Fixed Nitrogen Laboratory of the United States Department of Agriculture. The discovery has an important bearing on the heat capacities of solids. The structure of silicate crystals was solved by William L. Bragg of the Victoria University of Manchester and Professor Pauling."

"We are faced with a stagnation of industry such as we thought would never occur again. The country is full of remedies, but in trying to think our way through the difficulty we must remember that we are living in a 20th-century industrial world and not an 18th-century one. Our new world is a world of interdependence and solidarity. It is a world that machines have woven together with thousands of criss-crossing threads. It is a world in which the relations between cause and effect have been so lengthened that on any given day the Egyptian planter cannot know what his cotton or sugar is worth until he receives the quotation from Galveston or Cuba. A heavy frost in the Mississippi valley will affect prices on the Liverpool exchange, and the disturbance will reverberate in Australia and India. French savings, through the channel of a loan to Argentina or Chile, contribute to the development of Belgian or German industry. Prosperity in Czechoslovakia, by increasing the consumption of chocolate, results in the stimulation of the plantations of Venezuela. The fundamental difficulty with our present situation is that two distinct principles are struggling for mastery. In spite of the fact that the economic tides are overflowing the world, we are still trying to maintain our old nationalistic watertight compartments. Our political conceptions have not caught up with our machines. We still cling to the idea that we can maintain political isolation in a world in which economic isolation has long since gone by the board."

SCIENTIFIC AMERICAN

MARCH, 1881: "Vicomte Ferdinand de Lesseps is certainly the very personification of indomitable energy and perseverance. Not content with having successfully canalized one isthmus in the face of every possible obstacle, geographical, political and financial, he has now, in his 76th year, undertaken to execute a similar work on the other side of the Atlantic, which arouses perhaps even a fiercer opposition than did his scheme to pierce the Isthmus of Suez. In a recent conversation he remarked, 'I have greater confidence than I had for Suez. The Atlantic and Pacific breezes blowing down the isthmus will make it the healthiest region in the world. We were there for months—my wife, children, friends and laborers—and we had not a single death.' In the United States, where the Monroe Doctrine of no European interference with the American continent is very popular, M. de Lesseps encountered the fiercest opposition, but he nevertheless succeeded in gaining over a considerable segment of the community to his views. Thus if energy, perseverance and confidence will carry the enterprise through all its difficulties, there is every chance of its success."

"A field for the manufacture of steam engines specially adapted to the propulsion of dynamo electric machines has been opened by the recent extensive and rapid development of the electric light. It is the aim of inventors and manufacturers of electric lamps to provide automatic adjustments that will secure the greatest possible uniformity in the light, and these adjusting devices are called on to compensate not only for the unequal combustion of the carbons but also for the irregularities of the propelling power, every variation of which produces a corresponding variation in the strength of the electric current. This effect is more strikingly illustrated in electric lamps of the incandescent variety, by whose regular fluctuations the strokes of the engine may sometimes be counted. The highest measure of success in electric illumination demands the employment of high-speed engines running with great uniformity. In view of the progress that this kind of illumination is making, together with the great variety of excellent automatic governing valve gear in use, it would pay some of our best engine builders to give attention to this special class of work. The field is large and constantly growing, and offers rich promise to enterprise."

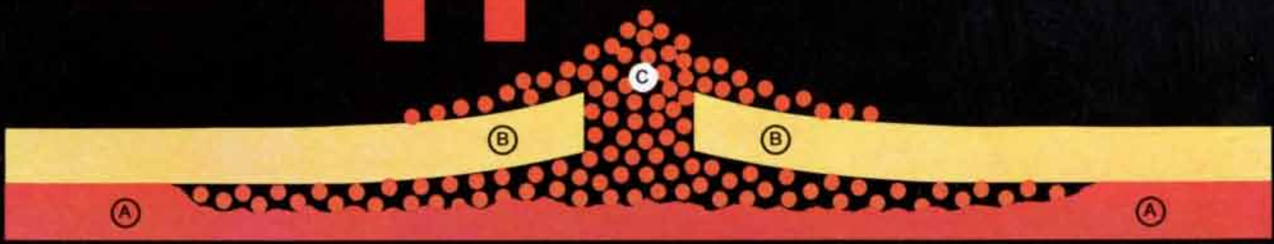
"The records kept by the Hartford Steam Boiler Inspection and Insurance Company show that 170 steam boilers exploded in the United States last year, killing 259 persons and injuring 555. The classified list shows the largest number of explosions in any class to have been 47, in sawing, planing and woodworking mills. The other principal classes were in order: paper, flouring, pulp and grist mills, and elevators, 19; railroad locomotives and fire engines, 18; steam boats, tugboats, yachts, steam barges, dredges and dry docks, 15; portable engines, hoisters, thrashers, piledrivers and cotton gins, 13; ironworks, rolling mills, furnaces, foundries, machine and boiler shops, 13; distilleries, breweries, malt and sugar houses, soap and chemical works, 10."

"The dust boxes of New York are industriously searched by the Italian poor population before they can be removed, and they furnish many a meal, though it is not nice to think of the way such meals are got at. How far more humane would it be were a certain plan of gathering carried out and were these remnants to come in a wholesome condition to poor people? It is not only in the faraway East that Lazarus sits at the gate, sorrowing and suffering, but here in the proud metropolis of the United States that sends food out to millions abroad and yet cannot use up its own broken pieces to best advantage for the poor that sit at its gates waiting for the morsels that come not but are in pure wantonness cast to the dogs or, worse still, on the waters of the broad Atlantic."

"On January 29 the French steamer *Lafayette* arrived at Colón with Messrs. Armand Reclus, G. Blanchet and about 40 others who are to be employed upon the Panama Canal. It is understood that the work is to begin at once."

Consumer Orientation
No. 11 in a Series
Subject: Longlife Design
The Antithesis of
Planned Obsolescence

11 Porsche 911



At Porsche, we design our cars for the long term. We don't believe in the concept of planned obsolescence. Each new model is expected to have a production run of at least 10 years. (The 911 has been in production for more than 15 years.) In addition, every car we build is further expected to bring its owner many years of continued enjoyment.

Our goal of the longlife car has shown, however, that components such as the engine and drive train can have a far longer life span than the body. Road salt, acid rain, and other aggressive airborne matter have created increasing demands on the body's protective coating. Now more than ever, there is the danger of rust.

Paint alone is no protection against rust. (See diagram above.) Harmful elements can attack a car's sheet steel (A) through pinholes in the paint skin (B). Unchecked, rust (C) can expand, mar the finish, and weaken an ordinary steel body. So in addition to a 4-step paint process, the Porsche 911—like all Porsches—is protected by a hot dip galvanizing process.

All of the sheet steel used in all Porsches is hot dip galvanized—on both sides—in a 500°C liquid zinc bath. This produces (see diagram below) a protective zinc coating (D) on the sheet steel (E) that actually *grows* into the damaged portion (F) of the paint skin (G)—plugging pinholes and preventing further corrosion.

As a result of the effectiveness of hot dip galvanizing—beginning with the 1981 model year, we offer a limited warranty on the *entire* body of every new Porsche—911, 924, 924 Turbo, 928—against rust perforation for 7 years:

Porsche Audi represents that Porsche vehicles will be free from rust perforation for the duration of 7 years. If rust perforation develops under normal use and service and the vehicle has been maintained in accordance with manufacturer's requirements, any Porsche Audi authorized dealer will replace or repair the defective parts free of charge.

The hot dip galvanizing process is expensive. But it represents the most comprehensive anti-corrosion protection available. At Porsche, excellence is expected.

For more information or your nearest dealer, call toll-free: (800) 447-4700. In Illinois, call (800) 322-4400.

PORSCHE + AUDI
NOTHING EVEN COMES CLOSE

At the Porsche Research and Development Center (Entwicklungzentrum) in Weissach, the steel body of a Porsche 911 has been standing outside in the elements—protected only by hot dip galvanizing—for 7 years. To date, there has not been one speck of red rust.

What if you chose
as a technical

"At Nikon, HP computers save us $750,000 annually on our production line alone."

Nippon Kogaku K.K. (Nikon) is a leading Japanese optical company and maker of world-famous Nikon cameras. Nikon's Camera Division uses 13 HP 1000's in its Computer Aided Manufacturing process.

Shigehide Segawa, General Manager of Camera Production Engineering, says, "The HP 1000 is our capital partner in manufacturing cameras noted for precision and ease of use.

"With integrated circuits automating the camera's operation, our acceptance testing of IC's, adjustment and inspection of circuits during assembly, and final inspection of the end product depend heavily on computer control.

"Our HP 1000's have saved us $750,000 a year in production. But their total contribution to quality and reliability cannot be expressed by a number."

HP can be your business computer partner too!

The new, top-of-the-line HP 3000 Series 44 computer makes it easy for novices to enter, process, and retrieve data from up to 96 terminals. Thus, it's a powerful tool for high-volume distributed processing. As a member of the compatible HP 3000 family, the Series 44 uses HP's award winning data base management software, and can be networked for instant information access and resource sharing. Prices start at $110,000.* Update kits for smaller 3000's are available.

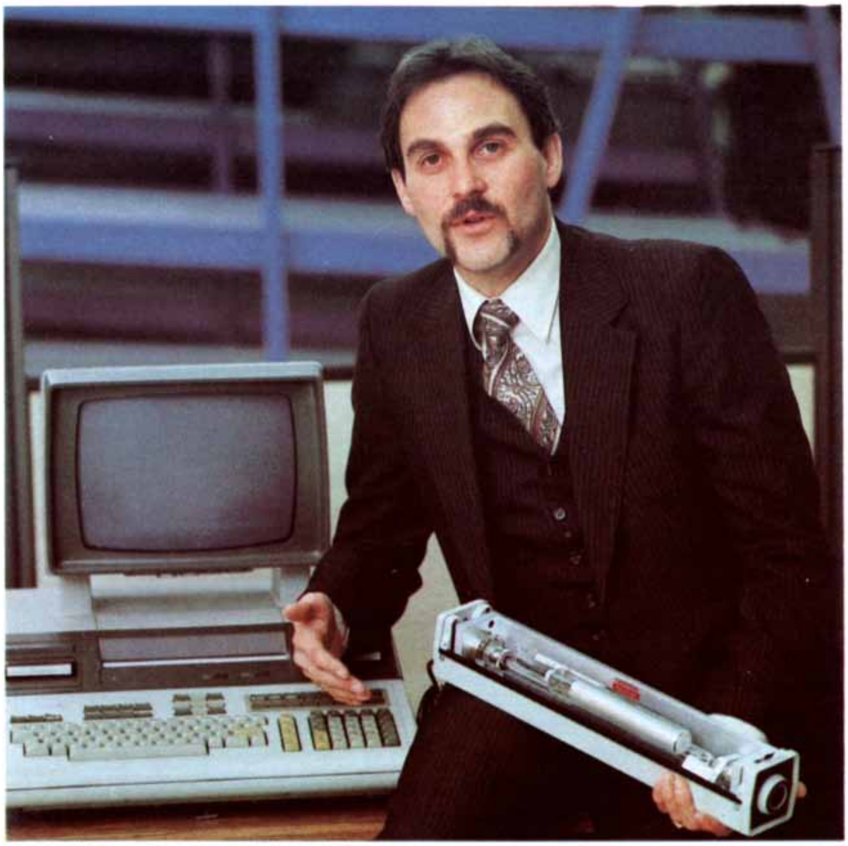
99-percent uptime service guarantee!

This unprecedented guarantee is available under full-service maintenance plans on Series 44's within 100 miles of any of the 43 HP service centers throughout the U.S.

New printing system reduces information cost.

The new HP 2680 Laser Printing System eliminates many information bottlenecks. Operating via HP 3000 computers, the 2680 provides

Hewlett-Packard computer partner?



"At Spectra-Physics, an HP computer tests laser life, and has increased sales by $140,000 a year."

Spectra-Physics' Component Laser Division, in Mountain View, Calif., manufactures low-power lasers for the OEM market. An HP 9845 computer system runs a test program to determine life expectancy and performance of a variety of laser tubes.

"We chose the HP 9845 system because it gives us the data handling and documentation capabilities we need," says Richard Rudy, Manager of Quality Assurance. "Its printouts are used by our marketing people as sales tools.

"With information called up from our test program, we can show customers exactly what the performance characteristics and life expectancy of our laser tubes are. On the basis of the increased sales this documentation has generated, as well as early warnings of potential production problems, our HP system paid for itself in the first 11 months."

versatile, high-speed forms generation on letter-size paper with high print quality. Forms can be merged with data base entries to produce completed documents. As it comes with 32 type fonts—and can reproduce special characters, logos, and symbols—you can tailor forms to your needs. Prices start at $121,000.*

HP's new microcomputer: good news for OEMs.

Newly available 64K RAMs make possible a 512-K byte main memory in the new HP 1000 L-Series two-

board microcomputer. So much main memory and computation power in so little space makes the L-Series easily adaptable to many applications. Prices start at $13,250.*

*Domestic U.S. prices only.

A working partnership with HP.

HP offers a free, 75-page catalog of computer products that provide solutions for Original Equipment Manufacturers. For your copy, call Dept. 304A toll free, (800) 547-3400 (except from Alaska and Hawaii). Oregon residents call 758-1010. Or write A.P. Oliverio, Vice-President, Marketing, Hewlett-Packard Co., 1502 Page Mill Rd., Palo Alto, CA 94304.

When performance must be measured by results

HEWLETT PACKARD

00142

THE AUTHORS

ELI GINZBERG and GEORGE J. VOJTA ("The Service Sector of the U.S. Economy") have a common interest in the analysis of economic issues. Ginzberg is A. Barton Hepburn Professor Emeritus of Economics at the Columbia University Graduate School of Business. He did his undergraduate and graduate work at Columbia, receiving his doctorate in 1934. An expert on labor statistics, he currently serves as director of Columbia's program on the conservation of human resources and as chairman of the National Commission for Employment Policy. Vojta is executive vice-president of Citicorp/Citibank. A graduate of Yale University, he joined Citibank in 1961; he now heads the company's Office of Corporate Strategy and Development. (The views expressed in this article, he writes, are his and Ginzberg's and do not necessarily reflect those of Citicorp.)

KEITH R. PORTER and JONATHAN B. TUCKER ("The Ground Substance of the Living Cell") address the subject of their article from quite different backgrounds. Porter, a pioneer in the application of the electron microscope to the study of cellular structures, is professor of cell biology at the University of Colorado at Boulder. A native of Nova Scotia, he obtained his B.S. from Acadia University in 1934. He then went to Harvard University, where he got his master's and doctor's degrees. For more than two decades he was at the Rockefeller Institute for Medical Research, where he attained the rank of professor and head of the cytology department. In 1961 he took the position of chairman of the biology department at Harvard, and in 1974 he became chairman of the department of molecular, cellular and developmental biology at Boulder. Tucker, who majored in biology as an undergraduate at Yale University, is a free-lance writer.

ROBERT DECKER and BARBARA DECKER ("The Eruptions of Mount St. Helens") have collaborated on several books and articles about volcanoes. They were both engaged in on-site studies of Mount St. Helens in the spring and summer of 1980; during part of that time Robert Decker was the acting scientist in charge of the U.S. Geological Survey's Mount St. Helens Project. He studied geology and geophysics at the Massachusetts Institute of Technology and the Colorado School of Mines, obtaining his D.Sc. from the latter institution. He is a past president of the International Association of Volcanology and Chemistry of the Earth's Interior. Before he joined the Geological Survey in 1979 he was professor of geophysics at Dartmouth College. Barbara Decker, who has a B.A. in journalism from the University of California at Berkeley, is a professional writer. The Deckers (husband and wife) now live in Hawaii, where he heads the Geological Survey's Hawaiian Volcano Observatory.

BART J. BOK ("The Milky Way Galaxy") is professor emeritus of astronomy at the University of Arizona. Born and educated in the Netherlands, he received his Ph.D. at the University of Groningen. He went to Harvard University in 1929 as Robert Wheeler Willson Fellow in astronomy and stayed on to become Robert Wheeler Willson Professor of Astronomy and associate director of the Harvard Observatory. In 1957 he left Harvard to become professor of astronomy at the Australian National University and director of the Mount Stromlo Observatory. In 1966 he was appointed head of the department of astronomy and director of the Steward Observatory of the University of Arizona. Since his retirement as a professional astronomer in 1974, he says, he has returned to his original childhood hobby: amateur astronomy. Bok recently completed work on the fifth edition of his classic text *The Milky Way;* the first edition, coauthored with his late wife Priscilla F. Bok, was published in 1941.

JOHN PETTITT, SOPHIE DUCKER and BRUCE KNOX ("Submarine Pollination") are botanists. Pettitt is a member of the scientific staff of the British Museum (Natural History). He got his B.Sc. and Ph.D. from University College London and spent a postdoctoral year at the University of Michigan. A specialist in the reproductive mechanisms of plants, he has gone to Australia to work with his coauthors on two occasions. Next year Pettitt will be working as a visiting scientist at the Botanical Institute of the University of Stockholm. Ducker retired as a member of the undergraduate teaching faculty at the University of Melbourne in 1975, but she has continued to work with graduate students. She was born in Berlin and was educated in Germany, Britain and Switzerland. Finding the Nazi regime "not congenial," she and her husband left Germany "on Christmas Eve in 1938 and arrived in Australia in 1941." For a number of years she worked at Melbourne on the possible antibiotic activity of soil fungi, receiving her B.S. and M.S. in 1953 and 1956 respectively. A long-standing interest in algae led her to found the first course in marine botany at Melbourne in 1962. In 1978 the university awarded Ducker a doctoral degree for her published work. Knox is professor of botany at Melbourne. Born

in Scotland, he studied botany as an undergraduate at the University of Edinburgh. He earned his Ph.D. from the University of Birmingham in 1962 and joined the faculty of the Australian National University in 1964. Knox was a visiting research fellow at the Institute of Plant Development of the University of Wisconsin in 1969 and at the Royal Botanic Gardens at Kew in 1973.

IRVIN ROCK ("Anorthoscopic Perception") is professor of psychology at Rutgers University, where for many years he has been with the Institute for Cognitive Studies at the university's Newark campus. He will be transferring next fall to the New Brunswick campus, where he will be joining the Program in Cognition. A graduate of the City College of New York, he got his Ph.D. from the New School for Social Research in 1952. Rock taught at the New School and at Yeshiva University before joining the Rutgers faculty in 1967.

MILTON J. FRIEDMAN and WILLIAM TRAGER ("The Biochemistry of Resistance to Malaria") began their collaboration a few years ago at Rockefeller University, where Friedman did his graduate work and where Trager is professor and head of the department of parasitology. Friedman is now an assistant research biologist at the Cancer Research Institute of the University of California at San Francisco. As an undergraduate at Reed College he experimented on the biochemistry of hemoglobin at the University of Oregon Medical Center and also on the biochemical genetics of simple eukaryotic organisms both at Reed and at the University of Washington. Friedman combined these two interests and an interest in international health, he reports, when he went to Rockefeller to begin graduate work on malaria parasites with Trager. Trager has been on the faculty at Rockefeller since 1934, the year after he received his Ph.D. from Harvard University. During World War II he studied malaria in the southwest Pacific as a captain in the U.S. Army Sanitary Corps. He has been a guest investigator at institutions around the world, and he has served as an adviser to many national and international bodies concerned with health.

I. BERNARD COHEN ("Newton's Discovery of Gravity") is Victor S. Thomas Professor of the History of Science at Harvard University. Awarded his Ph.D. by Harvard in 1947, he was the first American to be given a doctorate in this field. Formerly editor of *Isis*, the quarterly journal of the History of Science Society, he is a past president of that society and also of the International Union of the History and Philosophy of Science. In addition to his teaching at Harvard he has lectured in England, France, Italy and Japan.

SELECT

AMERICAN MOTORS, the world leader in 4-wheel drive technology, introduces another breakthrough. The American Eagle now has Select Drive...a spectacular new piece of engineering that lets you switch back and forth between full-time 4-wheel drive traction and 2-wheel drive economy.

No other automobile in the world gives you everything the Select Drive Eagle does.

1. You get 4-wheel

THE AMERICAN EAGLE

DRIVE

drive confidence and security at highway speeds, through any kind of weather. 2. You get 4-wheel drive traction and control when you're heading off the beaten path. And...3. With a simple flick of a switch, you get the high gas mileage of 2-wheel drive when 4-wheel drive isn't desired.

22 / 29*
EPA EST MPG / EST HWY

So, while others are sitting at home, tow a trailer, climb hills, get through bad weather. All this plus an EPA mileage estimate matching Chevy Citation.**

Select Drive...the American Motors invention that fuses high mileage and high traction technology together in one automobile...is now available on all American Eagles. And the only place to get it is your American Motors Dealer's.

Experience driving in the 4th Dimension.

American Motors

*Use these figures for comparison. Your mileage may vary due to speed, weather and trip length. Actual highway mileage will probably be less. Calif. figures lower. **Vehicles with manual 4-speed transmission.

METAMAGICAL THEMAS

The Magic Cube's cubies are twiddled by cubists and solved by cubemeisters

by Douglas R. Hofstadter

Cubitis magikia, *n.* A severe mental disorder accompanied by itching of the fingertips that can be relieved only by prolonged contact with a multicolored cube originating in Hungary and Japan. Symptoms often last for months. Highly contagious.

What this stuffy medical-dictionary entry fails to mention is that contact with the multicolored cube not only cures the itchiness but also causes it. Furthermore, it fails to point out that the affliction can be highly pleasurable. I ought to know; I have suffered from it for the past year and still exhibit the symptoms.

Bűvös Kocka—the Magic Cube, also known as Rubik's Cube—has simultaneously taken the puzzle world, the mathematics world and the computing world by storm. Seldom has a puzzle so fired the imagination of so many people, perhaps not since Sam Loyd's famous 15 Puzzle, which caused mass insanity when it came out in the 19th century and is still one of the world's most popular puzzles. The 15 Puzzle and the Magic Cube are spiritual kin, the one being a two-dimensional problem of restoring the scrambled numbered pieces of a 4 × 4 square to their proper positions and the other being a three-dimensional problem of restoring the scrambled colored pieces of a 3 × 3 × 3 cube to their proper positions. The solutions of both demand that the solver be willing to seemingly undo precious progress time and time again; there is no route to the goal that does not call for a partial but temporary destruction of the order achieved up to that point. If this is a difficult lesson to learn with the 15 Puzzle, it is much harder with the Magic Cube. And both puzzles have the fiendish property that well-meaning bumblers or cunning rogues can take them apart and put them back together in innocent-looking positions from which the goal is absolutely unattainable, thereby causing the would-be solver indescribable anguish.

This Magic Cube is much more than just a puzzle. It is an ingenious mechanical invention, a pastime, a learning tool, a source of metaphors, an inspiration. It now seems an inevitable object, but it took a long time to be discovered. Somehow, though, the time was ripe, because the idea germinated and developed nearly in parallel in Hungary and Japan and perhaps even elsewhere. A report surfaced recently of a French government official who remembers having encountered such a cube, made out of wood, in 1920 in Istanbul and then again in 1935 in Marseilles. Without confirmation the claim seems dubious. In any event Rubik's work was completed by 1975, and his Hungarian patent bears that date. Quite independently, however, Terutoshi Ishige, a self-taught engineer and the owner of a small ironworks near Tokyo, came up with much the same design within a year of Rubik and filed for a Japanese patent in 1976. Ishige also deserves credit for this wonderful insight.

Who is Rubik? Ernő Rubik is a teacher of architecture and design at the School for Commercial Artists in Budapest. Seeking to sharpen his students' ability to visualize three-dimensional objects, he came up with the idea of a 3 × 3 × 3 cube any of whose six 3 × 3 faces could rotate about its center, yet in such a way that the cube as a whole would not fall apart. Each face would initially be colored uniformly, but repeated rotations of the various faces would scramble the colors horribly. Then his students had to figure out how to undo the scrambling.

When I first heard the cube described over the telephone, it sounded like a physical impossibility. By all logic it ought to fall apart into its constituent "cubies" (one of the many useful and amusing terms invented by "cubists" around the world). Take any corner cubie—what is it attached to? By imagining rotating each of the three faces to which it belongs you can see that the corner cubie in question is detachable from each of its three edge-cubie neighbors. How then is it held in place? Some people postulate magnets, rubber bands or elaborate systems of twisting wires in the interior of the cube, yet the design is remarkably simple and involves no such items.

In fact, the Magic Cube can be disassembled in a few seconds [*see bottom illustration on page 25*], revealing an interior structure so simple that one has to ponder how it can do what it does. It actually does fall apart. To see what holds it together first observe that there are three types of cubie: six center cubies, 12 edge cubies and eight corner cubies. The center cubies have only one "facelet," the edge cubies have two facelets and the corner cubies have three. Moreover, the six center cubies are really not cubical at all—they are just façades attached to axles that issue from a sixfold spindle at the middle. The other cubies are nearly complete cubes, except that each one has a blunt little "foot" reaching toward the middle of the cube and some curved nicks facing inward.

The basic trick is that cubies mutually hold one another in by means of their feet, without any cubie's being attached to any other. Edge cubies hold corner cubies' feet, corner cubies hold edge cubies' feet. Center cubies are the keystones. As any layer, say the top one, rotates it holds itself together horizontally and is held in place vertically by its own center and by the equatorial layer below it. The equatorial layer has a sunken circular track (formed by the nicks in its cubies) that guides the motion of the upper layer's feet and helps to hold the upper layer together.

In his definitive treatise "Notes on Rubik's 'Magic Cube'" David Singmaster, professor of mathematical sciences and computing at the Polytechnic of the South Bank in London, defines the "basic mechanical problem" as that of figuring out how the cube is constructed. I sometimes wonder whether Rubik's intended visualization task for his students was to solve the unscrambling problem (Singmaster calls it the "basic mathematical problem") or to solve the mechanical problem. I suspect the latter is the harder of the two. I myself must have put in more than 50 hours of work, distributed over several months, before I solved the unscrambling problem, and I never did solve the mechanical problem until I saw the cube disassembled. Singmaster informally estimates that people who eventually solve the unscrambling problem (without hints) take on the average two weeks of concentrated effort. Of course, it is hard for anyone who has done it to say exactly how long it took (how can you tell play from work?), but it is safe to say that if you are destined to solve the unscrambling problem at all, it will take you somewhere between five hours and a year. I trust this is reassuring.

An important fact many people fail to appreciate at first is that restoring the cube to the Start position (the state

Introducing the only direct flights from Dallas/Ft. Worth to Tokyo and Bangkok.
Our splendid Royal Orchid 747B service. With a convenient wake-up call in the Orient.

Rockwell International is the builder of America's Space Shuttle.

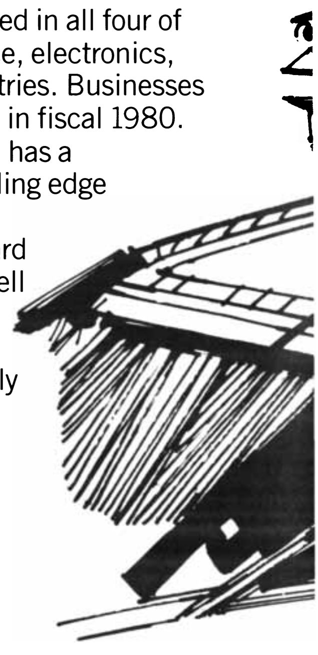
NASA's Space Shuttle is America's most important manned step into space since our last Apollo mission five years ago.

Rockwell International is the prime contractor for the Shuttle and its main engines and assists NASA in the integration of the Space Transportation System. The Shuttle is designed to carry more than 30 tons of useful payload into earth orbit and is capable of making at least one hundred missions into space and back.

The development of the Space Shuttle reflects a Rockwell International strength: the capability to bring together many diverse technological and scientific disciplines to solve complex problems.

This strength is reflected in all four of our core businesses: aerospace, electronics, automotive and general industries. Businesses with total sales of $6.9 billion in fiscal 1980.

Rockwell International has a heritage of working at the leading edge of technology. And today, that heritage is being carried forward by 108,000 dedicated Rockwell employees, including 16,000 engineers and scientists who develop, apply and innovatively manage technology to meet many of the world's unprecedented challenges.



Aerospace.

(Sales, fiscal 1980: $2.0 billion)

Our record for developing and managing aerospace technology is unmatched. In addition to the Space Shuttle, Rockwell International is working on the B-1 bomber and its multi-role bomber derivative, the fourth stage of the MX missile, the Navstar satellites for the Global Positioning System, and a revolutionary forward-swept-wing concept for an Air Force fighter. In our commercial aircraft business, we will build two large fuselage sections for the new Boeing 757 jetliners.

Electronics.

(Sales, fiscal 1980: $1.7 billion)

In defense electronics, Rockwell International is a leader in electronic and electromechanical products and systems for integrated command, control, communications and intelligence equipment, and has a major role in strategic programs of the highest national priority. In the commercial electronics area, Rockwell is a worldwide manufacturer of avionics, microwave transmission systems, data and voice switching equipment and microelectronic devices.

Automotive.

(Sales, fiscal 1980: $1.7 billion)

A prime example of Rockwell International's automotive technology is our SQHP™ tandem truck axle. It's designed to go a half-million miles without a major overhaul, up to a million miles in line-haul service — and it's more than 100 lbs. lighter than competition in its class. Out of one of the country's largest independent automotive research facilities, we develop many such product improvements for on-highway and off-highway vehicles and passenger cars.

Here's mileage you can be proud of in a car you can be proud of.

The V6 Cadillac.

With 29 hwy. est.... [18] EPA est. mpg.*

The V6 Cadillac is the successful merger of two beautifully compatible propositions. Namely, that your next car can (1) combine the inherent quality and value of a Fleetwood Brougham or DeVille with (2) the inherent efficiency of a V6.

New Automatic Overdrive is standard.

The four-speed automatic transmission includes a fuel-saving overdrive gear that helps the V6 Cadillac achieve a most impressive 29 mpg highway estimate.

725 miles est. hwy. range... [450] miles est. driving range.*

The V6 Cadillac also saves time because you don't have to stop so often to refuel.

Or combine a V6 with front-wheel drive...the V6 Eldorado and the V6 Seville.

Their scores of standard features embrace everything from front-wheel drive and four-wheel disc brakes to three-speed automatic transmission and four-wheel independent suspension.

The Diesel Cadillac.

With 33 hwy. est.... [21] EPA est. mpg.*

The Diesel Cadillac boasts the best EPA mileage estimates ever for Fleetwood Broughams and DeVilles.

891 miles est. hwy. range... [567] miles est. driving range.*

The more you drive, the more you need a Diesel Cadillac.

Cadillac outselling Mercedes in Diesel cars.

Figures for the 1980 model year show Cadillac Diesel models outselling Mercedes Diesel models. Source: Ward's Automotive Reports.

Our front-wheel drive Diesels...the Diesel Seville and the Diesel Eldorado.

The Diesel Seville is the only car in the world equipped with a Diesel V8 as standard.

Now you must decide which Cadillac is best for you...with available Diesel or V6. To buy or lease, see your Cadillac dealer.

***Use estimated mpg for comparison. Your mileage and range may differ depending on speed, distance, weather. Actual highway mileage and range lower. Mileage and range lower in California. Range estimates obtained by multiplying EPA est. mpg and hwy. est. by standard fuel tank rating...25 gallons for V6 Fleetwood Broughams and DeVilles...27 gallons for Diesel Fleetwood Broughams and DeVilles.**

Cadillacs are equipped with GM-built engines produced by various divisions. See your Cadillac Dealer for details.

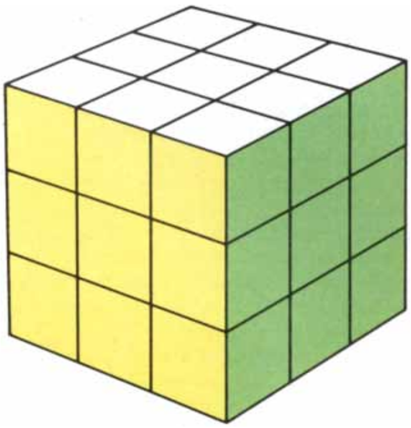
GM

Cadillac

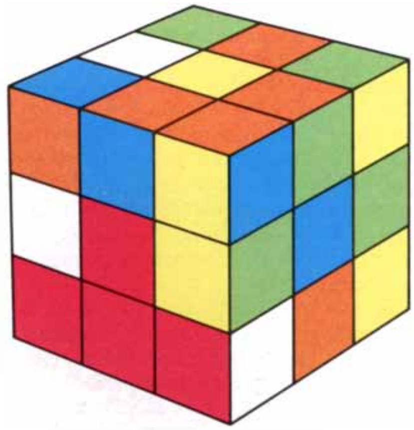
CADILLAC MOTOR CAR DIVISION, U.S.A.

where each face is a solid color) is so hard that it is necessary to find a general algorithm for doing it from any scrambled state. No one can restore a messed-up Magic Cube to its pristine state by mere trial and error. Anyone who gets back to the Start position has built up a small science.

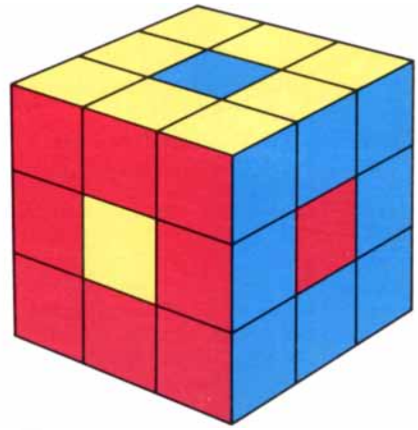
A word of warning: Proposed solutions to the mechanical problem are often lacking in clarity, having either too much detail or too little. It is certainly a challenge to come up with a mechanism that has the multifaceted twistability of the Magic Cube, but it is perhaps no less of a challenge to describe the mechanism in language and diagrams other people can readily comprehend. By the same token, to convey algorithms that restore the cube to Start calls for a good, clear notation. Singmaster himself has an excellent notation that is now considered standard; I shall present it below. A second word of warning: I am not a "cubemeister" (defined as one who has contributed to the profound science of cubology). I am a mere cubist, an amateur who is amazed by the cube and by the virtuosos who have mastered it. Therefore I am not a suitable recipient



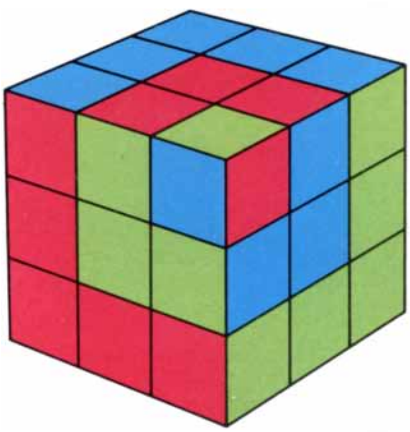
The Magic Cube in the Start position



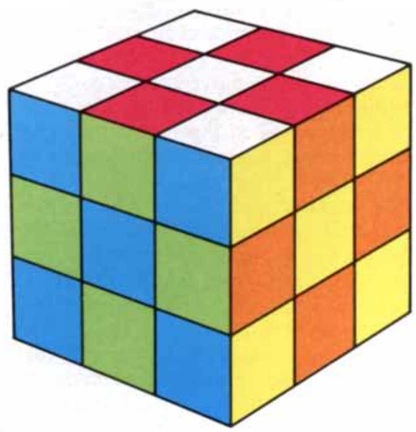
A scrambled cube



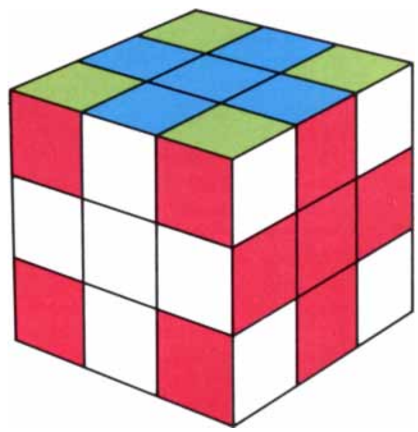
The "pretty pattern" called Dots



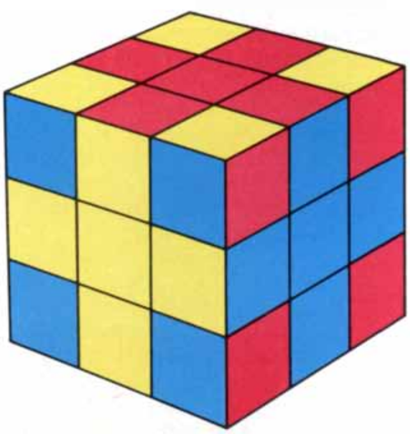
The Giant Meson with Cherries position



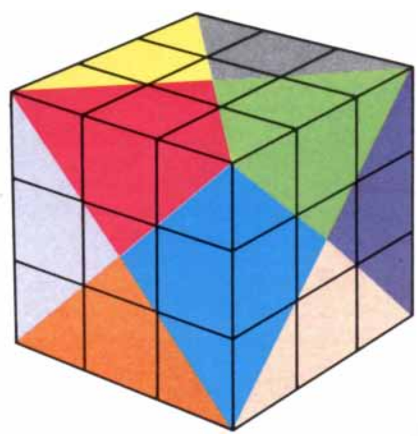
The Pons Asinorum configuration



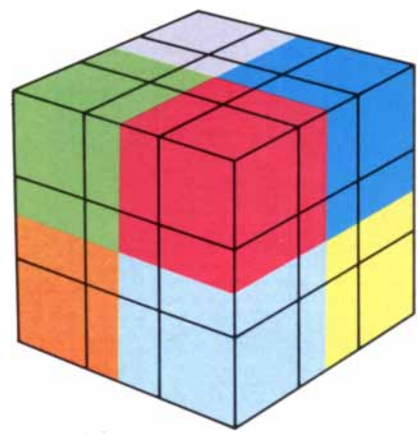
The Christman-cross configuration



The Plummer-cross configuration



An alternate coloring requiring 12 colors

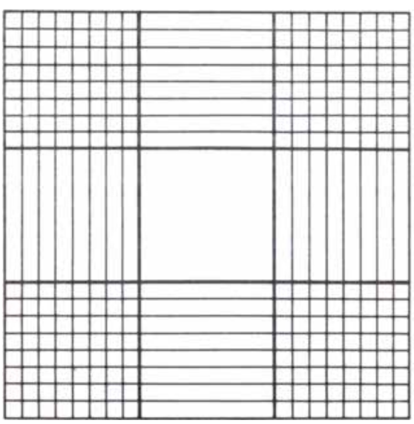


An alternate coloring requiring eight colors

F *stands for face,* **E** *for edge,* **C** *for corner*

of novel solutions to the mechanical problem or the unscrambling problem, of which by now there are hundreds. I recommend that readers who believe they have some novel insight communicate it to Singmaster, who periodically updates his booklet. It will soon be distributed to bookstores in this country by Enslow Publishers of Hillside, N.J., at $5.95. Singmaster's address is Department of Mathematical Sciences and Computing, Polytechnic of the South Bank, London SE1 0AA, England.

The reader's appetite has by now, I trust, been whetted to the point where immediate possession of a Magic Cube is an urgent priority. Fortunately this can be arranged quite easily. Magic Cubes are being manufactured both by Logical Games, Inc., and by the Ideal Toy Corporation. The name Rubik's Cube applies only to the Ideal version, but in all intrinsic ways either version is equally "magic." Cubes are available by mail order from Logical Games (4509 Martinwood Drive, Haymarket, Va. 22069) for $9 postpaid in the U.S. and at many toy and department stores for about $10. It is likely that many people will buy them, little suspecting the profound difficulty of the "basic mathematical problem." They will innocently



Pattern on any face of the Slice Group

turn four or five faces and suddenly find themselves hopelessly lost. Then, perhaps frantically, they will begin turning face after face one way and then another, as it dawns on them that they have irretrievably lost something precious. They feel a little like a child watching a toy balloon drift into the sky.

It is a fact that the cube can be randomized with just a few turns. Let that be a warning to the beginner. Many beginners try to claw their way back to Start by first getting a single face done. Then, a bit stymied, they leave the partially solved cube lying around where a friend may spot it. The well-known "Don't touch it" syndrome sets in when the friend picks it up and says, "What's this?" The would-be solver, terrified that the hard-won progress will be destroyed, shrieks, "Don't touch it!" Ironically, victory can come only through a more flexible attitude allowing precisely that destruction.

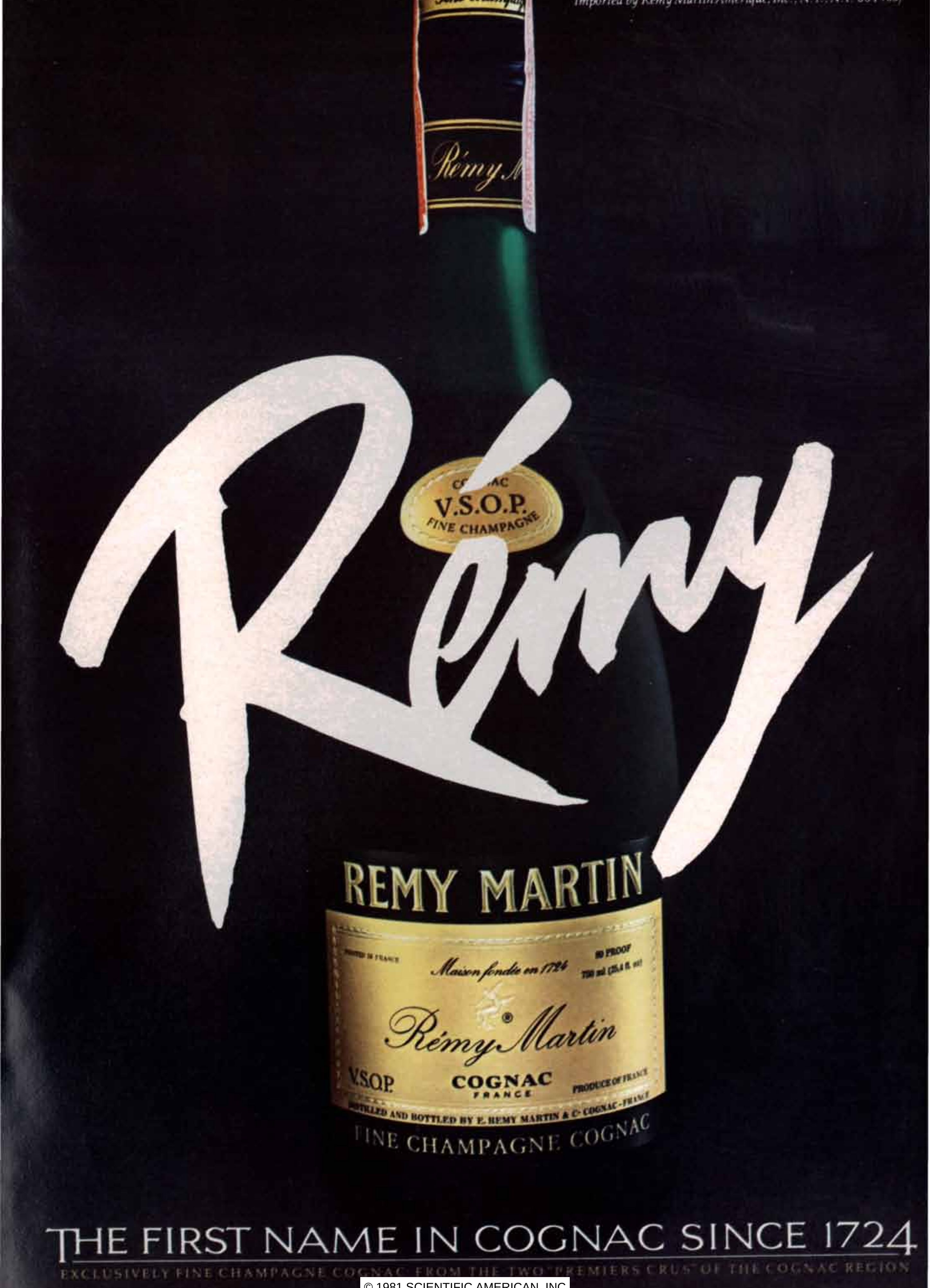
For the beginner there is an awesome sense of irreversibility about destroying the Start position, a kind of fear of tumbling off the edge of a precipice. When my own first cube (I now have five) was first messed up (by a guest), I felt both relieved (because it was inevitable) and sad (because I feared the Start position was gone forever). The physicist in me was reminded of entropy. Once the Start position had become irretrievable each new twist of one face or another seemed irrelevant. To my naive eye there was no distinguishing one messed-up state from another, just as to the eye there is no distinguishing one plate of spaghetti from another, one pile of fall leaves from another and so on. The details meant nothing to me, hence they did not register. As I performed my "random walk" the vastness of the space of possible shufflings of the little cubies became vivid.

As with a deck of cards, one can calculate the exact number of possible rearrangements of the cube. An initial estimate would run this way. The first observation—a rather elementary one—is that on the rotation of any face each corner goes to another corner, each edge to another edge and the center of the face stays put (except for its invisible rotation). Therefore corners mix only with their own kind, and the same goes for edges. There are eight corner cubies and eight corner "cubicles" (the spatial niches, regardless of their content). Cubies and cubicles are to the cube as children and chairs are to the game of musical chairs. Each corner cubie can be maneuvered into any of the eight corner cubicles. This means that we have eight possible fillers for cubicle No. 1, seven for cubicle No. 2, six for cubicle No. 3 and so on. Hence the corners can be placed in their cubicles in $8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$ (=8!) different ways. Each corner, however, can be in any one of three orientations. Thus one would expect a further factor of 3^8 from the corners. One would expect the same for the 12 edge cubies: 12 objects can be permuted among themselves in 12! different ways, and then if each of them has two possible orientations, that gives another factor of 2^{12}. The center cubies never leave their Start positions (unless the cube is rotated as a whole), hence they do not contribute. If we multiply the numbers out, we get 519,024,-039,293,878,272,000 possible positions, about 5×10^{20}.

But there is an assumption here: that any cubie can be got to any cubicle in any orientation, regardless of the other cubies' positions and orientations. As we shall see, this is not quite the case. It turns out there is a constraint on the orientation of the corner cubies: any seven can be oriented arbitrarily, but the last one is then forced, thus removing a factor of three. Similarly, there is a constraint on edge cubies: of the 12 any 11 can be oriented arbitrarily, but the last one is then determined, so that another factor of two is removed. There is one final constraint on the permutations of cubies (disregarding their orientations) that says you can place all but two of them wherever you want but the last two are forced. This removes a final factor of two, reducing the estimate above by a factor of 12, bringing the possibilities down to a mere 43,252,-003,274,489,856,000, about 4×10^{19}. Still, it must be said, this does slightly exceed the assertion on Ideal's label: "More than three billion combinations."

Another way of thinking about this factor of 12 is that if you begin at Start, you are limited to a twelfth of the "obvious" states, but if you disassemble your cube and reassemble it with a single corner cubie twisted by 120 degrees, you are now in a formerly inaccessible state, from which an entire family of 43,252,-003,274,489,856,000 new states is accessible. There are 12 such nonoverlapping families of states of the cube, usually called orbits by group theorists.

Speaking of impossible twists, I should like to mention a lovely parallel in particle physics that was pointed out by Solomon W. Golomb (a familiar name to many of Martin Gardner's readers). It is impossible to make any sequence of moves that will leave just one corner cubie twisted a third of a full turn and everything else the same. Now, recalling the famous hypothetical fundamental particle with a charge of $+1/3$ and its antiparticle with a charge of $-1/3$, Golomb calls a clockwise one-third twist a quark and a counterclockwise one-third twist an antiquark. Like their cubical namesakes, quark particles have proved to be tantalizingly elusive, and many theoretical physicists now believe in quark confinement: the principle



Rémy
COGNAC
V.S.O.P.
FINE CHAMPAGNE
REMY MARTIN
Maison fondée en 1724
Rémy Martin
V.S.O.P.
COGNAC
FRANCE
FINE CHAMPAGNE COGNAC
THE FIRST NAME IN COGNAC SINCE 1724
EXCLUSIVELY FINE CHAMPAGNE COGNAC FROM THE TWO "PREMIERS CRUS" OF THE COGNAC REGION

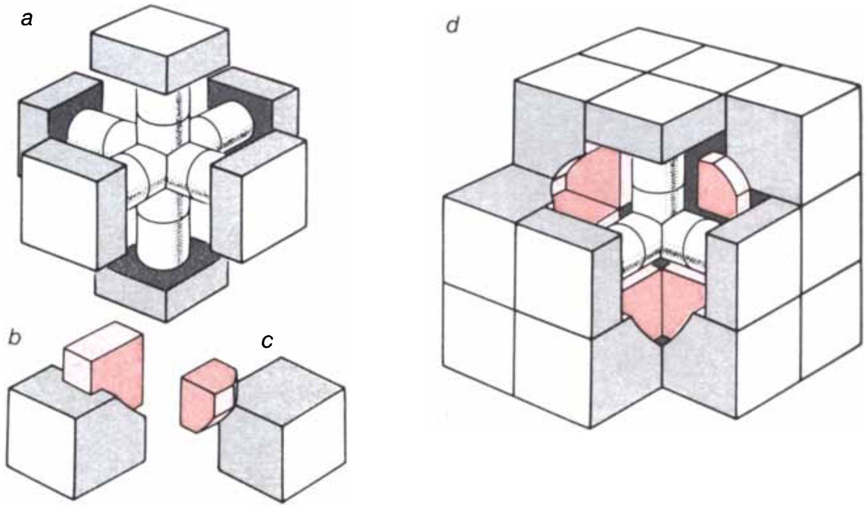

FORD'S KING OF CLUBS ADDS AN ACE!
INTRODUCING THE 1981 FORD LUXURY CLUB WAGON. INTERIOR LUXURY AND COMFORT, WITH MORE ROOM THAN ANY STATION WAGON!
Ford's van achieves an outstanding new level of luxury...in a Club Wagon with Chateau trim, tri-tone paint, super-thick carpeting, low-mount mirrors, dual recline/swivel Captain's Chairs, and much more. Plus Ford's roomy out-front design. And ask your Ford Dealer about Ford's Extended Service Plan.
Tows more. Ford Club Wagon offers light- and heavy-duty towing options that can handle up to 8,000 lbs. of trailer and equipment...more than any station wagon.
STANDARD FEATURES
● Color-keyed: sport wheel covers, grille, headlight doors ● Automatic transmission ● Gauge package ● Tilt wheel ● Speed control ● Privacy™ glass ● WSW radials ● Power steering and door locks ● Ford's tough body-on-frame construction
WIDE CHOICE OF ENGINES
● 4.9L (300 CID) 6-cylinder
● 5.0L (302 CID) V-8
● 5.8L (351 CID) V-8
Built Ford Tough
FORD
FORD DIVISION
Ford

that it is impossible to have an isolated free quark (or antiquark). This correspondence between cubical quarks and particle quarks is a nice one.

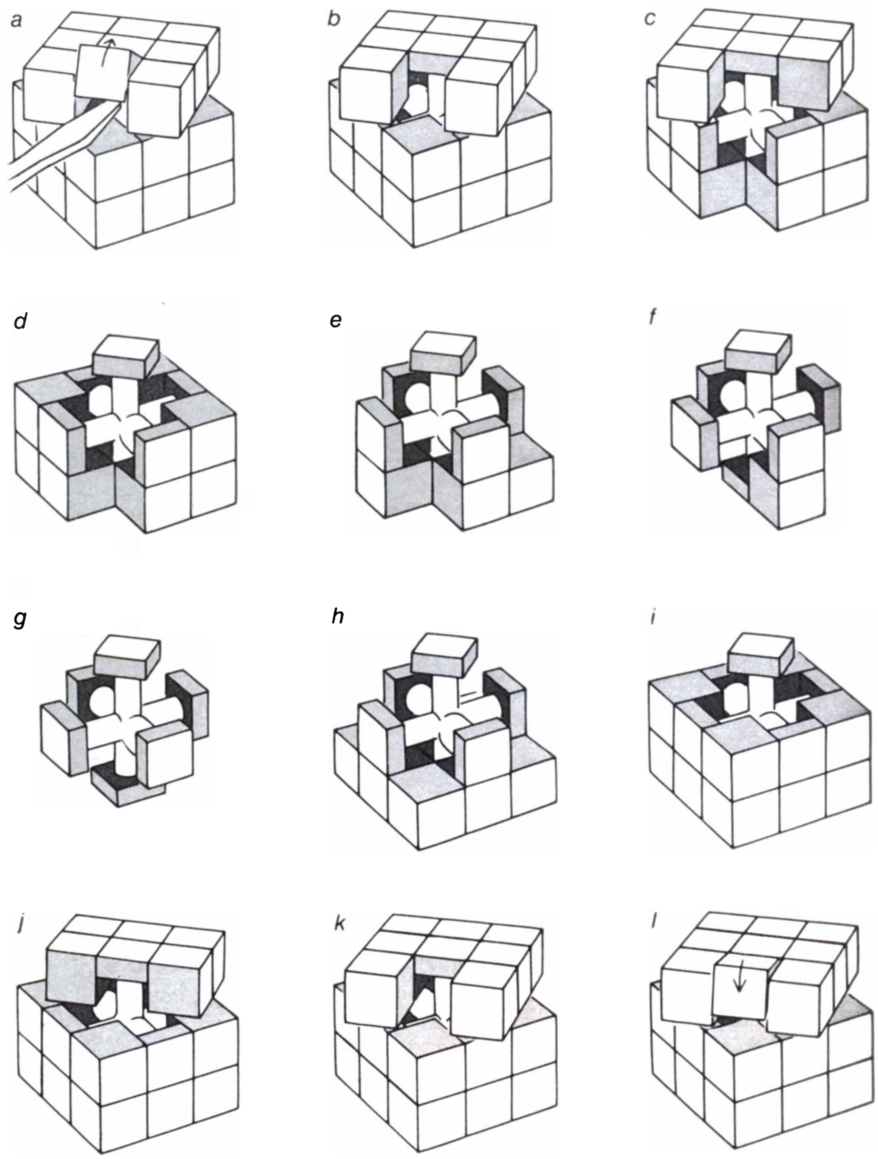
Actually the connection runs even deeper. Quark particles cannot exist free, but they can exist bound together in groups: a quark-antiquark pair is a meson, and a quark trio with integral charge is a baryon. (An example is the proton, with a charge of +1.) Now, in the Magic Cube, amazingly enough, it is possible to give two corner cubies one-third twists, provided they are in *opposite* directions (one clockwise, the other counterclockwise). It is also possible to give three corner cubies one-third twists, provided they are all in the *same* direction. Thus Golomb calls a state with two oppositely twisted corners a meson and one with three corners twisted in the same direction a baryon. In the particle world only quark combinations with an integral amount of charge can exist. In the cubical world only quark combinations with an integral amount of twist are allowed. That is just another way of saying the orientation of the eighth corner cubie is always forced by the first seven. In the cubical world the underlying reason for quark confinement lies in group theory. There may be a closely related group-theoretical explanation for the confinement of quark particles. That remains to be seen, but in any event the parallel is provocative and pleasing.

If we have a pristine cube (one in the Start position), what kind of move sequence will create a meson or a baryon? Here we have an example of the most powerful idea in cubology: the idea of "canned" move sequences that accomplish some specific reordering of a few cubies, leaving everything else untouched (invariant, as group theorists say). There are many different terms for such canned move sequences. I have heard them called operators, transforms, words, tools, processes, maneuvers, routines, subroutines and macros, the first three being group-theoretical terms and the last three being adapted from computer-science argot. Each term has its own flavor, and I like to use them all.

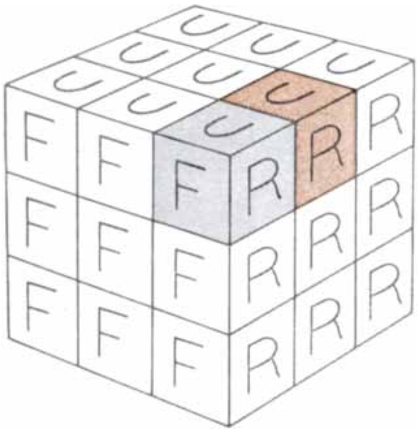
In order to talk about processes we need precision, and that means a good technical notation. I shall therefore present Singmaster's notation now. First we need a way of referring to any particular face of the cube. One possibility is to use the names of colors as the names of the faces, even after the colors of the cubies have become mixed up. It might seem that calling a face white, say, would be meaningless if white is scattered all over the place. Remember, however, that the white center cubie never moves with respect to the five other center cubies, and so it defines the "home face" for white. Then why not use color names for faces? The problem



The spindle **(a)**, *an edge cubie* **(b)**, *a corner cubie* **(c)** *and a view of the internal mechanism* **(d)**



Disassembly **(a–f)** *and assembly* **(g–l)**. *Always reestablish original Start position!*



The urf *cubie* (gray) *and the* ur *cubie* (color)

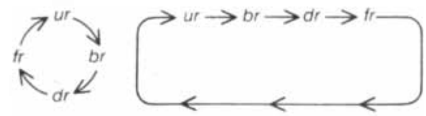
is that different cubes come with their colors arranged differently. Even two cubes from one manufacturer may have different Start positions. A more general convention is to refer to faces simply as "left" and "right," "front" and "back" and "top" and "bottom." Unfortunately the initials create conflicts. Singmaster resolves the conflict by replacing "top" and "bottom" with "up" and "down." Now we have six faces: *L, R, F, B, U, D.* Any particular cubie can be designated by lowercase letters naming the faces it belongs to. Thus *ur* (or *ru*) stands for the edge cubie on the right side of the top layer and *urf* for the corner cubie in front of it [*see illustration above*].

The most natural move for a right-handed cubist seems to be to grasp the right face with the thumb pointing up along the front face and to move the thumb forward. Seen from the right side this maneuver causes a clockwise quarter twist of the *R* face. This move will be designated *R* [*see illustration below*]. The mirror-image move, where the left hand turns the *L* side counterclockwise (as seen from the left), is L^{-1}, or for short *L'*. A clockwise twist of the *L* side is called, naturally, *L*. A 90-degree clockwise turn of any face (from the point of view of an observer looking at the center of that face) is named by the letter for that face, and its inverse—the counterclockwise quarter turn—has a prime mark (') or a superscript -1 following the face's initial. Quarter turns will henceforward be called *q* turns.

With this nomenclature we can now write any move sequence, no matter how complex. A trivial example is four successive *R*'s, which we shall write as R^4. In the language of group theory this is the identity operation: it has zero effect. An equation expressing it is $R^4 = I$. Here *I* stands for the "action" of doing nothing at all.

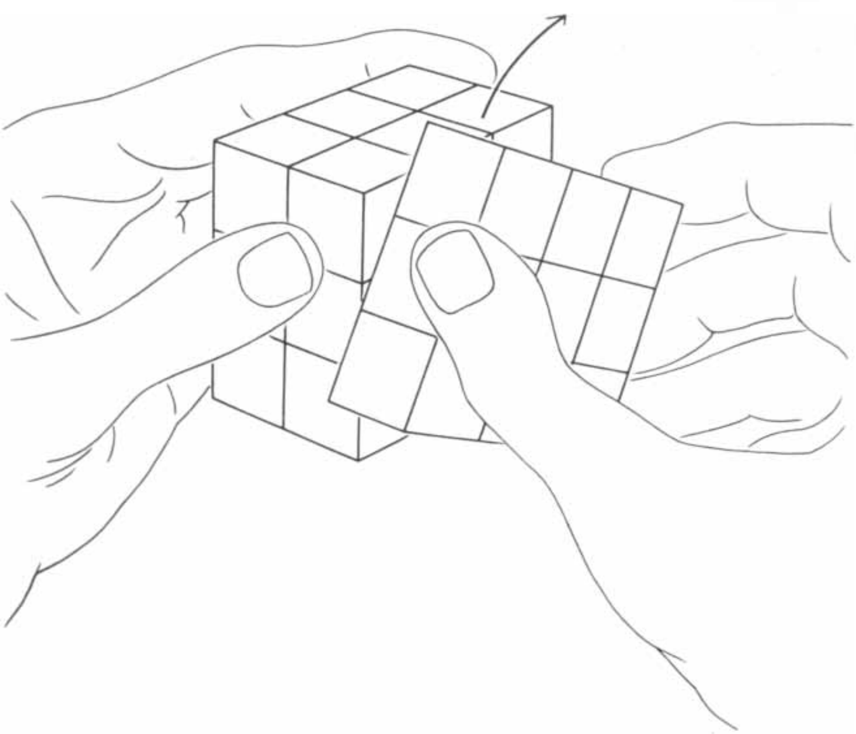
Suppose we twist two different faces, say *R* first, then *U*. We shall transcribe that as *RU*, not as *UR*. Notice first of all that *RU* and *UR* are quite different in their effects. To check this out, first perform *RU* on a pristine cube, notice its effects, then undo it, try *UR* and see how its effects differ. The inverse of *RU* is quite obviously *U'R'*, not *R'U'*. (Incidentally, this strategy of experimenting with move sequences on a pristine cube is most helpful. Very early I found it useful to buy a second cube so that I could work on solving one while experimenting with the other, never letting the second one get far from Start.)

What is the effect of a particular "word"? That is to say, which cubies move where? To answer the question we need a notation for the motions of individual cubies. The effect of *R* on edges is to carry the *ur* cubie around to the back face to occupy the *br* cubicle. At the same time the *br* cubie swings around underneath, landing in the *dr* position, the *dr* cubie moves up like a car on a Ferris wheel to fill the *fr* cubicle and the *fr* cubie comes to the top at *ur*. Symbolically this becomes either of the following:

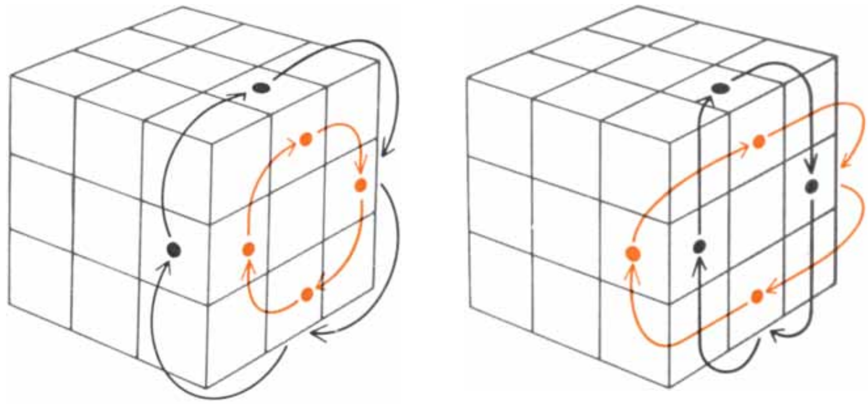


This is called a 4-cycle, and we shall write it in a more compact way: (*ur, br, dr, fr*). Of course, it does not matter where we start writing; we could equally well write (*br, dr, fr, ur*). On the other hand, the order of the letters in cubie names does matter. We can reverse all of them or none of them but not just some of them. If you think of the letters as designating facelets, it will become clear. For example, if we wrote (*ur, rb, dr, rf*), it would represent a 4-cycle involving the same four cubicles, but one in which each cubie flipped before moving from one cubicle to the next. Of course, such a cycle cannot be accomplished by a single *q* turn, but it may be the result of a *sequence* of *q* turns of different faces (an operator). Or consider the following 8-cycle: (*ur, uf, ul, ub, ru, fu, lu, bu*). This has length eight but involves only four cubicles. Each cubie, after making a full swing around the top face, comes back flipped. After two full swings it is back as it started. Each facelet has made a "Möbius trip." We can designate this "flipped 4-cycle" as $(ur, uf, ul, ub)_+$, where the plus sign designates the flipping. The designation $(ru, fu, lu, bu)_+$ and numerous others would do as well. Thus the cycle notation tells you not only where a cubie moves but also its orientation with respect to the other cubies in its cycle.

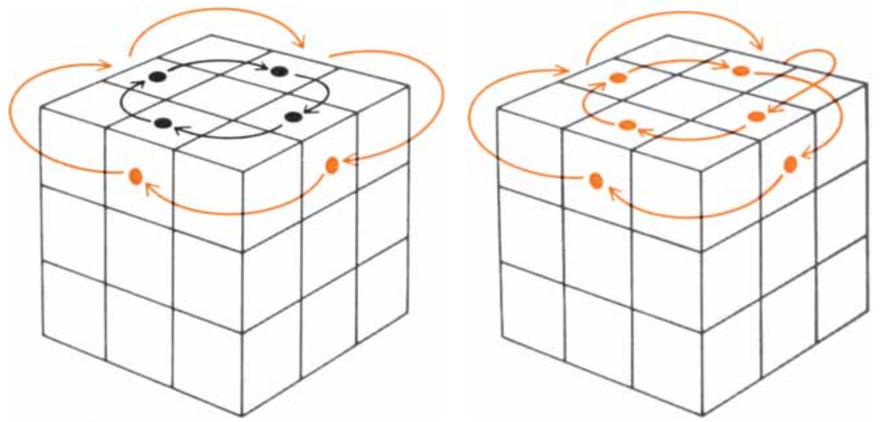
To complete our description of the effect of *R* we must transcribe the 4-cycle of the corners. As with edges we have the freedom to start at any corner we want, and once again we must be careful to keep track of the facelets so that we get the orientations right. Still, *R* has a rather trivial effect on corners: (*urf, bru, drb, frd*), which could also be written (*rub, rbd, rdf, rfu*) and many other ways. Summing up, we can write *R* = (*ur, br, dr, fr*)(*urf, bru, drb, frd*). This says that *R* consists of two disjoint 4-cycles. (If we



R, *a natural move for a right-handed cubist*



The 4-cycle (ur, br, dr, fr) *is at the left, the 4-cycle* (ur, rb, dr, rf) *at the right*



The 4-cycle (ur, uf, ul, ub) *is at the left, the flipped 4-cycle* $(\text{ur, uf, ul, ub})_+$ *at the right*

wanted to, we could throw in a term standing for the 90-degree rotation of the *R* face's center, but since such rotation is invisible, we shall not do so.)

What about transcribing a move sequence such as *RU?* Well, take a pristine cube and perform *RU.* Then start with some arbitrary cubie that has moved and describe its trajectory. For example, *ur* has moved to *br*. Therefore *br* has been displaced. Where has it gone? Find the new location of that cubie (it is *dr*) and continue chasing cubies around and around the cube until you find the one that moved into the original position of *ur*. You will find the following 7-cycle: (*ur, br, dr, fr, uf, ul, ub*) [*see illustration on next page*].

What about corners? Well, suppose we trace the cubie that originated in *urf.* Where did *RU* carry it? The answer is: nowhere; it took a round trip but got twisted along the way. It changed into *rfu*. We can designate this clockwise twist—this quark—$(urf)_+$. This "twisted unicycle" is shorthand for the following 3-cycle: (*urf, rfu, fur*). You can even see this as cycling the three letters *u, r* and *f* inside the cubie's name. If the cycle had been an antiquark, we would have written $(urf)_-$ and the letters would cycle the other way.

What about the other seven corners? Two of them—*dbl* and *dlf*—stay put and the other five *almost* form a 5-cycle: (*ubr, bdr, dfr, luf, bul*). It is unfortunate that the cycle does not quite close, because *bul,* although it gets carried into the original *ubr* cubicle, does so in a twisted manner. It gets carried to *rub,* which is a counterclockwise twist away from *ubr*. This means we are dealing with a 15-cycle. It is so close to the 5-cycle above, however, that we shall just add a minus sign to represent the counterclockwise twist. Our twisted 5-cycle is then $(ubr, bdr, dfr, luf, bul)_-$, and the entire effect of *RU,* expressed in cycle notation, is $(ur, br, dr, fr, uf, ul, ub)(urf)_+(ubr, bdr, dfr, luf, bul)_-$.

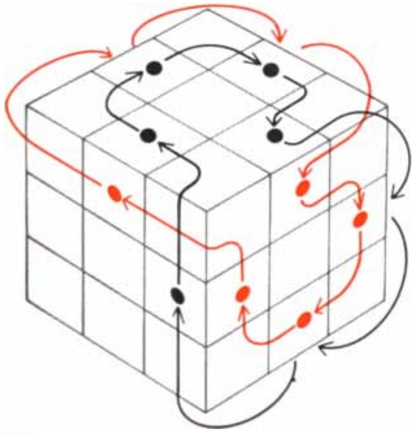
Now that we have *RU* in cycle notation we can perform rotations mentally by sheer calculation. What would be the effect, for instance, of $(RU)^5$? Edge cubie *ur* would be carried five steps forward along its cycle, which would bring it to *ul.* (This can also be seen as moving two steps backward.) Then *ul* would go to *fr* and so on. The 7-cycle is replaced by a new 7-cycle: (*ur, ul, fr, br, ub, uf, dr*). Let us now look at the twisted 5-cycle. Corner cubie *ubr* would be carried five steps forward along its cycle, which brings it back to itself negatively twisted, namely *rub*. Similarly, all the corner cubies in the 5-cycle would return to their starting points, but negatively twisted; thus on being raised to the fifth power a negatively twisted 5-cycle becomes five antiquarks. But if that is so, how is the requirement of integral twist satisfied? Do we not have one quark—$(urf)_+$—and five antiquarks, and does that not add up to four antiquarks, which have a total twist of $-1\frac{1}{3}$? Well, I have slipped something by you here. Can you spot it? To gain facility with the cycle notation you might try to find the cycle representation of various powers of *RU* and *UR* and their inverses.

Any sequence of moves can be represented in terms of disjoint cycles of various lengths: cycles with no common elements. If you are willing to let cycles share members, however, any cycle can be further broken up into 2-cycles (called transpositions, or sometimes swaps). For example, consider three animals: an alligator, a bobcat and a camel. They initially occupy three ecological niches: *A, B* and *C* [*see illustration on page 30*]. The effect of the 3-cycle (*A, B, C*) is to put them in the order camel, alligator, bobcat. The same effect can be achieved, however, by first performing the swap (*A, B*) (what was in *A* goes to *B* and vice versa) and then performing (*A, C*). Of course, this can also be achieved by the two successive swaps (*A, C*)(*B, C*) and (*B, C*)(*A, B*). On the other hand, no sequence of three swaps will achieve the same effect as (*A, B, C*). Try it yourself and see. (Note that a niche is like a cubicle and an animal is like a cubie.)

An elementary theorem of zoo theory (a field we shall not go into here) states that no matter how a given permutation of animals among niches is reduced to a product of successive swaps (which can always be done) the parity of the number of such swaps is invariant, that is to say, a permutation cannot be expressed as an *even* number of swaps one time and an *odd* number another time. Moreover, the parity of any permutation is the sum of the parities of any permutations into which it is broken up (using the rules for the addition of even and odd numbers: odd plus even is odd and so forth).

Now, this theorem has repercussions for the Magic Cube. In particular you can see that any *q* turn consists of two disjoint 4-cycles (one on edges and one on corners). What is the parity of a 4-cycle? It is odd, as you can work out for yourself. Hence after one *q* turn both the edges and the corners have been permuted oddly, after two *q* turns evenly, after three *q* turns oddly and so forth. The edges and corners stay in phase, in the sense that the parities of their permutations are identical. Now, clearly the null permutation is even (zero swaps). Thus if we have a null permutation on corners, the permutation on edges must also be even. Conversely, a null permutation on edges implies an even permutation on corners. Imagine a state identical with Start except for two interchanged edges (that is, one swap). The state is even in corners but odd in edges, hence it is impossible. The best we could do would be to have *two pairs* of interchanged edges. The same argument holds for corners. In short, single swaps are impossible; swaps must always come in pairs. (This is the origin of one of our factors of two in the calculation above of the number of states of the cube.) There are processes for exchanging two pairs of edges, two pairs of corners and even for exchanging one pair of edges along with one pair of corners. (This last process necessarily involves an odd number of *q* turns.)

To round out the subject of constraints let us ponder the origin of the constraints on corner twisting and edge flipping. Here is a clever explanation provided by the mathematicians John Horton Conway, Elwyn R. Berlekamp and Richard K. Guy, building on ideas of Anne Scott. The basic concept is that we want to show that the number of flipped cubies is always even and that the twist is always integral. In order to determine what is flipped and what is twisted, however, we need a frame of reference. To supply it we shall define two things called the "chief facelet of a cubicle" and the "chief color of a cubie." (Remember that a cubicle is a niche and a cubie is a solid object.) The chief facelet of a cubicle will be the one on the up or down surface of the cube, if there is one; otherwise it will be the one on the left or right wall [*see top figure in illustration on page 32*]. There are nine chief facelets on *U*, nine on *D* and four on the equator. We can forget about centers, however, because they never can be flipped or twisted. The chief color of a cubie is defined as the color that should be on the cubie's chief facelet when the cubie "comes home" to its proper cubicle in the Start position.



The 7-cycle **(ur, br, dr, fr, uf, ul, ub)**

Now suppose the cube is scrambled. Any cubie that has its chief color in the chief facelet of its current cubicle will be called sane; otherwise it will be called flipped (if it is an edge cubie) or twisted (if it is a corner cubie). Obviously there are two ways a cubie can be twisted: clockwise ($+1/3$ twist) and counterclockwise ($-1/3$ twist). The "flippancy" of a cube state will be defined as the number of flipped edge cubies in it, and the "twist" as the sum of the twists of the eight corner cubies. We shall say that the flippancy and the twist of Start are both zero.

Next consider the 12 possible *q* turns out of which everything else is compounded. Performing *U* or *D* (or their inverses) preserves both the flippancy and the twist, since nothing leaves or enters the up or down face. Performing *F* or *B* (or their inverses) leaves the twist constant, by changing the twist of four corners at once: two by $+1/3$ and two by $-1/3$. It also leaves the flippancy alone [*see middle figure in illustration on page 32*]. Performing *L* or *R* will likewise leave the twist constant (four corner twists again cancel in pairs) and will change the flippancy by 4, since always four cubies will change in flippancy [*see bottom figure in illustration on page 32*]. The conclusion is what we stated above without proof: the eight corner cubies are always oriented to make the total twist a whole number, and the 12 edge cubies must always be oriented to make the total flippancy even.

After this discussion of constraints you should be convinced that no matter how you twist and turn your Magic Cube you cannot reach more than a twelfth of the conceivable "universe," beginning at Start. It is another matter, however, to show that every state within that one-twelfth universe is accessible from Start (or what amounts to the same thing, only backward: that Start is accessible from every state in the one-twelfth universe). For this we need to show how to achieve all even permutations of cubies and how to achieve all orientations that do not violate the two constraints described above. What it comes down to is that we have to show there are operators that will perform seven classes of operations: (1) an arbitrary double edge-pair swap, (2) an arbitrary double corner-pair swap, (3) an arbitrary two-edge flip, (4) an arbitrary meson, (5) an arbitrary 3-cycle of edges, (6) an arbitrary 3-cycle of corners and (7) an arbitrary baryon.

Of course, each of these operators should work without causing side effects on any other parts of the cube. With these powerful tools in our kit we would be able to cover the one-twelfth universe without any trouble. In the case of the overlapping swaps of animals, I showed that a 3-cycle is really two overlapping 2-cycles. This means that classes 5 and 6 can be made out of the first four classes. Similarly, a baryon can be made from two overlapping mesons. Hence all we really need is the first four classes.

To show that all the operators belonging to these four classes are available we use another of the most crucial and lovely ideas of cubology: that of conjugate elements. It turns out that all we need is one example in each class; given one example, we can construct all the other operators of its class from it. How does this work? The idea is very simple.

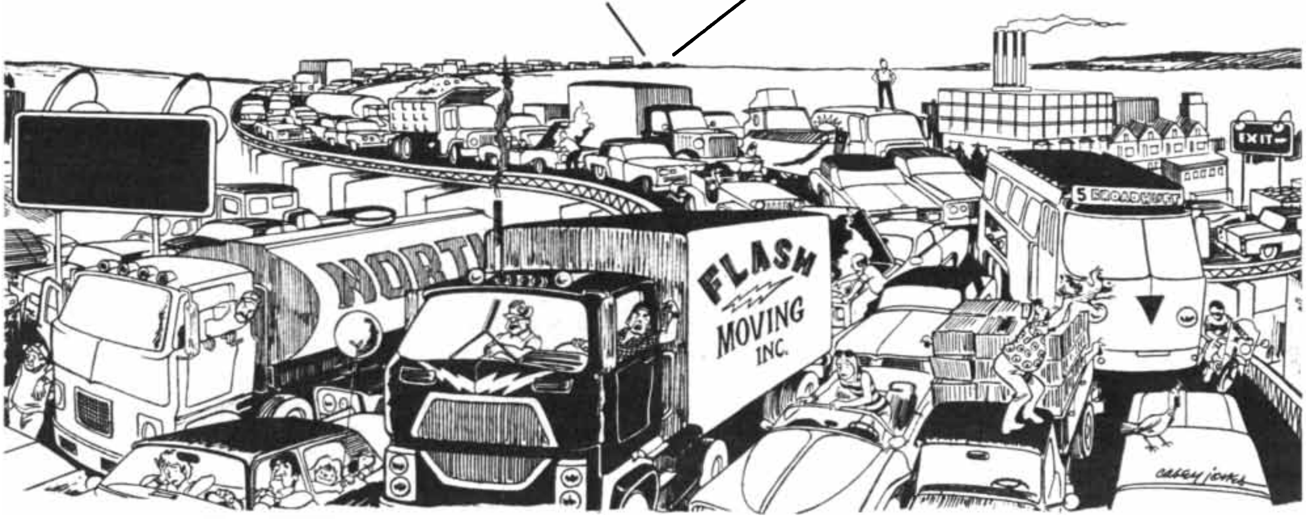
Suppose we had found one operator in class 1 that swapped, say, *uf* with *ub* and *ul* with *ur*, leaving the rest of the cube undisturbed [*see colored arrows in illustration on page 35*]. Let us call it *H*. Now suppose we wanted to swap two entirely different pairs of edge cubies, say *rf* with *df* and *rb* with *dr* [*see black arrows in same illustration*]. One daydreams: "If only those cubies were in the 'four magical swapping spots' on the top surface. . . ." Well, why not just put them up there? It would be fairly simple to get four cubies into four specific cubicles. The obvious objection is: "Yes, but that would have an awful side effect—it would totally mess up the rest of the cube." There is, however, a clever retort. Let the destructive maneuver that gets those four cubies into the magical swapping spots be called *A*. Suppose we were smart enough to transcribe the move sequence of *A*. Then right after performing *A* we perform our double swap *H*. Now comes the clever part. Reading our transcript in reverse order and inverting each *q* turn, we perform the exact inverse of *A*. This will not only unmaneuver the four cubies back into their old cubicles but also undo the side effects *A* created. Does that restore the cube intact? Not quite. Remarkably, since we sandwiched *H* between *A* and *A'*, the four edge cubies go home permuted, that is, each one winds up in the home of its swapping partner! Otherwise the cube is restored, and so we have accomplished precisely the double swap we set out to accomplish.

When you think this through, you see that it is flawless in conception. The inverse maneuver, *A'*, does not "know" we have exchanged two pairs of edges. As far as it is concerned, it is merely putting everything back where it was before *A* was executed. Hence we have "snuck" our swaps in under *A'*'s nose, which is to say we have "fooled the cube." Symbolically, we have carried out the sequence of moves *AHA'*, which is called a "conjugate" of *H*.

It is this kind of marvelously concrete illustration of an abstract notion of group theory that makes the Magic Cube one of the most amazing things ever invented for teaching mathematical ideas. Normally in group-theory courses the examples of conjugate elements are either too trivial or too ab-

“By the time we get out of this traffic, my 6-month auto policy will be up for renewal. And they may raise my rate, too!”

“6 months? Next time, call GEICO and get a 12-month policy.”



Unlike most companies, GEICO still offers the 12-month auto policy...a full year's coverage at a low rate. With a choice of convenient payment plans to fit your budget. Up to 9 payments if you wish.

And this is possible because GEICO tries to insure only good drivers. Good drivers are the best inflation-fighters around. You have fewer accidents. You cost us less. So you are one reason we can afford to charge you less.

"Good" doesn't mean "perfect." Even good drivers may occasionally have an accident. That doesn't necessarily bar you from GEICO...or from being renewed once you're a GEICO policyholder.

If you call, you'll talk directly to a GEICO sales insurance counselor, not a salesman. So there's no pressure. And no extra sales expense to inflate the cost.

GEICO has been saving money for good drivers since 1936. Today, nearly 1,500,000 people are satisfied GEICO policyholders. And that makes GEICO the 5th largest of the more than 1,000 stock companies insuring autos in the U.S.

GEICO has fast and efficient claim service with toll-free and collect-call telephone numbers throughout the U.S. So if you have to report an accident or loss, you can call our claim adjusters immediately. We also have over 3,500 claim representatives countrywide to assist in handling your claim if necessary.

NOTE: While you're checking GEICO, why not find out about our rates for homeowner/renter and boatowner coverage. And—even if you do not meet GEICO's auto policy standards—you may still obtain the same quality auto coverage and service from our affiliate, Criterion Insurance Company, at somewhat higher rates. Criterion operates in the District of Columbia and every state except Massachusetts, New Jersey and South Carolina.

GEICO

Where good drivers get the low rates they deserve.

Call toll-free **800-841-3000** (8 a.m.-9 a.m. East Coast Time)

In Georgia call toll-free 800-342-2800 or mail coupon. No obligation. No salesman will call.

Send to: GEICO & Affiliates, Consumer Information Services, GEICO Center, Macon, GA 31295

Name ______

Address ______ Apt. # ______

City ______ State ______ Zip ______

Employer ______ # Yrs. ______

Govt. or Military Grade ______ Have you previously been insured with GEICO? ☐ Yes ☐ No

Location of car if different from address above:

Car # 1: City ______ State ______

Car # 2: City ______ State ______

Car # 3: City ______ State ______

Name of current insurance co. ______ Month/Year Current Policy Expires ______

List all drivers	M or F	Relation	Age	Marital Status	Occupation	Yrs. lic.	Driver Training Yes	Driver Training No	% of Use of Cars #1	#2	#3	Accidents in Past 5 Years* Yes	No	Violations in Past 3 Years* Yes	No	License Suspension* Yes	No
		SELF															

If student away at school, how many miles to school? ______ miles.

	CAR 1	CAR 2	CAR 3
Year & Make ▶			
Model (Nova, Granada, etc.) ▶			
Body type: 2/4 dr. sdn, hardtop, sta. wag ▶			
No. of cylinders ▶			
Est. Annual Mileage ▶			
Days per week driven to work or parking area ▶			
One way driving distance ▶			
Is car used in business (except to/from job)? ▶			

If "yes" explain ______

*For accidents, violations, or license suspension, give dates and complete details, including cost of damages, on a separate sheet.

Check for Information on ☐ Homeowners/Renter's Insurance ☐ Boatowner's Insurance

668

GEICO and Affiliates. Corporate Headquarters: Washington, D.C. 20076

Not available in New Jersey or Massachusetts.

stract to be enlightening or exciting. The Magic Cube provides a vivid illustration of conjugate elements and of many other important concepts of group theory.

Suppose you want to get a quark-antiquark pair on opposite corners but know how to do so only on adjacent corners. How do you do it? Here is a hint: There are two nice solutions, but the shorter and prettier one involves using a conjugate. Incidentally, any maneuver that creates a quark on one corner (with other effects, of course) might be called a quarkscrew.

What we have shown for edges goes also for corners: the ability to swap two *specific* corners enables you to swap *any* two corners. Conjugation allows you to build up an entire class of operators from any single member of that class. Of course, the question still remains: How do you find some sample operator in each of the four classes? For example, how do you find an operator that creates a meson on two adjacent corners (a combination of a quarkscrew and an antiquarkscrew)? How do you find an operator that exchanges two edge pairs both of which are on the top surface? I shall not give the answer here but shall follow Singmaster, who points the way by suggesting quasi-systematic exploration of some small "subuniverses" within the totality of all cube states, that is, he suggests you look at *subgroups*. This means limiting your set of moves deliberately to some special types of move. Here are five examples of interesting subgroups created by various kinds of restriction:

1. The Slice Group. In this subgroup every turn of one face must be accompanied by the parallel move on the opposing face. Thus R must be accompanied by L', U by D' and F by B'. The name comes from the fact that any such double move is equivalent to rotating one of the three central slices of the cube. Singmaster abbreviates the slice move RL' by R_s, $R'L$ by R_s' and so forth. With this restriction faces cannot get arbitrarily scrambled. Each face will have a pattern in which all four corners share one color. A special case is the pattern called Dots, wherein each face is all one color except for its center [*see illustration on page 21*]. Can you figure out how to achieve Dots from Start? How many different ways are there of arranging the dots? How does the Dots pattern resemble a meson? (Note: The reader will find answers to all these questions, along with much else, in Singmaster's book. Please do not send me answers. Send novel ideas to Singmaster.)

2. The Slice-squared Group. Here we restrict the Slice Group further, allowing only squares of slice moves, such as R_s^2 (which is the same as R^2L^2) or F_s^2 (which is the same as F^2B^2).

3. The Antislice Group. Here instead of rotating opposing faces always in parallel we rotate them always in antiparallel, so that R is accompanied by L, F by B and U by D. An antislice move has a subscript a, as in R_a, which equals RL. (Of course, the Antislice-squared Group is no different from the Slice-squared Group.)

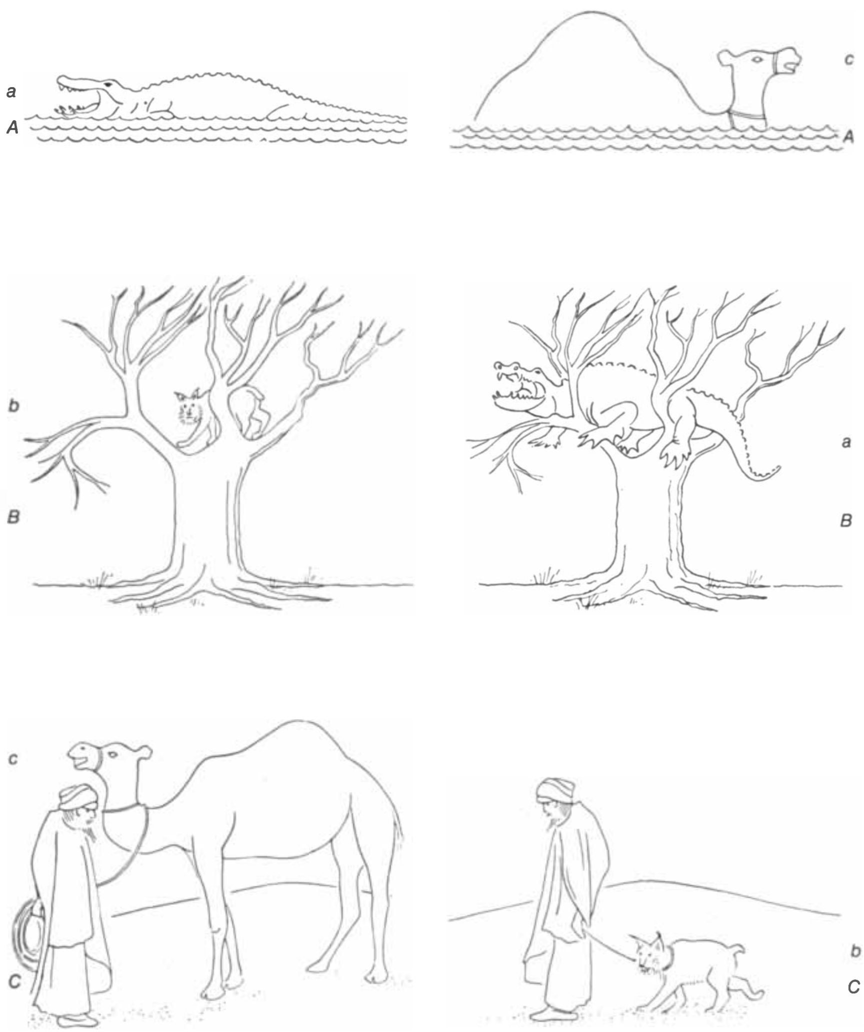
4. The Two-Faces Group. Allow yourself to rotate only two faces, say F and R.

5. The Two-Squares Group. As in the Two-Faces Group, you may rotate only two faces, using only 180-degree turns at that. This is a very simple subgroup.

If you limit your attention to just the Two-Faces and Two-Squares groups, you will be able to find processes that achieve double swaps, some of edges and others of corners. It is a remarkable fact that these processes alone, together with the notion of conjugation, will allow us—in a theoretical sense—to solve the entire unscrambling puzzle.

Why do we not also need a meson maker and a double edge flipper? Well, consider how we might make a double edge flipper from the two classes of tools one may assume will be found. In order to flip two edges without creating any side effects we shall perform two successive double edge-pair swaps, and both times they will involve the same pairs! For example, we might swap *uf* with *ub* and *df* with *db* and then reswap them. This seems to be an absolute "nothing process," but that need not be the case. After all, just as before, we can sandwich the second swap between a process X and its inverse X', where we carefully choose the process X so as to... (Oh dear, I totally lost my train of thought there. I am sure you can finish it up, though. I do remember that it was not too tricky, and that I thought the idea was elegant. I am sure you will too.)

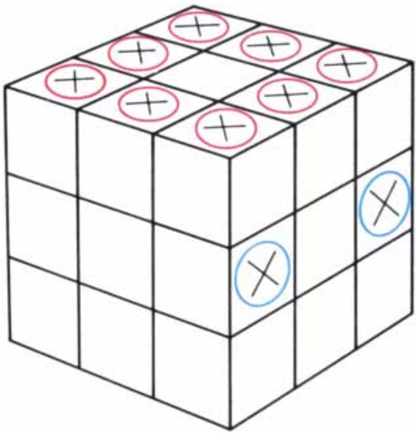
The same kind of thinking will show how you can build up a meson maker out of mere corner-swapping processes and conjugation. Given mesons you can build up baryons. And with mesons and baryons, double edge flippers and double edge-pair swappers and double corner-pair swappers you have a full set of



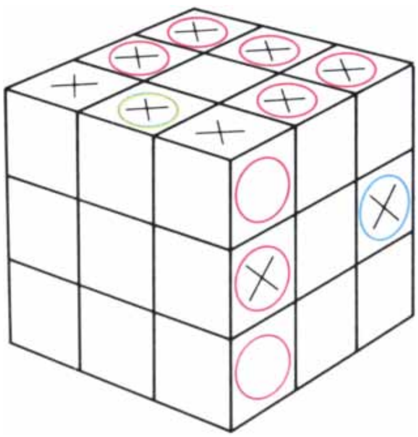
An alligator, a bobcat and a camel (a, b, c) *are permuted in their ecological niches* (A, B, C)

tools with which to restore any scrambled cube to Start, as long as it belongs to the same orbit as Start. This is, needless to say, a highly theoretical existence proof, and any practical set of routines would be organized quite differently. The type of solution we have described has the advantage of being compact in description, but it is enormously inefficient. In practice a cube solver must develop a fairly large and versatile set of routines that are short, easy to memorize and highly redundant. There is an advantage to being able to carry out required transformations in a variety of ways: one can choose whichever tool seems best adapted to the situation at hand, instead of, for instance, using some theoretical tool that can make a baryon in several hundred q turns.

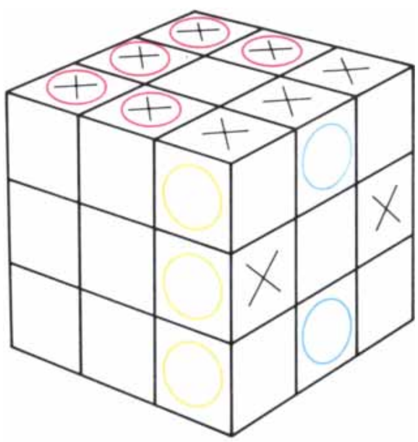
The typical solver evolves a set of transforms partly by intuition, partly by luck, sometimes with the aid of diagrams and occasionally with abstract principles of group theory. One principle nearly everyone formulates quite early is that of "getting things out of the way." This is once again the idea of conjugates, only in a simpler guise. The typical patter that goes along with it is something like this (I have included sound effects of a sort): "Let's see, I'll swing *this* out of the way [flip, flip] so that I can move *that* [flap, flap, flap], and now I can swing *this* back again [unflip, unflip]. There—now I've got *that* where I wanted it to be." You can hear the conjugate structure inside the patter.



The cube is in the Start position. The chief facelets of the cubicles are shown by the black *X*'s and the chief colors of cubies by the colored *O*'s. (The centers can be ignored, since they are stationary.) Think of the *X*'s as floating in space and the *O*'s as being attached to the cube, so that when the faces turn, the *X*'s stay where they are but the *O*'s move.



The q turn *F* has been executed. The two empty colored circles indicate that those two facelets have lost their "sanity." For them to be returned to sanity one would have to be twisted one-third clockwise and the other one-third counterclockwise. The same thing has happened on the invisible left-hand face.



The q turn *R* has been executed from the Start position. Empty circles occur in pairs. The top and bottom corner yellows have canceling twists. The blue edges from a pair, and the yellow edge is paired with a flipped edge on the invisible back face.

Proof that flippancy is even and twist is integral

The only problem with being conscious of why it all works as you carry it out is that it may be too taxing. My impression is that most cubemeisters do not think in such detail about how their tools are achieving their goals, at least not while they are in the midst of restoring some scrambled cube. Rather, expert cube solvers are like piano virtuosos who have memorized difficult pieces. As one M.I.T. cubemeister said to me, "*I* have forgotten how to solve the cube, but my *fingers* remember."

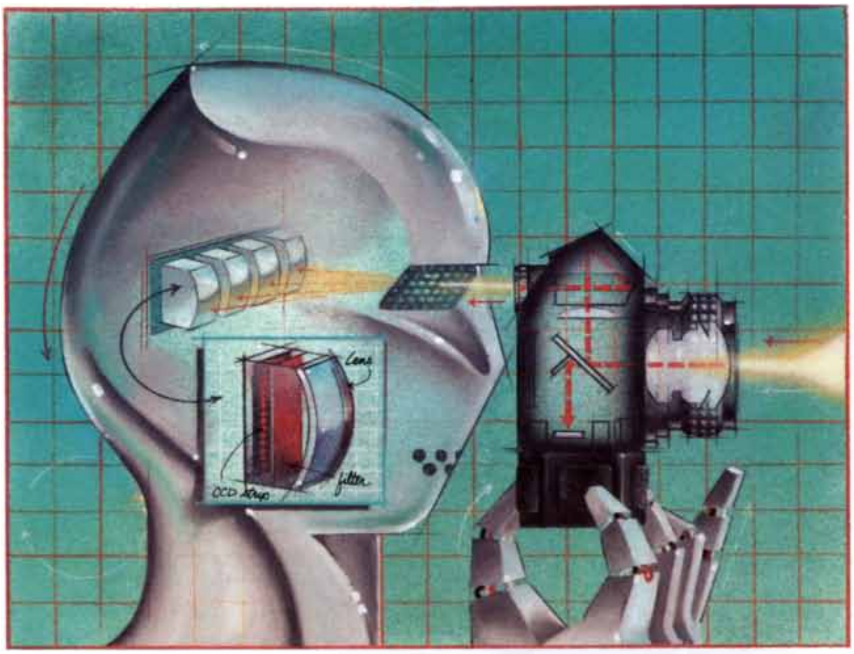
The average operator seems to be about 10 to 20 q turns long. You do not ever want to get lost in mid-operator, because if you do, you will have a totally scrambled cube on your hands, even if you were carrying out your final transform. As cubemeister Bernie Greenberg, who with Dave Christman is responsible for the beautiful computer graphics of the cube on the cover of this issue, said to me once, "If I were solving a cube and somebody yelled 'Fire!' I would finish my transform before clearing out."

My own style is probably overly blind. Not only do I not think about why my operators work as I am carrying them out; I must also admit that with some of them I have not the foggiest idea why they work at all. I found these "magic operators" through a long and arduous trial-and-error procedure. I used some heuristic approaches: "Explore various powers of simple sequences," "Use conjugates a lot" and so on. One thing I hardly used at all—alas, poor Rubik—was three-dimensional visualization. I do, however, know one Stanford cubemeister, Jim McDonald, who can give the reason for every q turn he makes. His operators do not seem magic to him because he can see what they are doing at every moment along the way. In fact, he does not have them memorized as I do mine; he seems to reconstruct them as he unscrambles cubes, relying on his "cube sense." He is like an expert musician who can improvise where a novice must memorize. For interested readers the central idea of Jim's method is first to solve the top layer, minus one corner, and then to let a vertical column of three cubicles act in the manner of a driveway one uses as an aid in turning a car around. The other two layers are cleaned up by shunting cubies in and out of the "driveway."

Perhaps not coincidentally, the abstract approach has been carried to its extreme by Singmaster's officemate Morwen B. Thistlethwaite (I am not jok-


IMPORTED BY CALVERT DIST. CO. N.Y.C.
PRODUCED IN SCOTLAND
IMPORTED
PASSPORT SCOTCH
BLENDED SCOTCH WHISKY
Because you enjoy going first class.
In Rome or at home, life's more satisfying when you're enjoying the best. That's Passport.
Enjoyed worldwide because it's made of Scotland's finest whiskies. Ask for Passport—go first class.
Passport Scotch.

OUR RESEARCH IN CCDs HAS GIVEN US A VERY PRACTICAL VISION OF THE FUTURE.



THE INTEGRATED SENSORS IN OUR NEW ELECTRONIC AUTOFOCUS SYSTEM FOR SLR CAMERAS COULD ALSO TURN ROBOTS INTO VISUALLY SMART MACHINES.

In 1980, Honeywell will include more than 80 million of its own integrated, solid-state sensors into the products it produces. Given that volume, it is not surprising to find an ongoing program in each of Honeywell's four Minneapolis research centers—programs which focus on the latest developments in silicon-based technologies.

Charged-Coupled Devices (CCDs) represent one of the areas in advanced microelectronics where Honeywell research teams have had considerable success. And where there appears to be the promise of even bigger breakthroughs in the years to come.

Our Solid State Electronics Center's work in monolithic CCDs has resulted in a new electronic autofocus system for SLR cameras (U.S. patent No. 4,185,191 is held by Honeywell's Visitronic Venture Group). Called TCL (for Through Camera Lens), this system has utilized the high speed optical scanning and signal processing capabilities of CCDs to create the first electronic focusing system which is able to compensate for interchangeable lenses.

We use 48 sets of CCDs. One CCD element in each set "sees" from one-half the camera lens. The other element "sees" from the other lens half. When the light to each element is the same, the image is in focus.

At our Technology Strategy Center, which concentrates on industrial and commercial research, another group is looking at the possibility of using CCDs as a way to provide robots with optical intelligence. An array of CCDs, no larger than a postage stamp, could contain enough sensors and circuits to give a robot capabilities in optical sensing, pattern recognition, and distributed control processing.

CCD technology is also at the leading edge of infrared imaging systems. Our Systems & Research Center is currently working to use CCDs to process the electronic readout in low-cost, high-resolution infrared staring focal planes for detection and recognition systems designed for aerospace and defense applications.

Supporting these efforts is the basic materials research being done at our Corporate Technology Center. A research team there is developing the technology needed to grow large crystals of mercury cadmium telluride. These infrared sensitive crystals will then be bonded to a silicon CCD chip, creating a hybrid CCD with integral sensing and processing capabilities.

If you are interested in learning more about the research work on CCDs being done in any one of these four research centers, or if you have an advanced degree and are interested in a career in solid-state electronics, sensors, or material sciences, please write to Dr. W. T. Sackett, Vice President, Honeywell Corporate Technology Center, 10701 Lyndale Avenue South, Minneapolis, MN 55420.

Honeywell

Although Honeywell engineering is world-wide, the bulk of research and applied research is done in Minneapolis. The most recent Quality of Life Study conducted by Midwest Research Institute shows Minnesota to be one of the best places in the country to live and work considering cultural, social, economic, educational and political factors.

This ideal research environment is further enhanced by Honeywell's affiliations with universities across the country. We have an ongoing program of seminars with Berkeley, MIT, Stanford, Carnegie Mellon, the University of Illinois, Cornell, Waterloo of Ontario and the University of Minnesota.

ing), who currently holds the world record for the shortest unscrambling algorithm. It requires at most 52 "turns." (A turn is either a *q* turn or a half turn, that is, a 180-degree turn of one face.) Thistlethwaite has used ideas of group theory to guide a computer search for special kinds of transforms. His algorithm has the property of not seeming to converge toward the solved state at all—until the last few turns.

This must be contrasted with the more conventional style. Most algorithms begin by getting one layer, say the top layer, entirely correct. (In saying "top layer" rather than "top surface" I mean that the "fringe" has to be right, too: the cubies on top must be correct as seen from the side as well as from above.) This represents the first in a series of "plateau states." Although further progress requires any plateau state's destruction, that state will later be restored, and each time this happens more order will have been introduced. These are the successive plateau states.

After getting the top layer the solver typically works on corners on the bottom layer, or perhaps on getting the horizontal equator slice all fixed up. Most algorithms can, in fact, be broken up into about five natural stages, corresponding to natural classes of cubies that get returned to their home cubicles. My personal algorithm, for instance, goes through the following five stages: top edges, top corners, bottom corners, equator edges and bottom edges. In the first two of my stages placement and orientation are achieved simultaneously. Each of the last three stages breaks up into a placement phase and then an orientation phase. Naturally the operators of any stage must respect all the accomplishments of preceding stages. This means they can damage the order built up as long as they then repair it. They are welcome, however, to indiscriminately jumble up cubies scheduled to be dealt with in later stages. I find that other people's algorithms are usually based on the same classes of cubies, but the order of the stages can be wildly different.

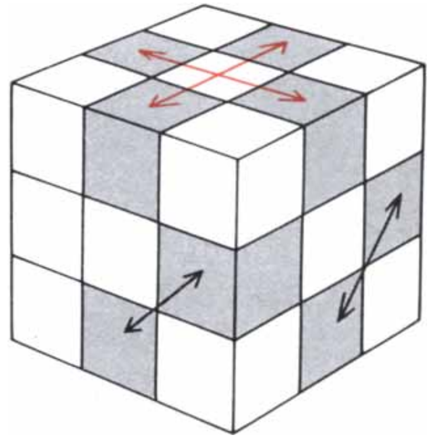
Virtually all algorithms have the property that if one were to take a series of snapshots of a cube at the plateau states, one would see entire groups of cubies falling into place in patterns. This is called monotonicity at the operator level, that is, a steady, visible approach toward Start, with no backtracking. Of course, there is backtracking at the *turn* level, but that is another matter.

Very different in spirit is Thistlethwaite's algorithm. Instead of trying to put particular classes of cubies into their cubicles he makes a "descent through nested subgroups." This means that starting with total freedom he makes a few moves, then clamps down on the types of move that will thenceforward be allowed, makes a few more moves, clamps down a bit more and so on until the constraints become so heavy that nothing can move any more. Just at this point, however, the Start position has been achieved! Each time, the clamping down amounts to forbidding *q* turns on two opposite faces, allowing only half turns thenceforward in their stead. The first faces to be thus "clamped" are *U* and *D*, then come *F* and *B* and finally *L* and *R*. The strange thing about this approach is that you cannot see Start getting nearer, even if you take a series of snapshots at carefully chosen moments. All of a sudden there it is.

This raises a natural question. Is there any easy way to tell how far you are from Start? It could obviously be quite useful. For example, it is rather embarrassing if one has to resort to the full power of a general unscrambling algorithm to undo what some friend has done with four or five casual twists. For that reason alone it would be nice to be able to assess quickly if some state is "really random" or is close to Start. But what does "close" mean? Distance between two states in this vast space can be measured in two fairly natural ways. You can count either the number of *q* turns or the number of turns needed to get from one state to the other (where "turn," as above, means either a *q* turn or a half turn). But how can one figure out how many turns are needed to get to Start without doing an exhaustive search? A reliable and at least fairly accurate estimate would be preferable, one that could be carried out quickly during a cursory inspection of the cube state. A naive suggestion is to count the number of cubies that are not in their home cubicle. This estimator, however, can be totally fooled by the Dots position, in which nearly all cubies are on the "wrong" side. That position is only eight *q* turns away from Start. Perhaps the flippancy and the number of quarks could also be taken into account by a better estimator, but I do not know of one.

There are sophisticated group-theoretical arguments estimating that the farthest one can get from Start is 22 or 23 turns. This is quite amazing, considering that most solvers' early algorithms take several hundred turns, and highly polished algorithms take a number somewhere in the 80's or 90's. Indeed, many mere operators take considerably more turns than Thistlethwaite's entire algorithm does. (My first double edge flipper was nearly 60 turns long.)

One thing we know, and it can be demonstrated easily, is that there exist states at least 17 turns away from Start. The argument goes as follows. At the outset there are 18 possible turns we might make: L, L', L^2, R, R', R^2 and so on. After that there are 15 reasonable turns to make. (One would not move the same face again.) The number of distinct turn sequences of length 2 is therefore 18×15, or 270. Another turn will add another factor of 15, and so on. How long does it take before we have reached the number of accessible states? It turns out that 17 turns will create more turn sequences than there are distinct states, and that 16 turns are too few. Now, not every turn sequence leads to a distinct state, not by a long shot, and so we have not shown that 17 turns will reach every accessible state. We have simply shown that *at least* 17 turns are needed if you want to reach every state from Start. Hence conceivably no two states are much more than 17 turns away from each other. But which 17 turns? That is the question.

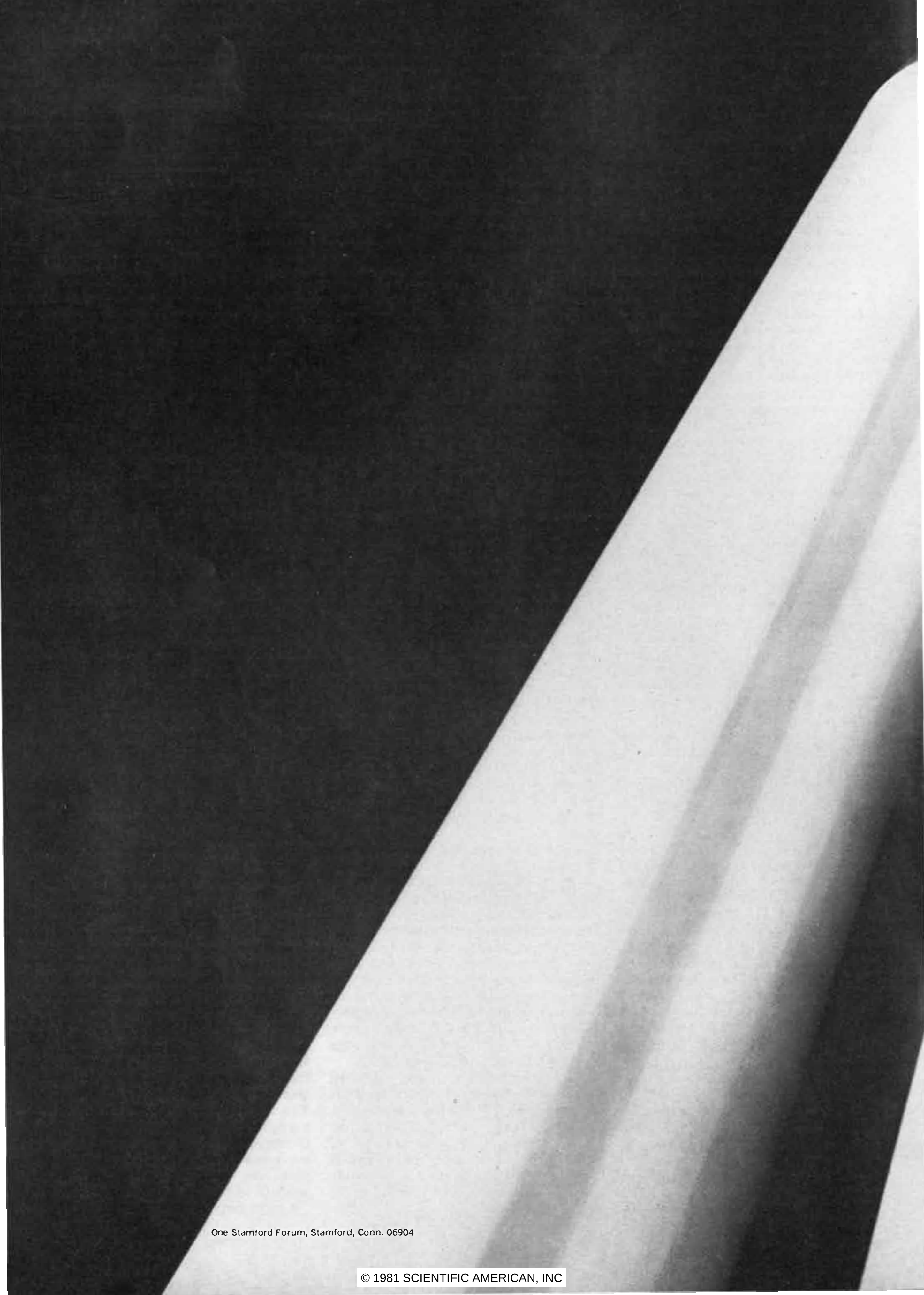


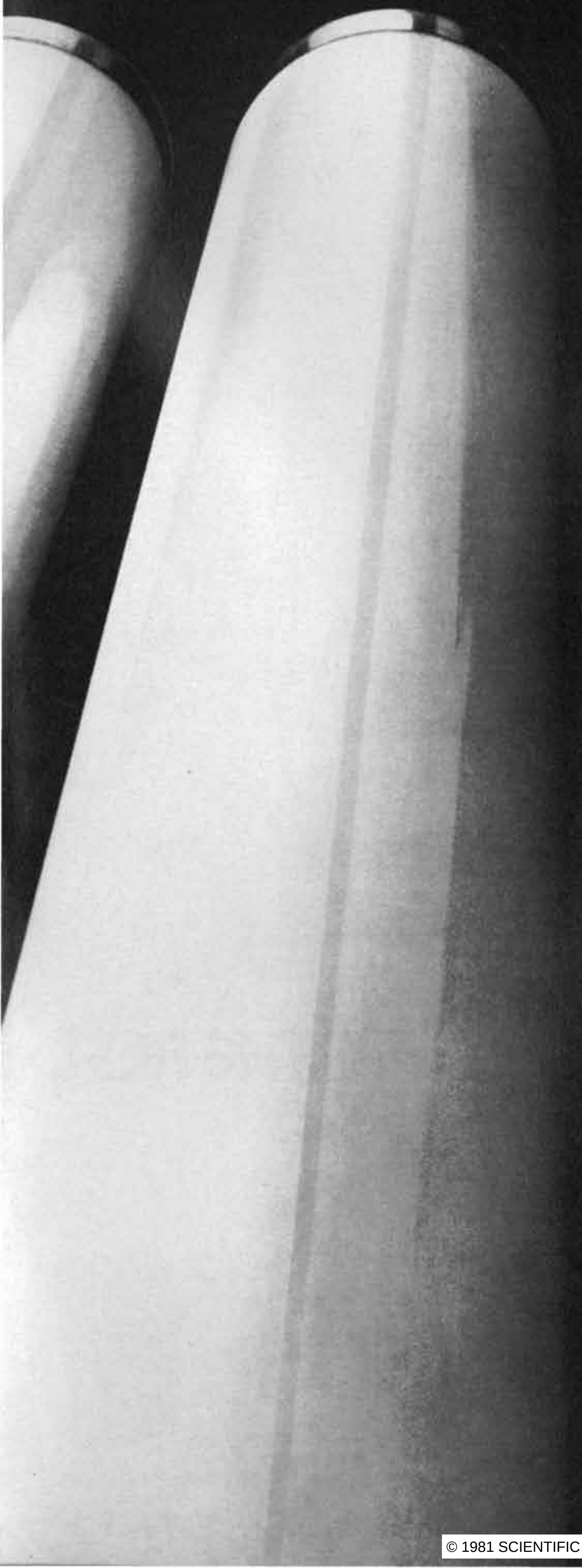
How to swap arbitrary edge pairs

So far only God knows how to get from one state of the Magic Cube to another in the fewest turns. "God's algorithm" is the one that would tell you how to do it. A burning question of cubology is: Is God's algorithm just a gigantic table without any pattern in it, or is there a significant amount of pattern to it, so that an elegant and short algorithm based on it could be mastered by a mere mortal?

If God were to enter a cube-solving contest, he might encounter some rather stiff competition from a few prodigious mortals, even if they do not know his algorithm. I am told there is a young Englishman from Nottingham named Nicholas Hammond who has got his average solving time down to close to 30 seconds! Such a phenomenal performance calls for several skills. The first is a deep understanding of the cube. The second is an extremely polished set of operators. The third is to have the operators down so cold that you could do them in your sleep. The fourth is sheer speed at executing twisty hand motions. The fifth is having a well-oiled "racing cube": one that turns at the merest twitch of a finger, eagerly anticipating every operator before it is needed. In short, the racing cube is a cube that *wants* to win.

I have not heard of people naming

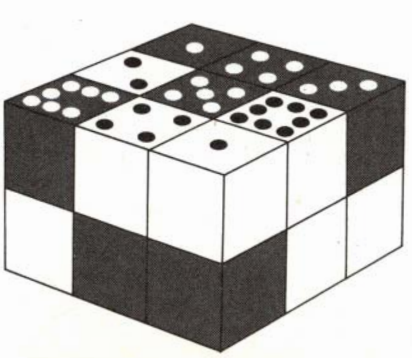




Shown here with optional CX-10 Cabinet and SX-10 Loudspeakers their racing cubes, although it is sure to come. It would seem, however, that there is an association between colorful names and major contributors to cubology. Apart from Singmaster and Thistlethwaite there is Dame Kathleen Ollerenshaw (recently Lord Mayor of Manchester), who has discovered many streamlined processes, has written an article on the Magic Cube and has the distinction of being the first to report an attack of cubist's thumb, a grave form of the disease mentioned at the beginning of this column. Then there is Oliver Pretzel, the discoverer of a delicious twisted 3-cycle and the creator of a lovely "pretty pattern" called the 6-*U* state, which can be reached from the start position by way of the word $L'R^2F'L'B'UBLFRU'RLR_sF_sU_sR_s$.

Pretty patterns are of interest to many cube lovers, but I cannot do them justice here. I shall mention only a few of the best. My favorite is the Worm, whose "genotype," or turn sequence, is $RUF^2D'R_sF_sD'F'R'F^2RU^2FR^2F'R'U'F'U^2FR$. Then there is the Snake, a similar sinuous pattern that winds around the cube: $BR_sD'R^2DR_s'B'R^2UB^2U'DR^2D'$. If you cut off the Snake's tail R^2D' and stick on instead $B^2R_aU^2R_a'B^2D'$, you will create a curious bi-ringed pattern. All of these are from pretty-pattern-meister Richard Walker.

A beautiful pattern is the Giant Meson, made from a giant quark (a 2 × 2 × 2 corner subcube rotated 120 degrees) and a giant antiquark. To top it off you can use quarkscrews to twist a standard-size quark and an antiquark onto the corners of the giant quark and antiquark, like cherries on top of sundaes [*see illustration on page 21*]. I shall let you figure this one out. A good warm-up exercise is to figure out how to make the Pons Asinorum (Bridge of Asses) state [*see same illustration*], so called because, as one M.I.T. cubemeister remarked to me, "If you can't hack this one, forget about cubing." Then there are two kinds of cross, known to the M.I.T. cube-hacking community as the Christman cross and the Plummer cross [*see same illustration*]. The Christman cross involves three pairs of colors (*U-D, F-R* and *L-B* in the illustration); the Plummer cross involves two triples in the quark-antiquark style.



The Magic Domino in a scrambled state

I should like to leave the reader with a set of hints and some things to think about. A difficult challenge, good for cubists at all levels of cubistry, is for someone to do a handful of turns on a pristine cube, for him to give it to you in this mildly scrambled state and for you to try to get it back to the Start position by finding the exact inverse word. Cubemeisters will be able to invert a bigger handful of turns than novices. Apparently Kate Fried can invert seven turns regularly, and once after a full day of staring at the cube she undid 10. (I can undo about four.)

My royal road to discovering an algorithm is based on two challenging exercises involving corner cubies only. The preliminary exercise is as follows. Maneuver the four corner cubies with white on them to the top face with their white facelets pointing upward. Do not worry about which cubie is in which cubicle. Simultaneously do the same thing on the bottom face (of course with *its* color pointing downward). The advanced exercise is to do the preceding one but in addition to make sure that all the corner cubies are in their proper cubicles. This amounts to solving the 2 × 2 × 2 Magic Cube puzzle, and it will take you a long way toward mastery of the Magic Cube.

To help you with your edge processes, here is a wonderful trick discovered by David Seal, based on a type of operator called a monoflip. I shall give it to you as a puzzle. How can you make a double edge flipper out of a process that messes up the lower two layers but leaves the top layer invariant except for flipping a single edge cubie? I shall give you a hint: The answer involves the important group-theoretical idea of a commutator, a word of the form $PQP'Q'$. I shall also leave it to you to find your own monoflip operator. After I found out about it I incorporated the trick into my method.

Here is a small riddle: Why do 5- and 7-cycles crop up so often in an object whose symmetries all have to do with numbers such as 3, 4, 6 and 8? Where do cycle lengths such as 5 and 7 come from? A somewhat related question is: What is the maximum order a word can have? The order of a word is the power you have to raise it to in order to get the identity. (For example, the order of *R* is 4.) You can show that the order of *RU*, for instance, is 105 by inspecting its cycle structure.

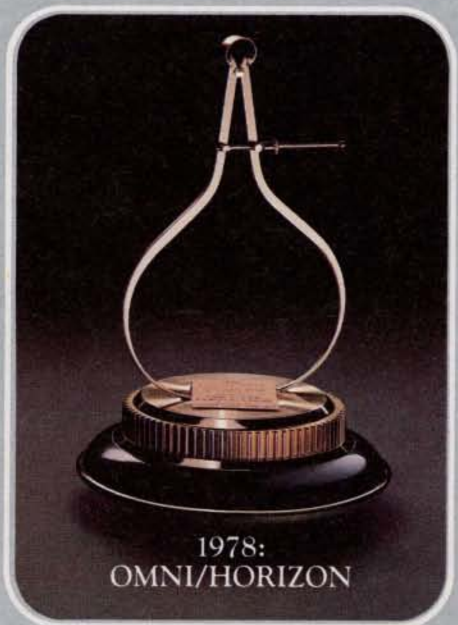
Where do we go from here? I must mention that I have only scratched the surface of cubology in this column. Rubik and others are working on generalizations of various types. There already is a Magic Domino, which is like two-thirds of a magic cube: two 3 × 3 layers. You can rotate it only by half turns about two of its three axes and by *q* turns about the third axis. In its Start position one face is black, the other is white and both faces have numbers from 1 through 9 in order. It resembles the 15 Puzzle even more strongly than the cube does. Various people have made 2 × 2 × 2 cubes, and such cubes may go on sale one day. You can make your own by gluing little three-cornered hats over each of the eight corners of a 3 × 3 × 3 cube. Readers will naturally wonder about such enticing possibilities as a 4 × 4 × 4 cube. Rest assured, it is being developed in the Netherlands, and it may be ready soon. Inevitably there is the question of both higher and lower dimensionalities. Cube theorists are beginning to discuss the properties of higher-dimensional cubes.

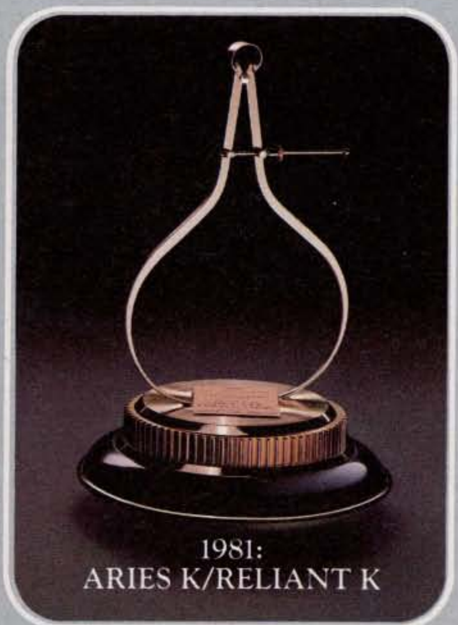
The potential of the 3 × 3 × 3 cube is not close to being exhausted. One rich area of unexplored terrain is that of alternate colorings. This idea was mentioned to me by various M.I.T. cube hackers. One can color the cubies in a variety of ways. Each new coloring presents a different kind of unscrambling problem. In one variant coloring, edge-cubie orientations become irrelevant and center-cubie orientations take on a vital importance. In another variant, corner-cubie orientations are irrelevant and centers matter. Then, moving toward simplicity, one can color two faces the same, thereby reducing the number of distinct colors by one. Or one can paint the faces with just three colors. An extreme would be to have three blue faces meet at one corner and three white ones meet at the corner diagonally opposite. The French government official I mentioned above says that on the cubes he saw, five faces had one color and the sixth face had another color!

Who knows where it will all end? As Bernie Greenberg has pointed out: "Cubism requires the would-be cubist to literally invent a science. Each solver must suggest areas of research to himself or herself, design experiments, find principles, build theories, reject them and so forth. It is the only puzzle that requires its solver to build a whole science." Could Rubik and Ishige have dreamed that their invention would lead to a model and a metaphor for all that is profound and beautiful in science? It is an amazing thing, this Magic Cube.

I should like to thank the many cubemeisters and cubists who have contributed directly or indirectly to my knowledge and enthusiasm. Among them are Clark Baker, Alan Bawden, Jim Boyce, Larry Breed, Charles Brenner, Bob Filman, Carl Hoffman, Scott Kim, Bill McKeeman, Jeannine Moseley, Richard Pavelle, Dave Plummer, David Policansky, David Singmaster, Ann Trail, Allen Wechsler, Dan Weise and John Woodcock. My warmest thanks, however, are reserved for Greenberg, who set up a cube-in for my benefit at M.I.T. and who infected me with his enthusiasm for the beauties of cubology.

Chrysler wins the Motor Trend Car of the Year award. Again.



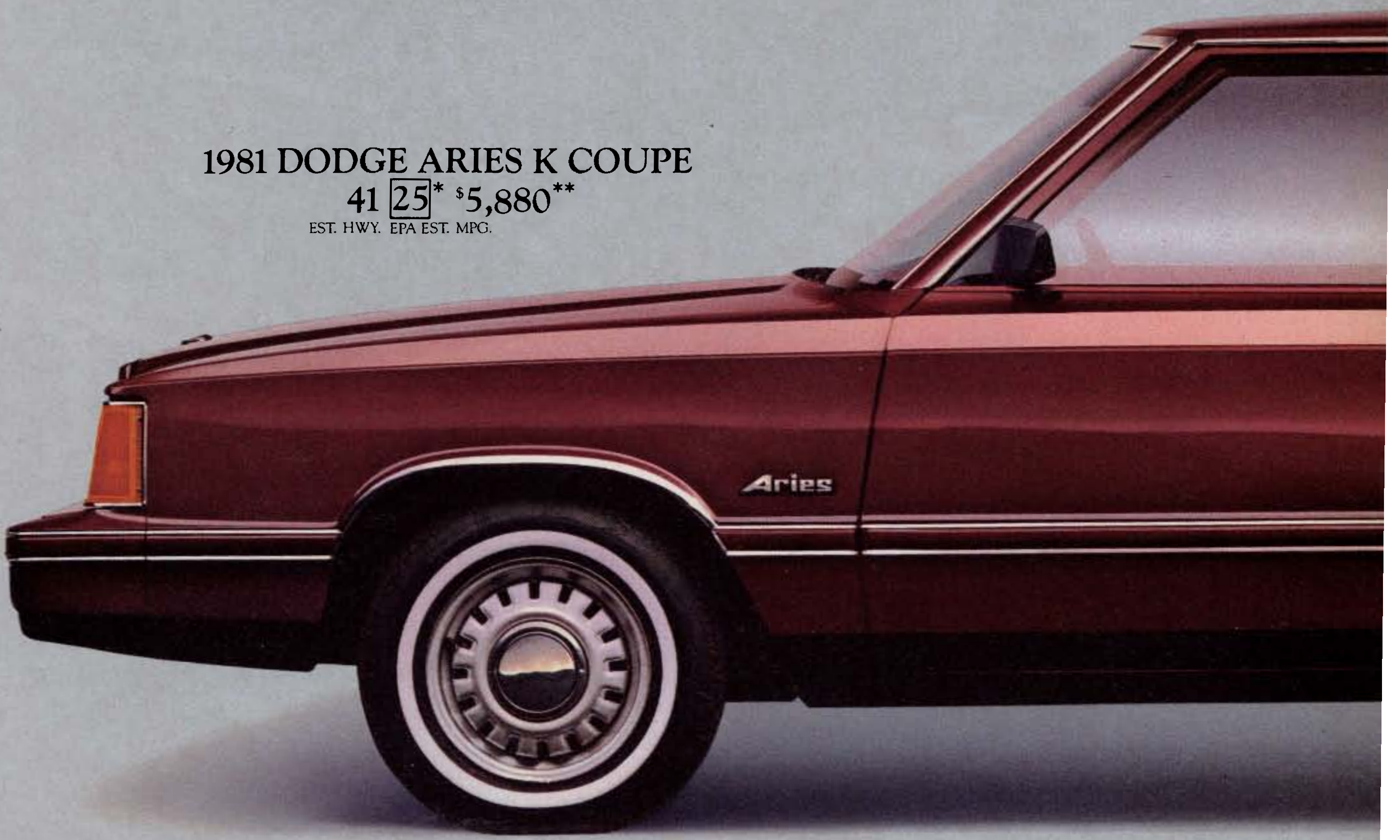


Chrysler's K cars win the 1981 Motor Trend Car of the Year award. It's the second time in four years Chrysler's front-wheel-drive cars have won a top automotive industry award—Omni and Horizon were chosen Car of the Year in 1978. Only Chrysler gives you four front-wheel-drive winners to choose from.

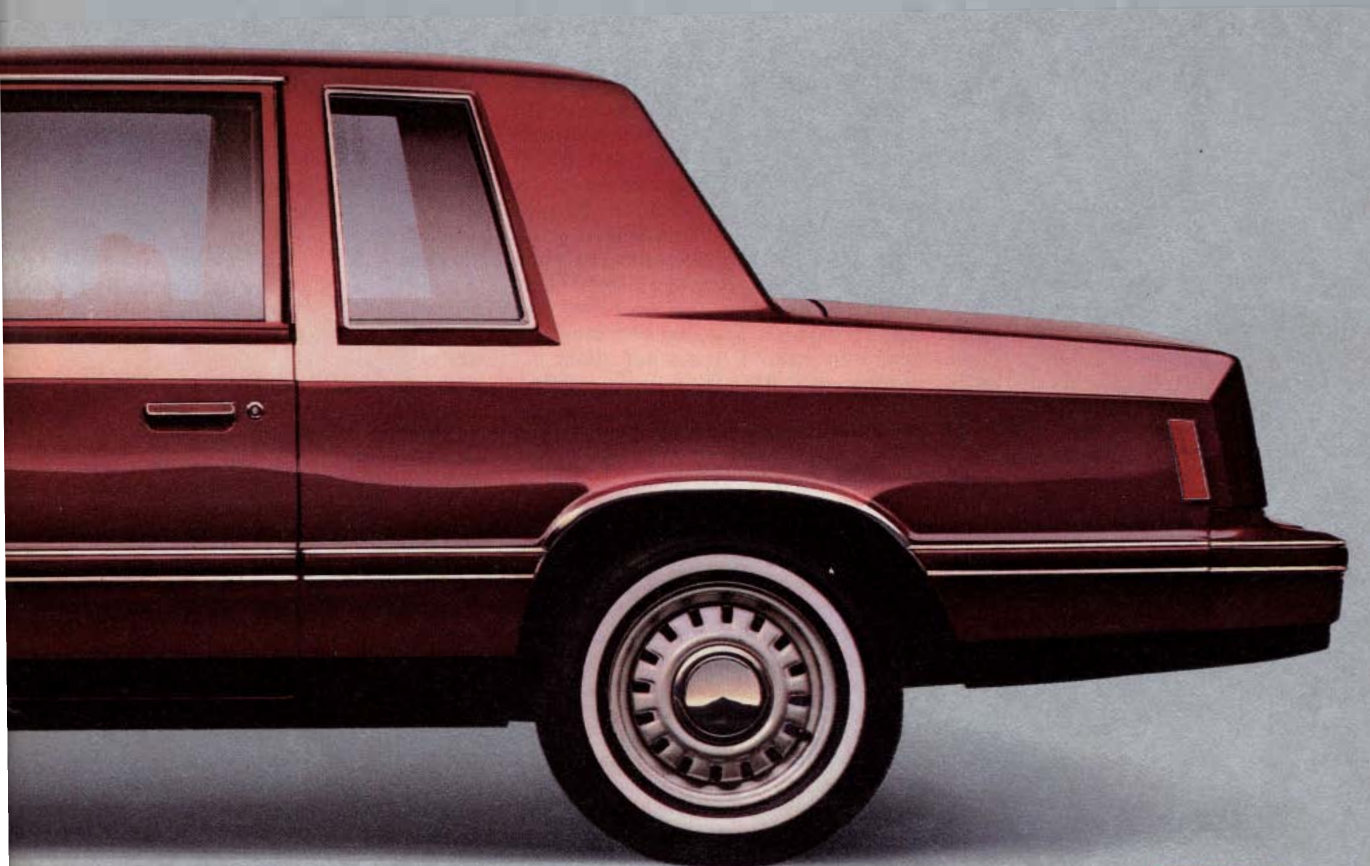
Chrysler's K cars beat Ford Escort, Ford Granad
after 50 tough Motor

This year, the Motor Trend Car of the Year finalists were rated on more than 50 counts of performance, quality, comfort, and design. A 75-mile torture course, with every driving condition from freeways to twisting mountain roads, was used to evaluate the quality of ride and drive. And a panel of

1981 DODGE ARIES K COUPE
41 | 25* | $5,880**
EST. HWY. EPA EST. MPG.



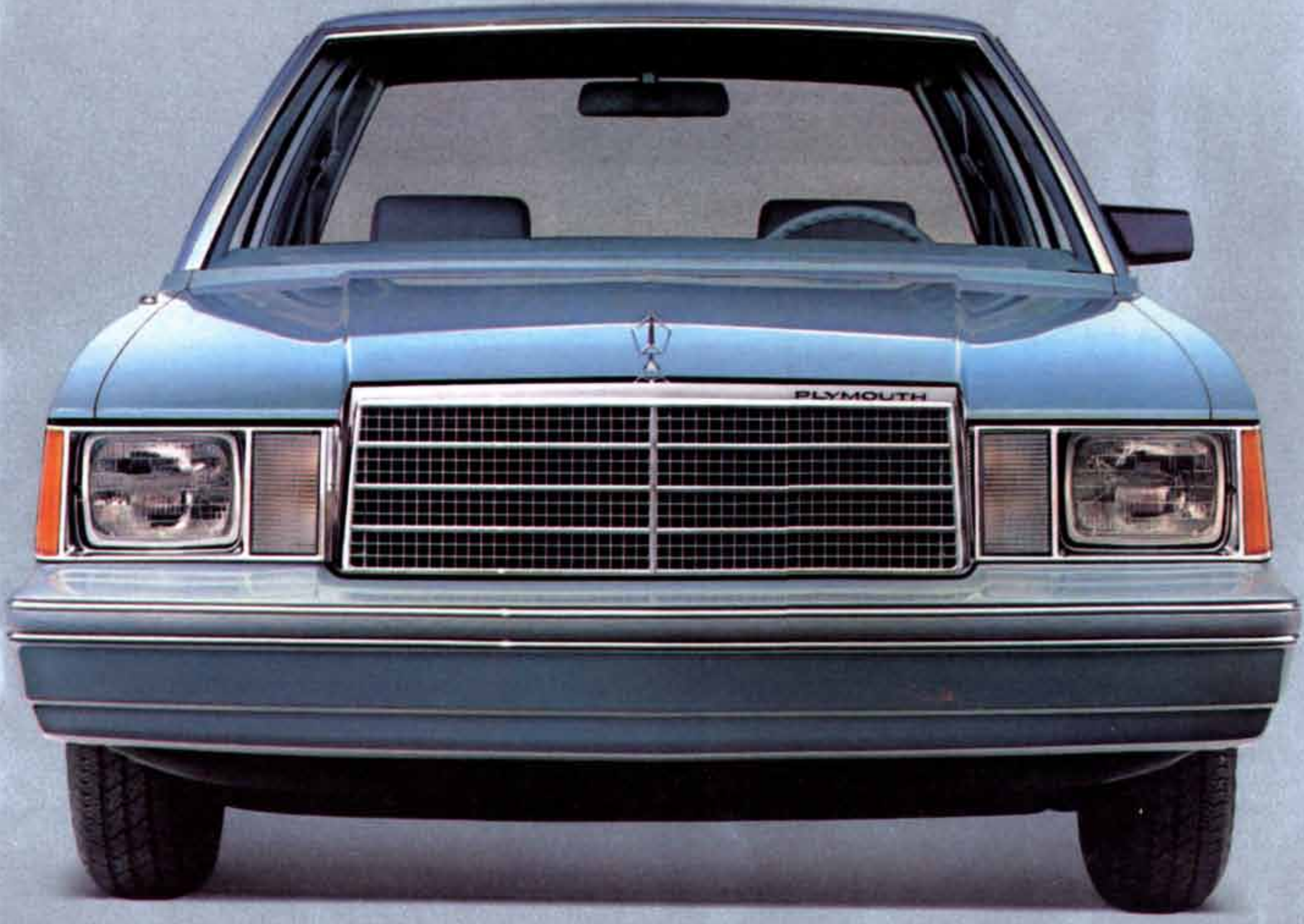
*Use EPA est. mpg numbers for comparison. Actual mileage may differ depending on speed, trip length and weather. Actual highway mileage probably lower. CA mileages lower.
Base sticker prices. Sticker prices differ in the states of CA, HI, NB, LA, NY, MA, WA, PA, MS. See your dealer for details. **Standard equipment levels vary among cars. Destination charges, title and taxes extra. WSW tires $51 extra on Relia



"In this part of the new decade, passenger comfort, ride and drive, styling and design...and most important, dollar value...are more relevant than all-out performance. In these classes, the K car simply outscored its competition; and in three cases out of four, the competition turned out to be another K car. This accomplishment alone was enough to make it Car of the Year."

The Editors, Motor Trend Magazine, February issue

1981 PLYMOUTH RELIANT K SEDAN
41 [25]* $5,980**
EST. HWY. EPA EST. MPG.



BOOKS

Lightning, whales, a reminiscence of modern physicists and the architecture of the Incas

by Philip Morrison

LIGHTNING AND ITS SPECTRUM: AN ATLAS OF PHOTOGRAPHS, by Leon E. Salanave. University of Arizona Press ($25). BALL LIGHTNING AND BEAD LIGHTNING: EXTREME FORMS OF ATMOSPHERIC ELECTRICITY, by James Dale Barry. Plenum Press ($29.50). Folk wisdom notwithstanding, lightning mostly strikes twice. First of all, down from the negatively charged thunderhead a faint leader steps its way, 50 meters at a time. Fading out at the end of each step, it brightens again as it essays the next segment of a halting, tortuous journey to ground. That first path, once complete (only in this phase does lightning send out dead-end branches inconclusively into the air), marks out the route for the swift, brilliant upward return. Again and again the sequence continues, slowly enough so that a hand-held camera, swept back and forth without hurry, the shutter open for a minute or two under storm clouds on a dark night, will beautifully record the multiple flash, from a handful up to a dozen or more strokes easily separated in the print.

A typical event lasts for at least half a second. Down the quick frames of high-speed photography that initial path wanders; the cascading pathfinder spends up to a millisecond on each step, and the strong current stream returns the visit at a good fraction of the speed of light, appropriate for ionization travel. Subsequent leaders, which can follow the first one down to initiate succeeding strokes, called dart leaders, move an order of magnitude faster than the first stepped leader in the freshly ionized gas, although they tend to slow down as they travel. If the next stroke is delayed more than a tenth of a second, there are no darts. The process reverts to the slower pace of the stepped beginning; the ionic channel has decayed, its ions recombined.

Succeeding return strokes can surge upward from different ground points, to rejoin the older channel just where the dart leader lost its swift way. The lightning has then forked; a fork is the sign not of a divided simultaneous current path but rather of two successive journeys, conflated by the eye but distinct to the moving camera. (About one in four flashes forks.) Some flashes are initiated by a positive leader that moves up from near ground level; such flashes are less branched, smoother, single rather than multiple, draining the cloud charge in one slow, strong stream instead of in a dozen swifter strokes. All these processes and much more are plainly revealed to the eye in *Lightning and Its Spectrum*. Its photographs are made with equipment ranging from sophisticated, rotating, dispersive instruments to the familiar hand-held 35-millimeter camera.

Time is only one dimension of the lightning phenomenon entered by these images. Indeed, the time scale of the multiple stroke is marginally perceptible to the unaided eye; the flicker seen in some flashes led 100 years ago to the first simple swinging-camera resolution of the lightning flash. One photograph shown here from New Mexico disclosed 26 strokes making up one swollen flash, almost two seconds in duration. The wind can sweep some glowing section of channel along to make the visibly broadened path called ribbon lightning. It too was first clarified a century ago.

Some pictures explore the three dimensions of space. One stereoscopic pair shows lightning vivid in depth, something we can hardly see because of the large scale of the phenomenon. Eyes are too close and head motion too small and slow to perceive lightning in depth without cameras spaced in the kilometer range. With them, however, we can see the long, twisted path. One wonderful cross view explains the tight reentrant figure drawn by one flash, a Chinese character coiling in the air. That arc did not in fact cross itself over and over again, like the brush stroke it resembles; it was a normally twisted, extended single path, by chance seen more or less end on. The crossfire camera view disentangles the calligraphy.

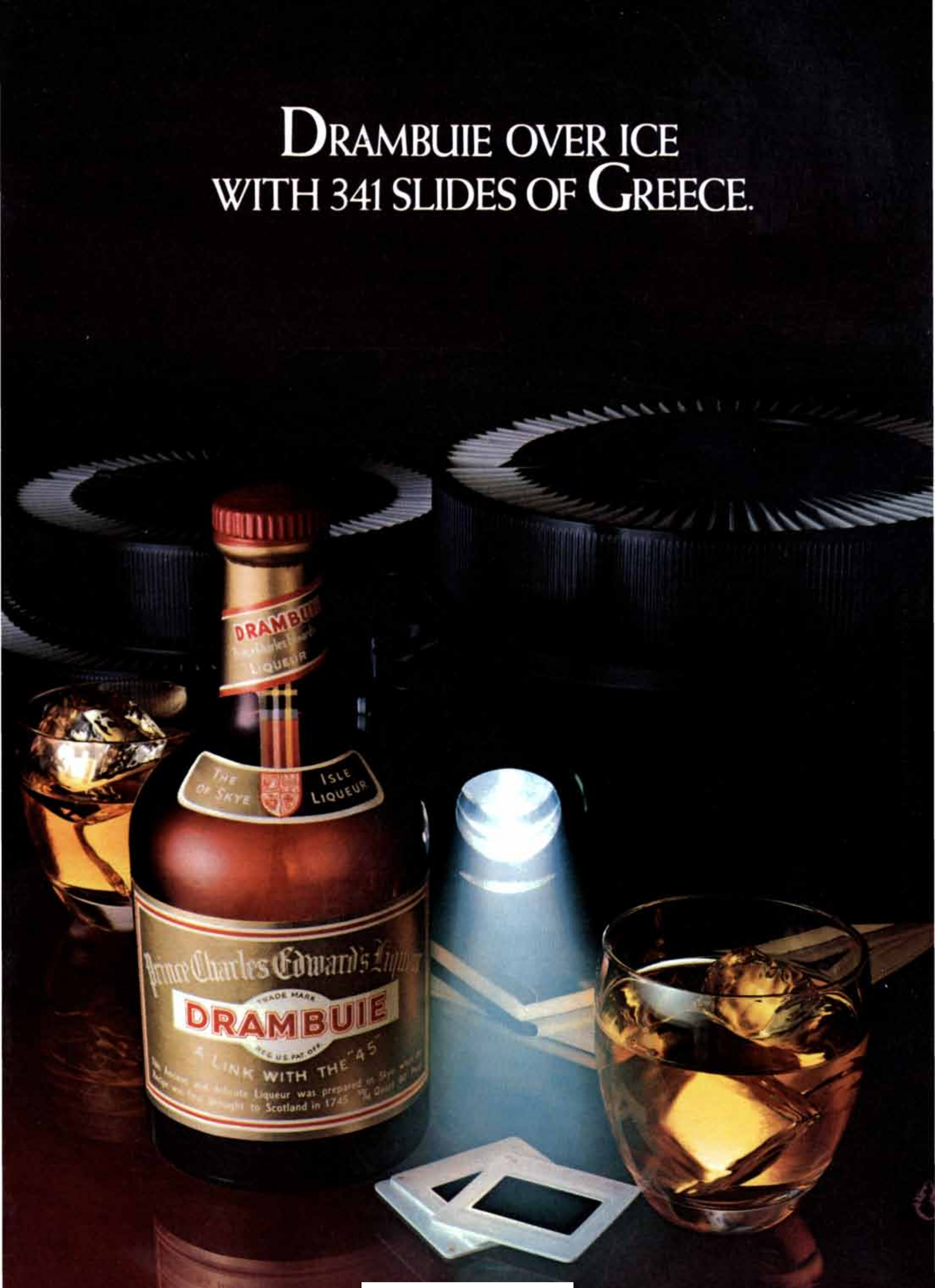
The subtlest fact recovered by photography is atomic: the composition of the glowing gas in the transient channel. Some dispersing device—a prism or a grating—simply diverts each color present in the glow into a distinct direction. Without a slit or a collimating lens the instrument yields a photograph on which the lightning is repeated in its own inimitable forms, zigzag by identical zigzag standing in uniformed procession, ranked in the colors of the rainbow. This, of course, is the slitless spectrogram of the astronomer, and such a photograph is the best possible introduction to the meaning of the spectrum. The spectral line is revealed as the optician's artifact it is.

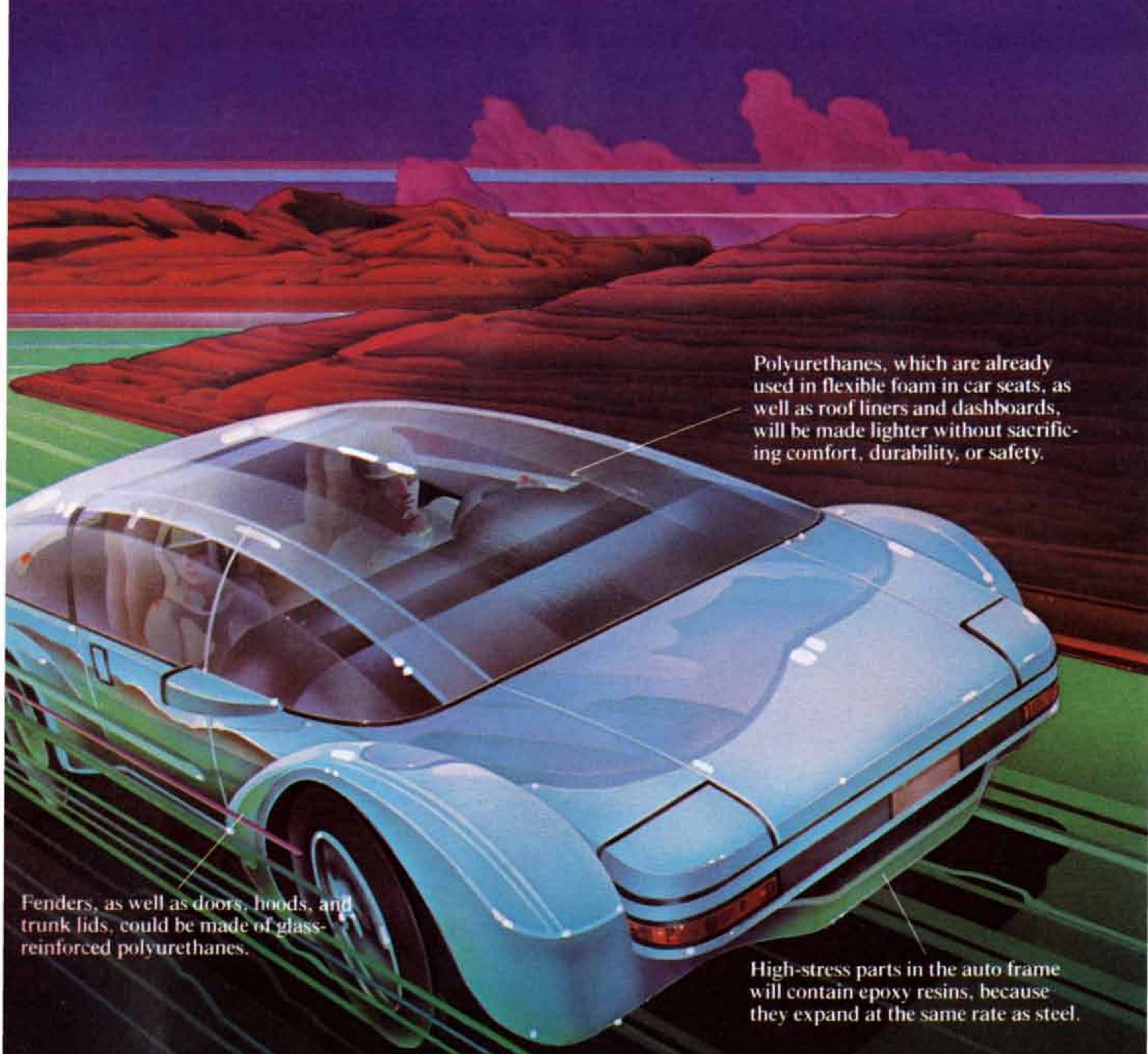
It is no surprise that the lightning flash is the luminous output of oxygen and nitrogen atoms and ions. They, of course, comprise the air. Atmospheric argon shows up with a strong infrared line, and some N_2 and O_2 molecular lines are conspicuous at both ends of the spectrum. There is a weak continuous spectrum, across it the molecular absorption bands of the cool intervening air. Even single strokes can be analyzed today, with the new gratings, fast lenses and films.

The peak temperature of flashes has been measured; for a few microseconds as the big return surge begins the temperature reaches about 30,000 degrees Kelvin. It is a little more unexpected to learn that the red end of the visible spectrum is dominated by that familiar deep red line of hydrogen called $H\alpha$. The air has no free hydrogen atoms to speak of, and even water is usually a weak constituent. In a thunderstorm matters are different: the storm world is watery, and the dissociation of H_2O to release atomic hydrogen is no trick at all for the energies in play. Lines of titanium, calcium and even copper show up above where the lightning strikes the southern Arizona desert.

Ball Lightning and Bead Lightning is entirely different in aim and nature. It is a comprehensive report on the strange phenomena of its title (although there is not much to be said about the bead case). Another such survey appeared a decade ago, by Stan Singer of Pasadena, with whom the present compiler, another Los Angeles physicist, is in close touch. The later book attempts a comprehensive literature list: more than 1,600 references are cited back to 1850 and some 50 before then. The earlier book, learned and rather brief, sought more to present the entire history of the topic, with a friendly exposition of every theory proposed. It was not critical.

James Dale Barry, the author of this volume, explicitly sifts every available photograph published, setting out stiff criteria for accepting them as evidence. A good many, of course, are chance finds in photographs made for other purposes, some unattended camera set with an open lens. Most of the pictures seem to be the result of streetlights, unexpected camera motion, unusual examples of ordinary lightning and so on. Of two dozen published photographs of natural ball lightning Barry admits only three as being reasonably representative. Not one can be regarded as definite. The better part of the evidence is anecdotal, but it is often crisp. Above all one forms the impression of a phenomenon


DRAMBUIE OVER ICE
WITH 341 SLIDES OF GREECE.
THE ISLE OF SKYE LIQUEUR
TRADE MARK
DRAMBUIE
A LINK WITH THE "45"


Polyurethanes, which are already used in flexible foam in car seats, as well as roof liners and dashboards, will be made lighter without sacrificing comfort, durability, or safety.
Fenders, as well as doors, hoods, and trunk lids, could be made of glass-reinforced polyurethanes.
High-stress parts in the auto frame will contain epoxy resins, because they expand at the same rate as steel.

of modest scale, intimate rather than daunting, perhaps a little dangerous but not on the explosive scale of even a hand grenade.

Ball lightning is a luminous sphere about the size of a toy balloon, drifting or hovering nearby for a second or two, with a marked odor, glowing usually with a red or whitish yellow light but of no intolerable brightness. There is a famous engraving showing the death of Dr. Georg Richmann in Saint Petersburg, when in 1753 that early electrical martyr sought to improve on Benjamin Franklin, collecting storm-borne charge from an ungrounded lightning rod on the laboratory roof. The drawing shows a luminous ball passing straight to the head of the unfortunate physicist. The cut was first published in Paris around 1870, and it looks too anachronistic in detail to offer hard evidence about a tragedy during the reign of the daughter of Peter the Great. Poor Richmann was certainly electrocuted; the autopsy was clear. The blue white lightning ball was not, however, so plain in more contemporary reports. In the 300-year record the peak of ball-lightning reports came just a century ago, "when observational investigation of atmospheric electrical phenomena was popular."

A few chapters are given to a skeptical calculational study of the general physics of the events. The calculations yield convincing, if not quite certain, results. The ball hovers freely, so that in pressure and density it cannot differ much from the ambient air. It answers to no strong forces. Energy estimates based on a dozen reports are distributed almost too neatly along a single normal curve of error; the mean energy in the ball is like that of a few matchsticks in 1,000 times the volume. The ball may vanish suddenly or decay more slowly. Its light is not thermal but of quite another kind.

The best clues seem to have been found in the shipyard and the laboratory. A glowing fireball rather like the accounts can be "repeatedly produced under quasi-controlled conditions," floating off the big switches that make and break the 100-kiloamp currents from the main batteries of a diesel-electric submarine. That ball looks green: plenty of copper vapor. Even more to the point, the plasma physicist David Finkelstein and his colleagues have learned how to make a long-lived glowing ball in ordinary air as the low-efficiency result of a powerful radio-frequency discharge. "The ball occasionally disappeared suddenly with a loud noise." Its color depended on the ambient gas; in air it was yellow white, then it turned orange, brightened and lasted the longer as more nitrogen oxides were present.

The books are not yet closed, but it seems a good bet that ball lightning is neither some ionized plasma wonder nor any other very powerful energy concentration. It is rather a local puff of fluorescing gas, the energy stored as long-lived molecular excitation in some plausible constituent of disturbed air and fed thereto rarely and at low efficiency from some strong natural electrical discharge. Watch for the bright ball, but stay calm and observant; listen and sniff, but it would be best not to touch it.

THE BOOK OF WHALES, written and illustrated by Richard Ellis. Alfred A. Knopf ($25). The epigraph is from *Moby-Dick:* "It is some systematized exhibition of the whale in his broad genera that I would now fain put before you. Yet it is no easy task.... Any way you may look at it, you must needs conclude that the great Leviathan is that one creature in the world which must remain unpainted to the last." So disarmingly begins this book by a well-known marine painter of New York. In it he offers big full-color reproductions of his easel paintings of 13 of the best-known whale species, which he has done over the past five or six years. They are colorful, probably heightened beyond life yet natural in form and pose, but most of them lack the sense of great size. His right whale looks like some sleek curled carp. All the animals are shown, alone or paired, against the bottom texture of a transparent sea or against a blander area evoking the colors of seawater. But absent those ponderous smooth motions amidst revealing veils of sea surface and spray, how can the prodigious be caught on canvas? Melville was right: "There is no earthly way of finding out precisely what the whale really looks like" except to see one.

Still, Ellis presents us with a fine visual approximation. Even better is his lively, documented and remarkably up-to-date text. It is the first of two volumes; the second will treat the dolphins and porpoises, along with the orcas and the pilot whales. He includes 33 species—all the great whales—in this book, which soon grew far beyond his plan for a field guide, a few paragraphs to a species, into full-scale essays, each with its drawings of the entire animal and important details.

The center of the text is the interaction of each species of whale and its human students and predators. We learn the general size and habitat, the look and actions and then, particularly for the important species, the ebb and flow of those relationships. Not a general biology of whales, the book is valuable for the personal quality of the writing and the degree to which Ellis stays in touch with the scientific and the popular literature, with the whaling industry and its growing controls, with the whole of the cetological news. He refers to more than one result so new that it is not yet published.

We still know whales poorly. One species (its description is illustrated by the only photograph in the book, a picture of a whale skull) is known solely from two skulls, one found 50 years ago on a beach in Australia, the other found more recently in a fertilizer factory near Mogadishu, where fishermen who had picked it up on a Somalian beach had sent it. That Indo-Pacific beaked whale is the poorest known of all the large animals of whose existence we are certain. No beaked whales are truly well known; what is one to make of a certain species, the strap-toothed whale, a six-meter form found all over the Southern Hemisphere? The adult males cannot open their mouths "to any significant degree" because their front teeth grow into curved straps a foot long that grow out and around to lock the upper jaw

The humpback whales are well described, including their famous and remarkable songs, complex patterns of sound lasting for half an hour. Not only are these common to all whales of a given breeding population—humpback dialectal music—but also they change over the years, and each year "all the whales seem to have learned the changes." No full publication of this remarkable observation, made by Roger and Katharine Payne, is yet cited.

It is noteworthy that the humpbacks, the most playful of the great whales, are easy to visit in our years. Their protected populations are small compared with the great herds hunted down in the cold North Pacific even in the 1960's, as in the North Atlantic a century ago. They perform now in groups of dozens in the rich waters off the tip of Cape Cod, each spring and fall as they pass in migration. A lucky visitor by boat can see them lobtailing, spy-hopping, fin-flapping and breaching, the animals still "the most gamesome and lighthearted of all whales." Their presence turns the coastal waters of Massachusetts into a nature watcher's wonder, a sea rival to the green plains of the Serengeti. A similar humpback park is found in Hawaiian waters. The long-protected eastern Pacific grays are a famous tourist sight off the coast of California and Baja California, but their Korean counterparts, unprotected, seem to have vanished.

The sperm whale, wide-ranging in the open seas, has survived our dogged century-long chase, now all but ended by law. The animals may still be numbered worldwide at some half million or even a million individuals, living chiefly on squids. Each year they take about 100 million tons, double the total human catch of all kinds of fish. Some of the squids are giants too; there is one record of a 10-meter specimen found in a sperm stomach, and the tall tales of battle between whales and squids of whale size are not to be held mere fantasy; perhaps we shall yet know the big squids of the dark depths, well hunted now by the sperm whale alone.

Even in our cetacean ignorance there

is a sense of growth, which does not so clearly attend the search for those lonely monsters dear to mysterymongers. Edward A. Wilson, the young artist-naturalist with Robert Falcon Scott who was to die of hunger and fatigue in his sleeping bag beside his dying captain on that fatal expedition to the South Pole, had earlier drawn some unknown whales along the Ross Ice Shelf. The whales were handsome, rather like the common killer whale but with a much higher dorsal fin, "erect, pointed and sabre-shaped." In 1964, 15 or 20 animals were seen and photographed on the coast of Chile that "agreed well" with Wilson's drawing, and in 1967 a fishing captain saw some again near Tobago, elusive yet not beyond our hopeful quest.

FROM X-RAYS TO QUARKS: MODERN PHYSICISTS AND THEIR DISCOVERIES, by Emilio Segrè. W. H. Freeman and Company ($20). The 20th century opened for physicists a few years early. "People, when they found out about it, would say, *Der Röntgen*... has really gone crazy. On the first of January [1896] I mailed the reprints, and then the devil was to pay!" The discoverer of X rays, whose find makes a reasonable marker for the rise of what we now call particle physics, the continuing elucidation of the ultimate structure of matter and radiation, thus wrote to a friend in a letter cited here in full. Professor Segrè implicitly offers the highest credentials for the validity of his lucid account of that unfolding: he was there.

One document is a snapshot showing him with that youthful team in shirt sleeves before a shuttered window of the Institute of Physics in Rome, where neutron-induced radioactivity was first found (and fission missed). Only Enrico Fermi, a little the senior, is wearing a tie. We see Segrè next in wartime, this time out for a walk, dark glasses against the New Mexico sun, along with Fermi, Hans Bethe, Victor Weisskopf and a few other companions. Then he is standing on the steps of the Delta Kappa Epsilon house at Rochester in 1960, along with I. I. Rabi, Werner Heisenberg, T.-D. Lee, C. N. Yang and a few other Nobel prizewinners.

The reader must not take away any hint that the account is self-centered; rather, it centers on the exciting times, people and places where this prizewinner, codiscoverer of the antiproton, was at home. Its terse appraisals and its acute and definite opinions are just what the author's reputation would lead one to expect. He has supported his ready phrases with precise scholarship, presenting with very few errors of omission or commission an entirely nonmathematical "impressionistic view of the events as they appeared to me" over some 50 years of action. Facsimile papers and personal photographs, with much apparatus and many laboratories, enrich the book, a fresh and pointed collection of illustrations.

Segrè periodizes the history of modern physics, if not explicitly, then at least in feeling. First came the Edwardian world, shared by the theorists Planck and Einstein with their experimental counterparts, such as the Curies and Ernest Rutherford. Then enters the young Bohr, by way of Rutherford's own laboratory, his quantum-state insight built on the Rutherford nuclear atom ("the highest form of musicality in the sphere of thought," as Einstein described it). Rutherford and Einstein could hardly even have exchanged ideas on physics; their viewpoints were utterly distinct. Yet Rutherford and the theorist Bohr remained close friends. The young Bohr came back to Manchester that summer to defend his manuscript (the marvel of 1913) obstinately line by line against Rutherford's criticism. He won his case, and the paper bears at length the rather stilted, meticulously phrased tentativeness of Bohr's lifelong style.

Then there grew in splendor the "true quantum mechanics at last," fully powerful by 1932, that year of particle wonders: the neutron, the positron and deuterium (and by 1933 Fermi's neutrino). The next phase sees the accelerators and a rich nuclear physics, first in peace, then terribly in war. By 1951 Fermi, "with an ironical smile," set about renewing himself, with mesons. The next 25 years (although alas not for Fermi) went "beyond the nucleus." Bigger and still bigger machines, immediate complex numerical processing, novel and abstract algebraic theories, all still resting on the sturdy foundations set before 1932, span the remaining story, told briefly with some allusions to condensed matter and to astrophysics. The book tapers rapidly to its close around 1975, although there are some later points. (That was the writing date of the first version, published in Italian.)

We learn how the science spread far beyond Europe, particularly to Hideki Yukawa's generation in Japan and now to all the world. Segrè sees the forerunner of all our great physics laboratories, with their heavy engineering, big staff and multiple researches, in the famed low-temperature laboratory of Heike Kamerlingh Onnes at Leiden soon after 1890. "Even if... Onnes's work is the first example of large-scale physics, it was not particularly familiar to... accelerator builders," who came many years later. (The great observatories and the big organic chemistry laboratories had begun even earlier.)

The book ends with a reflective chapter, all too brief, in which the author tries to untangle the structure of the living being, physics, with its complex interactions and its many indispensable organs of thought and experience. It is no linear block diagram that can generate discovery; rather, we see an evolving interconnection of diverse activities, places for fruitful work by all kinds of personalities. The record is clear on that human diversity and on the continued evolution of insight. "I do not believe Galileo, Newton and Einstein have been the last of their ilk" is the concluding sentence of this affirmative and never ingenuous text.

There is a series of appendixes on a more mathematical level, a full set of indexes and a candid and helpful annotated bibliography.

INCA ARCHITECTURE, by Graziano Gasparini and Luise Margolies. Translated by Patricia J. Lyon. Indiana University Press ($32.50). "The long rosary" of the Royal Highway of the Incas stretches from Cuzco north along the high valley of the Cordillera all the way to equatorial Quito. Along it are strung its beads: scores of the imposing stone settlements of the Inca empire. Alexander von Humboldt wrote of them 150 years ago: "It is impossible to examine a single edifice of the Incas... without recognizing the same type... throughout the Andes, extending over four hundred leagues.... It seems as if a single architect built this great number of monuments." These authors, an architect and an anthropologist in partnership at the Universidad Central de Venezuela in Caracas, entirely agree. They have visited, drawn, photographed and pondered the sites, their eyes keen for both structure and function. They have carefully compared what is there on the ground with the rich but confusing testimony of those chroniclers of the 16th century who saw the monuments of the Inca state, Tawantinsuyu, before they were unmade by the conquistadors.

The authors conclude that indeed there was a single architect, that extraordinary state itself. Its chief tool was its thoroughgoing organization of its subjects. Their technology, their ancient mastery of stonemasonry (with stone tools) achieved eight centuries before the first Incas, are witnessed in the massive, ornate ruin of Tiwanaku at the southern end of Lake Titicaca. The Incas left no moldings, cornices, pilasters or other ornament, no geometric or representational carvings, no inscriptions, no reliefs. Civil servants were trained in Cuzco to direct the great public works. They applied the same austere norms with diligence everywhere. Trapezoidal niches, doorways and windows always mark Inca work; even along the northern coast of Peru, where the Incas ruled over old and populous seaside cities, we find Inca structures with trapezoidal niches, although here they were built not of stone but of sun-dried adobe, the local material.

The Inca state endured for only about 80 years. Its founder, Pachakuti, built his planned capital, Cuzco, a uniform one-storied city in worked stone, out of


nature

a highland village in a generation. Its great open plazas gave the town plan a particular Renaissance quality, a feature older in this and other centers of the Incas than it was in Spain. The wide Plaza Mayor of Madrid was started only in 1617. Cuzco was partly burned during the war of conquest, but the most important change came after 1534, when the Spanish distributed private house lots within the original open public spaces. Superposed on the first-floor Inca stonework, so severe in its monumentality and precision of workmanship, is now a Spanish colonial second-story city in plastered adobe and wood.

The Inca settlements were not real cities except near Cuzco. The triumphant Inca strategy for wide conquest and rule was "archipelagic . . . rather than conversion of every last peasant to the solar cult." Up and down the Inca lands the imposing stone settlements grew, but their forms disclose their nature. Terraced fields for crops are few; storage cells for imports of grain and potatoes are numerous. There are no cemeteries; the pottery is state ware, not the same as that of the farmers round about. The settlements appear suddenly, without archaeological roots; they were abandoned rapidly once the state fell. No one was at home in them. The most striking buildings are long halls, stone-walled and pitch-roofed. They seem meant not for resident families but for transients, conscript soldiers passing through, or the crowds of laborers who were fulfilling their mita, the universal levy of forced labor demanded as the chief tribute to the Inca ruler. The workers were called up in rotation, never permanently, and often served hundreds of miles from home.

The great blocks of local limestone in Cuzco's wondrous walls, particularly in the zigzag platforms of Saqsaywaman on the heights just north of the town, were never fortifications. The Incas had no powerful enemies in their central lands. Those immense stones, "more than twenty feet long, and others thicker than an ox, and all set so delicately that between one and another they could not fit a *real*," amaze us still. They were the top of that hierarchy in stone the Incas built to house and to mirror their regime. The more important the building, the greater the excellence of its walls. The chief structures of Cuzco are marvels of size and fit; practical structures in some provincial way station are most often of dry-laid fieldstone. True, it is hard to explain the "lapidary fanaticism" of those huge fitted blocks, puffy and elegant, wildly polygonal, artful, lavishly laborious. They were not mere practical solutions; the masons weary of their tax labor may have taken delight in them, and Leviathan, then as now centered on its power, may have "permitted the . . . justifying byword": anything at all, as long as it keeps them working.

INTRODUCING

Electrical and Electronics Engineering Book Club

Ideas and answers at your fingertips

Take any 3 books for only $1.00 each (values to $74.40)

You simply agree to buy 3 more books—at handsome discounts—within the next 12 months.

Now you can build a valuable professional library easily, conveniently and at great savings! Select your introductory books now and mail in the reply card today.

(Publishers' Prices shown)

36915. BUCHSBAUM'S COMPLETE HANDBOOK OF PRACTICAL ELECTRONIC REFERENCE DATA. *Walter H. Buchsbaum.* Over 600 pages with hundreds of charts and diagrams. **$19.95**

52170. HANDBOOK OF DIGITAL IC APPLICATION. *David L. Heiserman* **$22.95**

84545. TRANSISTOR CIRCUIT ANALYSIS AND APPLICATION. *Ben Zeines.* **$18.95**

84555-2. TRANSFORMER AND INDUCTOR DESIGN HANDBOOK. *Wm. T. McLyman.* Manual for aerospace and electronics industries. *Counts as 2 of your 3 books.* **$35.00**

42175. DIAGNOSIS AND RELIABLE DESIGN OF DIGITAL SYSTEMS. *Breuer and Friedman.* **$23.95**

83380-2. THEORY AND APPLICATION OF DIGITAL SIGNAL PROCESSING. *Rabiner and Gold. Counts as 2 of your 3 books.* **$29.95**

80270. STARTING AND MANAGING YOUR OWN ENGINEERING PRACTICE. *John A. Kueken.* Models for projecting income, taxes and savings **$13.95**

62672. MICROELECTRONICS: Processing and Device Design. *Roy A. Colclaser.* Includes pattern generation and photo-engraving, substrate preparation, and more. **$24.95**

42033-2. DESIGN OF SOLID-STATE POWER SUPPLIES. *Eugene Hnatek.* Covers everything from oscillators and transistors to finished converters/inverters. *Counts as 2 of your 3 books.* **$27.50**

77790-2. SEMICONDUCTOR DEVICES AND INTEGRATED ELECTRONICS. *Arthur George Milnes.* Covers Schottky barrier diodes, thyristors, microwave applications and much more. *Counts as 2 of your 3 books.* **$26.50**

44900-3. THE ENCYCLOPEDIA OF COMPUTER SCIENCE. *Edited by Ralston and Meek.* Monumental 1550-page volume covers topics ranging from theory and applications to programming and Boolean algebra. *Counts as 3 of your 3 books.* **$60.00**

70240. PRINCIPLES OF TECHNICAL MANAGEMENT. *William A. Cohen.* Establishing schedules, hiring, budgeting, and more. **$19.95**

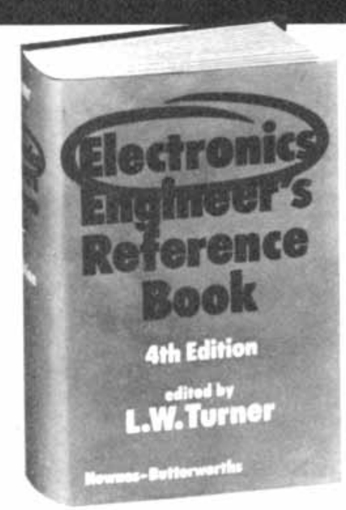
32265-2. ADVANCED LOGICAL CIRCUIT DESIGN TECHNIQUES. *Svoboda and White.* Presents major breakthroughs in sequential circuit theory and its application to LSI chip design. *Counts as 2 of your 3 books.* **$27.50**

88008. WRITING FOR RESULTS IN BUSINESS, GOVERNMENT, THE SCIENCES AND THE PROFESSIONS. *David W. Ewing.* Management reports, letters and more. **$19.95**

52237. HANDBOOK OF MICROPROCESSORS, MICROCOMPUTERS AND MINICOMPUTERS. *John D. Lenk.* **$17.95**

45435. ENVIRONMENTAL CONTROL IN ELECTRONIC MANUFACTURING. *Edited by P. W. Morrison.* **$24.50**

Special Selection



44145-3. ELECTRONICS ENGINEER'S REFERENCE BOOK. *Edited by L. W. Turner.* Up-dated, covers electronic materials and components, solid state devices, integrated circuits, microelectronics, telecommunications and electronics in industry plus much more. 1600 pages. *Counts as 3 of your 3 books.* **$78.00**

79163. SOFTWARE QUALITY MANAGEMENT. *Cooper and Fisher.* **$20.00**

79167. SOFTWARE RELIABILITY GUIDEBOOK. *Robert L. Glass.* **$18.95**

52214. HANDBOOK OF ELECTRONIC FORMULAS, SYMBOLS AND DEFINITIONS. *John R. Brand.* **$15.95**

69750. A PRACTICAL GUIDE TO COMPUTER METHODS FOR ENGINEERS. *Terry E. Shoup.* **$17.95**

70760. PROJECT MANAGEMENT: A Systems Approach to Planning, Scheduling and Controlling. *Harold Kerzner.* **$22.50**

35455. AUTOMATION, PRODUCTION SYSTEMS, AND COMPUTER-AIDED MANUFACTURING. *Mikell P. Groover.* **$24.95**

39825-2. COMPLEX DIGITAL CONTROL SYSTEMS. *Guthikonda V. Rao. Counts as 2 of your 3 books.* **$36.50**

52210-2. HANDBOOK OF ELECTRONIC DESIGN AND ANALYSIS PROCEDURES USING PROGRAMMABLE CALCULATORS. *Bruce K. Murdock. Counts as 2 of your 3 books.* **$26.50**

73865-2. RELIABILITY AND MAINTAINABILITY OF ELECTRONIC SYSTEMS. *Edited by Arsenault and Roberts.* 597 pages. *Counts as 2 of your 3 books.* **$28.95**

44125. ELECTRICAL ENGINEERING LICENSE REVIEW. *Jones and Lima.* **$16.95**

00470-3. VAN NOSTRAND'S SCIENTIFIC ENCYCLOPEDIA. *Edited by Douglas M. Considine.* Outstanding reference for professional engineers with over 7,200 articles on scientific and engineering subjects. 2,382 pages; 2,500 illustrations and charts. *Counts as 3 of your 3 books.* **$69.50**

48437. FIBER OPTICS AND LIGHTWAVE COMMUNICATIONS STANDARD DICTIONARY. *Martin H. Weik.* **$18.50**

60388. MAN-MADE RADIO NOISE. *Edward N. Skomal.* Analyzes all sources of man-made radio noise in industrial areas. Packed with charts and diagrams. **$22.50**

82727-2. TELECOMMUNICATION SYSTEM ENGINEERING: Analog and Digital Network Design. *Roger L. Freeman. Counts as 2 of your 3 books.* **$32.50**

83400. THEORY AND DESIGN OF SWITCHING CIRCUITS. *Friedman and Menon.* **$22.95**

If the reply card has been removed, please write:
Electrical and Electronics Engineering Book Club, Dept. M-AC3, Riverside, N.J. 08075
to obtain membership information and application

Established 1845 SCIENTIFIC AMERICAN March 1981 Volume 244 Number 3

The Service Sector of the U.S. Economy

It is now the dominant sector, replacing the goods-producing sectors. Other changes are the growth of the nonprofit sector, the increase in human capital and the further internationalization of the economy

by Eli Ginzberg and George J. Vojta

A proposition that has gained popularity with the change of administration in Washington holds that the Government should speed the "reindustrialization" of the U.S. economy. The economy, it is said, has lost momentum because of a failure to keep modernizing its basic industries. The same failure is said to have reduced the nation's competitiveness in exports. This proposition and its supporting rationale warrant close inspection and perhaps second thought.

Public discussion of the economy tends to focus on the short run: the next quarter, the next year, the next turn in the business cycle. Implicit is the assumption that the economy of today is basically the same as the economy of yesterday. Over the years, however, changes that go unperceived in the short run accumulate. A dynamic, growing economy inevitably undergoes structural transformation.

In 1929 the U.S. gross national product was just over $100 billion. By 1980 it had grown 24 times to $2,400 billion. Allowing for inflation, the growth (in constant dollars of purchasing power) was almost fivefold. Profound changes in the structure of the economy that accompanied this growth are reflected in the distribution of the labor force and the gross national product. The data show that the provision of services has displaced the production of goods as the country's principal economic activity. Since much of this service activity is conducted by the Government and by private nonprofit institutions, a vast not-for-profit sector, encompassing Government and nonprofit institutions, has emerged. Concurrently the growth of the economy has come to depend on human capital more than on physical capital. Finally, the growth of imports and exports as a percentage of the gross national product demonstrates that the economy has also become increasingly internationalized.

Hence in the perspective of a half century a fourfold transformation of the U.S. economy becomes apparent. We shall attempt here an assessment of these changes and their implications for national economic policy.

In defining services we observe the convention of national accounting that allocates to services all output that does not come from the four goods-producing sectors: agriculture, mining, manufacturing and construction. The service sector thus embraces distributive services such as wholesale and retail trade, communications, transportation and public utilities; producer services such as accounting, legal counsel, marketing, banking, architecture, engineering and management consulting; consumer services such as restaurants, hotels and resorts, laundry and dry-cleaning establishments, and nonprofit and government services such as education, health, the administration of justice and national defense.

The U.S. economy in 1929 employed 45 percent of the working population in the production of goods; by 1977 that sector employed only 32 percent. Employment in the service sector therefore increased from 55 to 68 percent of the working population. Most of the shift came in the three decades after the end of World War II. Between 1948 and 1977 employment in goods production declined by 12 percentage points; the decline in manufacturing employment accounted for two-thirds—8 percentage points—of that decline. In 1948, when the nation was restocking its homes and garages after the long depression of the 1930's and the wartime restrictions on the output of consumer goods, manufacturing employed one worker out of three; by 1977 it employed fewer than one worker out of four. Over this period the service sector absorbed not only the percentages displaced from manufacturing but also those displaced from agriculture and mining.

The gross-national-product figures in both current and constant dollars exhibit the same massive shift from goods to services. Current-dollar figures, although they grossly overstate the real rate of growth, have the advantage of reflecting more accurately the changes in relative prices and consumer preferences. Again the major transformation came in the postwar years. In 1948 the goods-producing sector accounted for 46 percent of the gross national product and the service sector for 54 percent. In 1978 the figures were 34 percent and 66 percent. The major contributors to the total decrease of 12 percentage points in goods production were agriculture with 6 percentage points and manufacturing with nearly 5.

The disparity between the shares of the service sector in employment (68 percent) and in the gross national product (66 percent) reflects differences between services and goods production in the type of labor employed and in the utilization of capital; it also reflects certain national accounting conventions. The share of government measured in terms of employment is much larger than it is in terms of gross national product because of the large number of relatively low-paid wage earners (particularly women) on the public payroll and also because of the national-income accounting convention that measures the value of government output solely in terms of wages and salaries. On the other hand, producer services, with their many highly paid professional, technical and managerial people, are more prominent in the economy as a share of the gross national product than as a share of employment.

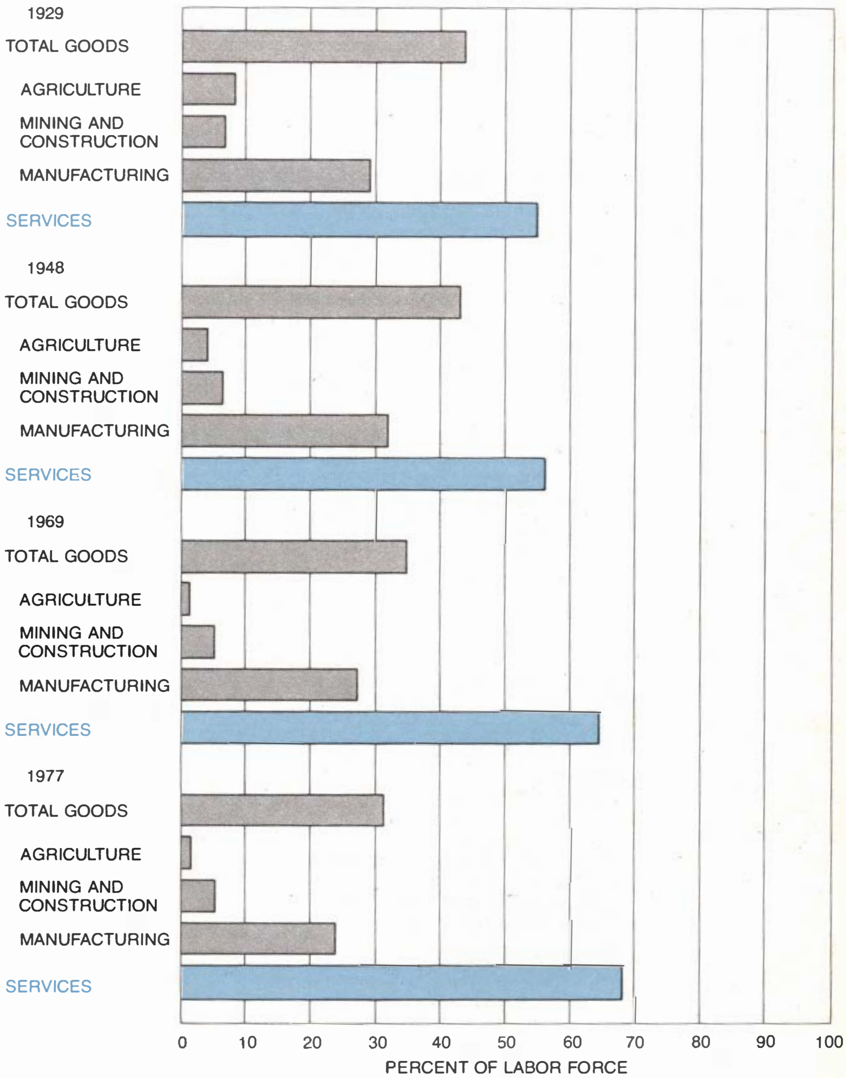
Perhaps more dramatically than the relative comparisons, the absolute employment and dollar figures demonstrate the transformation of the structure of the economy over the three postwar decades. In 1948, 48.1 million people were employed, 20.9 million in goods production (15.5 million in manufacturing) and 27.2 million in services. The 1977 economy employed 79.5 million people, an increase of more than 30 million, 25.1 million in the goods sector (19.1 million in manufacturing) and 54.4 million—more than the total payroll of the 1948 economy—in services. Comparison here shows that manufacturing employed only 3.6 million more people (to account for a much bigger product, as will be seen below) and the expanded service sector employed 27.2 million more people, roughly 7.5 people for every additional worker employed in manufacturing.

In 1948 the gross national product had reached $259 billion, with the goods sector contributing $118.4 billion ($74.3 billion from manufacturing) and the service sector $140.6 billion. By 1977 the gross national product had increased to $1,900 billion, and the goods sector was contributing $649.8 billion ($456 billion from manufacturing) and the service sector $1,250.2 billion. Of the $1,600-billion increase in the gross national product manufacturing generated $382 billion and services generated $1,100 billion, three times as much.

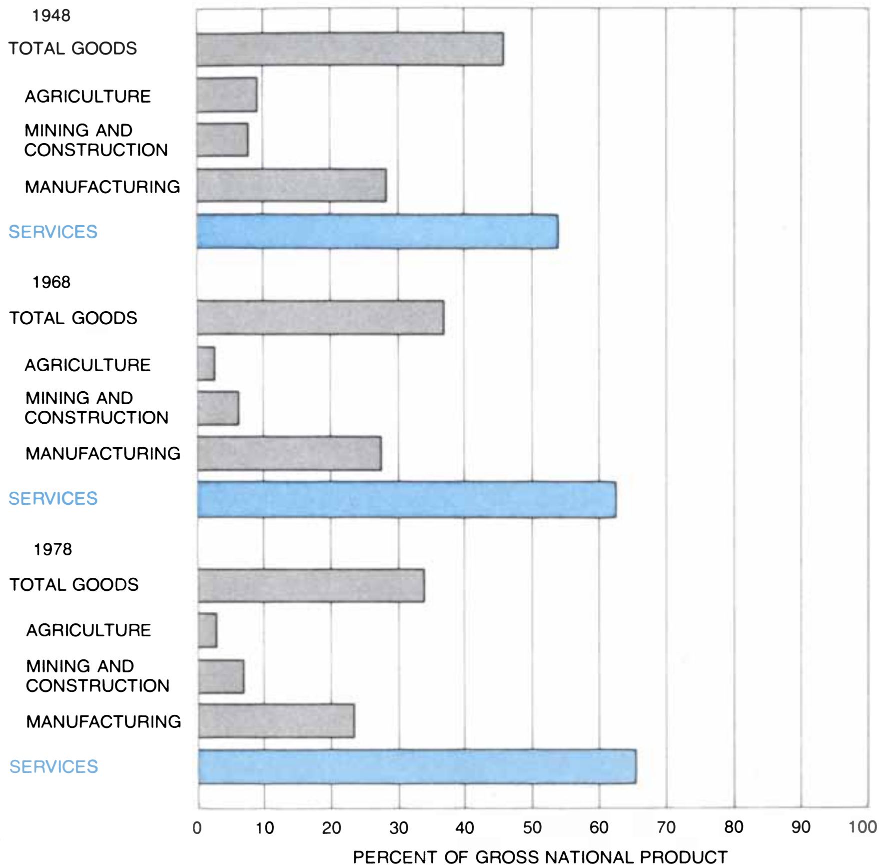
Accounting conventions here conceal an even more striking contribution to economic growth from the increasing number of people engaged in service occupations. The official figures indicate that manufacturing in 1978 contributed 24 percent of the gross national product and producer services 19 percent. The latter figure represents, however, only the contribution from independent accounting, architectural, advertising, investment banking, law and management consulting firms. A more complete statement of the contribution from producer services would credit the input from people employed in the same kind of work on the direct payroll of goods-producing enterprises. Between 1948 and 1978 the number of nonproduction workers, including those in producer services, increased from 4.7 to 12.7 million. In manufacturing this element in the payroll increased from 2.7 to 5.7 million: three million of the total 3.6 million increase in manufacturing employment noted above. If a reasonable adjustment is made for the people engaged in producer-service functions, the remarkable result is that the value added by producer services now approximates the "value added by manufacture" in the U.S. economy.

This finding leads to our second thesis, which is that human capital, defined as the "skill, dexterity and knowledge" of the population, has become the critical input that determines the rate of growth of the economy and the well-being of the population. We contend that the competence of management and the skills of the work force, particularly of those engaged in producer services, determine the ability of enterprises to obtain and utilize effectively the other essential resources, such as physical capital, materials and technology.

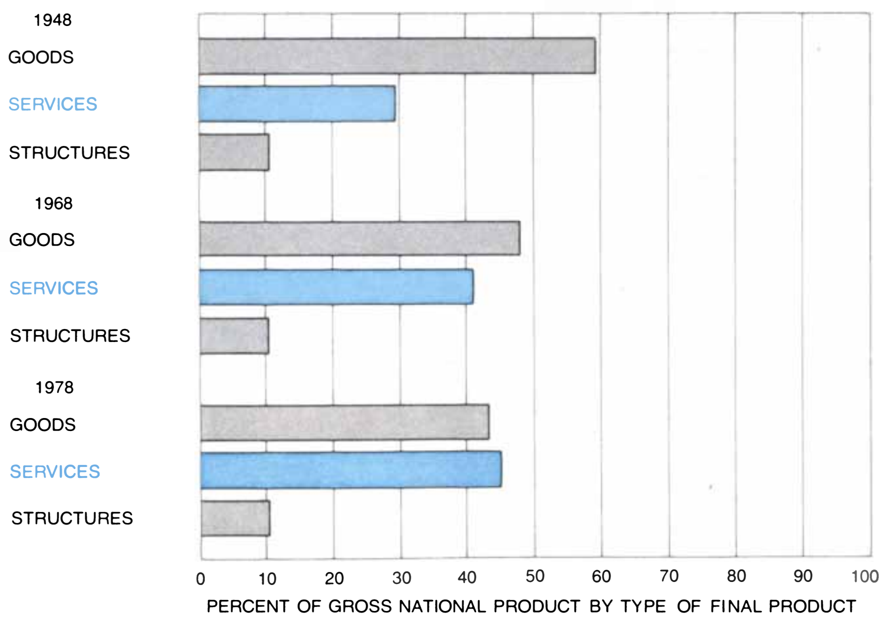
The concept of human capital is not new, but it has not generally been accorded priority in economic analysis or in business management. An insight into the dominance of human resources in the economy is provided by the finding that in 1947 the compensation of employees plus half of the returns (reflecting the input of labor) to farm and nonfarm proprietors amounted to 75 percent of the national income; in 1979 it had risen to 79 percent. The return



SHIFTS IN EMPLOYMENT since 1929 are charted for the goods-producing industries (*gray*) and the service sector (*color*). The service sector includes distributive services such as communications, utilities and wholesale trade; retail trade; consumer services such as restaurants, dry cleaning and recreation; producer services such as accounting, banking and legal work, and nonprofit and government services including health, education and national defense.



SHARES OF GROSS NATIONAL PRODUCT have shifted among the major sectors in the past 30 years. The contribution from services (*color*) has increased, whereas the contribution from the four goods-producing sectors (*gray*) has declined. Taken together, the goods-producing sectors accounted for 46 percent of the gross national product in 1948 and 34 percent in 1978. The share of services rose from 54 percent to 66 percent of the gross national product.



CHANGING IMPORTANCE of goods with respect to services when the gross national product is classified according to type of final product is reflected in this chart. The share of the service sector is actually understated, since by a convention of national-income accounting the output of nonprofit, government and certain producer services is measured only by what the workers are paid. The total value of what is produced is not taken fully into consideration.

to labor—labor of a new kind—now amounts to essentially four out of five dollars of national income.

The economist Edward F. Denison of the U.S. Department of Commerce has made a comprehensive study of the comparative contribution to the growth of the economy in this period of the several factors of production. Of the total growth between 1948 and 1973, he calculated, 15.4 percent could be credited to "more capital." The increase in the number and the education of the work force and the greater pool of knowledge available to the workers account for about two-thirds of the increased growth of the economy during this period. Even the remaining fraction of the growth still to be accounted for reflects the improved utilization of the human-resource factor by improved management and the shift of workers to more productive sectors.

The Denison calculations are buttressed by figures from other sources that show human capital formation proceeding at an impressive rate throughout the period under study. Taking education as a measure of improvement in human resources, the figures show that the median of the years of school completed by people in the civilian labor force stood at 10.6 in 1948. In 1978 the median was 12.6, a substantial gain within a generation and a half. The figures for secondary and higher education show still more impressive gains. As recently as 1957 fewer than half of the labor force had finished high school and only 9 percent had finished college. By 1978 three out of four workers had finished high school, and a third had attended college, had obtained a college degree or had attended graduate school.

Improvement in the educational preparation of the labor force is reflected in the upgrading of the occupations in which the workers are employed. Between 1959 and 1978 total employment increased by about 30 million. More than half of the increase—16 million new jobs—was in the higher-level professional, technical, managerial-administrative, sales and crafts occupations. The white-collar work force rose from 42.7 percent of the total in 1959 to 50 percent in 1978. According to a 1975 survey by the Department of Labor, 42.1 percent of the white-collar employment in the service sector must be reckoned in the top white-collar occupational groups; in manufacturing the figure is 24.9 percent. The addition of craftsmen to these totals brings the respective percentages to 49.3 and 36.1. Considering the absolute number of jobs in the most rapidly growing producer and nonprofit service sectors, we find they employ more than double the high-level manpower employed in manufacturing. The upgrading of the labor force is also evidenced by the disappearance of 800,000 private-household jobs and a 2.5-million decline in farm employment.

The foregoing analysis does not deny that physical capital continues to play a prominent role in the country's economic growth. It does, however, support our thesis that human capital has come to play an increasingly dominant role. Simply put, it is the expansion of knowledge, skills, imagination, ideas and insights of working people that creates the margins from which physical capital is accumulated, leading through productive investments to the further accumulation of capital.

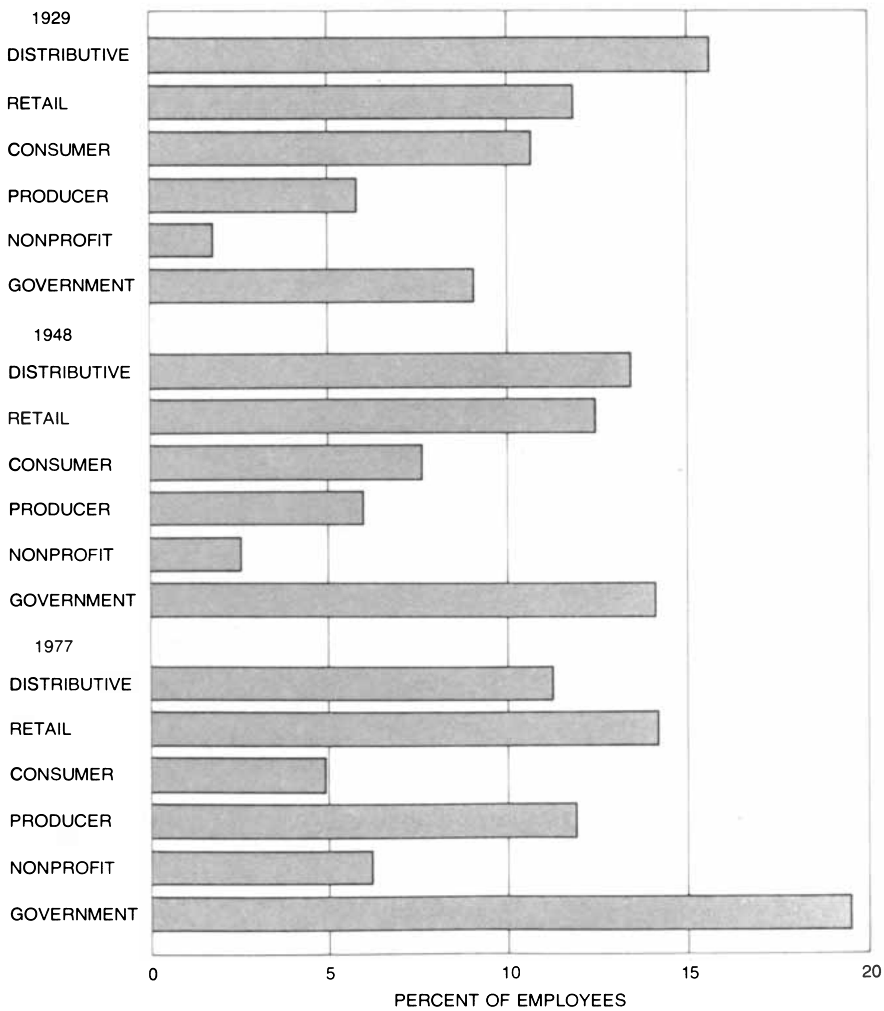
We turn now to consideration of the not-for-profit sector, which includes private nonprofit institutions and all government, and of the contribution the growth of this sector has made, particularly over the past 30 years, to the growth of the entire economy. To begin with, it is to the not-for-profit sector that the country has looked for the enhancement of its human capital. From this sector there have come also the investment in infrastructure, specific demands for goods and services and even outright subsidies that have given major impulsion to the growth of the private for-profit sector.

At first it is difficult to understand how the national political ethos, expounded by Republicans and Democrats alike, continues to maintain that five out of every six jobs are created by the private sector. That was true in 1929, when government expenditures were about 10 percent of the gross national product. It ceased to be true, however, in the depression years of the New Deal and in World War II, and it is surely not true today. The misconception arises in part from the classification of such nonprofit institutions as Columbia University, the Metropolitan Museum of Art and the Jet Propulsion Laboratory as private-sector enterprises and from categorizing the production of military aircraft by the Lockheed Corporation and nuclear submarines by the General Dynamics Corporation as private-enterprise activity. These classifications obscure the true situation.

To comprehend the actual dimensions of the government sector one must count not only the people employed on the public payroll but also the people in the private sector who are employed because of government purchases from or grants to private-sector enterprises. Civilian employment on government payrolls increased from 11 million in 1962 to 16.5 million in 1978. Practically all these new jobs were created by state and local governments. In addition, during this same period purchases by government at all levels brought an increase in the direct employment in the private sector occasioned by such purchases from 6.3 to 8.2 million. Counting also employees in such government enterprises as the Tennessee Valley Authority, state liquor stores and municipal power plants, the number of people employed by government increased from 18.4 to 26.2 million in those 16 years. When the contribution of the private nonprofit sector is added to that of government, the not-for-profit sector accounts for more than a third of the total employment and nearly a third of the gross national product.

The argument that such employment and expenditure are at cost to the "productive" side of the economy is belied by the historical record of the very period that has seen the rapid expansion of the not-for-profit sector. For the past several decades agriculture and the automotive industry were the movers and shakers of the economy, both at home and abroad. The role of Federal Government expenditures in research and development and in the operation of the agricultural extension service, which was in effect a technology-transfer enterprise, was crucial to the enormous gains of output per man-hour and per acre of the country's farms. Similarly, it was Federal, state and local outlays for highway building, greatly enhanced by the Eisenhower Administration's interstate-highway legislation of 1956, that contributed to the prolonged prosperity of the automotive industry that has just recently foundered on fuel-price increases and competition from imported automobiles.

Today the aircraft and the electronics industries continue to contribute significantly, along with agriculture, to the country's exports. The facts are clear: huge Government expenditures for defense and space, for research and development and for the education of highly specialized personnel have underpinned and sustained the rapid progress of these

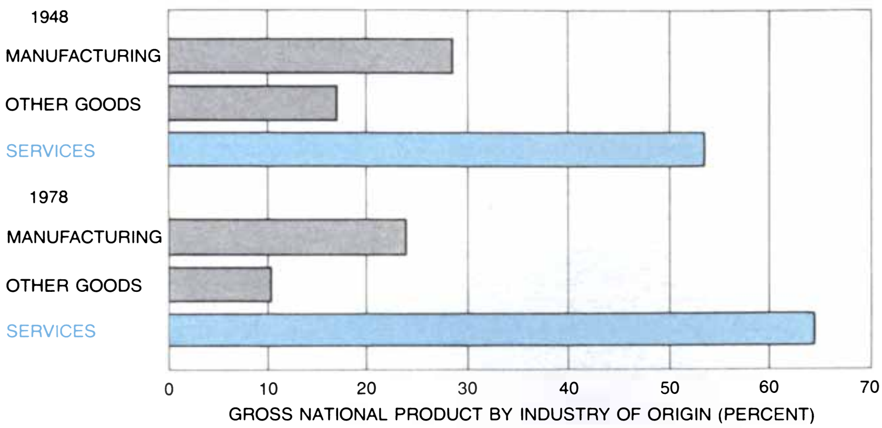


EMPLOYMENT BY TYPE OF SERVICE shows several shifts over a period of almost 50 years. The bars represent employment in service industries as a percentage of total employment. Most of the growth was in government at all levels, mainly between 1948 and 1968, and in producer services and the nonprofit sector, where the gains have continued to the present.

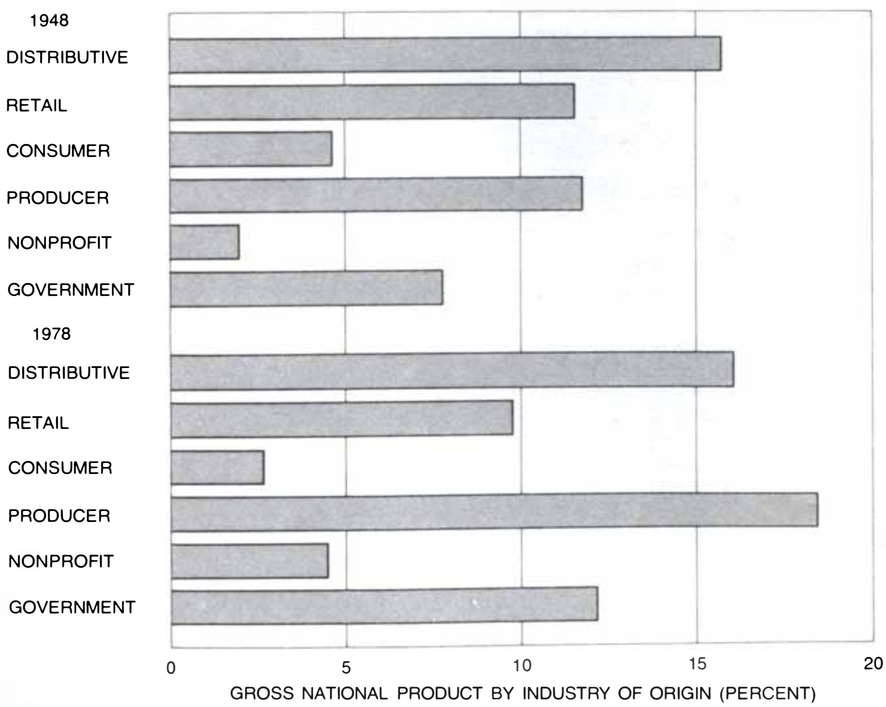
industries. Hence although the private sector continues to dominate the U.S. economy, the Government and the private nonprofit sector perform a large number of functions, as both investors and purchasers, that are critically important to the continued expansion of the private sector.

None of these functions is more important than the acknowledged responsibility and role of the not-for-profit sector in the nurturing and enhancement of the country's human capital. Between 1950 and 1979 total expenditures for education increased from $8.3 to $151.5 billion, of which all but $25 billion came from public treasuries. Expenditures for higher education alone increased nearly 30 times, from $1.7 to $48.9 billion, with the public sector covering two-thirds of the bill. In parallel with the growth of the educational enterprise, expenditures for research and development increased from $13.5 to $51.6 billion during the same period, with the Federal Government's share two-thirds of the total in 1960 and just under one-half in 1979. The resulting transformation and upgrading of the labor force contributed measurably to the increased wealth of the nation.

Another major contribution to human-capital formation took the form of an expansion of the country's health-care system, stimulated alike by expenditures by government and by nonprofit institutions such as Blue Cross. In 1950 expenditures for medical care were $12 billion, with consumers paying $4 of every $5. In 1980 these expenditures had climbed to $240 billion, with the Government providing $2 out of every $5 and private and nonprofit insurance most of the rest.



INDUSTRY OF ORIGIN of the major components of the gross national product is indicated for 1948 and 1978. The contribution of services now accounts for almost two-thirds of the total.



SERVICE COMPONENTS of the economy contributed to the economy in 1948 and 1978 as is shown here. The chart provides more detail on the data in the bars labeled "Services" in the illustration above this chart. The contribution of producer services is much higher when it is measured in terms of gross national product than when it is measured in terms of employment because the category includes many professionals and specialists who are highly paid.

The fourth dimension of the continuing transformation of the U.S. economy is the trend toward internationalization, which is reflected in the increased flow abroad of American trade, investments and grants and loans by the Government. Between 1950 and 1979 total U.S. exports increased from $10.2 to $175.3 billion; exports and imports as a share of the gross national product rose from 6.8 to 15.4 percent. Taken by itself the trade account understates the real involvement of the U.S. in the world economy. In 1950 American private assets abroad totaled $19 billion. By 1978 they were $377 billion, 20 times the 1950 figure. In the single decade 1966–76 the yearly sales of foreign affiliates of American companies increased from $98 to about $515 billion.

Direct foreign investment in the U.S. meanwhile rose from $3.4 to $40.8 billion (12 times). In the late 1940's the U.S. was earning annually about $1 billion net on overseas investments. In 1978 it earned $43 billion on overseas investments and paid out $22 billion to foreign governments and individuals, leaving a net balance of $21 billion. These "invisible exports," preponderantly the yield from human capacity, particularly organizational and managerial capabilities, nearly offset the increased expenditure for petroleum imports that put the foreign-exchange account $7 billion in the red.

The increasing participation of the U.S. in the global economy has been accomplished largely by the nation's leadership role in international affairs since World War II and by the transformation of many large American corporations into multinational enterprises interested in expansion and diversification. Corporations could frame and carry out such ambitions because they had the human capacity to plan, organize and control global production, distribution and financing systems. Critical to the success of their penetration abroad was a growing pool of specialized and experienced managerial and technical-professional people on their payrolls.

Therefore four mutually reinforcing changes—the displacement of goods by services at the cutting edge of economic growth, the growth of the not-for-profit sector, the increasing importance of human capital and the internationalization of the business system—have transformed the U.S. economy over the past 30 years. We now consider the linkages and interrelations among these developments. The discussion will examine how structural changes in the household, in the company and in government speeded the transformation.

Between 1950 and 1980 the participation of women in the labor force rose from about a third to slightly more than half. Three out of every five new workers were women. Most of them found their jobs in the expanding service sector. Most of them also were married women living with their husband at home, and so their employment brought new affluence to the American household. Of the 47 million families (in 1977) consisting of at least husband and wife the husband was the only worker in 12.7 million, but both spouses worked in 23.4 million. The median income of the former group was $15,400, of the latter $20,400.

If consumer spending is allocated to (1) basic necessities, (2) comfort and convenience and (3) products and activities related to self-identity or way of life, the newly affluent households are spending less on the first two categories and more on the third, which includes leisure, travel, recreation, education and so on. Between 1948 and 1978 consumers reduced their outlays for goods from 68.4 to 54.1 percent of their total outlays. Expenditures on services rose accordingly from 31.6 to 45.9 percent, reflecting primarily changes in taste associated with the continuance of education, higher income and the need for support services (which are more in demand when the wife works), together with a relatively larger inflationary increase in the cost of services compared with goods.

The affluence of the years after World War II has substantially broadened and deepened the national market. In the 1930's President Roosevelt identified the South as the nation's primary economic problem. The region's agriculture was in decline; its industries were concentrated in low-wage manufacturing, and its poor whites and blacks were underemployed. By the mid-1950's, as a result of large-scale emigration of poor and unskilled people and of substantial expenditures by the Federal Government for regional development and defense that attracted a reverse migration of people with higher skills, the South had begun to boom.

The Southwest and the West similarly responded to Federal outlays that placed the center of the new defense industries there and accelerated the population growth of those regions. Federal, state and local investment in infrastructure underwrote the relocation of American households in suburbia, in the new "Sun Belt" and elsewhere in the country.

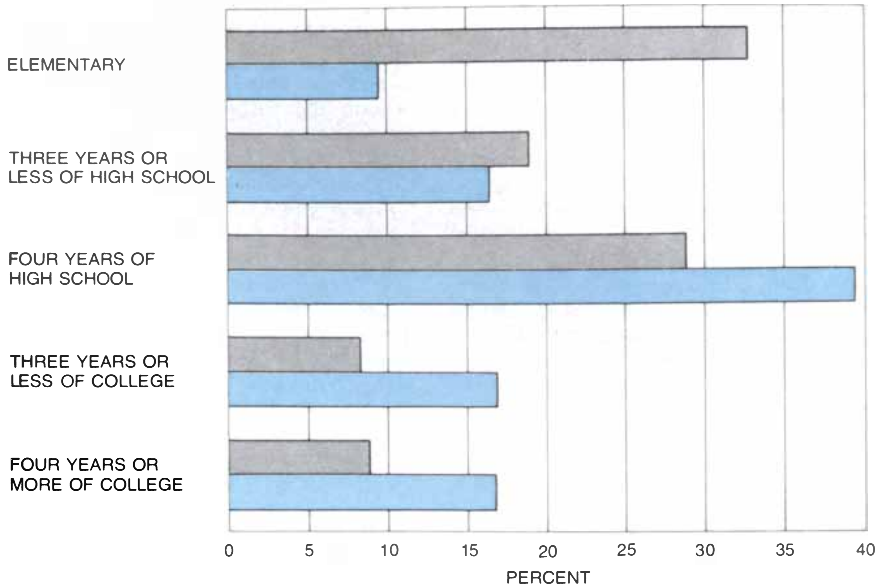
These changes in the national market brought the establishment of many new business enterprises in the South and the West, and many already established in the East and the Middle West expanded to beyond the Rockies. It is not surprising that for most of the postwar period the South and the West have had above-average rates of gain in employment and income whereas in the Middle Atlantic and East North Central regions the rate of gain has been considerably lower.

At the end of World War II multiunit manufacturing companies employed only a third more workers and produced only a third more of the value added than single-unit companies. By 1972 the multiunit companies had three times as many employees as the single-unit ones and accounted for four times the value added. Responding to the growth of the national market, these large manufacturing companies had restructured their organization and had also changed their ways of doing business. They did these things primarily by establishing one form or another of decentralized operating structure and by building up decentralized groups of service functions—research and development, marketing, advertising, employee and government relations, legal, financial, accounting—to help them plan and control their multiunit enterprises.

Manufacturers of high-priced items such as automobiles, refrigerators, color-television sets and mobile homes recognized that they stood to gain almost as much from extending credit (a service function) as from the markup on their products. Between 1950 and 1978 the total amount of consumer credit outstanding rose from $25.6 to $340.3 billion, or from about 12 to 23 percent of disposable personal income.

Big companies could diversify into services as well as goods production, decentralize their operations and plan and control them on a national scale because they could call on a wide variety of specialized knowledge and talent in their growing cadres of managerial-administrative and technical-professional personnel. Once an aggressive company had found its way onto the national scene its managers and professionals would naturally look for new worlds to conquer. As early as the 1950's many companies entered the rapidly recovering European market before the Common Market erected barriers, and many others recognized that during the second half of the century there would be a quickening of economic activity in other areas: Asia, Australia, Latin America and even Africa. The experience these companies had acquired at home in establishing close links with their customers and in tying the sale of goods to financing and servicing gave them confidence that they could surmount the special problems they would encounter in expanding abroad. Their ventures abroad were also facilitated, directly or indirectly, by foreign loans and grants extended by the Federal Government in the amount of $234 billion between 1945 and 1978.

Since services for consumers have to be provided where the consumers are, economists have long assumed that

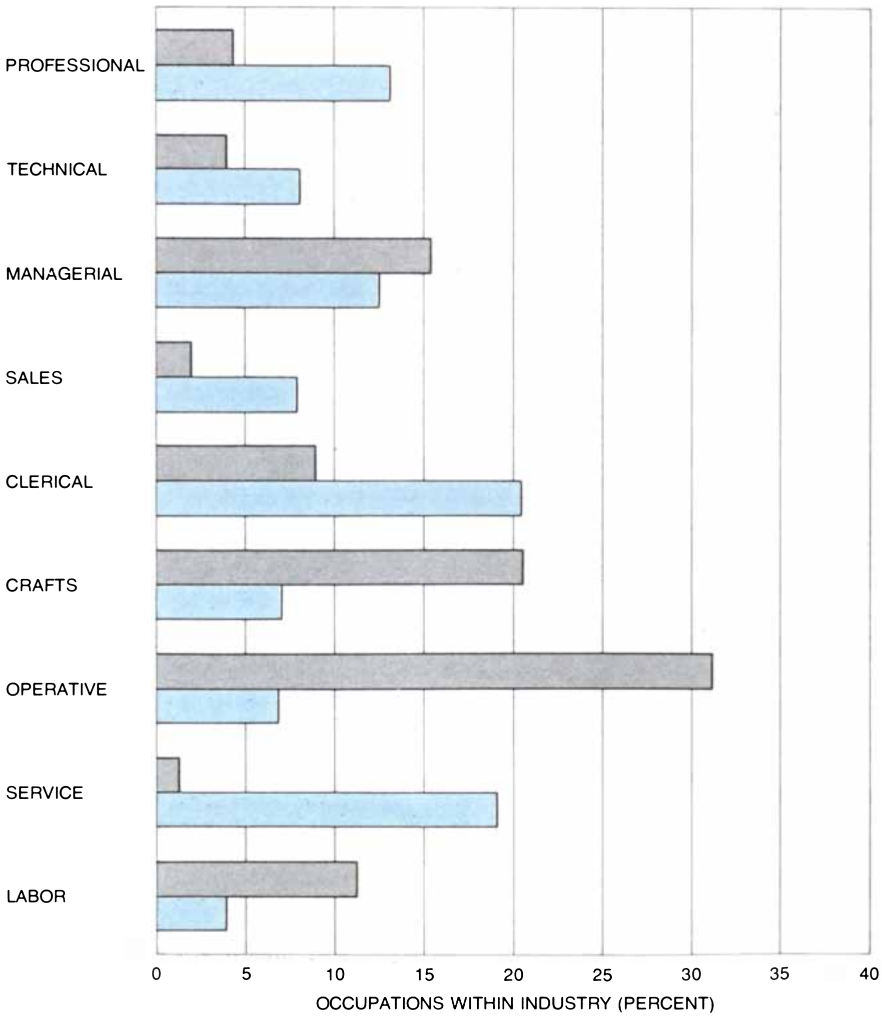


INCREASES IN EDUCATION have contributed to the expanding role of human capital in the national economy. The chart shows the percentage distribution of the civilian labor force in terms of years of school completed in 1957 (*gray*) and 1978 (*color*). In 1948 the median number of years of school completed by members of the civilian labor force was 10.6; in 1978 it was 12.6. One out of three workers had at least attended college and 16.9 percent had finished.

the economies of scale characteristic of manufacturing could not be achieved in service enterprises. Services cannot be produced for inventory and cannot be shipped. That, however, is not the entire story. Improvements in communications, particularly in processing and transmitting numerical data, facilitated the growth of large service companies in the postwar decades by linking together in single enterprises large numbers of small service establishments. Major banks were among the first to develop worldwide systems of branches. Now multiunit hotel chains, automobile-rental companies and fast-food-franchise enterprises have followed the example set by the banks.

The economics of these arrangements are based on the gains that the large service company can achieve through integrated planning, financing, accounting, marketing and similar functions. Even large producer-service firms in law and accounting have increasingly expanded overseas through the establishment of branches, partnerships or franchises. This development helps to explain the surprising fact that legal services have recently emerged as the largest export industry in New York City, outranking its apparel industry.

Government contributed to this transformation as an investor in human and physical capital, as a purchaser of goods and services and as both a direct and an indirect employer. It also contributed through a variety of other activities: as a stimulator of foreign trade by granting loans to foreign purchasers and guaranteeing loans of American exporters, through its expanded regulatory activity, which had a substantial impact on the flow of private investment for such purposes as the construction of nuclear power plants, and through large income transfers, subsidies and interest payments that flowed from one group of citizens to another, amounting in 1978 to $648 billion. These large transfers had uncertain but assuredly large effects on investment, consumption and work.



OCCUPATIONAL DISTRIBUTION in the major sectors of the economy, goods (*gray*) and services (*color*), is charted on the basis of data for 1975. The data are based on the *Survey of Income and Education,* published by the U.S. Department of Labor. In services 42 percent of the workers are in the high-level white-collar classifications, in manufacturing 25 percent.

In conclusion let us consider the implications for policy that stem from the transformation of the economy. To begin with, the proposed reindustrialization of the country collides with the hard economic fact that traditionally structured manufacturing activity cannot be maintained intact while the service sector continues to expand. The loss of competitiveness in some companies and industries is inevitable in an increasingly open world economy characterized by changing technologies, enterprises, labor forces and markets. Import quotas, tariff barriers, export subsidies and the large-scale commitment of Federal credit to shore up weak and failing enterprises would threaten the U.S. with the same consequences as the British policy that has subsidized labor and capital in dying industries and communities. Potentially profitable sectors of the economy would necessarily be deprived of resources they need to optimize their growth.

The U.S. cannot maintain its position among the industrial nations of the world unless it pursues policies that encourage the greater use of resources in which the country has gained a comparative advantage as a result of its generous investment in human capital, such as research and development, management and organization, the development of new products and improved services ranging from financing to marketing. What is needed is not reindustrialization but revitalization of the U.S. economy.

Since people represent the principal input of an advanced economy, and since their contributions will vary according to their endowment, their developmental opportunities and their motivation, the nation should pursue policies aimed at strengthening its human resources. The present scanting of public expenditure on education at every level of government constitutes a reversal of a public policy that has served the nation well over the past 30 years. Adverse demographic and financial circumstances are weakening the nation's major research universities, with the result that the pool of knowledge and the supply of future talent are likely to be diminished. The consequences for the continued growth of the economy may be serious.

The family has a strategic role in the process of human development. Many poor and disorganized families in this country are not able, however, to provide their children with the discipline and motivating experience they need, and the schools are not able to compensate for these shortcomings. As a result many a young person reaches working age unable to get or hold a regular job. In spite of the large Federal effort directed to training and employing disadvantaged adults and youths, which involved expenditures of $64 billion between 1962 and 1978, many of these people have been unable to cross the employment threshold; at best they continue to move in and out of marginal jobs. Only a few industrial companies and labor unions have recognized that they must help the Government's effort to bring the excluded into the work arena. Unacceptable delays in achieving this priority are reflected in the riots, arson and crimes against individuals and property that beset, in varying degree, every American community with large numbers of disadvantaged people.

The obligation and opportunity to seek the enhancement of human capital do not rest alone with government and the schools, nor are they limited to the disadvantaged. The big public, private and nonprofit organizations that loom large in the country's life can do more to improve the work environment in order to motivate their employees to develop their potential and use their skills for both their own satisfaction and the good of the enterprise. Participatory management, incentive rewards, flexible schedules and opportunities for outside assignments are all appropriate working conditions for high-capacity personnel. As the U.S. economy comes to depend even more on trained manpower this challenge becomes greater. It will be met only if large organizations learn to treat their human capital with the concern they have long devoted to their physical capital.

	MANUFACTURING	DISTRIBUTIVE	RETAIL	PRODUCER	CONSUMER	NONPROFIT	GOVERNMENT
PROFESSIONAL	5.4	3.3	1.3	13.5	.5	36	10.2
TECHNICAL	5.2	3.4	.7	9	4.4	12.9	25.1
MANAGERIAL	11.5	17	21.5	17.3	9.6	4	12
SALES	2.8	9.5	25.8	11.7	.6	.2	.1
CLERICAL	11.3	21.2	20.3	34.1	7.4	17.9	33.5
CRAFTS	14.9	15.2	11.1	2.6	7.8	2	7.2
OPERATIVE	42.2	21.3	9.1	2.2	6.8	1.1	2
SERVICE	1.9	2	2.9	8.2	58.6	24.8	6.5
LABOR	5	7	7.4	1.5	4.3	1	3.7

DISTRIBUTION OF OCCUPATIONS in manufacturing and the six service-sector categories (*color*) is presented in terms of percentages of the respective work forces. For example, in the distributive segment of the service sector 3.3 percent of the employees are professionals.

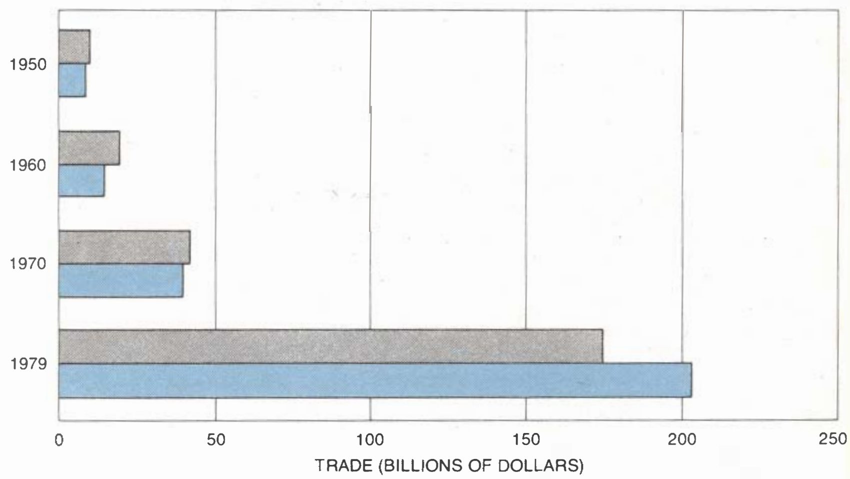
One of the unique characteristics of the U.S. economy is the important role of its private nonprofit institutions and the extent to which government accomplishes many of its missions by contracting with such institutions as well as with companies organized for profit. Although this pluralistic arrangement has served the nation well, the public is becoming aware of shortcomings. A pronounced shortcoming is the absence of reliable data that would enable legislators and the public to judge the effectiveness with which the nonprofit sector employs the large resources it commands.

Current estimates place the unreported income (from transactions kept off the books, from barter, from illicit arrangements and so on) at about 10 percent of the gross national product. Moreover, it is widely agreed that unreported economic activities are growing disproportionately fast. To the extent that these conditions exist the American record of growth, productivity and well-being may be substantially understated.

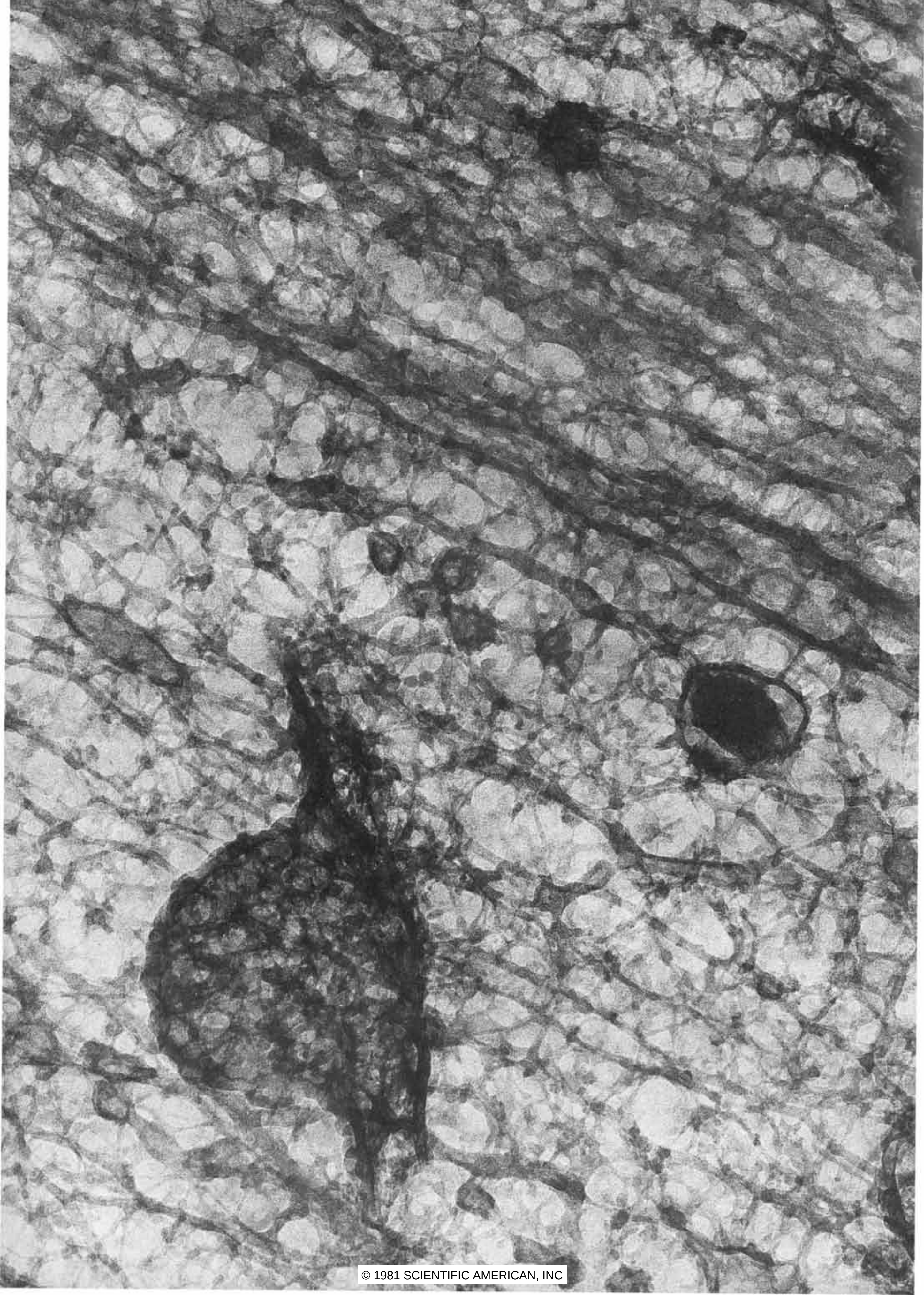
Even if the economy is doing better than the reporting systems suggest, it faces many serious challenges, the solutions to which will not come easily. The shift from manufacturing to services is making it difficult for many poorly educated young people to find a regular job with career prospects. With the Federal budget out of balance and defense expenditures going up the Federal Government will be hard pressed to do more to increase investments that are critical for the nation's continuing economic growth.

Another challenge is that neither government with its monetary and fiscal policies nor the private sector with its wage and price policies has been able to control inflation or to find an answer to the accelerating costs of energy. Moreover, no mechanisms are in place or in sight that will help to keep wage payments in step with gains in productivity. Although the trend toward further internationalization of the world's economy is likely to be beneficial in the long run, many industries and many workers face a bleak future because of the dislocations that accompany the process.

The proposed reindustrialization of an economy dominated by services is an exercise in futility. Americans must unshackle themselves from the notion, dating back to Adam Smith, that goods alone constitute wealth whereas services are nonproductive and ephemeral. At the same time, they should act on Smith's understanding that the wealth of a nation depends on the skill, dexterity and knowledge of its people.



INTERNATIONAL TRADE has a growing role in the U.S. economy, in large part because of the increasing internationalization of American companies. Here exports (*gray*) and imports (*color*) are charted for four different years in terms of current dollars, that is, dollars with the value they had during each of the specified years. Between 1950 and 1979 exports and imports as a percentage of the nation's gross national product rose from 6.8 percent to 15.4 percent.



The Ground Substance of the Living Cell

The high resolution that has been achieved with the high-voltage electron microscope reveals the microtrabecular lattice: a system of gossamer filaments that support and move the cell organelles

by Keith R. Porter and Jonathan B. Tucker

Half a century ago the living cell was regarded as a tiny bag of fluid with only a few structures in it. With the application of the electron microscope to the study of the cell in the 1940's it became apparent that the cytoplasm of the cell is more rigorously organized, having an assortment of fibers, membrane-enclosed vesicles and organelles: among others the mitochondria (the particles that make ATP, the chemical energy currency of the cell), the ribosomes (the particles that coordinate the synthesis of proteins in the cell) and the endoplasmic reticulum and the Golgi apparatus (the systems of vesicles that package proteins for export from the cell). Electron microscopy revealed little, however, about the nature of the indistinct, apparently structureless medium in which the various cellular components are suspended: the cytoplasmic ground substance.

The ground substance nonetheless has several features that immediately suggest a complex internal structure. It displays viscoelastic properties more akin to those of a gel than those of a fluid, it influences the shape of the cell, it regenerates lost components and it incorporates molecular machinery for the directed transport of vesicles, granules and chromosomes. Moreover, the organelles do not flow freely as they would in an unstructured solution but retain a radial distribution around the center of the cell, suggesting that they are restrained by some kind of internal scaffolding.

In the 1970's the understanding of cell structure made a quantum leap with the realization that the cell has an elaborate and dynamic network of skeletal elements. There are at least three chemically distinct cytoskeletal systems: the microtubules, the microfilaments and the intermediate filaments. These structural elements are implicated in the ability of cells to move and to maintain their shape. No structure in the cytoplasmic ground substance was found, however, that could integrate and control the assembly of these different fiber networks so that they could account for the known behavior of cells.

Further insight into this mystery called for an analytic tool more powerful than the standard electron microscope: the high-voltage electron microscope. Although the first high-voltage instrument was built in 1947, it was not until the early 1960's that Gaston-Léopold Dupouy of the Electron Optics Laboratory in Toulouse and V. E. Cosslett of the University of Cambridge developed models for scientific work. Current versions of the high-voltage electron microscope are some 30 feet high, weigh more than 20 tons and are capable of accelerating electrons across a potential drop of a million volts, 10 times that of standard electron microscopes. That voltage gives the electrons sufficient kinetic energy to penetrate thick specimens or even intact cells up to several micrometers thick. (A micrometer is a thousandth of a millimeter.) The resulting image is analogous to an X-ray plate, revealing the internal structure of the cell without the necessity, as is the case at lower voltages, of slicing the cell into sections thinner than .2 micrometer.

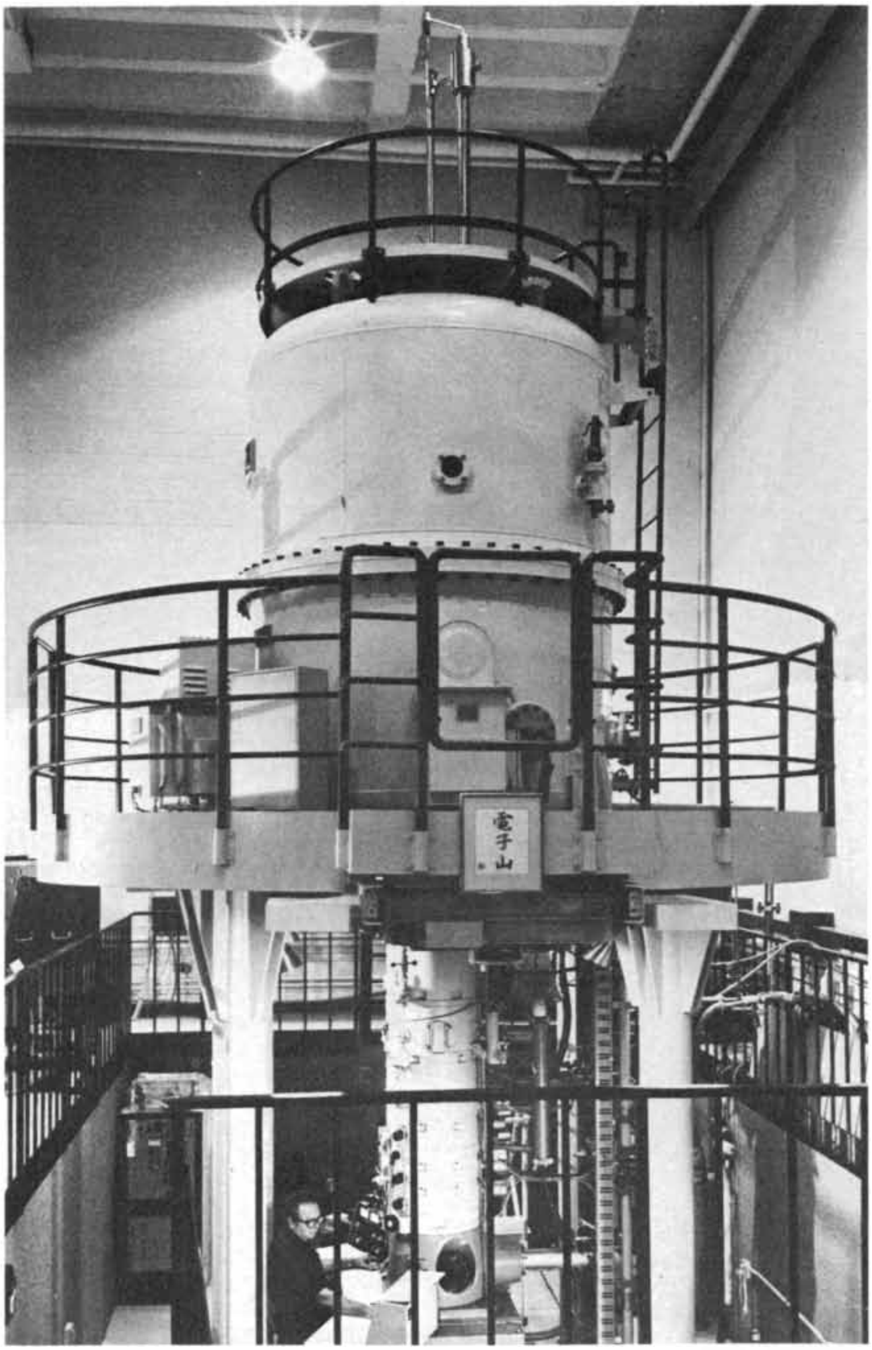
The chief advantage of high-voltage electron microscopy is that it provides information in depth, giving an unparalleled view of the three-dimensional organization of cells. In addition resolution and depth of field are excellent throughout the entire thickness of the specimen. In spite of these advantages biologists have been slow to exploit the full potential of the high-voltage electron microscope; the instrument is currently employed primarily by metallurgists for studying thin films of metals. Although there are several high-voltage electron microscopes in Europe and Japan available for biological research, there are only three being utilized for this purpose in the U.S.: one at the University of Colorado at Boulder, another at the University of Wisconsin at Madison and a third in the laboratories of the New York State Department of Public Health in Albany.

In the early 1970's one of the authors of this article (Porter), working with Ian K. Buckley and John J. Wolosewick, employed the high-voltage electron microscope at Boulder to search for structure in the cytoplasmic ground substance. Cells grown in culture were chosen for study because they tend to spread out on a flat surface and hence can be viewed throughout their volume. Cells were cultured directly on three-millimeter gold grids that had been coated with a plastic film; they were fixed with glutaraldehyde, stained and further fixed with osmium tetroxide and dehydrated with alcohol and acetone. They were then transferred to liquid carbon dioxide and dried by the critical-point method; this method avoids the distorting effects of surface tension, which causes the cells to collapse when they are dried in air. Finally the grid was placed in the vacuum chamber of the

MICROTRABECULAR LATTICE is visible as a meshwork of interlinked filaments in this high-voltage electron micrograph of a tiny area near the edge of a rat cell grown in culture. The lattice is seen only at the very high magnifications of the high-voltage instrument because the microtrabeculae, the individual filaments, are exceedingly fine: about six nanometers, or six millionths of a millimeter, thick. Here the three-dimensional lattice suspends a cell organelle, the lysosome (*small dark body at the right*), and cross-links a bundle of microfilaments (*top*). The lattice also supports a microruffle: an extension that reaches up like a mitten from the cell surface (*bottom*). The small round bodies at a few intersections of the lattice are ribosomes, organelles that mediate protein synthesis. The micrograph, made by Karen L. Anderson of the University of Colorado at Boulder, enlarges the structures some 100,000 diameters.

microscope and the cells were viewed at high magnification.

When the Boulder investigators examined small regions of the cytoplasmic ground substance near the edge of the cell, they noticed an irregular, three-dimensional lattice of slender strands throughout the cytoplasm and coextensive with the outer membrane of the cell. The interlinked filaments appeared to suspend the various cell systems and organelles. Moreover, larger cytoskeletal elements such as the microtubules and the microfilaments were coated with a matrix material coextensive with the individual lattice filaments, indicating that the larger elements were integral with the lattice and essentially suspended in it. The lattice was reminiscent of the trabecular structure of spongy bone, and so it was named the microtrabecular lattice. Since then this lattice morphology has been observed in the ground substance of all eukaryotic (nucleated) cells that have been examined.

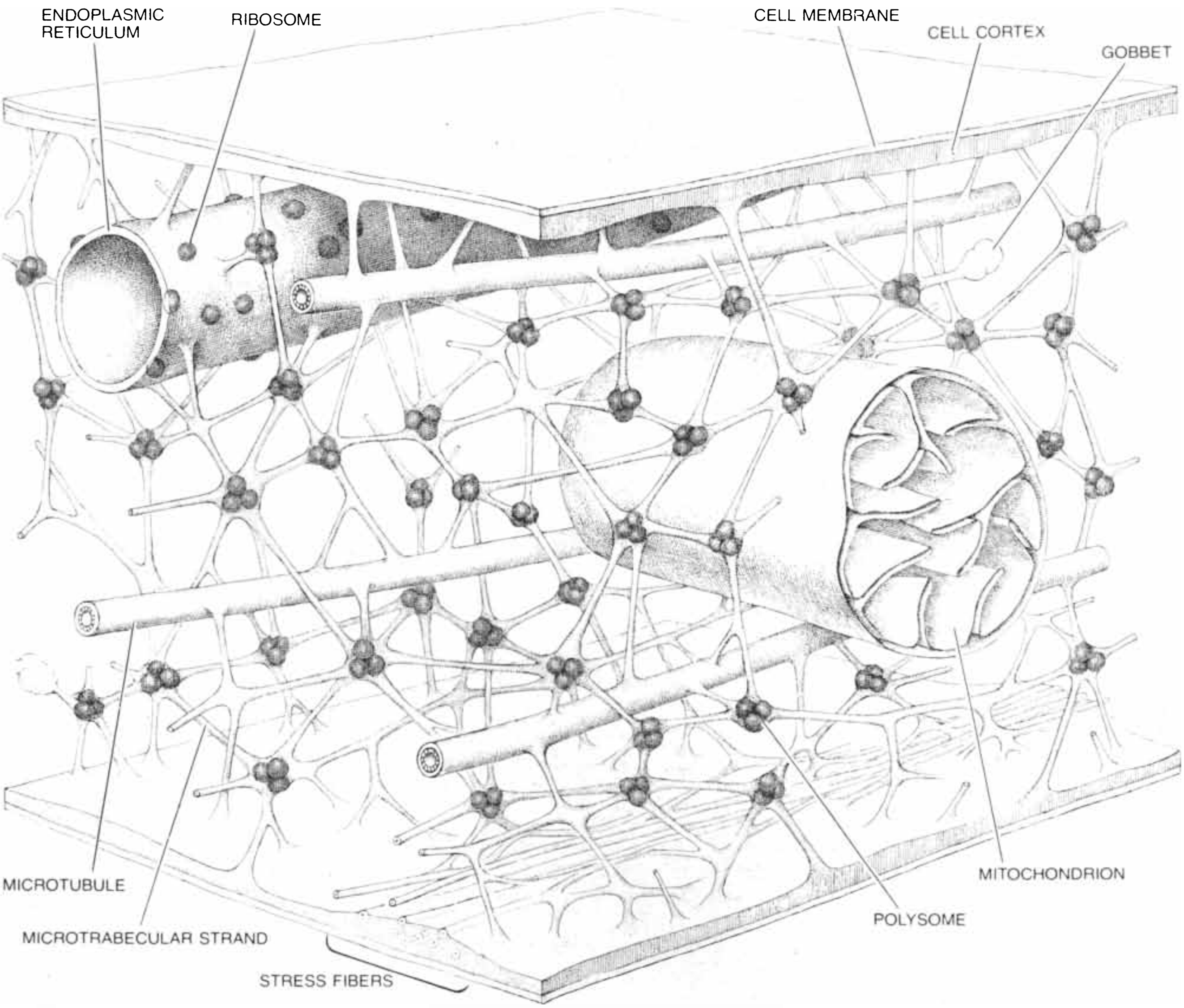


HIGH-VOLTAGE ELECTRON MICROSCOPE, installed in 1973 at Boulder, is one of only three such microscopes in the U.S. devoted to biological research. The instrument is 32 feet high and weighs 22 tons. In order to reduce vibration it has its own foundation independent of that of the building. Electrons are injected into an accelerator tube, enclosed in the large pressure tank in the center, that increases their energy to a million volts. These electrons are capable of producing images of thick sections and even intact cells without significant loss of resolution, which cannot be done with electrons at lower voltages. The electron-optical column must be large to accommodate the powerful electromagnetic lenses required to deflect and focus the high-energy electrons. Because of the intensity of X rays generated where the electrons strike a surface the microscope column and base are shielded with lead to protect the operator (in this case George Wray). The Japanese characters on the sign translate as "Electron Mountain."

When the first paper fully describing the microtrabecular lattice appeared in *The American Journal of Anatomy* in 1976, a number of investigators were skeptical. They contended that the harsh conditions required to fix and stain the cells for microscopy had caused proteins in solution in the cytoplasm to condense around the various organelles and fibers, creating the misleading impression of a structured matrix. The Boulder workers accordingly began intensive studies to validate their conclusion, employing several different methods of fixation and dehydration to prepare whole cells for high-voltage electron microscopy. In every instance the lattice morphology was present in the ground substance. Moreover, protein-rich solutions exposed to the same treatments failed to show any artifact of fixation closely resembling the microtrabecular lattice.

The lattice does vary somewhat in the finer details of its morphology depending on the method of fixation. For example, in cells fixed with glutaraldehyde the individual microtrabeculae can be more than 10 nanometers thick at the ends but only two or three nanometers thick in the middle. (A nanometer is a millionth of a millimeter.) When cells are fixed in milliseconds by freezing them in propane cooled to the temperature of liquid nitrogen (−196 degrees Celsius), the microtrabeculae are thicker and have a more constant diameter of about 15 nanometers. This difference in lattice morphology is apparently due to the fact that glutaraldehyde is a fairly slow fixative, requiring seconds to exert its full effect, whereas freezing is virtually instantaneous. Hence during glutaraldehyde fixation the microtrabeculae have time to contract or elongate, altering their diameter and their relation to one another and to the cytoplasmic organelles.

If the microtrabecular lattice is real, why had it not been seen in conventional electron micrographs? Actually it had been, but too indistinctly to be recognized as such. In individual thin sections the microtrabeculae appear as wispy, amorphous structures that are easily mistaken for artifacts; the organization of the filaments into a lattice is nowhere apparent. Part of the problem is that the section is so thin the interconnections of the lattice elements cannot be seen clearly. Moreover, the epoxy-resin embedding material required for thin-sectioning has the effect of obscuring the



MODEL OF THE MICROTRABECULAR LATTICE, some 300,000 times its actual size, was derived from hundreds of images of cultured cells viewed in the high-voltage electron microscope. The model illustrates how the microtrabecular filaments are related to the other components of the cell cytoplasm: the substance of the cell outside the cell nucleus. In the model the microtrabeculae suspend the elongated structures of the endoplasmic reticulum (the system of interconnected channels within the cell where some of the proteins manufactured by the cell are sequestered), the mitochondria (the organelles that manufacture ATP, the universal fuel of the cell), the microtubules (the complex fibers that serve many functions of cell structure) and the microfilaments buried in the cell cortex (the layer of material just under the outer membrane of the cell). At junctions of the microtrabecular lattice are polysomes: organized clusters of ribosomes.

cellular structures having an electron-scattering behavior similar to that of the epoxy itself.

In an effort to resolve the problem the Boulder group began preparing thin sections of cells by embedding them in a water-soluble matrix: polyethylene glycol. After this material has been cut into sections for electron microscopy it can be dissolved away and the sections can be dried by the critical-point method to prevent their collapse. When such sections are viewed in the standard electron microscope, they possess a lattice morphology closely resembling the one seen in unembedded whole cells in the high-voltage electron microscope. Other less disruptive preparative techniques, such as the freeze-fracture method employed by John E. Heuser of Washington University, yield a similar lattice structure. Thus although some aspects of the microtrabecular lattice have turned out to be artifacts (and it needs to be said that the rigorous preparation methods of electron microscopy make everything seen in the electron microscope to some degree an artifact), enough evidence remains to indicate that there is indeed a structured ground substance in cells.

Perhaps the most compelling support for the existence of the microtrabecular lattice is that its structure varies in response to changes in cell shape or in the cellular environment. For example, at a low temperature (four degrees C.) cultured cells become spherical, their lowest-energy shape. This change is paralleled by the sequential dismantling of the internal skeletal structures of the cell. First the microtubules disassemble, followed by the microfilaments. Finally the microtrabecular lattice deforms but does not decompose completely. Some of the microtrabeculae sever their connections with the lattice and gather into minute "gobbets," creating large gaps in the lattice. These gaps allow for the Brownian motion (free diffusion) of cytoplasmic organelles, a phenomenon that is seen in chilled cells but not in those at physiological temperatures. If the cell is returned to body temperature (37 degrees C.) for as little as five minutes, there is a dramatic restructuring of the lattice to its former morphology.

Similarly dramatic effects on the structure of the microtrabecular lattice can be elicited by cytochalasin B, a drug that inhibits a wide variety of cell movements. After 10 minutes of exposure to the drug the lattice becomes very coarse: the microtrabeculae thicken and elongate and the intertrabecular spaces correspondingly enlarge. Soon after the drug is removed the lattice returns to its normal finely divided structure. Other agents and conditions (excessively high or low osmotic pressure, altered levels of specific ions such as magnesium or calcium and the presence of metabolic inhibitors) also induce unique and reversible structural changes in the lattice. It would be difficult to account for such wide fluctuations in morphology if the lattice were simply an artifact of cell fixation.

What are the implications of the microtrabecular lattice for cell structure and function? It seems clear that the lattice divides the cell into two phases: the polymerized, protein-rich phase of the lattice itself and the fluid, water-rich phase filling the interstices in the lattice (the intertrabecular spaces). With a water content of 50 percent, this construction appears to give the cytoplasm its gel-like consistency.

The reality of these two phases was demonstrated in experiments by the Boulder group (including Karen L. Anderson) in which the water content of cells was manipulated. The cells were immersed in solutions containing high or low concentrations of sugar molecules, causing water to enter or leave the cell by osmosis. When the cells were examined in the high-voltage electron microscope, it was found that intertrabecular spaces expanded when water entered the cell and shrank when water was sucked out of it. When water was removed from the cells and then replaced soon afterward, the cells survived. This observation may be explained by the ability of the lattice to preserve the cell's viability by protecting it from normal fluctuations in water content.

The microtrabecular lattice may also play a role in organizing the enzymes in the cytoplasm. In 1929 Rudolph Peters of the University of Cambridge contended on purely logical grounds that the intricate biochemical reactions known to take place in the cytoplasm were too rigorously choreographed to be ruled by the random collisions of enzymes and their substrates (the substances whose transformation they effect) diffusing through a structureless medium. He postulated the existence of "some tenuous network by the action of which the cell's enzymatic activities are coordinated." Peters' speculation was remarkably prescient. There is now good evidence that a number of enzymes involved in metabolic pathways (for example the pathway of glycolysis, the conversion of glucose to pyruvate) are bound to structures within the cytoplasmic ground substance. It follows that such enzymes may be incorporated into the microtrabecular lattice, perhaps in a nonrandom orientation such that one enzyme passes its substrate along to the next enzyme in the pathway. In this way enzymes that are not bound to cell membranes might nonetheless be spatially coordinated.

CONVENTIONAL ELECTRON MICROGRAPH of a thin section through the cytoplasmic ground substance of a cultured human cell shows little evidence of the microtrabecular lattice, although thicker filaments can be seen along with parts of the endoplasmic reticulum and mitochondria. The ground substance does show numerous randomly distributed wispy strands, some of which run between microtubules (*long rodlike structures*). The epoxy in which the structures are embedded for electron microscopy obscures small areas. The micrograph, which enlarges the structures some 14,000 diameters, was made by John J. Wolosewick at Boulder.

The watery phase that fills the intertrabecular spaces is enriched in small molecules such as glucose, amino acids, carbon dioxide and oxygen. Because of the pervasive lattice structure the distances through which molecules diffuse in the water-rich phase are short. The effect is to quicken the rate at which substrates travel from one enzyme to the next.

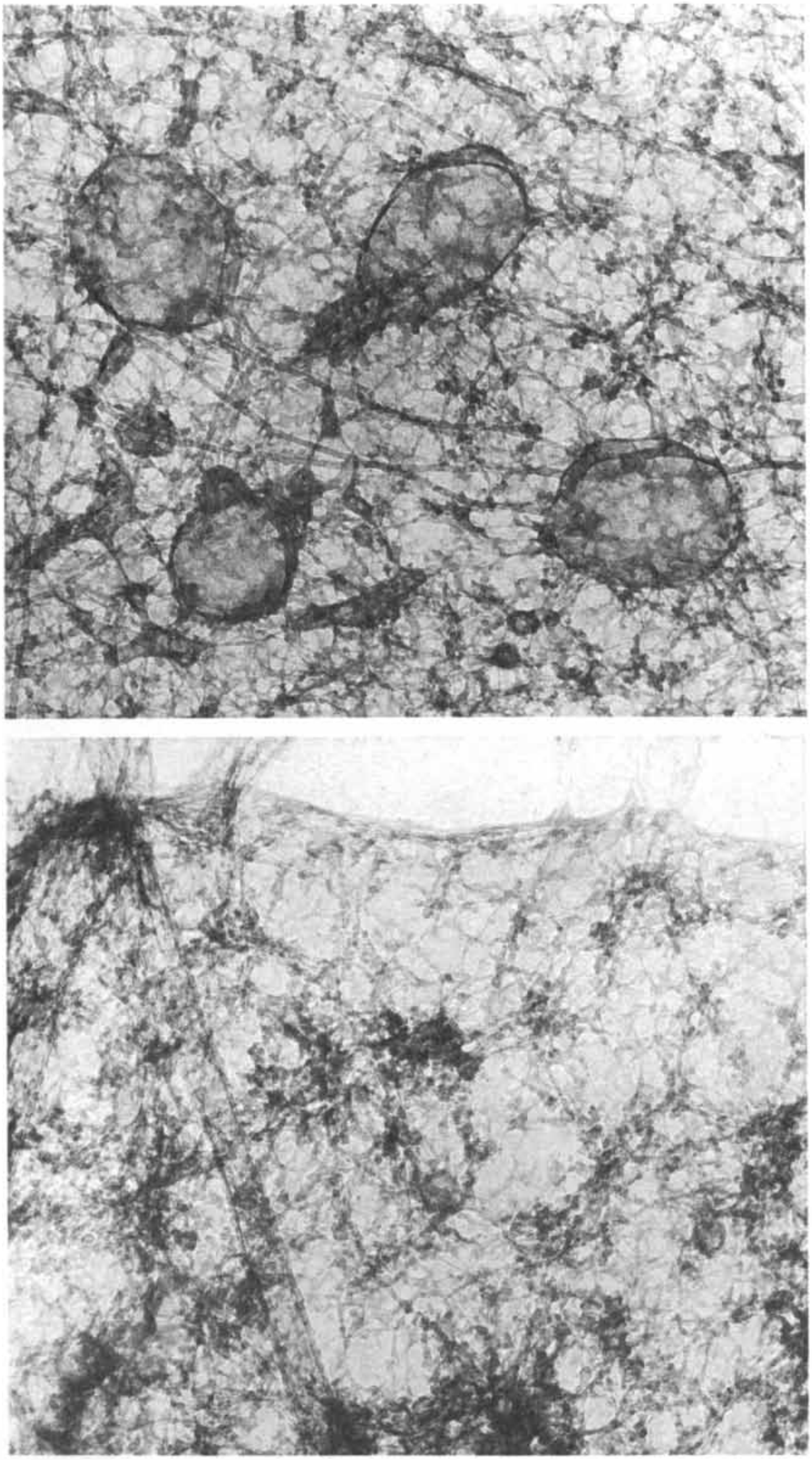
The lattice also appears to play an important role in cell differentiation and the associated protein synthesis: the process by which polysomes (clusters of ribosomes bound to a strand of messenger RNA) translate the genetic code into the specific sequence of amino acids in a protein molecule. On the basis of thin-section electron micrographs it has been generally believed the translation process is mediated by polysomes floating free in the cytoplasm. High-voltage electron micrographs reveal a strikingly different morphology: the polysomes lie at the intersections of the microtrabecular lattice, like spiders in a three-dimensional web.

This remarkable organization has significant implications, since it means that once the proteins have been manufactured they become part of the material associated with the microtrabecular lattice. The material is then available in a nonrandom way for the controlled assembly of microtubules, microfilaments and other cellular structures. For example, microfilaments can be organized into parallel bundles called stress fibers or into an amorphous meshwork. They are continually being assembled and disassembled in different parts of the cytoplasm as the cell goes about its other activities. Examining the role of the microtrabecular lattice in the assembly of cellular structures may be crucial for understanding how cells grow and specialize.

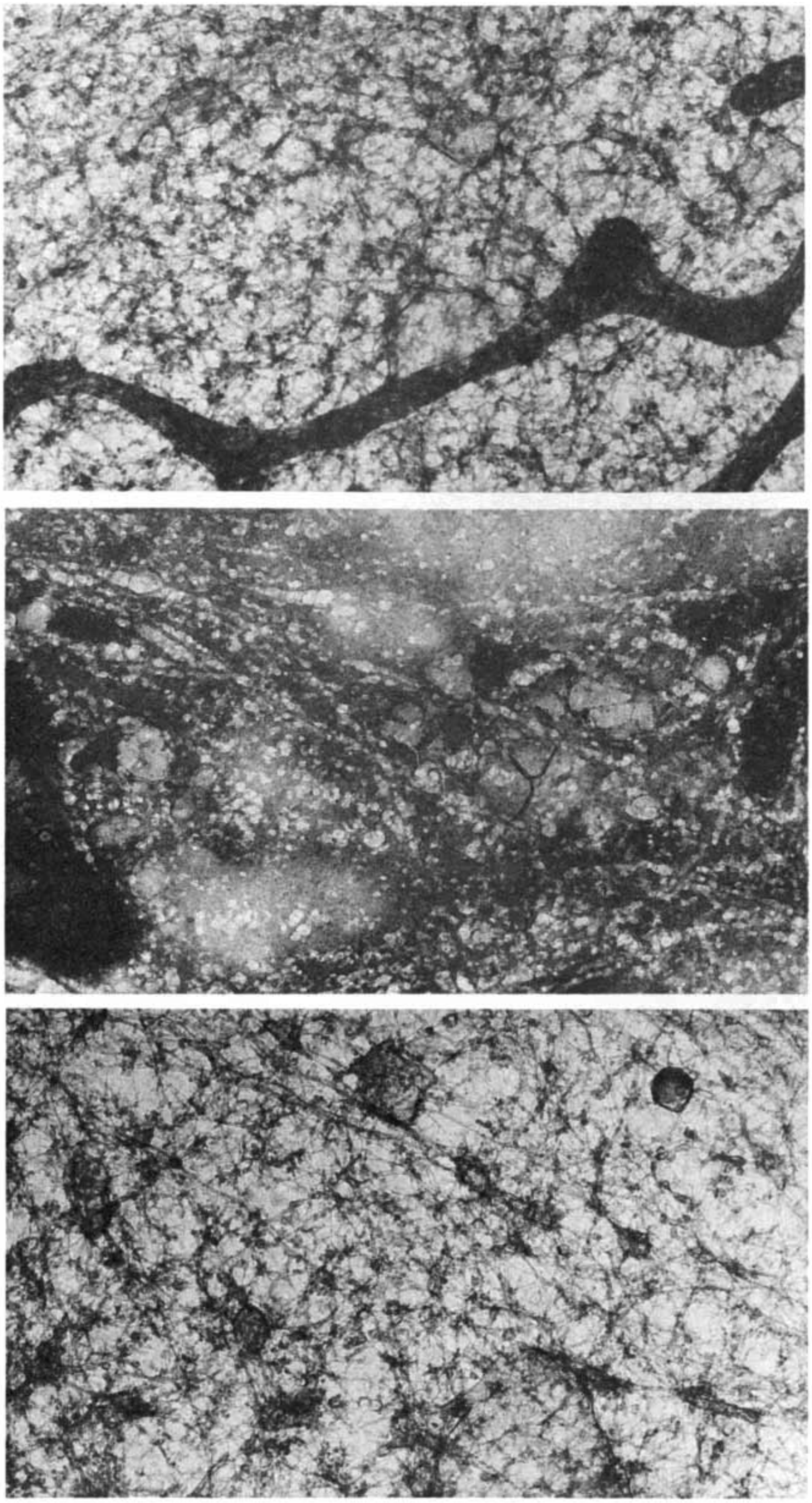
Since the microtrabecular lattice pervades the entire cell, it appears to link the disparate components of the cytoplasm—organelles and structural fibers—into a single structural and functional unit called the cytoplast. What is the relation between the lattice and the elaborate cytoskeletal network formed by the three major fiber systems? Manfred Schliwa, working at Boulder, recently explored this question by extracting cultured cells with a detergent (Triton X-100) in the presence of a buffer that stabilizes the cytoskeletal components. Treatment of the cells with the detergent for 60 seconds removed virtually all membrane-enclosed structures, including the outer cell membrane, the mitochondria, the endoplasmic reticulum and the envelope of the cell nucleus. In addition nearly all the microtrabecular strands were extracted, suggesting that they have physical properties and a chemical composition different from those of the stable cytoskeletal elements.

The material left behind was a filamentous framework approximating the shape of the cell and made up of microfilaments, microtubules and intermediate filaments. In the high-voltage electron microscope the three-dimensional organization of these filaments was well preserved and virtually identical with that of the intact cell. Moreover, the various cytoskeletal filaments appeared to be interconnected at multiple sites, explaining why the disruption of one filament system with specific inhibitors may induce changes in the organization of another filament system.

The microtrabecular filaments, together with the cytoskeletal elements, appear to act as a form-fitting support for the outer cell membrane and for highly asymmetric extensions of the cell surface. For example, cilia and flagella, the whiplike appendages that propel protozoa and spermatozoa, are complex structures made up of multiple sets of microtubules linked by "spokes" and "bridges." When these specialized structures are sectioned with a polyethyleneglycol matrix and examined in the high-voltage electron microscope, the spokes and bridges look exactly like microtrabeculae. Similarly, the microfilament bundles in microvilli, the extensions of the cytoplasm seen in cells of the intestine, appear to be associated with microtrabeculae. These observations suggest that the morphology within cilia, flagella, microvilli and other specialized ex-



EFFECT OF LOW TEMPERATURE on the microtrabecular lattice, perhaps due to a cold-induced shutoff of the cell's metabolic machinery, is dramatic. Normal cells chilled at four degrees Celsius for three hours assume a spherical shape. This change in shape is accompanied by a number of internal changes. First the microtubules disassemble within a few minutes. Then large gaps appear in the microtrabecular lattice, as is apparent in the high-voltage electron micrograph at the bottom. The normal morphology is shown in the micrograph at the top. The severed microtrabeculae appear to contract into large clumps as the individual microtrabeculae are transformed into tiny "gobbets." When the chilled cells are reincubated at 37 degrees C., the gobbets are transformed back into microtrabeculae; after 30 minutes of incubation the lattice is essentially back to normal. The sensitivity of the lattice to changes in the cellular environment suggests the lattice is a physiologically active component of the cell. The micrographs, which were made by Anderson, enlarge the structures some 78,000 diameters.



SPACES AMONG THE MICROTRABECULAE (*top*) contain bulk water enriched in molecules essential to cell metabolism. (The long dark structures are parts of mitochondria.) Since the dimensions of the spaces are small, the diffusion of these molecules from one enzyme to the next takes only microseconds. The existence of the intertrabecular spaces has been demonstrated by altering the concentration of the extracellular medium, thereby causing water to leave or enter the cell by osmosis. If the concentration outside the cell is decreased below that in the cytoplasm, water is sucked out of the cell and the intertrabecular spaces are reduced (*middle*). If the concentration is increased, water enters the cell and the spaces expand (*bottom*). The magnification of these micrographs, made by Anderson, is some 35,000 diameters.

tensions of the cell is the same as that of the cytoplast as a whole: it consists of microtrabeculae and intertrabecular spaces.

A cytoplasmic organelle that seems to be intimately associated with the microtrabecular lattice is the microtubule-organizing center (MTOC), which is near the center of the cell. A typical MTOC observed in animal cells is made up of a pair of centrioles (each consisting of a cylindrical array of microtubules) oriented at right angles to each other and surrounded by a complex of dense satellite proteins that vary in form, number and disposition from one cell type to another. The centrioles and their satellites serve as initiating sites for the assembly of microtubules, and it is the patterned distribution of these cytoplasmic elements that influences the shape of the cell.

MTOC's are continuous with the microtrabecular lattice; there is no sharp structural demarcation between them. Moreover, the information for the location of this complex in the cell center appears to be built into the ground substance. In the region of the cytoplasm immediately surrounding the MTOC the microtrabecular lattice is extremely dense. As a result this part of the cell, known as the centrosphere, is highly gelatinous. The density of the lattice falls off with distance from the MTOC toward the cell periphery.

In the early stages of cell replication the centrosphere divides. Its two progeny form the poles of the mitotic spindle, which direct the separation of the duplicated chromosomes in the course of mitosis, or cell division. At the end of mitosis each daughter cell contains a centrosphere surrounded by its own microtrabecular lattice. This observation raises a provocative question: What is the origin of the structural information stored in the lattice? Is it encoded in the genes or is it passed on physically from one generation to the next in the course of cell division during embryonic development? Remarkably, some experimental information suggests that the latter possibility may be the case.

In 1964 T. M. Sonneborn and his colleagues at Indiana University did some elegant experiments on the inheritance of cellular structures in the ciliated protozoan *Paramecium*. This single-cell animal is quite complex, possessing among other things numerous cilia attached to the cell cortex: the part of the microtrabecular lattice just under the outer cell membrane. Sonneborn and his co-workers performed microsurgery on these cells, altering the normal pattern of structure in various ways to yield, for example, cells with two or more mouths or anuses, or sections of the cortex having cilia with inverted polarity. The Indiana investigators expected that such imposed upsets of normal cell structure would either be lethal or be rapidly corrected by the cells' genes. Much to their surprise the induced abnormalities were inherited by the progeny at successive cell divisions! The changes were perpetuated during divisions for more than a year, through more than 700 generations of cells.

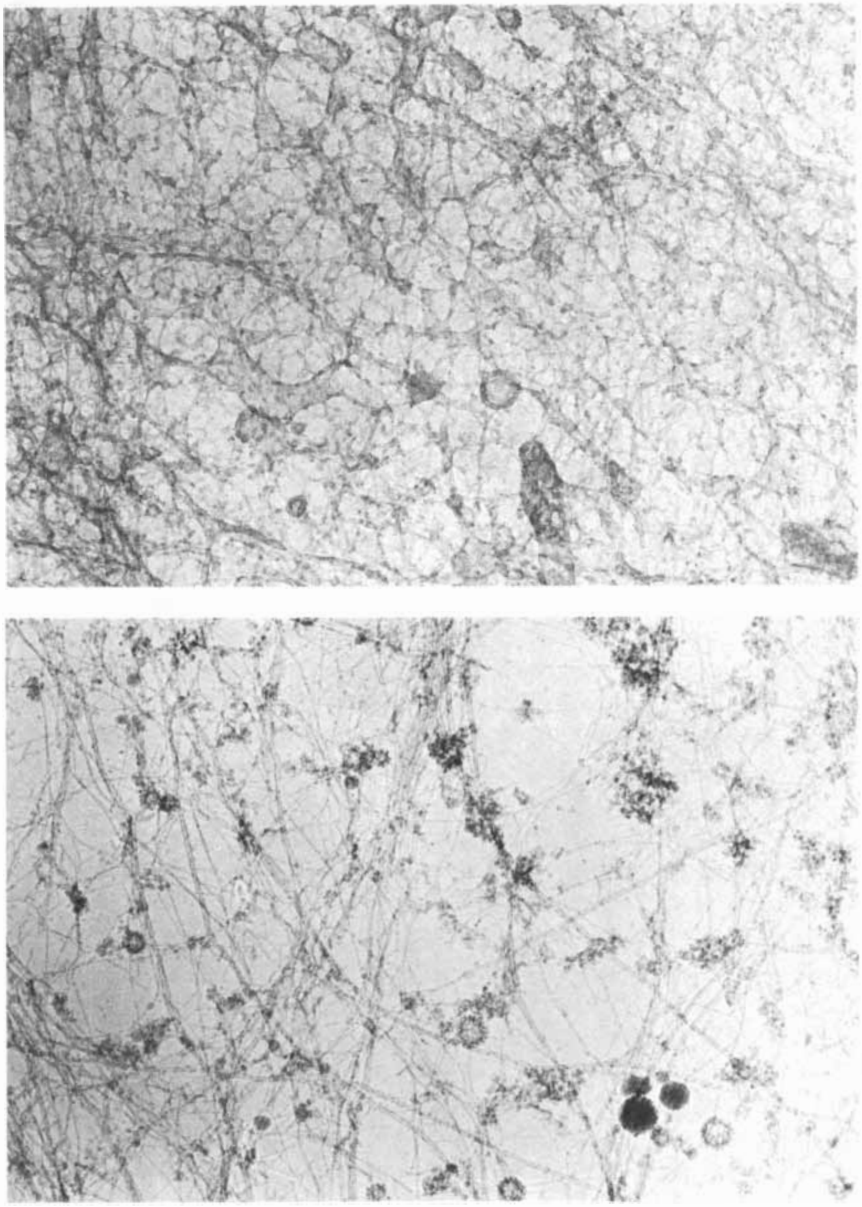
Sonneborn concluded that form information was being transmitted from one generation to the next through a nongenetic mechanism: the basic form information of the cell is stored in the structure of the cytoplasmic ground substance, which is passed on to both daughter cells at mitosis. This ordering and arranging of new cell structure under the influence of preexisting cell structure was termed cytotaxis. Sonneborn wrote: "While the genes determine the molecular building blocks and, through their properties, the kinds of molecular associations that can occur, the associations that actually do occur depend also on those that already exist."

The idea that cytoplasmic structure and motility are related goes back to the early French cytologist Félix Dujardin, who proposed in 1835 that all cells were composed of a material called sarcode, which had both structural and contractile properties. It now appears that the microtrabecular lattice plays such a role, serving as a kind of dynamic cellular musculature that continually redistributes and reorients the organelles and cytoskeletal fibers as the cell goes about its activities. The lattice is not in any sense static. Rather, it is assumed to be constantly changing, displaying local contractions and deformations with resulting increases or decreases in the size of the intertrabecular spaces.

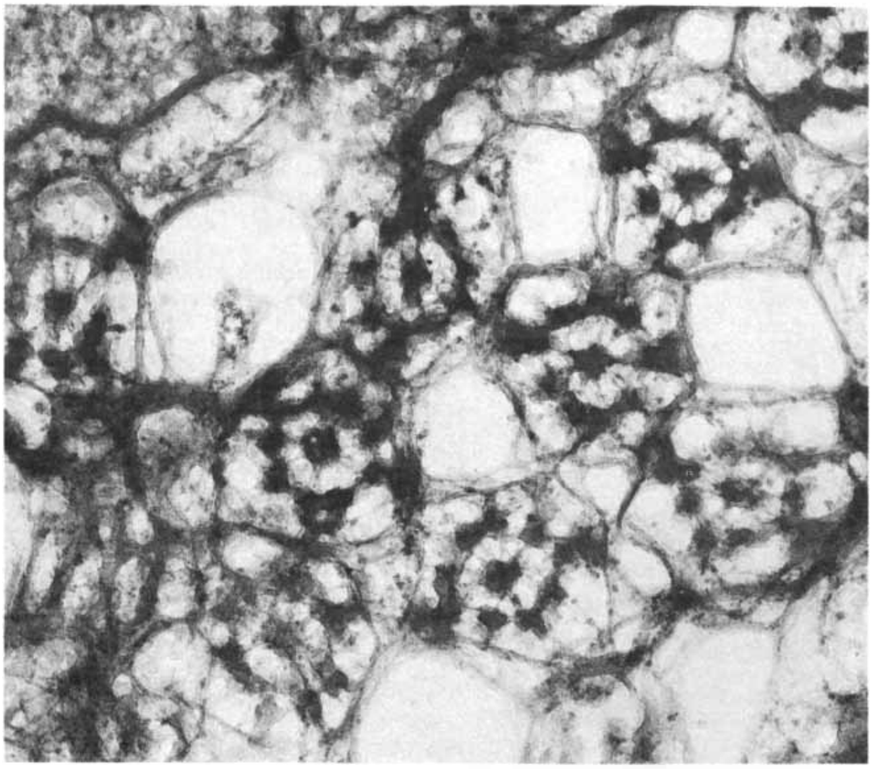
Cells are capable of many different modes of internal movement, ranging from the slow, continuous movement of chromosomes during mitosis to the discontinuous backtracking movement of vesicles and granules along nerve fibers. Chromatophores, or pigment cells, are of particular interest to cell biologists because they display both smooth and jerky intracellular movements. The conversion of chemical and electrical energy into the mechanical work of transporting pigment granules within pigment cells is therefore an important topic for investigation.

Chromatophores, situated in the skin of certain amphibians and the scales of many fishes and reptiles, operate under hormonal and nervous control and enable these animals to darken or lighten their skin color to match the color of their surroundings, the better to escape predators. In experiments by the Boulder workers the red pigment cells called erythrophores, obtained from the squirrelfish (*Holocentrus ascensionis*), were chosen because they can be removed easily and maintained in isolation in tissue culture. Thousands of bright red granules, dispersed throughout the erythrophore's cytoplasm, aggregate in the center of the cell in response to the hormone epinephrine. Each granule accelerates to a velocity of between 15 and 20 micrometers per second and traverses the entire distance in a straight line and without stopping.

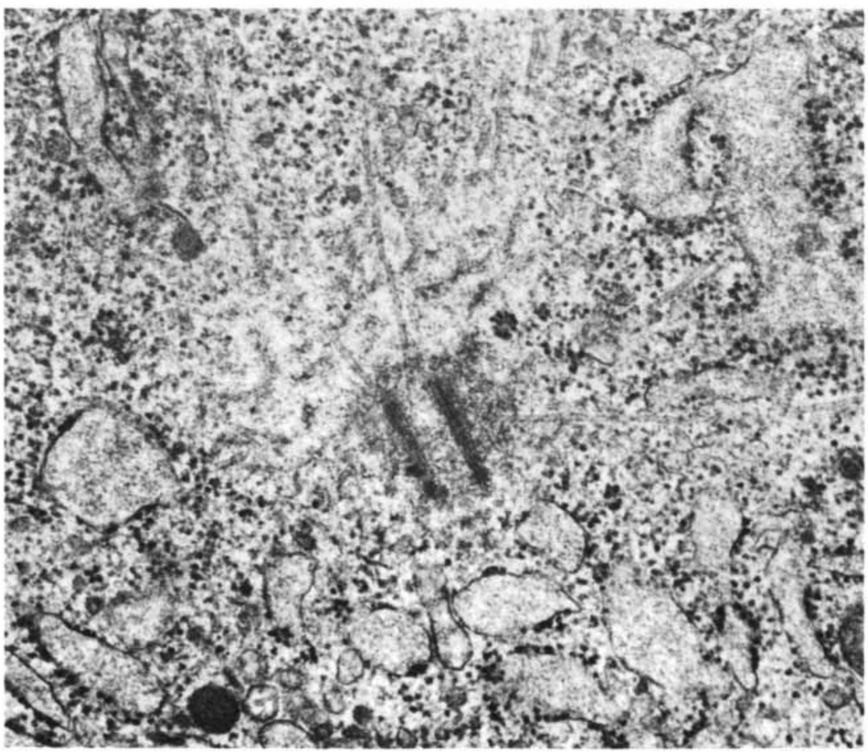
The constant velocity of the aggregating granules is analogous to the movement of chromosomes in cell division but is 500 times faster. In three to five seconds aggregation is complete; the granules are all clustered in a bulging mass at the cell center surrounded by clear cytoplasm, with the cell retaining its original nearly circular outline. The cell nucleus and the mitochondria do not move with the granules and either are left behind in the peripheral cytoplasm or are independently displaced toward the cell center. When aggregation occurs simultaneously in millions of erythrophores, the fish can change its



TREATMENT WITH DETERGENT illustrates the relation between the microtrabeculae and the filaments of the cytoskeleton: the microtubules, the microfilaments and the intermediate filaments. The micrograph at the top shows the untreated cell with the microtrabecular lattice and the organelles intact. The micrograph at the bottom shows a similar cell treated with the detergent Triton X-100, which removes the lattice material covering the cytoskeletal filaments, organelles and membranes. What remains is the cytoskeletal elements, which form a network that approximates the shape of the cell. The microtubules, microfilaments and intermediate filaments can be distinguished by their different thicknesses. The micrographs, which enlarge the structures some 40,000 diameters, were made by Manfred Schliwa at Boulder.



FLAGELLA OF MOUSE SPERMATOZOA, seen end on in this high-voltage electron micrograph, are made up of sets of microtubules bridged by spokelike structures closely resembling microtrabeculae. Such evidence indicates that specialized extensions of the cytoplasm such as flagella, cilia and microvilli share the lattice structure pervading the rest of the cell. The micrograph, which enlarges the structures some 55,000 diameters, was made by Wolosewick.



CENTRIOLE, which incorporates the two parallel structures at the lower right center, appears in this conventional electron micrograph of a thin section of a cell. It is one of a pair, oriented at right angles to each other. The centrioles and their "satellite" material form a complex that directs the assembly and orientation of microtubules. Three microtubules are seen projecting from the satellite material surrounding the parallel structures of the centriole. The lattice is dense in the vicinity of the complex. The enlargement is some 36,000 diameters.

color from bright red to off-white in a matter of seconds. Pigment dispersion is slower than aggregation, requiring six to 10 seconds for completion, and each granule travels in a discontinuous manner, with many starts and stops and even backtracking motions over short distances.

Erythrophores aggregate their pigment under conditions of low cellular energy, such as when the cell dies or when it has had prolonged exposure to a metabolic inhibitor. Therefore the aggregation process is powered by the release of stored potential energy. In contrast, dispersion consumes ATP and stores potential energy. The slow and discontinuous nature of the dispersion may result from its energy-dependent character.

It is well known from standard electron micrographs that the erythrophore contains a radial array of microtubules extending out from the cell center like rays from the sun. On the basis of this morphology some investigators proposed in the late 1960's that pigment aggregation is generated by some kind of sliding interaction of the microtubules and the pigment granules. This hypothesis is now considered an unlikely one, because there are few if any close contacts between the microtubules and the granules.

In 1977, in an effort to explore the possible connection between the microtrabecular lattice and pigment migration in the erythrophore, the Boulder group (including H. Randolph Byers) examined intact isolated erythrophores in the high-voltage electron microscope. As in other cells, they found a space-filling network of fine, elongated microtrabecular strands connecting the pigment granules, the microtubules and the cell membranes. In order to study further the mechanism of pigment migration they fixed erythrophores in various stages of pigment aggregation and dispersion and examined them in the high-voltage instrument. It turned out that the movement of the granules is accompanied by dramatic structural changes in the lattice.

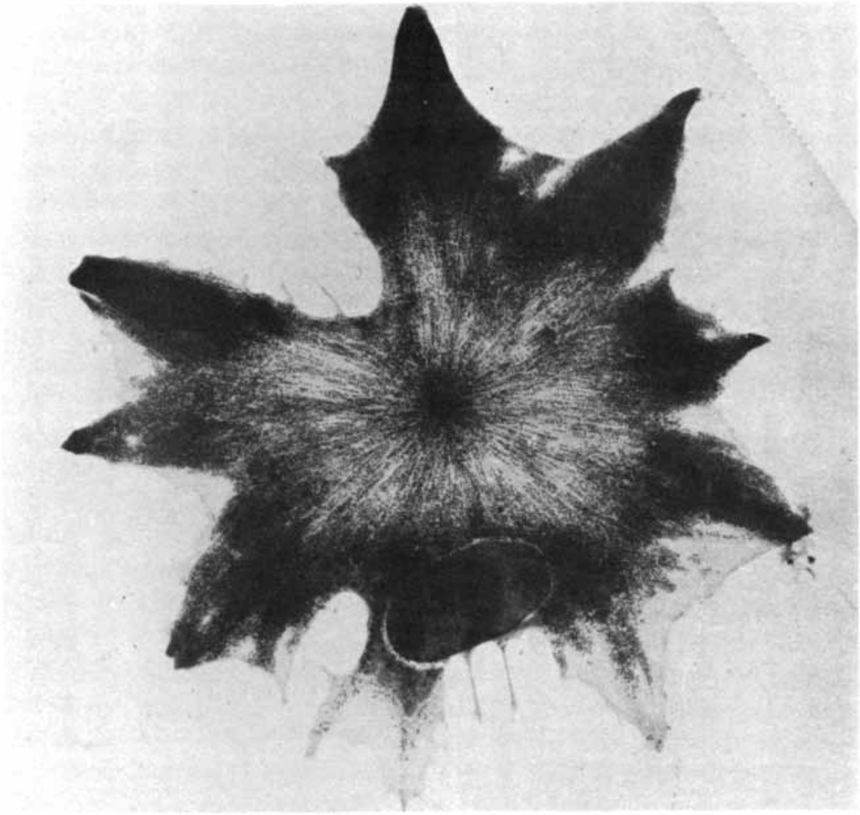
During aggregation the microtrabeculae shorten and thicken, drawing the granules together into clusters and transporting them passively toward the cell center along linear paths defined by the radial array of microtubules. As the lattice contracts, numerous microtrabeculae sever their connections with the granules and shrink back onto the surface of the microtubules and the cell cortex to form gobbets some 20 nanometers in diameter. After the granules and the associated lattice material have withdrawn from the cell periphery the upper cell cortex collapses onto the lower one. When this happens, the radially arrayed microtubules, which do not move with the granules, become readily apparent. The lattice, on the other hand, is now evident only in those regions of the cell immediately adjacent to the central granule mass and along some bundles of microtubules. In those areas the lattice is highly compact: the microtrabeculae are about half their normal length and are beaded at regular intervals with 20-nanometer gobbets.

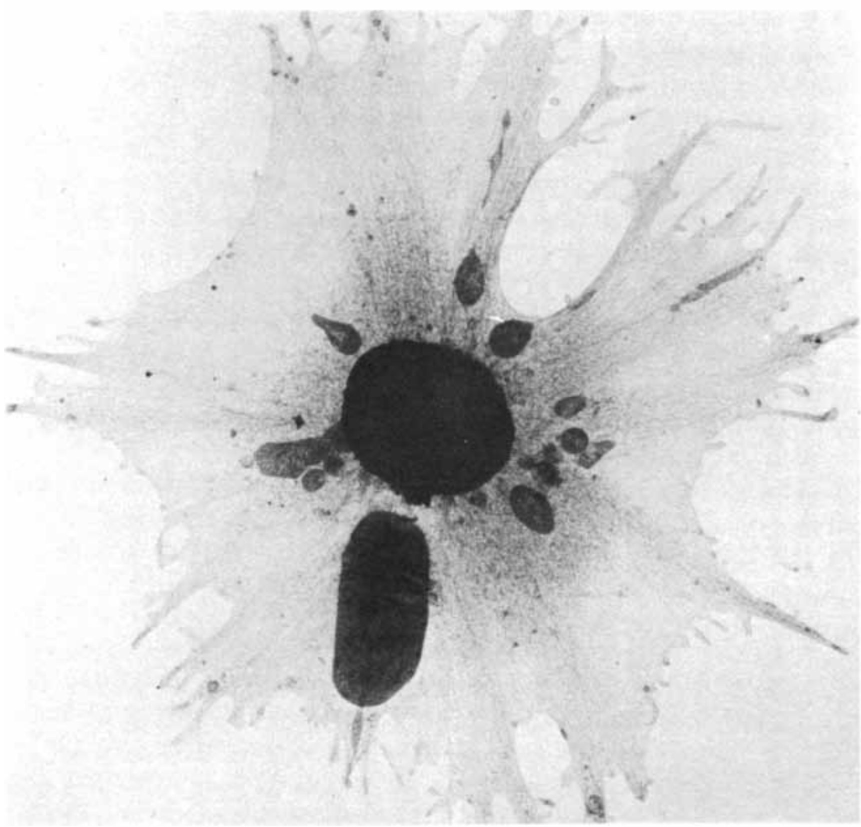
As dispersion is initiated the individual microtrabeculae elongate, carrying the pigment granules back toward the cell periphery and reconstructing the lattice. At the same time there is a lengthening of the gobbets that were left behind in the cell periphery. Together with the gobbets the microtubule array apparently serves as a structural "memory" to guide the rebuilding of the lattice. Indeed, when the population of microtubules is disrupted with the drug colchicine, the granules are disorganized and cannot disperse.

The work done by Katherine J. Luby-Phelps at Boulder has strengthened the hypothesis that the lattice mediates pigment migration. She studied the effects on pigment migration of utilizing metabolic inhibitors to deplete the erythrophore's supply of ATP. The treated cells exhibit a lattice with a disjointed appearance, characterized by a large proportion of incomplete or free-ending microtrabeculae. They can therefore disperse their pigment only partially. In the absence of ATP the individual microtrabecular filaments are apparently unable to elongate and restructure the lattice. It therefore seems that through the expenditure of chemical energy (ATP), potential energy is stored in the lattice during dispersion and converted into kinetic energy during aggregation. The shuttling motions of the pigment granules in the resting erythrophore may be the result of a continual tug-of-war between the entropic tendency of the lattice to contract and its ability to expand in response to energy input, like a rubber band that is being stretched.

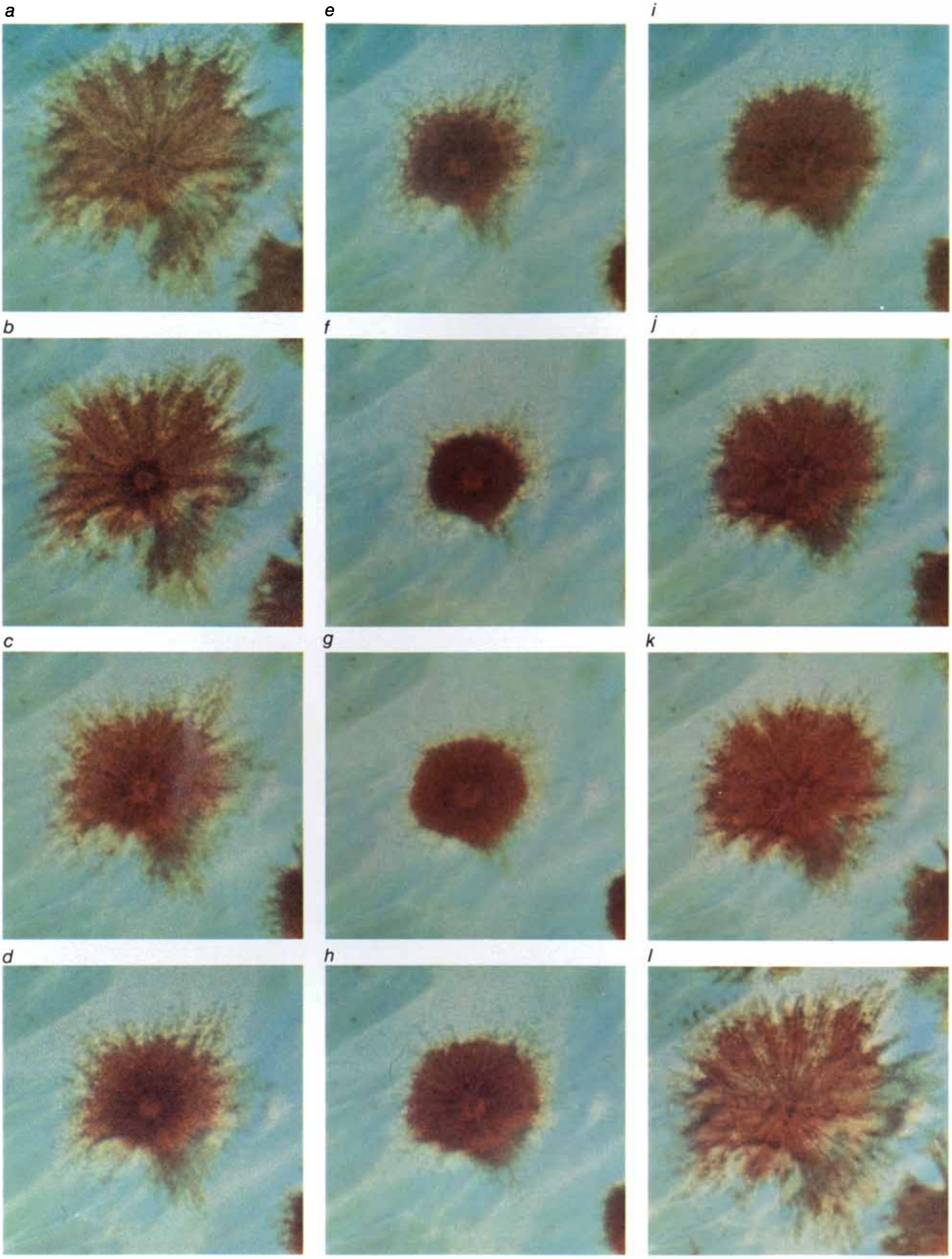
The contractile state of the lattice may be controlled by calcium ions, since they are known to regulate muscle contraction and a variety of motile activities in nonmuscle cells. Indeed, Luby-Phelps has found that an increase in the level of free calcium ions in the erythrophore triggers the contraction of the individual microtrabeculae and leads to pigment aggregation. The requirement for ATP during dispersion may be, at least in part, to power an uptake mechanism that removes free calcium ions from the intertrabecular spaces, so that the microtrabeculae can return to their elongated state.

The most probable source of free calcium ions is a membrane-enclosed system of tubular channels that threads its way among the pigment granules within the erythrophore. A polygonal system





MOVEMENT OF PIGMENT GRANULES in the erythrophore, or red-pigment cell, of the squirrelfish *Holocentrus ascensionis* is illustrated by these two high-voltage electron micrographs of whole cells. In the cell at the top the pigment granules are aligned along the radial array of microtubules. The cell nucleus (*lozenge-shaped object*) is situated toward the periphery of the cell and does not seem to interfere with the alignment of the granules. In the cell at the bottom the granules are all clumped into a dense mass at the center of the cell. The micrographs, which enlarge the cells 3,500 diameters, were made by Katherine J. Luby-Phelps at Boulder.



AGGREGATION AND DISPERSION of the red pigment granules in an erythrophore of the squirrelfish are seen in this series of frames from a motion picture made with a light microscope. The magnification is 150 diameters. The simultaneous aggregation of pigment in millions of erythrophores distributed throughout the scales of the squirrelfish enables it to change its color from red to off-white in a matter of seconds. Frames *a–f* show the aggregation process, which involves a rapid, continuous (uniform-velocity) motion of the granules toward the cell center and takes from three to five seconds. Frames *g–l* show the dispersion process, a slower, discontinuous (nonuniform-velocity) motion back to cell periphery and takes about eight seconds. The film was made by Judith L. Fleming at Boulder.

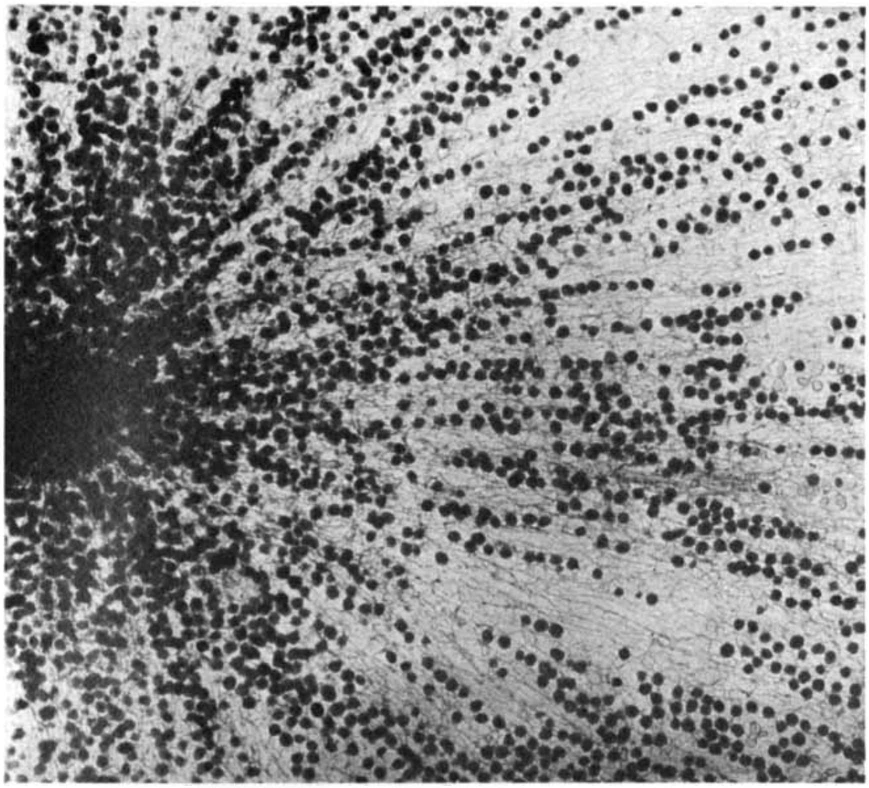
of these tubules extends throughout the cell, in contact with the lattice and in some places with the upper and lower cell cortex. Such a system could deliver the calcium to all parts of the cell in response to a nerve impulse. (The erythrophore has many nerve connections.) In this respect the tubule system of the erythrophore may function analogously to the sarcoplasmic reticulum of muscle, which responds to a nerve impulse by discharging calcium, thereby triggering muscle contraction.

A full explanation of how the microtrabecular lattice expands and contracts awaits a detailed understanding of its molecular composition and structure. This biochemical organization must clearly allow elastic behavior and at the same time encourage the formation of multiple interconnections, both with the lattice itself and with microtubules, microfilaments, intermediate filaments and organelles.

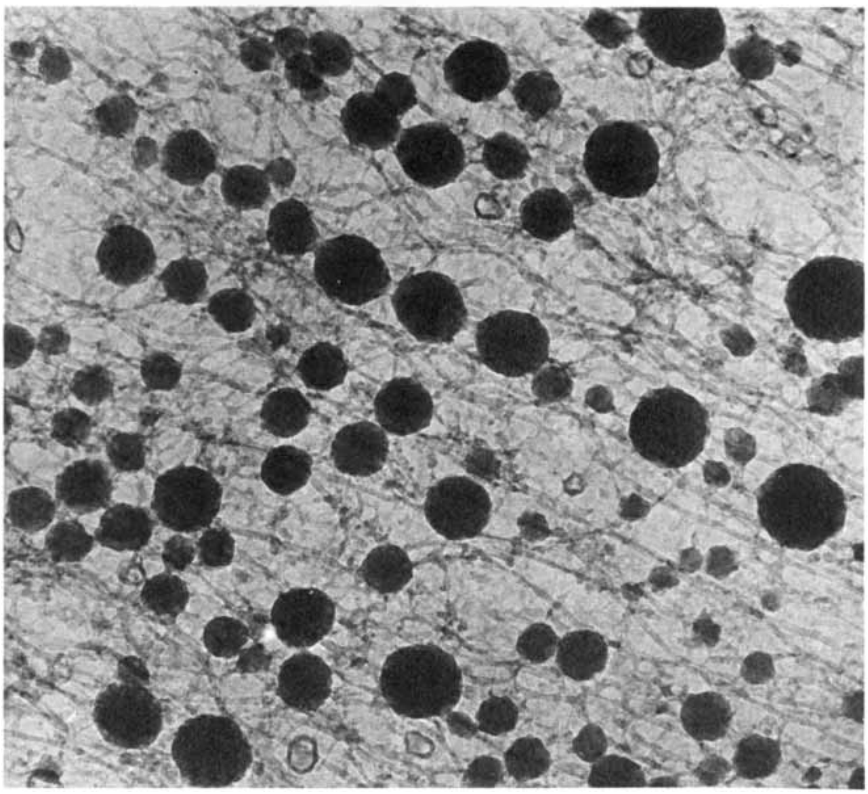
When cells are extracted with detergent and the extracted proteins are separated by two-dimensional gel electrophoresis, more than 100 proteins are separated. This indicates that the lattice is a complex polymer of many proteins. So far a few of the proteins have been identified, including actin and myosin (the major contractile proteins of muscle cells, also ubiquitous in nonmuscle cells), tubulin (the building block of microtubules) and two specific microtubule-associated proteins. Although much of the matrix structure may represent a loose association of proteins that are readily extractable, there is evidence for the existence of a true filamentous backbone in the lattice. Many of the microtrabeculae may have a core of filamentous actin that is sheathed by other proteins.

Another promising but still embryonic field of research concerns differences in the structure of the microtrabecular lattice in normal cells and in cancer cells. The lattice of cancer cells is denser than that of normal cells and shows signs of a lack of normal organization, particularly in the distribution of polysomes.

In conclusion, the cytoplasmic ground substance, long assumed to be a homogeneous, protein-rich solution, is now believed to possess a highly intricate structure and behavior of its own. The microtrabecular lattice organizes the diverse components of the cell into a functional unity—the cytoplast—and mediates regulated and directed transport within the cell. It is also becoming increasingly apparent that the control of cell shape and of cell movement depends on the integrated functioning of the microtubules, the microfilaments, the microtubule-organizing centers and the microtrabecular lattice. A new frontier in the study of cell structure and cell physiology has clearly been identified.



LINEAR ARRANGEMENT OF PIGMENT GRANULES in the dispersed erythrophore is apparent in this high-voltage electron micrograph. The dense cluster of granules at the left is at the cell center. Bundles of microtubules project toward the periphery, and the microtrabecular lattice of the ground substance is evident between the granules and the microtubules. The micrograph, which enlarges the structures some 17,500 diameters, was made by Luby-Phelps.



SPHERICAL PIGMENT GRANULES are seen embedded in the microtrabecular lattice in this high-voltage micrograph made by Luby-Phelps. The granules are coated with lattice material, and some of them are clearly linked by microtrabeculae. Microtubules course diagonally across the field from the upper left to the lower right. The magnification is 70,000 diameters.

The Eruptions of Mount St. Helens

The volcano's current cycle of activity is part of a pattern extending over 4,500 years. Indeed, its violent eruptions of last year were predicted by volcanologists on the scene

by Robert Decker and Barbara Decker

The violent eruption of Mount St. Helens on May 18 of last year was one of the most closely monitored, most intensively photographed and most fully documented volcanic eruptions in history. It was also the first volcanic eruption in the 48 contiguous states of the U.S. since the considerably milder eruptions of Lassen Peak from 1914 to 1917. The eruption of Mount St. Helens displaced 2.7 cubic kilometers of volcanic rock (including .5 kilometer of new magma, or liquid rock), devastating an area of more than 500 square kilometers and causing one of the largest avalanches in recorded history. In the past century the Mount St. Helens outburst was clearly surpassed in magnitude only by the eruptions of Santa Maria in Guatemala in 1902, Krakatoa in Indonesia in 1883 and Katmai in Alaska in 1912, which respectively expelled some five, six and 12 cubic kilometers of magma (with the volumes reduced to correspond to the density of solid rock).

Even Katmai is dwarfed by ancient eruptions that buried thousands of square kilometers under huge deposits of ash and rock tens to hundreds of meters thick in Japan, New Zealand, Central America, the western U.S. and many other volcanic regions of the world. The volume of material expelled by these enormous eruptions ranged from 100 cubic kilometers to more than 1,000. Did prehistoric volcanism exceed anything conceivable today, or has human experience provided too brief a perspective? Most geologists think the latter is true. The eruptions of the past century are probably only small samples of the volcanic energy still at the earth's command.

In the months since the eruption of Mount St. Helens the energy released by the eruption has often been described in terms of the energy released by nuclear explosions. The comparison is useful but somewhat misleading. Not only is the source of energy totally different; so too is the rate of energy release, or power. The thermal and mechanical energy released at Mount St. Helens on May 18, 1980, was about 1.7×10^{18} joules. Since a nuclear explosive rated at one megaton releases about 4.2×10^{15} joules, the Mount St. Helens eruption was equivalent to a 400-megaton explosion, one nearly eight times more powerful than that of the largest nuclear device ever detonated. The comparison is misleading because substantially all the energy of a nuclear explosion is transformed into thermal and mechanical energy in a flash, so that the almost instantaneous power in watts of a nuclear explosion is essentially the same as the energy in joules. (One joule for one second equals one watt.) In the Mount St. Helens eruption the 1.7×10^{18} joules was spread over nine hours, for an average power of about 5×10^{13} watts (1.7×10^{18} divided by 32,400 seconds). The volcano's sustained power output might therefore be compared to the serial detonation of some 27,000 Hiroshima-size bombs: nearly one a second for nine hours. For another comparison, the power generated by Mount St. Helens on May 18 was on the order of 100 times the generating capacity of all U.S. electric-power stations.

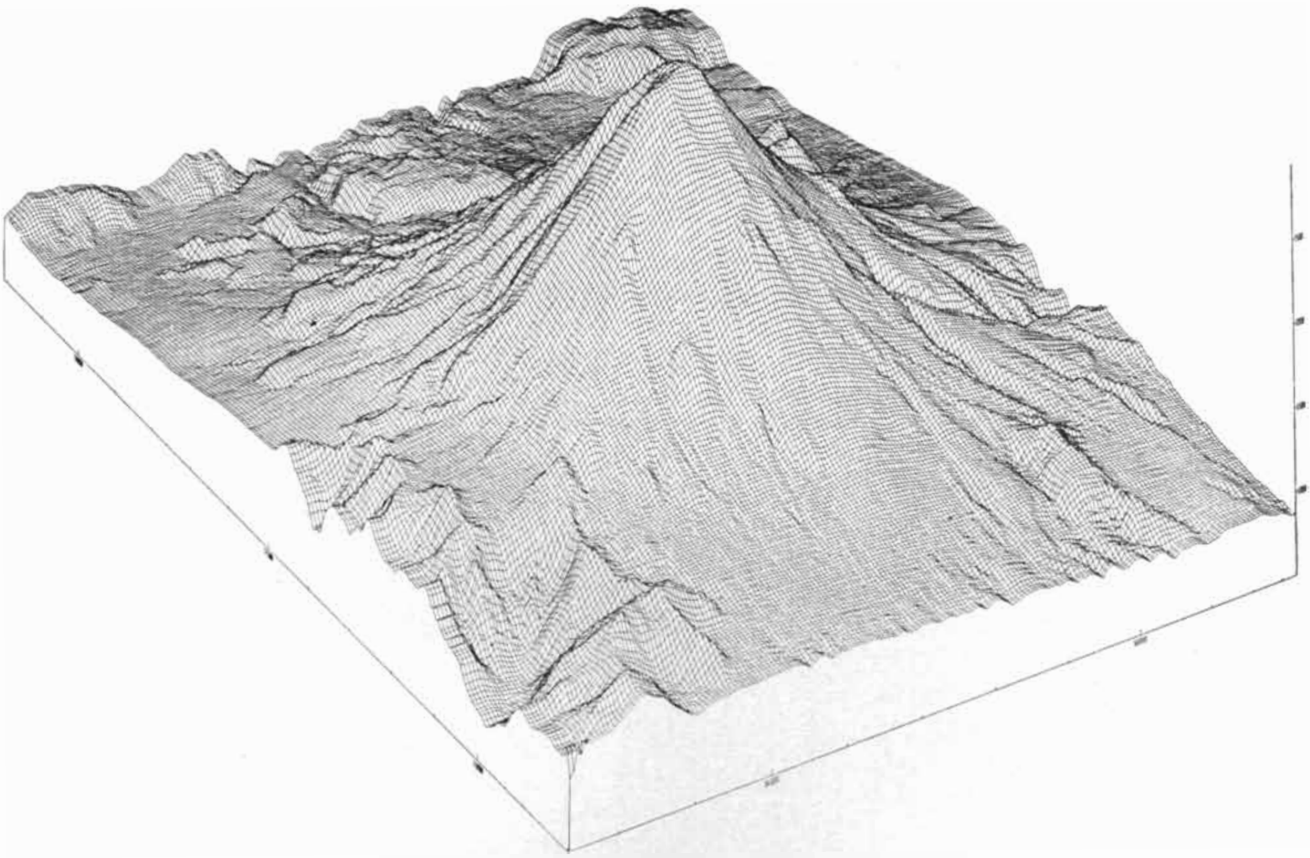
Mount St. Helens has erupted during at least 20 intervals over the past 4,500 years, often enough for the Indians of the American Northwest to know it as Loowit, the Lady of Fire. Before the 1980 eruptions Mount St. Helens was last active between 1831 and 1857. It is one of 15 major volcanoes in the Cascade Range, stretching northward from Lassen Peak in California to Mount Garibaldi in British Columbia. The Cascades are part of the "ring of fire," the volcanic ranges that nearly surround the Pacific. The mountains are thrust up at subduction zones where the moving tectonic plates of the Pacific Basin plunge under the confining plates to the west, the north and the east. The subduction of the Pacific plates generates great earthquakes and provides the molten rock and pressure that powers the volcanoes of the ring of fire. In historic times alone nearly 400 volcanoes have been active on the margins of the Pacific Basin.

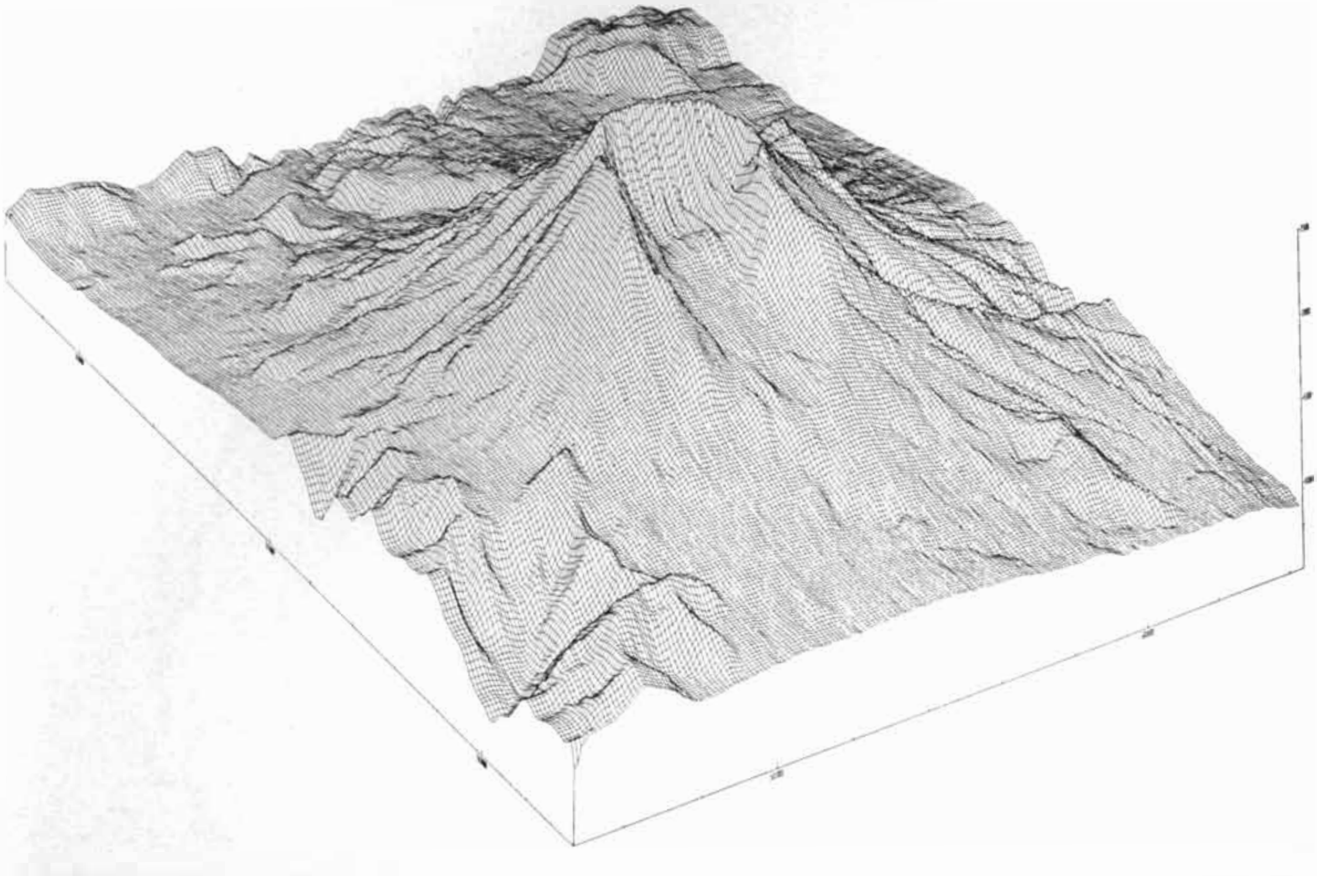
The level of volcanic activity seems to be governed in part by the rate at which the thrusting plates plunge under the bordering plates. In Indonesia and Japan, where the annual subduction rate is six or seven centimeters, there is usually at least one volcanic eruption per year. In the Cascades the lower incidence of eruptions appears to be related to the lower rate of convergence of the North American plate and the Juan de Fuca plate immediately to the west: two or three centimeters per year.

In the early 1960's an effort to evaluate the hazards presented by the dormant but potentially active volcanoes of the Cascade Range was undertaken by Dwight Crandell and Donal Mullineaux of the U.S. Geological Survey. They began their program with a study of volcanic deposits in the vicinity of Mount Rainier, 75 kilometers north-northeast of Mount St. Helens. Mount Rainier had last erupted sometime between 1820 and 1854. Guided by carbon-14 dates of earlier eruptions and following the geologist's dictum that what has happened before can happen again, Crandell and Mullineaux were able to forecast in a general way the potential hazards of each volcano they studied.

Crandell and Mullineaux' report on Mount St. Helens was issued in 1978. They concluded that Mount St. Helens had a bad record: over the preceding 4,500 years it had been more active and more explosive than any other volcano in the contiguous 48 states. In that period its eruptive products included lava domes too viscous to flow from their vent, large volcanic-ash falls full of lumps of pumice, flows of pyroclastic rock (hot fragments shattered and fluidized by eruptive activity), flows of lava and massive flows of mud in the stream valleys radiating from the volcano. The average interval between eruptive periods was 225 years.

On the basis of their study Crandell and Mullineaux stated: "In the future Mount St. Helens probably will erupt





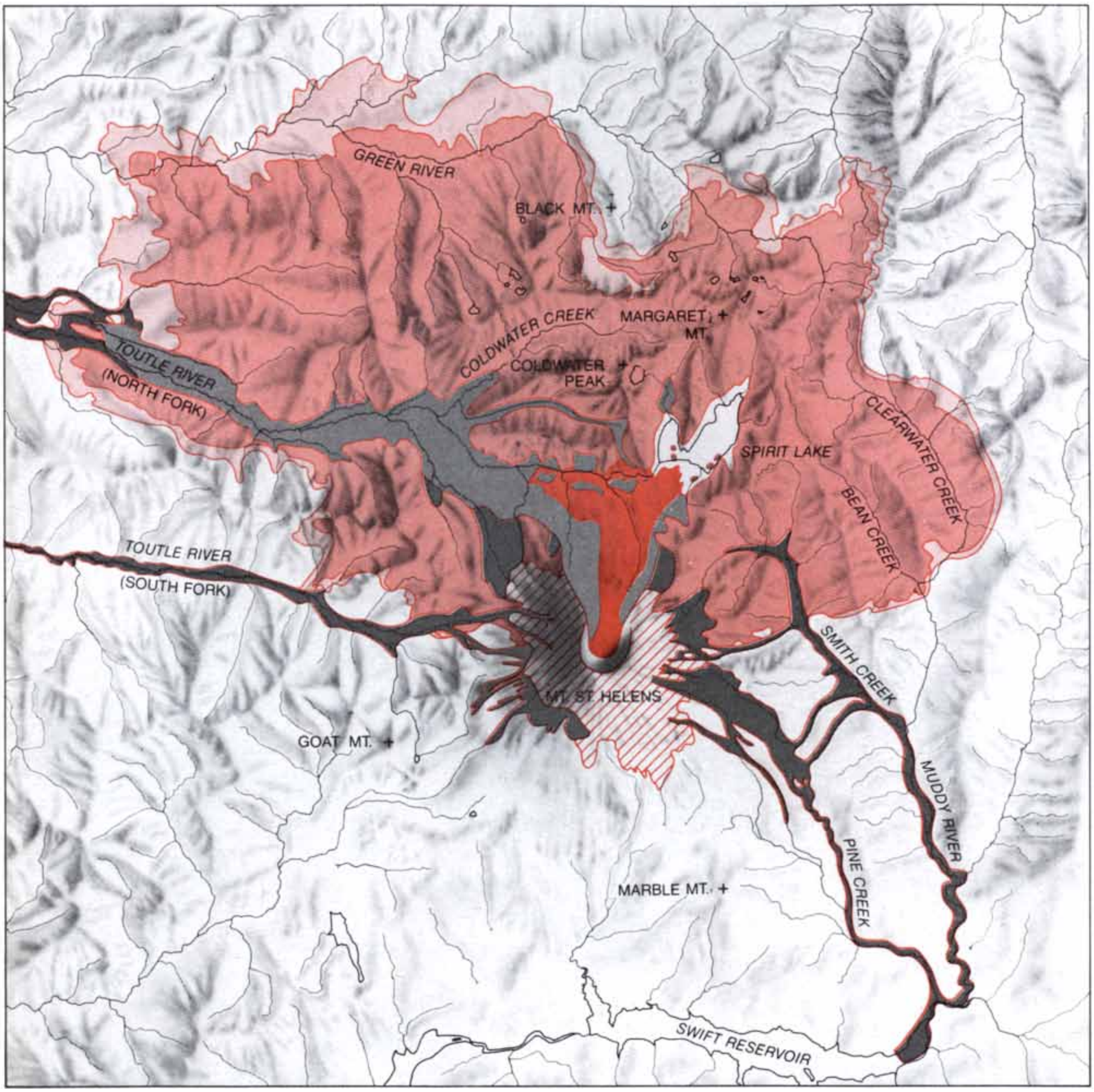
COMPUTER-GRAPHICAL MODELS of Mount St. Helens before the explosive eruption of last May 18 (*top*) and afterward (*bottom*) show how the eruption removed the top 400 meters of the mountain, leaving a crater 750 meters deep. In these views the mountain is seen from the northeast. They were created by the Digital Elevation Model program at Western Mapping Center of U.S. Geological Survey.

violently and intermittently just as it has in the recent geologic past, and these future eruptions will affect human life and health, property, agriculture and general economic welfare over a broad area.... The volcano's behavior pattern suggests that the current quiet interval will not last as long as 1,000 years; instead an eruption is more likely to occur within the next 100 years, and perhaps even before the end of this century."

Many volcanic eruptions are preceded by swarms of small earthquakes. Although not all the volcanoes of the Cascade Range are monitored by seismographs, Mount St. Helens fortunately was. The University of Washington had a seismograph on the west flank of the volcano that was linked to Seattle by telemetry. On March 20, 1980, at 3:47 P.M. Pacific Standard Time an earthquake of magnitude 4 on the Richter scale occurred under the mountain. When this unusual event was followed by an increasing number of local earthquakes, it became apparent that a major earthquake swarm was in progress. In order to improve the recording and locating of the earth tremors additional seismographs were installed.

On March 25 the seismic energy released by the swarm reached its peak rate: 47 earthquakes of magnitude 3 or more occurred within a 12-hour period at shallow depths under the mountain's north flank. The first small explosions of steam came two days later, beginning at 12:36 P.M. and forming a new crater about 70 meters across on the snow- and ice-covered summit. Large east-west cracks also developed in the snow and ice, indicating that a down-faulted block was forming across the summit area. A second crater opened on March 29. Blue flames, possibly burning hydrogen sulfide, were visible from the air at night. Ninety-three small eruptions of steam and ash were observed on March 30.

On April 1 the seismographs recorded the first volcanic tremor, a more or less



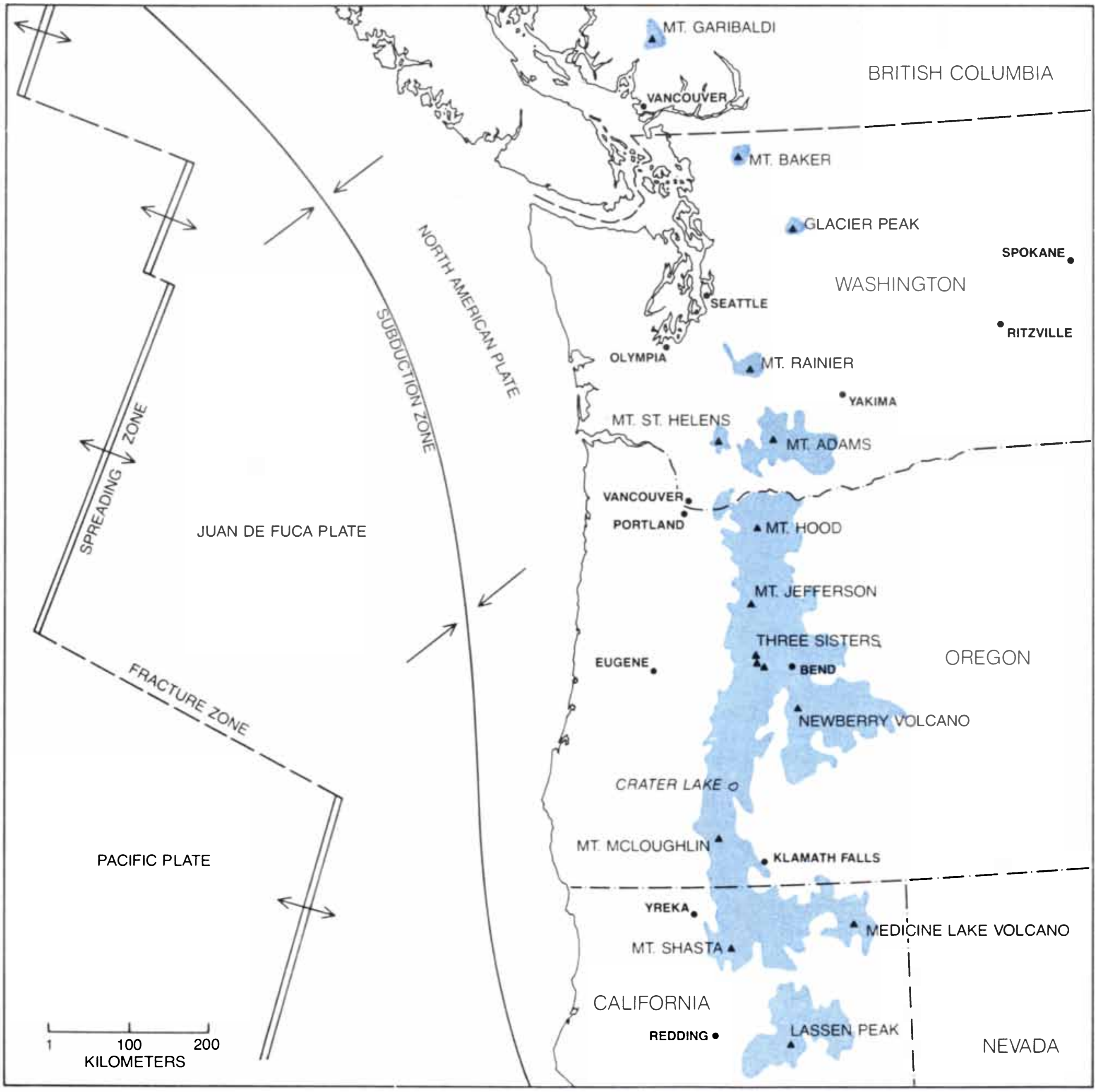
GEOLOGIC MAP outlines the region devastated by the May 18 eruption. The area in dark color is pyroclastic deposits: hot bits of fragmented rock. Hatched area is ash falls. The area in medium color is the one in which trees were blown down. The area in light color is the one in which trees were left standing but with their needles killed by heat. Light gray is the avalanche deposits; dark gray, mudflows.

continuous ground vibration observed at many active volcanoes. The precise cause of such a tremor is not known, but it presumably reflects the movement of magma or the rumbling release of gas previously dissolved in the magma.

By this time the new volcanic eruption, the first in the Cascade Range since Lassen Peak quieted down in 1917, had become a magnet for the curious. Nolan Lewis, Director of Emergency Services for Cowlitz County, Wash., in which Mount St. Helens is situated, reported that "Sunday [March 30], when the weather was clear, the roads up the mountain looked like downtown Seattle at rush hour." When the potential danger of violent eruptions became more and more apparent, Governor Dixy Lee Ray ordered additional blockades established on roads leading to the mountain.

The small steam and ash eruptions continued, some as single explosions, others as pulsating jets lasting for hours. Columns of steam and ash climbed to three kilometers above the summit. The crater of the mountain grew to a single oval basin 500 meters by 300 meters across and 200 meters deep. The ash consisted of fragments of old volcanic rock; the gas emissions included small amounts of carbon dioxide, sulfur dioxide, hydrogen sulfide and hydrogen chloride together with large volumes of steam. The small explosive eruptions apparently resulted from ground water high in the volcanic cone being heated above the subsurface boiling point and flashing suddenly into steam, much as a geyser does but with energy enough to incorporate ash particles and blast out a crater. The energy released by all the steam explosions up to May 18 is estimated at 10^{14} joules.

As the high level of seismic activity continued (about 50 earthquakes of magnitude 3 and greater per day) there was another ominous sign. Several observers had noted as early as March 27 that the down-faulting of the block across the summit of the mountain



TECTONIC MAP shows the relation between the North American tectonic plate and the Juan de Fuca and Pacific plates to the west. At the subduction zone the Juan de Fuca plate is plunging under the North American plate, giving rise to the volcanoes of the Cascade Range (*small black triangles*). Colored areas are volcanic deposits less than two million years old. Data are from the U.S. Geological Survey.

seemed to be accompanied by a bulge, or uplift, of the high north flank of the volcanic cone. The bulge continued to grow early in April and was manifested by a spreading network of large cracks in the cover of ice and snow. Photogrammetric maps made by the U.S. Geological Survey from aerial photographs taken on April 12 disclosed that the bulge was nearly two kilometers in diameter and had already moved up or out by as much as 100 meters.

Ground surveys conducted in late April and early May established that the bulge was continuing to expand northward more or less horizontally at a rate of about 1.5 meters per day. The rapidly deforming area was directly over the center of the earthquake zone two kilometers below. The consensus of observers on the site was that the continuing seismic swarm and major surface deformation were good evidence that magma was being injected under the volcano at a shallow depth. They expected a major eruption or an avalanche from the expanding north face. The only questions were how soon it would come and how violent it would be.

Robert Christiansen, the investigator in charge of the U.S. Geological Survey's monitoring program, reviewed the historic sequence of activity at other volcanoes, specifically Lassen Peak, the only Cascade Range volcano with a well-observed eruption, and Bezymianny on the Kamchatka Peninsula on the Pacific coast of the U.S.S.R., which had exploded violently in 1956. Christiansen concluded that an eruption comparable to that of Lassen in 1915 was the most likely course for Mount St. Helens. An eruption on the scale of Bezymianny, however, could not be ruled out. A third possibility was that all the activity might subside without a major eruption.

Meanwhile, on May 7, after about two weeks in which there had been little visible activity, small explosions of steam and ash began again. Although seismic activity had continued unabated and the bulge had grown steadily, the lack of dramatic visible activity led residents of the region to question the political authority that was keeping the area around the mountain closed. On May 15, 16 and 17 earthquakes and bulging continued, but there was no sign of steam or ash.

At 7:00 A.M. Pacific Daylight Time on May 18 Dorothy and Keith Stoffel, two Washington geologists, boarded a light plane at Yakima airport near Mount St. Helens and set off on their first reconnaissance flight. They made several passes around and over the volcano. No activity was visible, although the morning was bright and clear. At 8:32 the mountain was shaken by an earthquake of magnitude 5.1 centered below its north flank. At that moment the Stoffels were directly over the summit and looking down from a height of 400 meters. They noted several small ice falls starting on the steep sides of the crater. Fifteen seconds later they were the closest witnesses to the onset of a huge volcanic eruption triggered by one of the largest landslides of historic times.

"The whole north side of the summit crater began to move instantaneously as one gigantic mass," Dorothy Stoffel recalled. "The entire mass began to ripple and churn without moving laterally. Then the whole north side of the summit started moving to the north along a deep-seated slide plane."

Seconds later there was a vast explosion. Curiously, the Stoffels neither felt nor heard it even though they were just to the east of the summit. From their position the initial explosion cloud seemed to mushroom sideways to the north and then plunge down the slope. Their first concern was survival. Even though they dived at full throttle to gain speed, the expanding gray cloud gained on them. They finally escaped it by turning south. Behind them giant ash clouds boiled upward, thrusting north and northwest. To the east the clouds expanded into billowing mushroom shapes, illuminated by lightning bolts thousands of meters long. Half an hour later the Stoffels landed in Portland.

In the avalanche they had witnessed from the air more than two cubic kilometers of crushed rock and glacier ice plunged into Spirit Lake and the north fork of the Toutle River. Fluidized by exploding steam, the avalanche accelerated rapidly to velocities as high as 250 kilometers per hour. One lobe of the gigantic mass plowed through the west arm of Spirit Lake and northward into the valley beyond. An adjacent lobe swept across another valley with such momentum that it went over a ridge 360 meters high that bounded the valley on the north. The bulk of the fluidized debris funneled down the valley of the Toutle, forming a hummocky deposit 21 kilometers long, one to two kilometers wide and up to 150 meters deep. The gravitational energy of the avalanche was about 5×10^{16} joules (roughly the equivalent of 12 megatons).

As the great avalanche of rock and ice suddenly released the pressure within the volcanic cone, superheated ground water flashed into steam. Simultaneously dissolved gases exploded from the shallow magma body recently intruded into the upper core of the mountain. The steam blast, magma explosion and giant avalanche combined to form a lateral blast of hot (up to 300 degrees Celsius), dense, debris-filled steam clouds that surged northward from the breached mountainside at speeds of between 100 and 400 kilometers per hour. The steam blast and its fluidized charge of volcanic rock fragments devastated 550 square kilometers of mountain terrain northwest, north and northeast of Mount St. Helens. The ground-hugging black clouds rolled over four major ridges and valleys, going as far as 28 kilometers from their source.

The destruction was complete. For the first few kilometers entire trees one to two meters in diameter were uprooted and swept away with the roiling explosion cloud. Beyond this was a blowdown zone 10 to 15 kilometers across, where prime Douglas firs were snapped like matchsticks. At the outer limits of destruction trees were still standing, but their needles were scorched beyond recovery.

The first impression from viewing the blowdown zone was that a great shock wave or concussion front had knocked the trees down in a pattern radial to the exploding peak. This was not, however, substantiated by the evidence. Survivors near the edge of the devastated area heard only a moderately loud explosion or roaring sound two to three minutes before the black cloud with its hot hurricane winds descended on them. The velocity of the front of the steam blast clouds was well below the speed of sound. On closer inspection the pattern of felled trees showed turbulent eddies and curving streamlines. Near the edge of the devastated area the trees tended to be blown down in down-valley directions, even where that meant they fell toward the source of the surging clouds.

Gravity apparently energized the dense, fluidized mass as the energy of the initial steam blast waned. Then as the turbulent internal winds abated, ash and rock debris that were carried in the dense clouds settled to the surface in deposits that decrease in thickness, with distance from the source, from about a meter to a centimeter. Angular fragments up to several tens of centimeters in diameter of both old volcanic rocks and fresh, hot ones were carried in the blast clouds to distances of 10 to 15 kilometers. Trees were charred and blackened on the blast side out to seven kilometers on the northwest and north, and as far as 18 kilometers on the northeast.

The fluidized character of the masses of fragments carried by the blast is even more evident in the deposits on steep slopes. Here after the initial deposition secondary flows formed ponded deposits tens of meters thick in basins and valley bottoms. The total volume of the blast deposits is about .18 cubic kilometer. Of this amount about .06 cubic kilometer is magmatic: fresh volcanic rock. The heat energy provided by the magmatic component was approximately 2×10^{17} joules.

Another probable source of energy for the steam blast explosion was the superheated ground water inside the volcano. Assuming a porosity of 15 percent in a volume of two to three cubic kilometers filled with water at an aver-

BEFORE-AND-AFTER PHOTOGRAPHS of Mount St. Helens show the extent of the new crater. The photograph at the top was made by one of the authors (Robert Decker) in June, 1970. The elevation of the summit, seen here from the north-northeast, was 2,950 meters, rising from a base with an elevation of about 1,000 meters. The photograph at the bottom was made by Ray Foster of the Sandia Laboratories in July, 1980. The crater, seen here from the north, is two kilometers across. The elevation of the rim of the crater is between 2,400 and 2,550 meters, that of the floor of the crater between 1,800 and 1,900 meters. Pyroclastic flows cover much of foreground.

age temperature of 175 degrees C., this yields an additional 10^{17} joules of energy. Steam in the explosion cloud probably was generated in two ways. Heat in the flashing ground water would generate 4.4×10^{10} kilograms of steam, and the additional heat in the .06 cubic kilometer of magmatic fragments could convert another 8.8×10^{10} kilograms of water into steam. The total of 1.3×10^{11} kilograms of steam would expand to 220 cubic kilometers at 100 degrees C. and atmospheric pressure. The assumed mass of heated ground water inside the volcano (3.75×10^{11} kilograms) is nearly three times more than is needed to supply the calculated volume of steam. These energy and steam-volume figures are only rough estimates, yet they yield figures that are reasonable in terms of the 550-square-kilometer area devastated by the ground-hugging lateral blast clouds. The final blast deposit was a layer of wet ash up to six centimeters deep, containing in many places the pea-size mud balls volcanologists call accretionary lapilli. These accretions formed around nuclei provided by raindrops condensing out of the steam clouds.

By 9:00 A.M. on May 18 the worst of the eruption was over, but the vertical column roared on, reaching altitudes in excess of 20 kilometers for much of the day until it began to wane at about 5:30 P.M. The source of this almost continuously exploding and uprushing column

VERTICAL BEFORE-AND-AFTER VIEWS of Mount St. Helens were made on May 1, 1980 (*left*), and June 19, 1980 (*right*), by U-2 aircraft of the National Aeronautics and Space Administration. The peak is at the lower right. The film was infrared-sensitive and the

of gas and ash was the effervescing shallow magma body being progressively cored out to greater depths. The abrasive uprush continued to enlarge the horseshoe-shaped crater that had initially been formed by the avalanche and the lateral blasts.

High-altitude winds blew all day to the northeast, and ash began to fall on towns in central Washington by midmorning. At Yakima, 150 kilometers away, the first ash fall was a "salt and pepper" layer consisting of sand-size fragments of dark rock and of lighter-colored feldspar crystals. On top of this layer was deposited a thicker one of silt-size particles of volcanic glass. Thirty kilometers north of Yakima the ash fall was about 20 millimeters deep. To the east the fine ash was deeper, reaching more than 70 millimeters near Ritzville, Wash., 330 kilometers from Mount St. Helens. Here the texture of the ash was like talcum powder.

Near Spokane, Wash., 430 kilometers northeast of the volcano, the ash was only five millimeters deep, but by 3:00 P.M. visibility was reduced to three meters in near darkness. At about noon on May 19 a trace of ash fell on Denver. Within three days the ash cloud had crossed the U.S. The weight of the measured ash fall was equal to .15 cubic kilometer of magma and represented a dissipation of heat energy amounting to 5×10^{17} joules. Studies of the traces of

prints are in false color. The red areas are equivalent to the green of vegetation, mostly Douglas fir. Gray areas in the photograph at the right are those devastated in the eruption of May 18. Over the peak in the photograph at the right is a fume cloud issuing from the crater.

MOMENT OF ERUPTION on May 18 (actually 20 seconds after the eruption began) was captured by Keith and Dorothy Stoffel, geologists who were flying over the mountain at the time. The eruption was preceded by beginning of the avalanche. The Stoffels escaped.

DOME OF LAVA was formed in the crater between June 13 and June 20. It was 300 meters wide and 65 meters high and was destroyed by the eruption of July 22. This aerial photograph was made by Maurice and Katia Krafft of Centre de Volcanologie at Cernay in France.

ash that fell beyond the area of measured ash thickness and the very fine ash and sulfuric acid aerosols still suspended in the stratosphere suggest that an additional volume of .1 cubic kilometer of magma was dispersed.

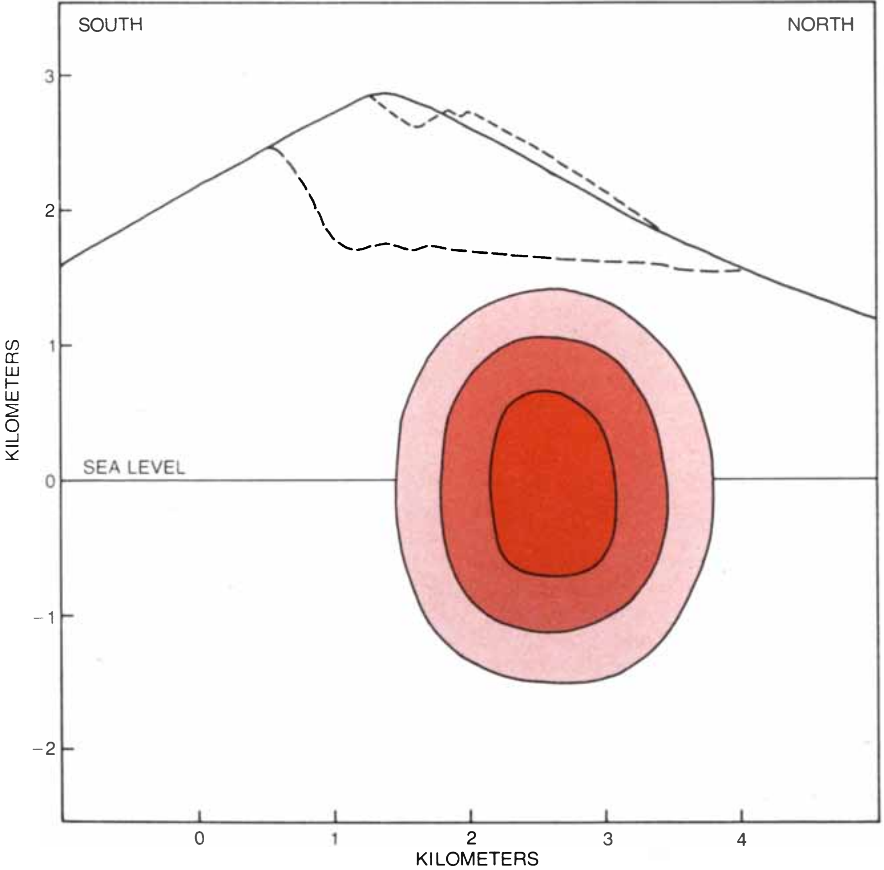
Judging from experience with other eruptions that have injected dust and aerosols into the stratosphere, the very fine particles will take a year or two to settle out. The effect on world climate, if any, is not yet evident.

Floods and mudflows were another major aspect of the eruption. The flows consisted of a slurry of volcanic ash and fine rock particles mixed with water, and they had the consistency of wet cement. The ash blanket near the mountain and the crushed rock in the avalanche deposit provided the solid matter; the extra water probably came from several sources: melting snow and ice, the water displaced from Spirit Lake and the north fork of the Toutle River by the avalanche deposit, the water from the breached hydrothermal system that did not flash to steam, and condensing steam.

The first mudflow crest came in the south fork of the Toutle River near Silver Lake at 10:50 A.M. on May 18. It exceeded by 30 centimeters the highest flood level of historic times. The largest mudflow came from the north fork of the Toutle and crested at about 7:00 P.M., destroying the gauge station near Silver Lake. High mud and water marks indicated that it had exceeded historic flood levels by nine meters.

Downstream from the Toutle, mudflow deposits clogged the channels of the Cowlitz River and caused severe shoaling of the navigation channel in the Columbia River. The volume of mud deposited was about .1 cubic kilometer. Roughly the same volume of water would have been involved in mobilizing the mudflows.

Sometime after the initial avalanche and steam blast eruption pyroclastic flows of fine ash and pumice blocks began rushing down the north slope of Mount St. Helens through the breach in the newly formed crater. These fluidized emulsions of hot rock and glass fragments mixed with hot volcanic gases, being denser than the ascending ash cloud, issued from the crater under the uprushing cloud. Successive flows poured down the north slope at speeds of up to 100 kilometers per hour, covering the earlier landslide and blast deposits and reaching the south edge of Spirit Lake. When these hot (300–370 degrees C.) pyroclastic flows came in contact with water, they set off secondary explosions, sending steam and ash clouds as high as two kilometers. The ash and pumice flows continued until the evening of May 18. Their total bulk volume was .25 cubic kilometer and their estimated thermal energy 3.3×10^{17} joules.



TOPOGRAPHIC PROFILES show Mount St. Helens as of August, 1979 (*solid line at top*), as of May 1, 1980, when a bulge had formed on the north slope of the mountain (*upper broken line*), and as of July 1, 1980 (*lower broken line*). The colored area at the right below the profiles is the general region of origin of the thousands of earthquakes in the swarm from March 20 through May 18. The darker the color, the higher the density of the earthquake locations. Data are from U.S. Geological Survey and Department of Geophysics of University of Washington.

The human cost of the eruption was commensurately great. Sixty-two people were killed or are missing. The economic loss, mostly to the lumber industry, was more than $1 billion. Perhaps the greatest damage was psychological. In the U.S. Northwest the Cascade Range volcanoes have been perceived as silent guardians; now they are a brooding threat.

The huge avalanche and eruption of May 18 was followed by smaller ash explosions on May 25, June 12, July 22, August 7 and October 16–18. Domes of viscous lava were extruded inside the crater on June 13–20, August 8–9 and October 18–19. The total volume of magma in the eruptions of May 25 through October 19 is about .05 cubic kilometer with an estimated energy content of 1.7×10^{17} joules, an order of magnitude less than the May 18 eruption. The May 25 eruption occurred at night and in poor weather, but a sudden increase in the volcanic tremor at 2:28 A.M. probably marked its beginning. By 2:45 National Weather Service radar showed that the ash column had reached an altitude of 14 kilometers. It diminished in height within an hour, but the volcano continued to emit smaller ash clouds throughout the day. New pyroclastic flows of ash and pumice blocks were also emitted, and they covered some of the same area on the north flank of Mount St. Helens swept by the earlier pyroclastic flows. Although the ash eruption was much less voluminous than the one of May 18, the wind directions were more variable and a thin layer of ash fell over wide areas of western Washington and Oregon, including the Portland metropolitan area.

The June 12 ash explosions were similar to the eruption of May 25. Volcanic tremor began in the afternoon, with the first ash emission at 7:05 P.M. reaching an altitude of four kilometers. A much larger ash eruption began at 9:09 P.M. and reached 15 kilometers. The eruption faded rapidly after midnight. Helicopter observation the next day revealed that additional pyroclastic flows two to 10 meters thick with temperatures of up to 600 degrees C. had descended toward Spirit Lake. A lava dome began to form in the explosion crater after the June 12 eruption, and when it was seen on June 15, it was 200 meters wide and 40 meters high and was covered with large, glowing cracks. The dome continued to rise about six meters per day and by June 20 had reached its greatest height of 65 meters.

It was quiet until July 22. That morning small, shallow earthquakes began under the crater area of the mountain. The earthquakes increased in number during the day, but no volcanic tremor was recorded. Suddenly at 5:14 in the afternoon of a clear summer day a large ash cloud began to roil up from the mountain. Radar registered the top of the cloud at 14 kilometers. A second ash cloud erupted at 6:25 P.M. and reached an altitude of 18 kilometers in just seven minutes 23 seconds, an average velocity of 2.2 kilometers per minute. The third and longest ash jet began at 7:01 P.M. and lasted for more than two hours, reaching a maximum height of 14 kilometers.

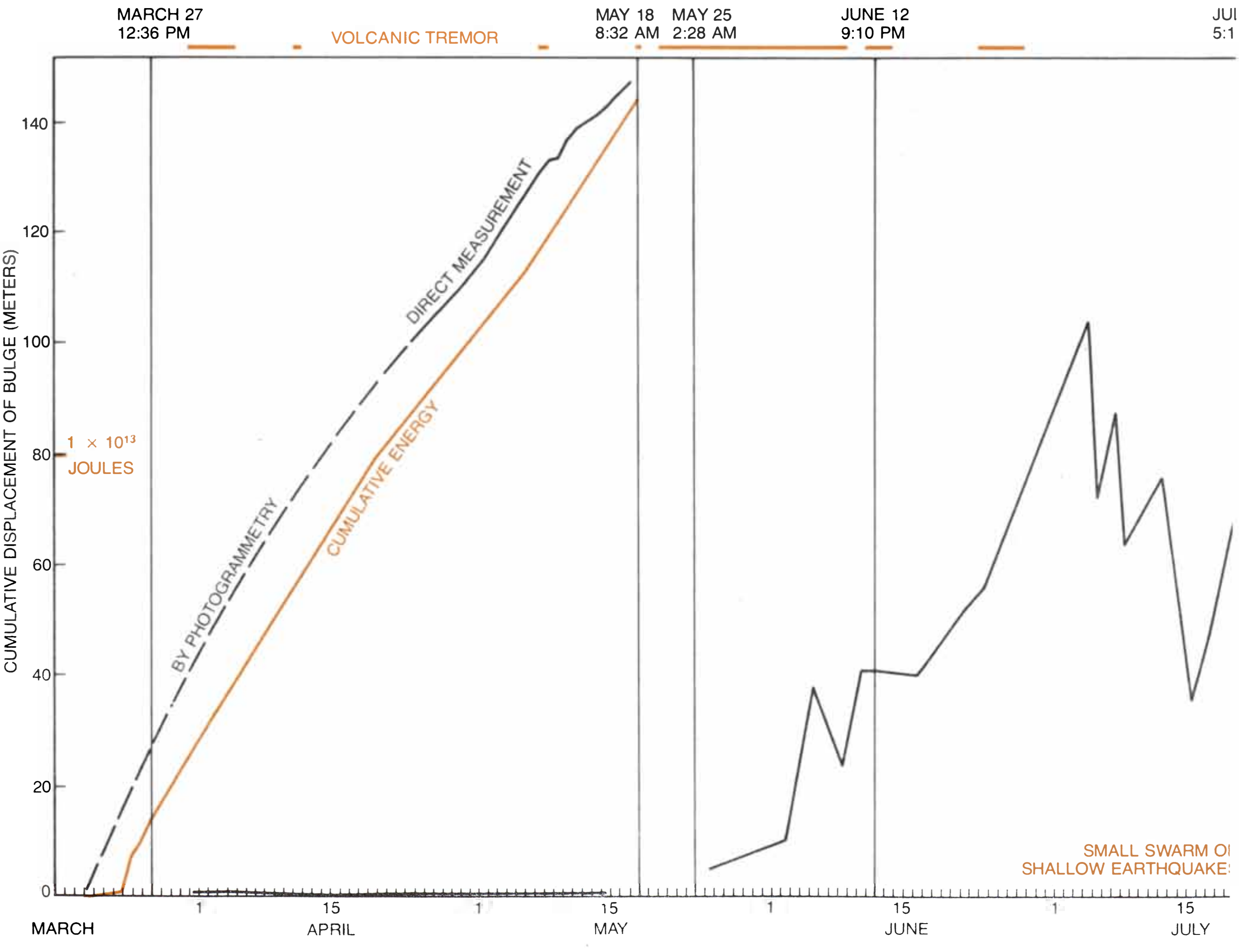
Although geologists and Forest Service fire fighters had left the area after being warned of the increase in small, shallow earthquakes, the eruptions were observed from helicopters and other aircraft. Pyroclastic flows poured from the vent during the second and third ash explosions and spilled down the north slope of the volcano toward Spirit Lake. Richard Hoblitt of the U.S. Geological Survey gave the following eyewitness account:

"We were flying from east to west approximately one mile north of the vent as the second eruption started. Following a period of a few seconds' duration during which the rate of gas emission increased, an ash fountain was ejected to about 1,500 feet above the vent. As the projections of the fountain arced over and reached the surface in the vicinity of the vent, they gave rise to a pyroclastic flow that began to rapidly flow northward out of the amphitheater. We exited to the west as quickly as possible." Not many people have seen a pyroclastic flow from this close and lived to describe it.

These new ash and pumice-block flows were one meter to two meters thick. Their temperature was measured the next day; the maximum temperature recorded was 705 degrees C. at a depth of 1.5 meters. The ash clouds drifted to the northeast on July 22, and only minor ash falls were reported from central and eastern Washington.

A change in the pattern of gas emission had preceded the July 22 eruption. The gas emissions changed again early in August, and volcanic tremor began just after noon on August 7. Warned by these two signs, investigators in the danger areas were evacuated. An eruption began at 4:26 P.M. and rapidly generated an ash cloud that reached a height of 13 kilometers. Small pyroclastic flows swept the area below the breach on the north side of the mountain. Smaller ash eruptions continued through the late afternoon and evening, with another large outburst at 10:32 P.M. A new lava dome formed in the crater on August 8–9.

Quiet prevailed for more than two months; then small earthquakes similar to those that had preceded the July 22 eruption began on October 16. As the swarm increased an alert was issued early in the evening. It was followed by an eruption that began at 9:58 P.M. Four ash explosions over the next two days sent clouds as high as 14 kilometers and destroyed the August lava dome. Some



FOUR TYPES OF ACTIVITY associated with the eruptions are plotted on this graph: earthquake swarms, the formation of the bulge, the emission of sulfur dioxide gas and volcanic tremor. Important eruptions are indicated by the vertical lines. The major deformation of the eruptions were accompanied by smaller pyroclastic flows that descended the north flank of the mountain. Light ash fell to the south and southeast, reminding the Portland metropolitan area that Mount St. Helens was still in action. A new lava dome, the largest yet to appear, arose on October 18–19.

The eruptions of volcanoes vary enormously in their explosiveness. In Hawaii fountains of incandescent lava spray spectacularly but harmlessly into the air and lava flows move slowly downslope from their vents. At the other extreme gigantic explosions destroy entire mountains when a substantial fraction of the heat energy in the magma is converted into mechanical work. One measure of the explosiveness of volcanoes is the nature of their products. Effusive eruptions are characterized by lava flows, explosive eruptions by fragmental products: volcanic ash, rock chips and blocks. In Hawaii about 98 percent of the eruption products are effusive lavas. With the volcanoes rimming the Pacific it is nearly the opposite: about 90 percent of the eruption products are fragmental debris.

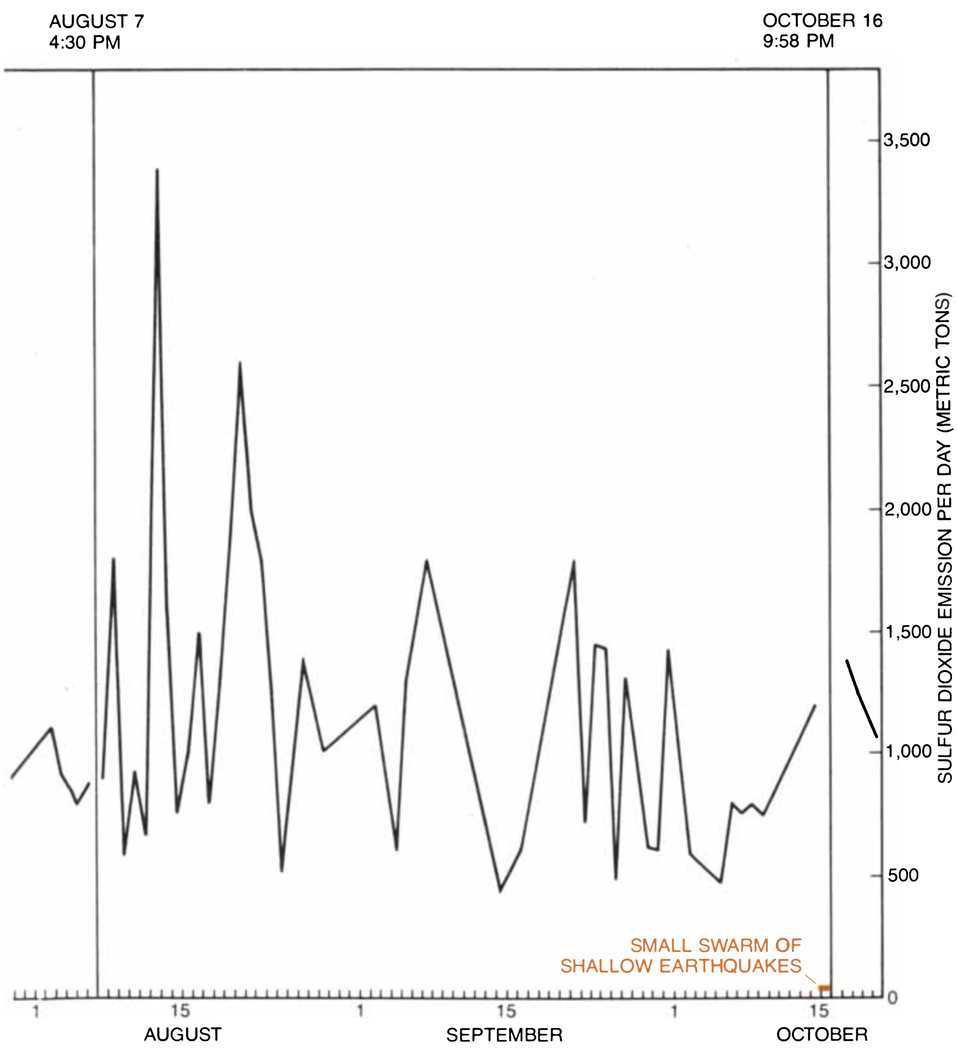
Volcanoes are found in three tectonic settings. The volcanoes in each of these settings differ remarkably in their explosive behavior from those in the others. As we have seen, the ring-of-fire volcanoes rise along subduction zones where tectonic plates converge. These volcanoes are generally explosive. Rift volcanoes occur where plates are diverging; their eruptions, particularly the deep submarine ones, are more effusive. Where rift volcanoes erupt in shallow water or through continental crust, however, they can be explosive. Hot-spot volcanoes, which penetrate tectonic plates, are generally effusive where they occur in areas of oceanic crust (Hawaii) and explosive in areas of continental crust (Yellowstone).

The key factors governing the explosiveness of volcanoes appear to be the viscosity of the magma, the amount of gas dissolved in the magma, the quantity of ground water near the vent area and the surface pressure. Volcanic explosions are caused by the rapid expansion of gas within the magma or in close contact with it. There is no release of energy as there is in a bomb; a better analogy is the explosion of a steam boiler. Viscosity is analogous to the strength of the boiler; the higher the viscosity, the larger the potential explosion. The availability of gas within the magma or in contact with it is comparable to the volume of the boiler. The difference between the gas pressure in the magma and the surrounding pressure is comparable to the pressure difference across the boiler walls. Steam is one of the major magmatic gases, but in volcanic explosions carbon dioxide is also important.

Consider two examples, the first a basaltic magma with a relatively low viscosity and a low content of dissolved gas. If this basalt rises through a fracture toward the surface, carbon dioxide and water begin to come out of solution as the pressure decreases. The bubbles expand against the low viscous forces and cause the foaming magma to expand. This process is most important in the last few meters before the magma reaches the surface: it gives rise to an effervescing fountain of incandescent lava at the vent. If the same basalt emerged under two kilometers of seawater, the decrease in pressure would not be sufficient for the dissolved water to come out of solution. Some carbon dioxide would probably exsolve, but the expansion of the mixture of magma and carbon dioxide would be small, and the mixture would well quietly from its submarine vent. In shallow water, depending on the rate at which the hot effusive lava mixed with the surrounding water, various degrees of explosiveness might result.

The second example is a siliceous magma with a relatively high viscosity and a high content of dissolved gas. If this dacitic magma rises toward the surface, the exsolved gas begins to form bubbles. Because of the high viscosity of the melt, however, the bubbles do not expand to their equilibrium size; they have an internal overpressure estimated to be as high as several hundred atmospheres. When enough bubbles form and the external pressure is low enough, the magma shatters into fragments that are accelerated out of a vent by the explosively released gases. The cooling of the expanding gases solidifies the magma fragments; they are ejected as hot but solid particles in an emulsion. These fluidized mixtures form eruption clouds or pyroclastic flows depending on their density and the direction and velocity of their ejection.

One of the closest analogues of the eruption of Mount St. Helens is the eruption of Bezymianny in Kamchatka in 1955–56. The Russian volcanologist G. W. Gorshkov has divided the Bezymianny eruption into five stages: (1) September 29 to October 21, 1955,



of the bulge and the earthquake energy are related to period before eruption of May 18. Data are from U.S. Geological Survey and Department of Geophysics of University of Washington.

pre-eruption earthquake swarm; (2) October 22 to November 30, 1955, strong ash eruptions; (3) December 1, 1955 to March 29, 1956, moderate ash eruptions with general uplift of the old lava dome; (4) March, 1956, gigantic explosion and lateral blast; (5) April, 1956, growth of new lava dome in the explosion crater.

The volume of rock and the area affected in the March 30, 1956, eruption of Bezymianny are similar to those of the May 18, 1980, eruption of Mount St. Helens. The major differences are that in the Bezymianny eruption the earthquake swarm preceding the first ash eruption lasted for 23 days compared with seven days at Mount St. Helens, the preliminary eruptions lasted for 160 days compared with 53 days and there were larger ash explosions in the early phase of the preliminary eruptions at Bezymianny. In the major explosion at Bezymianny most of the ejected rock was new magma, in contrast to the May 18 avalanche and explosion of Mount St. Helens, in which less than a fifth of the ejected rock was new magma. Perhaps significantly, Bezymianny has remained active since 1956. Its last major eruption occurred in 1979.

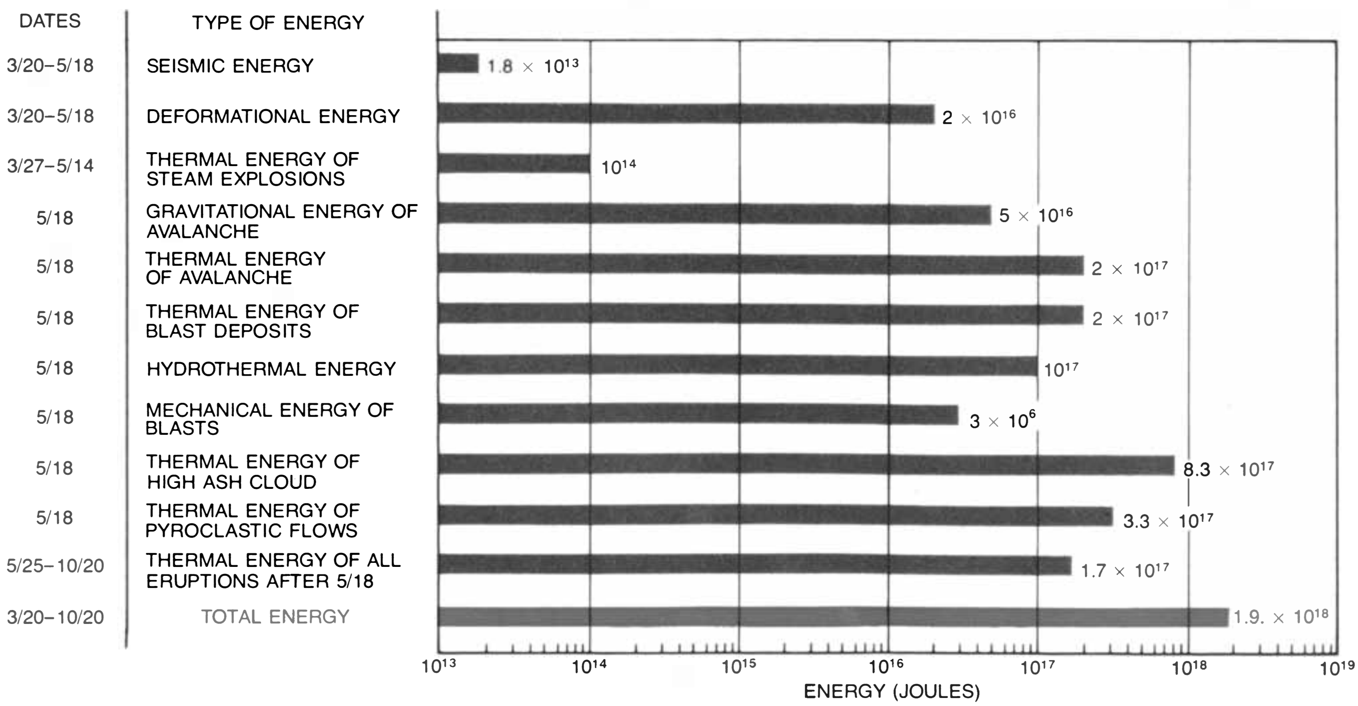
Mount St. Helens has provided a good test case for techniques of forecasting volcanic eruptions. There is no single key to success in such forecasting; all the factors must be evaluated and then interpreted in the light of geologic experience. Among the factors are the statistics of eruptions in historic times and the reconstruction of the statistics of eruptions in prehistoric times by geologic mapping and dating. The geophysical techniques involved include monitoring, on and near active and potentially active volcanoes, seismicity, ground-surface deformation, magnetic and electrical fields, and temperatures. The geochemical techniques include monitoring the volume and composition of gases, liquids and solids emitted by volcanoes. Repeated visual observations from the ground or the air provide important data on evolving volcanic activity. Most of the world's volcanoes do not get even this basic type of examination, let alone the more sophisticated instrumental monitoring.

At Mount St. Helens some methods worked better than others. The historic-activity data indicate eruptions in the period from 1831 to 1857, but apart from identifying Mount St. Helens as an active volcano the sample is too small to have any statistical meaning. The geologic mapping and dating identified 20 eruptive periods with diverse products over the past 4,500 years and established some important patterns. Dormant intervals formed two populations: 100 to 150 years and 400 to 500 years. The last two dormant intervals before 1800 were of the short type. The eruptive products also showed that ash and pyroclastic explosions were common and that prehistoric eruptions of Mount St. Helens affected large areas. It was this kind of analysis that led Crandell and Mullineaux to make their forecast that Mount St. Helens was dangerous.

In 1980 seismic monitoring of the increasing earthquake swarm under Mount St. Helens gave a week's warning of the first small explosive eruptions. Photogrammetry and electro-optical distance measurements documented the spectacular ground-surface deformation associated with the growing bulge on the north flank of the mountain. There was no change in the rate of seismicity or ground-surface deformation just before the May 18 eruption, but the continuation of these phenomena gave a general warning that something important was happening underground. That warning allowed Governor Ray and the U.S. Forest Service to stick to their order to keep the area closed in spite of insistent demands for freedom of access to it. Their firmness undoubtedly saved thousands of lives.

In the months since May 18, 1980, there have been successful forecasts of the smaller but still significant explosive eruptions of June 12, July 22, August 7 and October 16. Volcanic tremor preceded the June 12 and August 7 eruptions by several hours, and the unusual buildup of small earthquakes just under the mountain preceded the July 22 and October 16 eruptions by several hours. Anomalies in the pattern of the emission of gases and small deformations of the ground surface have also preceded some (although not all) of the eruptions since May 18 by hours or days.

Any change in the total monitoring pattern is suspicious, and experience helps in evaluating the change. There have been false alarms, but in any probabilistic forecasting system that seems to be inevitable. Technology has not solved the ancient problem of how sure one must be before crying wolf.



SIX TYPES OF ENERGY released in the eruptions are plotted. The dates of the energy release are given in the column at the far left. The seismic energy, the deformational energy and the thermal energy of steam explosions were released in the period before the eruption of May 18. The scale of the bars is logarithmic. The bar at the bottom gives total energy of all these releases through October of last year.

No! No! We shall NOT leave the molecular biologist up the creek without a paddle!

Dear Kodak:

Over the years I have enjoyed the advertising which Kodak has directed toward the scientific community. It has been interesting, informative, and well done. I find, however, that a recent ad (Science, March 7, 1980) expresses an idea which seems to be at variance with company policy.

My colleagues and I have been informed that Kodak will no longer supply its Kodak X-Omat R film because there is insufficient demand for it. I cannot overestimate the importance of this film for the current technologies of recombinant DNA, DNA sequencing, and gene mapping. Another film you offer requires a longer exposure. In the best case, an experiment which can be done overnight with Kodak X-Omat R film might require several days with the other film, while in the worst the experiment cannot be done at all because the half-life of the isotope limits the exposure which can be made.

It is thus disconcerting to me to read that Kodak still keeps up its payments on its collaborative bargain with science in the form of certain products short on appeal to the world at large. I would hope that Mr. Eastman's bargain today would range from stars, very large and very distant, to genes, very small and very near.

Sincerely,

The letter, one of many such, is right. The properties of Kodak X-Omat R film have indeed made it the choice for most[1] autoradiography of gel electrophoresis separations. Generations[2] of molecular biologists have been brought up on it. Then one day in a box of the favored product they find notice of its discontinuance in favor of products that better serve radiologists.

We hadn't appreciated how many of those boxes had been going to people who are not radiologists.

Let those people be advised that their film is now very, very available from the same sources in 35 x 43 cm, 8 x 10 in., and 13 x 18 cm sizes under the name Kodak X-Omat AR film.

The added "A" merely indicates that the blue tint in the base, which radiologists seem to prefer but which molecular biologists don't need, particularly when they want to make copies of their autoradiograms, is gone.

It can be processed manually, of course, but a laboratory that is still processing film by hand or depends on some other department's film processor really ought to ask about the table-top Kodak X-Omat M20 processor,[3] the lowest in price that we offer. Drop a line to Eastman Kodak Company, Department 740-B, Rochester, New York 14650. Might as well ask also for Kodak Pamphlet No. M3-508, "Autoradiography of Macroscopic Specimens."

[1] Kodak SB film gives equivalent or possibly a bit better results for fluorography with H^3, but it requires a longer processing cycle than given by up-to-date processing machines. Kodak no-screen film is more sensitive for direct recording of p^{32}, Ca^{45}, C^{14}, and S^{35}, but base fog is somewhat higher.

[2] A generation of molecular biologists is considerably shorter than the usually reckoned 30 years.

[3] We'll be showing this processor March 1-6, 1981, at the Dallas meeting of the American Society for Microbiology.

Kodak

SCIENCE AND THE CITIZEN

War of Nerves

For more than four years representatives of the U.S. and the U.S.S.R. have been meeting regularly in Geneva in an effort to work out the technical details of a treaty on chemical warfare. The treaty would prohibit the development, production and stockpiling of chemical weapons, including the lethal organophosphorus agents known as nerve gases. During most of this period and for several years before, repeated requests by the U.S. Army for funds to introduce a new generation of nerve-gas weapons, called binary munitions, were rejected by Congress, partly on the ground that such an action could impede the progress of the Geneva talks. The one exception came last year, when Congress approved a comparatively small sum—$3.15 million—to begin construction of a binary-munitions factory at the Pine Bluff Arsenal in Arkansas. A bill that would have provided an additional $19 million to equip the Pine Bluff plant was narrowly defeated.

The stage has thus been set for the Reagan Administration and the new Congress to redefine U.S. policy on the issue of chemical disarmament. Continued postponement of the Pine Bluff project would presumably be interpreted by advocates of arms control as an encouraging sign, perhaps indicating a revived interest on the part of the U.S. Government in pursuing arms-control measures outside the stalemated Strategic-Arms-Limitation Talks. Conversely, a prompt commitment to push ahead with the binary-nerve-gas facility would be viewed as a vote of no confidence in the bilateral chemical-disarmament negotiations and would probably be considered the first step in a major acceleration of the chemical-arms race.

Proponents of binary-nerve-gas munitions cite safety as the main justification for acquiring the weapons. The principle of the binary munitions is that two chemical precursors of the lethal agent would be manufactured and stored separately. Only when the precursors were mixed could they react to form the nerve gas itself. The product of the reaction would be one of the chemicals already in the U.S. arsenal, such as the nerve gas Sarin, or agent GB. In the proposed binary version of the 155-millimeter nerve-gas artillery projectile, for example, only one of the components would initially be present in the shell. The other component would be stored and shipped in a separate canister, which would be loaded into the shell at the gun site. While the projectile was in flight the contents of the two canisters would mix and react to form the lethal agent. Among the asserted advantages of such an arrangement are that it would decrease the likelihood of an unintended release of poison gas and that it would facilitate the disposal of the chemicals in the event of a future chemical-disarmament agreement or a unilateral decision to decommission the weapons.

Opponents of the adoption of binary munitions question both the utility and the safety of the new weapons. For service in combat, they argue, binary munitions offer no real advantages. Existing 155-millimeter nerve-gas projectiles, which contain either GB or the agent designated VX, can be fired from the most advanced artillery pieces, and they have been handled safely for decades.

In any case, the critics say, neither binary nor unitary ground-fired chemical munitions can have any military effect unless they are deployed near the scene of battle. Almost all the large U.S. stockpile of GB and VX artillery projectiles is currently in the U.S. There is some question of the willingness of the European allies of the U.S. to allow any increase in the number of chemical weapons deployed in their territory.

Although binary munitions may be safer in some respects, opponents of the weapons point out that they are by no means harmless. One of the two precursors of agent GB is methylphosphonyldifluoride, which is reported to be about as toxic as strychnine. The other precursor, ethyl 2-(diisopropylamino)-ethylmethylphosphonite (QL) reacts violently with water and oxygen and cannot be stored in rubber or in any plastic except Teflon. Contact with QL vapor is said to cause gastric upset, respiratory difficulty and a skin rash. The perceived greater safety of the munitions may even increase the probability of their accidental release by inviting a relaxation in the security precautions taken in their handling and storage.

According to a report prepared by the Center for Defense Information, a decision to proceed with the production of the binary-nerve-gas munitions would make the negotiation of a chemical-disarmament treaty more difficult in several ways. Because the component substances of the binary agents can be obtained from commercial sources a rapid buildup of chemical weapons by the U.S. would no longer be hindered by the need to build and operate a complex chemical plant. Verification of compliance with such a treaty would be further complicated by the fact that facilities for the production and storage of the binary chemicals would be harder to distinguish from commercial chemical plants. According to the report, the production of such weapons by the U.S. would also "legitimate and encourage the spread of chemical weapons to other countries," since the U.S. would have declared workable "a new technology for the most lethal chemical weapons which is well within the grasp of many countries."

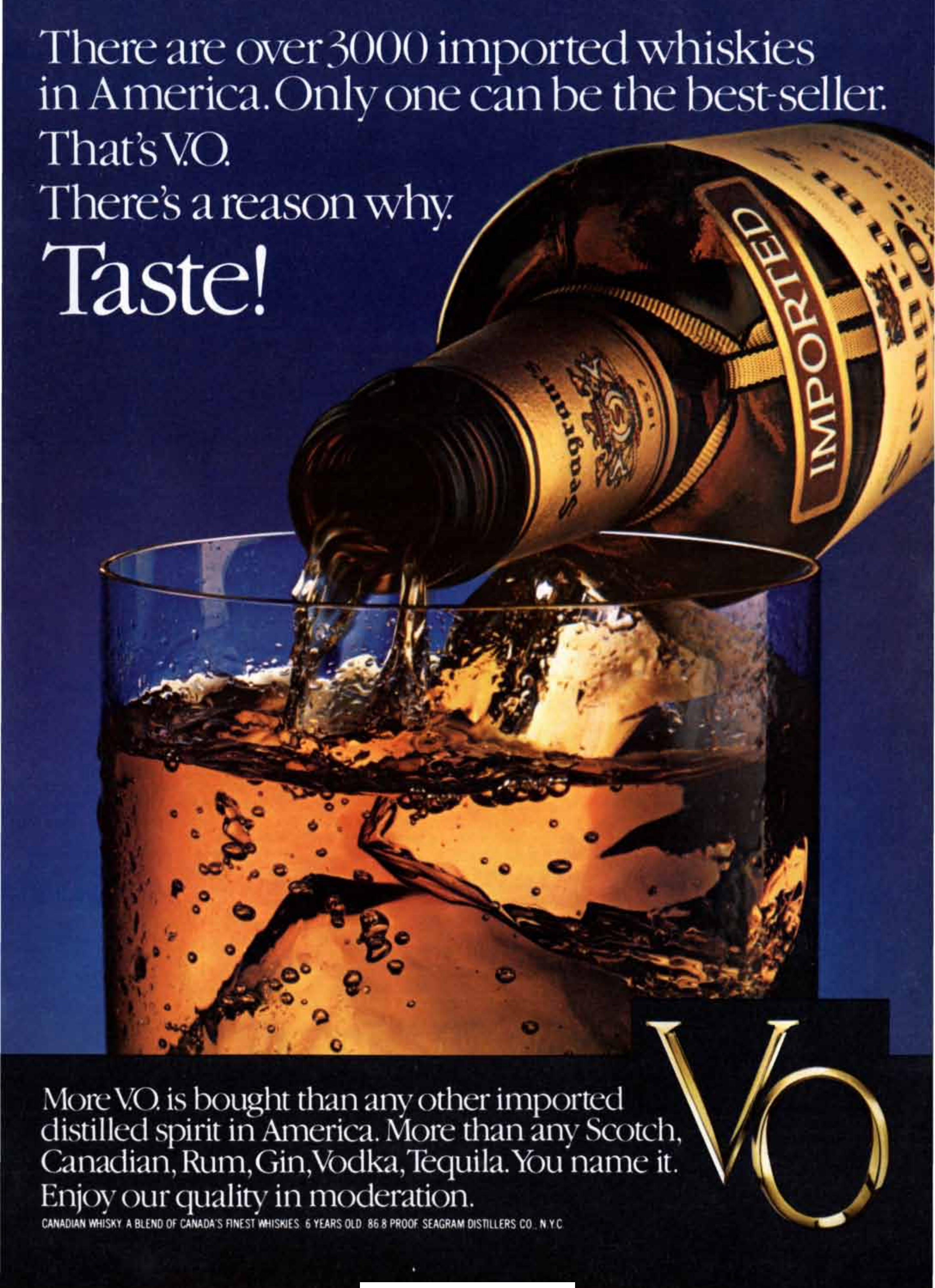
Head Starts

Programs for disadvantaged preschool children that attempt to overcome the self-perpetuating handicaps of poverty were begun in the U.S. about 20 years ago. Head Start, the largest of them, began in 1965. Some early reports on the programs, particularly an evaluation of Head Start done in 1969 by the Westinghouse Learning Corporation and Ohio University, concluded that the preschool intervention had not done much to improve the academic performance of disadvantaged children. Of necessity such early evaluations could follow the progress of the children for only a few years. It is now possible to ascertain how children who were in the programs in the 1960's are doing as teenagers and young adults. Two such longitudinal studies have recently been issued; both find that the preschool programs have been widely beneficial.

One report is *Young Children Grow Up: The Effects of the Perry Preschool Program on Youths through Age 15,* by Lawrence J. Schweinhart and David P. Weikart of the High/Scope Educational Research Foundation. The project was started in 1962 among preschool children in the attendance area of the Perry Elementary School in Ypsilanti, Mich. The 123 children included in the longitudinal study were black and from poor families. Over a period of four years 58 of the children attended (for either one year or two years) a preschool program that emphasized active learning, problem solving, motivation and communication with others in the group. The remaining 65 children received no special attention.

Tracing these children in recent years, Schweinhart and Weikart found several differences between the two groups. By age 15 the children who had attended the preschool program scored 8 percent higher than the children of the control group on tests of reading, mathematics and language. The preschool students required and received fewer years of special education as they progressed through school. Preliminary data indicate they are completing high school at a higher rate and are showing more interest in attending college. They also have better employment records and lower rates of arrest than the members of the control group.

A cost-benefit analysis done by Schweinhart and Weikart found that two years of preschool education for


There are over 3000 imported whiskies in America. Only one can be the best-seller.
That's V.O.
There's a reason why.
Taste!
Seagram's
IMPORTED
More V.O. is bought than any other imported distilled spirit in America. More than any Scotch, Canadian, Rum, Gin, Vodka, Tequila. You name it.
Enjoy our quality in moderation.
CANADIAN WHISKY. A BLEND OF CANADA'S FINEST WHISKIES. 6 YEARS OLD. 86.8 PROOF. SEAGRAM DISTILLERS CO., N.Y.C.
VO

one child cost $5,984 in 1979 dollars, whereas the benefits have a value of $14,819. The benefits include the reduced need for special education, an increase in the projected earnings of the students and the value of the mother's time as a wage earner when the child attended the preschool program.

The second report, "Head Start: A Successful Experiment," by Bernard Brown and Edith H. Grotberg of the U.S. Department of Health and Human Services, appears in the *Courrier* of the International Children's Centre in Paris. Head Start, which enrolled some 430,000 children in 1979, provides not only education but also assistance in health and nutrition. Parents are encouraged to participate, and social services are made available to the family. Brown and Grotberg summarize the conclusions of more than 700 studies of Head Start published between 1969 and 1977. The major findings are that the program has brought about gains in intelligence and academic achievement, has had a positive effect on the health and social behavior of the children and has benefited their families.

Brown and Grotberg also review the work of the Consortium for Longitudinal Studies, a group formed in 1976 to pool data from a number of studies of the long-term effects of preschool programs in many parts of the U.S. They find that "long-term positive effects were demonstrated in both cognitive and noncognitive areas." In the cognitive area "the preschool graduates had significantly fewer later school failure experiences" than control or comparison groups. Among noncognitive effects "preschool graduates had significantly higher achievement orientation as reflected in pride in work and school."

In sum, Brown and Grotberg write, "Head Start has brought about a quiet revolution in children's institutions.... Early childhood programs are now accepted and functioning. The acceptance can be seen not only in support for Head Start but also in the phenomenal rise in nursery schools, preschool programs in the public schools and quality day care for middle-class children. It is evidenced in the large numbers of books, records, toys and television programs for preschool children."

Sunny Disposition

The persistent human faith that the sun will rise tomorrow may have been tempered a year and a half ago by the suggestion that when the sun does rise, it might be slightly smaller. The suggestion came from a historical study completed in 1979, which indicated that the sun may have been shrinking steadily for at least a century. More recent studies and reinterpretations of the earlier ones now suggest that the rate of shrinkage is somewhat lower. Other workers report still smaller limits for the shrinkage and assert there has been no detectable long-term change. The controversy has not been settled. Yet it seems theories of the sun must now at least consider the possibility that the solar diameter is somewhat changeable.

The question of solar contraction was first raised by John A. Eddy of the High Altitude Observatory of the National Center for Atmospheric Research and Aram A. Boornazian of S. Ross and Co. in Boston. Eddy and Boornazian had examined data collected at the Royal Greenwich Observatory from 1836 through 1953. The data were records made at noon on each sufficiently clear day of the time at which first one edge and then the other edge of the image of the sun crossed a wire in the eyepiece of a telescope. Although the observations were made in order to fix the instant of local noon, each such pair of timings allows a calculation of the size of the sun. Analysis of the Greenwich Observatory records led Eddy and Boornazian to report that the sun may have been shrinking at a rate of two arc seconds per century in its diameter.

The diameter of the sun is approximately 1,920 arc seconds. Hence a simple extrapolation leads to the prediction that the sun will disappear in 960 centuries. A more complex calculation takes account of the fact that the shrinkage converts gravitational potential energy into kinetic energy and thereby augments the luminosity of the sun. A century ago Hermann von Helmholtz proposed that a shrinkage of the solar diameter by just .01 arc second per century would account for all the luminosity. (The release of energy in the sun by the thermonuclear fusion of hydrogen into helium was of course unknown.) What gave Helmholtz' speculation a particular appeal was that shrinkage at a rate of .01 arc second per century would be unobservably slow. On the other hand, shrinkage at a rate on the order of one arc second per century, as initially proposed by Eddy and Boornazian, might increase the luminosity by several hundred times if all the mass in the sun took part in the contraction.

A possibility that seems more likely (and less disturbing) is that only the outer layer of the sun is shrinking. No dramatic effect on the luminosity would then be expected. Small changes in the solar luminosity might nonetheless modify the climate of the earth. There is also an alternative to the prospect that the sun might dwindle away to nothing. The shrinkage may be part of a cycle of expansion and contraction. Oscillations of the sun that have a period of approximately five minutes are fairly well established. Oscillations with a longer period (for example 160 minutes) have also been proposed.

A group of investigators led by David W. Dunham of the International Occultation Timing Association in Silver Spring, Md., now reports determinations of the solar radius by a technique that exploits total eclipses of the sun. Some of the results appear in *Science;* others have not yet been published. During a total eclipse a circular shadow of the moon moves across the surface of the earth. The duration of the eclipse is maximal in the path traced by the center of the shadow; the duration is zero at the edge of the shadow or outside it.

The investigators in Dunham's group station themselves just inside the predicted path of the edge of the shadow. In such a place totality lasts for only seconds. For a period of several minutes, however, the mountains and valleys at the edge of the lunar disk break the crescent of the sun into the string of bright spots called Baily's beads. At the edge of the path of totality the beads near the north or the south pole of the moon are observed. The poles are the places on the moon least affected by the slow, rocking lunar motion called libration; thus the lunar contours that give rise to the polar beads are little changed from one eclipse to another. Each bead constitutes a mapping of a point on the moon onto a point at the edge of the solar disk. The observers time the appearance of each bead. From the timings and from the known configuration and motion of the earth and the moon with respect to the sun, the size of the sun can be calculated.

Dunham and his co-workers have made such observations themselves for solar eclipses on October 23, 1976, February 26, 1979, and February 16, 1980. They find no notable change in that brief interval. "Thanks to the promotional efforts of Sir Edmund Halley," Dunham and his colleagues write in *Science,* they also have data for "both edges of the path of the 3 May 1715 total solar eclipse in England." Halley organized observers throughout England and collected their reports.

Dunham and his colleagues note that the southern edge of the path of the 1715 eclipse is reliably fixed by two observations. The northern edge, however, must be determined from a single datum that has proved difficult to interpret. The datum was provided by Theophilus Shelton, Esq., who watched the eclipse from the village of Darrington in Yorkshire. He saw, according to Halley, that "the Sun at $9^h 11^m$ was reduced almost to a Point, which both in Colour and Size resembled the Planet *Mars;* but whilst he watched for the Total Eclipse, that Point grew bigger and the Darkness diminished; whence he argued the [northern] Limit [of the path of the eclipse] to have been very little more Southerly." In other words, Shelton thought he was just outside the edge.

Dunham and his co-workers suspect that Shelton was just inside the edge, so that the limit was north of his position;


An RCA Solid State Laser that fits through the eye of a needle can transmit 500 million bits of information per second through a thread of glass.
The entire contents of a 24-volume encyclopedia in 3 minutes. 2500 phone conversations or 20 TV programs at the same time. Staggering amounts of information – all made to pass through a tiny thread of glass by a solid state laser that radiates light from an area one-hundredth the size of the period at the end of this sentence.
This laser, developed by RCA scientists, emits a narrow, intense beam of light that can be efficiently coupled into the microscopic fibers that carry the information. Because of its incredible stability and small size, RCA's solid state "double dovetail" laser represents a significant step toward faster, smaller, more efficient fiber-optic communications systems.
This Solid State Laser is a product of RCA research – the kind of dedicated research that has been a tradition at RCA for over sixty years. Today, many years after Nipper started listening to "His Master's Voice," research and development of electronic products are funded at the highest level in our company's history. Our commitment is, as it has always been, to the advancement of technology – to the creation of products that expand the human horizon, whether through a wave of sound or a thread of light.
©1980 RCA Corporation
RCA
A tradition on the move

the redness of the pointlike image of the sun (that is, its resemblance in color to Mars) suggests that what Shelton saw was part of the chromosphere of the sun during a brief totality. With Shelton's position established, the radius of the sun in 1715 can ultimately be calculated. The finding reported in *Science* is a contraction in the radius of some .34 arc second between 1715 and 1979. That is equivalent to a shrinkage of the solar diameter at .26 arc second per century.

Eddy and Boornazian's results as well as the lesser contraction found by Dunham and his colleagues have been challenged by others. For example, John H. Parkinson of University College London, Leslie V. Morrison of the Greenwich Observatory and F. Richard Stephenson of the University of Liverpool cite numerous changes in the procedures at the Greenwich Observatory. Writing in *Nature,* they point out that the method of timing the observations was changed in 1854; the telescope's objective glass was repolished in 1891 and 1906; the telescope itself was replaced in 1851. Each change may have affected the trend of the observations. Moreover, "the observational reports from Greenwich show that weather and seeing conditions were often far from ideal." In addition the bias of each observer in judging the edge of the sun "would be expected to increase the scatter in the results." These difficulties make the data "unsuitable for investigating possible changes in the solar diameter." Parkinson and his colleagues add that the conclusion reached by Dunham and his associates "depends critically on the interpretation put on the only observation made near the northern limit of the 1715 eclipse."

Parkinson and his colleagues have studied total solar eclipses in order to determine the diameter of the sun. They base their findings on the duration of totality for the eclipses of 1715, 1842, 1851, 1878, 1900 and 1925. They also employ more than 2,000 timings of 30 occasions in the past 250 years when the planet Mercury passed behind the sun. They find "no detectable secular change in the sun's diameter since at least A.D. 1700." (A secular change is generally one that lasts for some hundreds of years.) On a shorter time scale, however, the eclipses and the transits of Mercury suggest that the radius of the sun is decreasing by .08 arc second per century. The value is not very different from that found by Dunham's group.

Dunham and his co-workers have now examined the eclipse of 1925. They are preparing to report a net decrease of a few tenths of an arc second in the solar radius since 1715. The eclipse of 1925 seems to suggest a brief upward excursion of the radius. The group intends to analyze 19th-century observations of total solar eclipses.

Eddy and Boornazian have been reexamining the data that started the controversy. They agree there are two reasons for making adjustments. In the first place an observer at the Greenwich Observatory looks south over an industrial section of London to see the sun at noon. The sky there loses transparency whenever pollution increases. In the second place Eddy and Boornazian now find that four workers at the Greenwich Observatory made measurements of the solar transit that deviate notably from those of the others. (More than 80 people made the measurements from 1836 through 1953.) Corrections made for both of these reasons now lead Eddy and Boornazian to conclude that the Greenwich data suggest a shrinkage of about one arc second per century in the diameter. It is a conclusion, Eddy says, "within range of many of the others."

Eddy considers the present data inadequate to determine decisively whether the sun has a constant size or alternately shrinks and expands. The data are largely historical. Moreover, none of the data constitute a direct measurement of the solar diameter or radius. He considers it possible, however, that the historical data have yielded the first evidence of a solar variability that may influence the climate of the earth. Three instruments are being readied to measure the diameter directly. The measurements will be made by groups at the Kitt Peak National Observatory, the University of Arizona and the National Center for Atmospheric Research. A result should be known in from three to five years.

The Moving Finger

Many animal cells are capable of amoeboid motion. They flatten out and retract; they develop transitory bumps or bubbles called blebs; they extend ruffled edges (lamellipodia) and wave threadlike extensions (filopodia). By alternately affixing lamellipodia or filopodia to a substrate and then detaching them the cells achieve a creeping locomotion that has a role in biological phenomena ranging from embryonic development to the metastasis of tumor cells. Even when the nucleus of a motile cell is removed experimentally, the enucleated "cytoplast" still ruffles, waves its filopodia and manages to get about; so do sizable pieces of cytoplasts. Movement, then, is evidently a property not only of the entire cell but also of parts of the cell. Is there some small structural unit of amoeboid motion in animal cells? A search for the smallest fragments of cytoplasm capable of the various movements has been undertaken by Guenter Albrecht-Buehler of the Cold Spring Harbor Laboratory. He reports his initial findings in *Proceedings of the National Academy of Sciences.*

Albrecht-Buehler grew fibroblasts (connective-tissue cells) from human skin in a plastic culture dish. When the cells had flattened themselves on the bottom of the dish, he treated them with cytochalasin B, a substance that causes fibroblasts to retract into thin, branching forms attached to the substrate at only a few points. He squirted the retracted cells vigorously with culture medium, breaking up most of the cells and leaving only small fragments fixed to the dish. The fragments healed to form spindly or triangular "microplasts," which were separated from the dish and replated on new surfaces for observation by light and electron microscopy.

The microplasts lived for eight hours or more. Once replated they generated waving filopodia, ruffling lamellipodia or blebs. Unlike whole cells or larger fragments of cells, however, each microplast had a characteristic, stereotyped mode of movement. Some ruffled or formed filopodia or did both; some developed blebs. None of the microplasts appeared to be capable of actual locomotion. They did display a vestige of the contact inhibition by which intact cells seem to repel one another. Instead of actually moving away from an abutting fragment, however, a microplast seemed only to stretch away from the point of contact, unable to pull its rearmost part along with its leading edge.

The size of the smallest microplasts displaying at least some form of autonomous movement was about 150 cubic micrometers, which is less than 2 percent of the volume of the original fibroblasts. The ultrastructure of the fragments, revealed in electron micrographs of thin sections, included at least a partial complement of typical cytoplasmic structures such as microfilaments, intermediate filaments and microtubules (see "The Ground Substance of the Living Cell," by Keith R. Porter and Jonathan B. Tucker, page 56). No correlation was apparent, however, between the ultrastructure of an individual microplast and its limited repertory of movements. It seems likely that each fragment carries some long-lived determinant of one type of movement or another, but the nature of the determinants is not known. It appears to Albrecht-Buehler that there may indeed be autonomous units of amoeboid motion. The coordinated movements of the whole cell, including locomotion, may eventually be understood as the result of communication and cooperation among such units under the influence of some supervising mechanism.

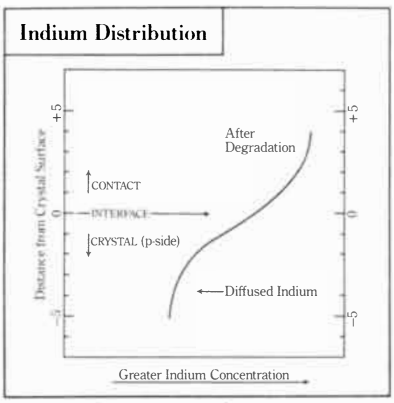
Going to the Roundies

Of all the visual arts the motion picture would seem to offer the most persuasive imitation of life, and yet it cannot fully satisfy the appetite for verisimilitude. What the movies lack most obviously is depth, and a number of methods have been tried for adding a third dimension to the projected image.

The Strategic Misalignment

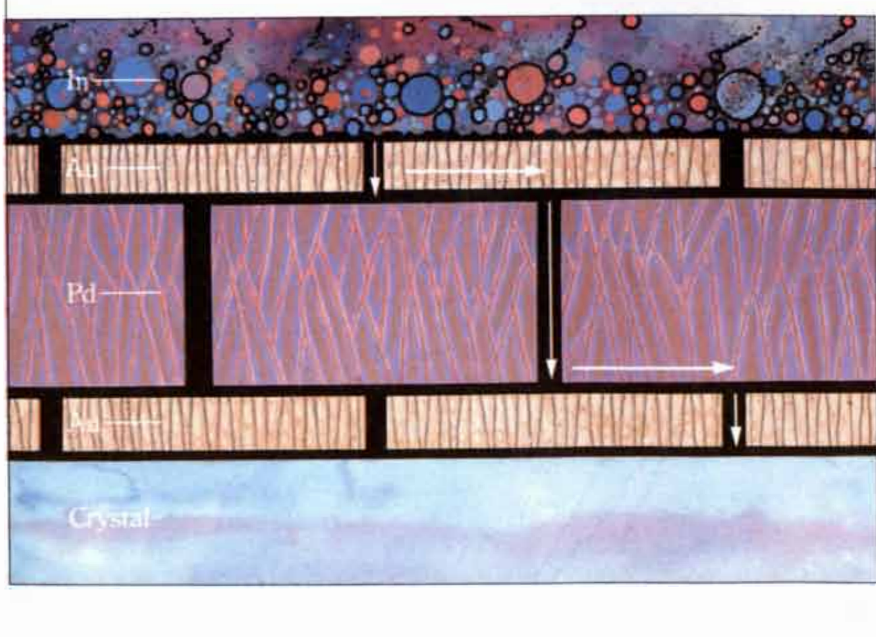
The Strategic Misalignment

Tunable semiconductor lasers can now measure specific gases in automotive exhaust with 25-millisecond response time. A successful strategy for improving laser reliability developed at the General Motors Research Laboratories makes this and other new spectroscopy capabilities practical realities.



Electron microprobe analysis of a crystal-contact interface, indicating indium penetration into the PbSnTe crystal.

Diagram of hypothetical indium diffusion paths for a three-layer contact structure of Au-Pd-Au.



THE ACHIEVEMENT of long lifetime and frequency stability makes the lead-tin-telluride diode laser a practical infrared spectrometer. Earlier innovations brought to this laser the characteristics of increased power, higher temperature operation, greater efficiency and wider tuning range.

Operating in the 5- to 10-micron range, the PbSnTe laser spectrometer can resolve the time-dependent emission of carbon monoxide, sulfuric acid vapor, methane and other species of interest in automotive exhaust. This permits measurement of transients in carbon monoxide to carbon dioxide gas conversion in a catalytic converter. This capability represents a significant advance over conventional spectroscopy instrumentation. The laser is also being tested by NASA for use in detecting the molecular species involved in chemical reactions in the stratosphere.

New knowledge of the process by which laser reliability is compromised has been revealed in fundamental studies conducted by Dr. Wayne Lo and his colleagues at General Motors. Dr. Lo's investigations have demonstrated that laser lifetime and stability are limited by the development of excessive electrical contact resistance. He has been able to stop increases in resistance by devising a multilayer ohmic contact consisting of different metal films. This configuration has extended laser operating lifetime to more than 1,000 hours and increased shelf-life to an estimated 25 years.

Slow degradation due to a gradual increase in contact resistance was observed in idle lasers stored at room temperature, but not in lasers maintained at a maximum temperature of 77 K, despite several hundred hours of continuous operation. These results suggested the temperature-dependent process of diffusion.

Degradation occurred primarily on the p-type side of the laser, where the contact consisted of a thin layer of gold followed by a layer of indium. Electron microprobe analyses revealed that indium, a semiconductor donor, was diffusing through the gold layer into the crystal, apparently causing a reduction in hole carrier concentration near the p-surface. This effect was counteracted to a great degree by sandwiching a thin layer of platinum between the layers of indium and gold. Laser reliability reached a full year.

When degradation was still observed, although to a reduced extent, Dr. Lo advanced the hypothesis that diffusion and transport were taking place along grain boundaries in the polycrystalline contact layers. He proposed replacing the Pt-Au barrier with a three-layer structure. Since palladium film structures have fewer grain boundaries than those of platinum, providing fewer leakage paths for the indium, Pd was tested in place of Pt.

DIODE LASERS composed of $Pb_{0.86}Sn_{0.14}Te$ and fabricated with a variety of contacts were maintained at 60°C in order to accelerate aging, with periodic interruptions for testing. The results showed that a multilayer structure of In-Au-Pd-Au, in which the grain boundaries tend to be misaligned, provides maximal reduction of indium penetration, confirming Dr. Lo's hypothesis.

The misaligned boundaries force diffusion to take place laterally, which slows transport into the crystal. The additional layer slows the process even further.

Solving the contact problem represents the culmination of efforts that began at General Motors with the development of an "ingot-nucleation" vapor transport method for growing crystals. The resulting crystals are of high purity, with a dislocation density of less than 1000 cm^{-2}. Lasers made from these crystals incorporate a low temperature cadmium-diffused p-n junction. This process, invented by Dr. Lo, increases the laser's output to five milliwatts.

A tuning range of 500 cm^{-1} and pulsed operating temperatures of up to 140 K are achieved by a two-step annealing process. This technique induces a graded carrier concentration that increases infrared light confinement in the laser structure, thus reducing losses and increasing output.

"These innovations," says Dr. Lo, "combine to produce a laser that allows us to make measurements previously impossible."

THE MAN BEHIND THE WORK

Dr. Wayne Lo is a Senior Research Scientist in the Physics Department at the General Motors Research Laboratories.

Dr. Lo was born in Hupei, China. He did his undergraduate work at Cheng-Kung University in Taiwan. He received an M.S. from the University of Rhode Island and a Ph. D. in electrical engineering from Columbia University in 1972. His doctoral thesis concerned the characterization of deep-level states and carrier lifetimes in gallium arsenide light-emitting diodes.

Before undertaking graduate studies, Dr. Lo was instrumental in setting up the first American transistor production plant in Taiwan. In 1973, he joined General Motors, where he is currently in charge of semiconductor laser and spectroscopy research.

GM General Motors
People building transportation to serve people

Most such methods have been based on stereoscopic principles, whereby slightly different images are presented to the two eyes. In the 1960's, however, the development of laser holography suggested a new approach. A hologram gives rise to a true three-dimensional image, which changes in appearance when the viewer changes position; one part of the image can even be eclipsed from view by another part. A holographic movie might therefore create a realistic illusion of solid objects moving in three-dimensional space. The techniques needed for making and projecting holographic movies have proved difficult to master, but some recent progress has been reported.

Investigators at the Cinema and Photographic Research Institute in Moscow are at work on a holographic motion-picture system for public theaters. In 1976 the group exhibited a 45-second holographic movie in which a young woman holding a bouquet of flowers in front of her face seems to walk out of the screen toward the audience. Each viewer, by moving his head horizontally or vertically, could look around the bouquet and see the woman's face. The effect was startling, although the registration of the image from frame to frame was not perfect and the color was unrealistic (the image was yellow). The movie and the way it was made are described by Tung H. Jeong of Lake Forest College in *Optical Spectra.*

In the years since the 45-second movie was shown the techniques of holographic cinematography have improved considerably. The recent developments are discussed by two members of the Moscow group, Victor G. Komar and O. I. Ioshin, in *SMPTE Journal,* a publication of the Society of Motion Picture and Television Engineers.

A hologram generates a three-dimensional image of an object by reconstructing the original light waves that were reflected from the object. In a basic method of making a hologram coherent light from a laser is split into two beams. One beam, the reference beam, passes directly to a photographic plate; the other beam is aimed at the object. The light reflected from the object combines at the plate with the reference beam to form an interference pattern, which is recorded on the plate. The developed plate is the hologram.

The hologram is viewed by illuminating it with laser light of the same wavelength as the light with which the hologram was made. The light waves passing through the hologram are transformed so that the emerging light waves reconstruct the ones that were reflected from the object when the plate was exposed. An observer perceives the waves as coming not from the hologram but from the object.

The theoretical and technological difficulties in extending holography to motion pictures are substantial. The Cinema and Photographic Research Institute is one of only a few institutions where the work has been attempted. Surprisingly, the most challenging problems may lie not in making holographic motion pictures but in projecting them to a large audience. The projection of the hologram onto an ordinary movie screen would destroy the reconstructive property of the image. It is necessary for the observer to intercept not a projection of the light rays emerging from the hologram but the rays themselves.

One hypothetical method of presentation, explained by Jeong, would employ a mirror rather than a diffusely reflective screen as the projection surface. For example, the mirror might be an elliptical one with two foci. The viewer would sit at one focus and the holographic image would be formed at the other focus. The drawback to this method is that it would work only for an audience of one. For a larger audience the screen would have to be equivalent to a superposition of as many mirrors as there are viewers, with the image at a common focus and each seat at the second focus of one of the mirrors.

It is impossible, however, to combine mirrors in this way. The investigators in Moscow, according to Jeong, overcame the problem by employing as the screen a specially constructed hologram that has the same multiple-focusing property as superposed mirrors would have. The hologram is made by exposing a photographic plate the size of the screen to beams diverging from each seat in the theater and from the point where the image will be formed. When light rays from the projector impinge on the holographic screen, they come to a focus at each seat. The movie of the woman and the bouquet was shown with a screen that could accommodate an audience of four. The investigators plan to construct a seven-foot-wide holographic screen for an audience of between 200 and 400 viewers.

The Moscow group points out that a holographic movie might have advantages besides visual realism. For example, the quality of the reconstructed image is not affected by motion of the film in a direction perpendicular to the projected beam. As a result the image would not be degraded if the film moved somewhat irregularly through the projector. Another advantage is that the size of the individual frames making up a holographic movie could be much smaller than the frame size of an ordinary movie. The reason is that information embodied in the light reflected from any part of the object is recorded over the entire surface of the hologram. A decrease in the frame size could lead to reductions in the cost of the film and in the size of the camera and the projector. Moreover, if a hologram is scratched or dirty, the flaw will scarcely be noticeable because its effect will be spread over the entire image.

Seismic Shift

The earthquake that shook San Francisco on the morning of April 18, 1906, devasted four square miles of the city. Although much of the damage was actually caused by the ensuing fire, it has long been the consensus among seismologists that the earthquake itself was an exceptionally large one. It is usually assigned a magnitude of 8.25, which suggests a geologic blockbuster. Hence the San Francisco earthquake has served as a benchmark against which other earthquakes are compared; it has also been adopted with a certain pride by the citizens of San Francisco as a measure of the city's resilient spirit.

The extent of the damage and suffering caused by the 1906 earthquake is beyond dispute, but it has now been suggested that the earthquake as a geologic event was not as extraordinary as had been thought. Paul C. Jennings and Hiroo Kanamori of the California Institute of Technology have recovered seismic instruments of the period and have reexamined old data in the light of the instruments' responses. They conclude that the local magnitude of the San Francisco earthquake was probably between 6.75 and 7.0 on the Richter scale. This magnitude is comparable to that of more recent large earthquakes in California. Because the magnitude of an earthquake represents the base-10 logarithm of the amplitude of the shocks, the revision suggests the force of the San Francisco earthquake had been overestimated by as much as a factor of 30.

Why is there such a large discrepancy? Jennings and Kanamori point out that the assigned magnitude of 8.25 is an estimate of the surface-wave magnitude: a measure of the amplitude of low-frequency (three per minute) ground vibrations. This amplitude depends mainly on the length of the fault line along which the plates of the earth's crust slip during an earthquake. The surface-wave magnitude of 8.25 tells the tale of a gigantic movement in the crust, running some 250 miles along the San Andreas Fault.

The slow surface-wave motions, however, are scarcely noticeable to those for whom an earthquake is a matter of personal or architectural concern. Most of the ground shaking near an earthquake epicenter results instead from higher-frequency (one per second) vibrations. The one-per-second waves are more destructive, but they are also more quickly attenuated by passage through the earth. Beyond a distance of some hundreds of miles the longer-period surface waves become dominant, and at extreme distances only the surface waves can be detected. The potential destructive intensity of an earthquake is best estimated, therefore, by local magnitude, which alone is capable of measuring the one-per-second waves. The Richter magnitude properly refers to this local magnitude, but common usage has extended it to refer to earthquake magnitude scales in general.

Various kinds of seismic instruments had been in use for some 50 years by the time of the San Francisco earthquake, but none were specifically tuned to one-per-second vibrations. One of the simplest instruments was the seismoscope, in which a massive pendulum is suspended over a smoked glass, with the pointed end of the pendulum bob just touching the glass. In an earthquake a record of the ground motion is scribed onto the glass. Seismoscopes were in operation near the epicenter of the San Francisco earthquake, but most of them were thrown off scale. Estimates of the earthquake magnitude have therefore been based entirely on measurements of distant surface-wave amplitudes. Indeed, until now the size of the San Francisco earthquake has been inferred primarily from data recorded by European seismic stations.

Jennings and Kanamori reasoned that the signature of the one-per-second waves associated with the 1906 earthquake might be reconstructed from recordings made by instruments that were close enough to the epicenter to respond to this rapidly fading vibration yet far enough away so as not to be thrown off scale. Recordings made on such early instruments, however, would bear no simple relation to local magnitudes. Hence the problem facing Jennings and Kanamori was not merely that of finding the appropriate instruments; their task amounted to the archaeological reconstruction of a broken tool and the theoretical reinterpretation of the markings made by it.

Jennings and Kanamori turned up two early seismoscopes in laboratory storerooms along with records of the 1906 earthquake made with those instruments. One of the seismoscopes was in Yountville, Calif., about 65 kilometers from the epicenter of the earthquake. The other was found in Reno, Nev., but had recorded the San Francisco earthquake from Carson City, Nev., some 291 kilometers from the epicenter. Both of the instruments were put in order and calibrated for period, damping and gain so that their recordings could be evaluated.

The revised estimate of the magnitude of the San Francisco earthquake might be seen as supporting the uniformitarian basis of modern geology: the catastrophes of 75 years ago were not different in kind from those of today. The revision does prompt one melancholy thought. Under 1906 building standards it did not take an 8.25 earthquake to bring San Francisco to the ground; a 6.75 or a 7.0 was quite enough.

The Milky Way Galaxy

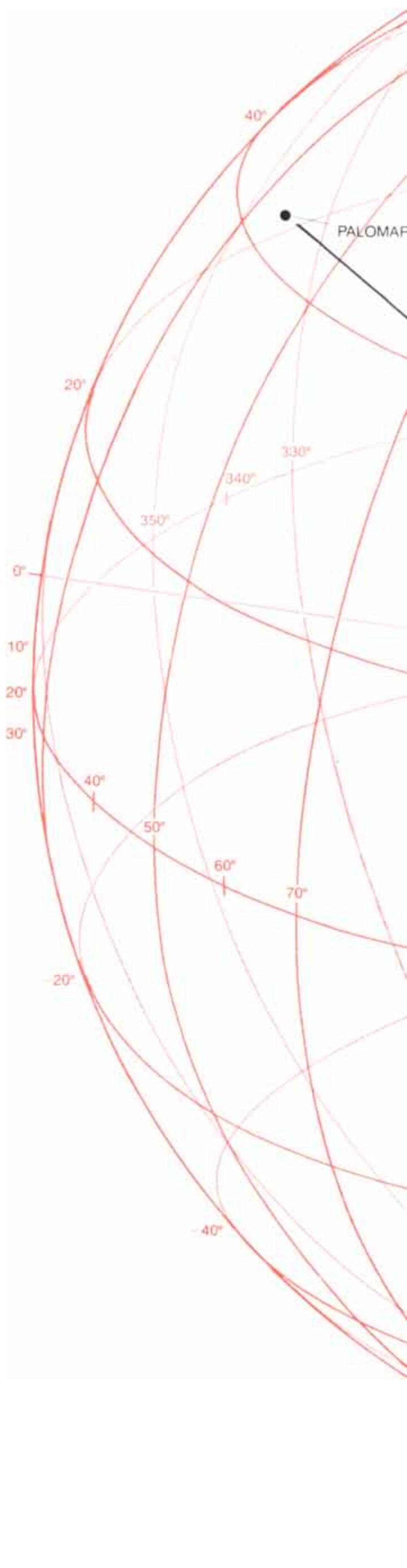
A few years ago the fundamental facts about it seemed to be fairly well established. Now even its mass and its radius have come into question

by Bart J. Bok

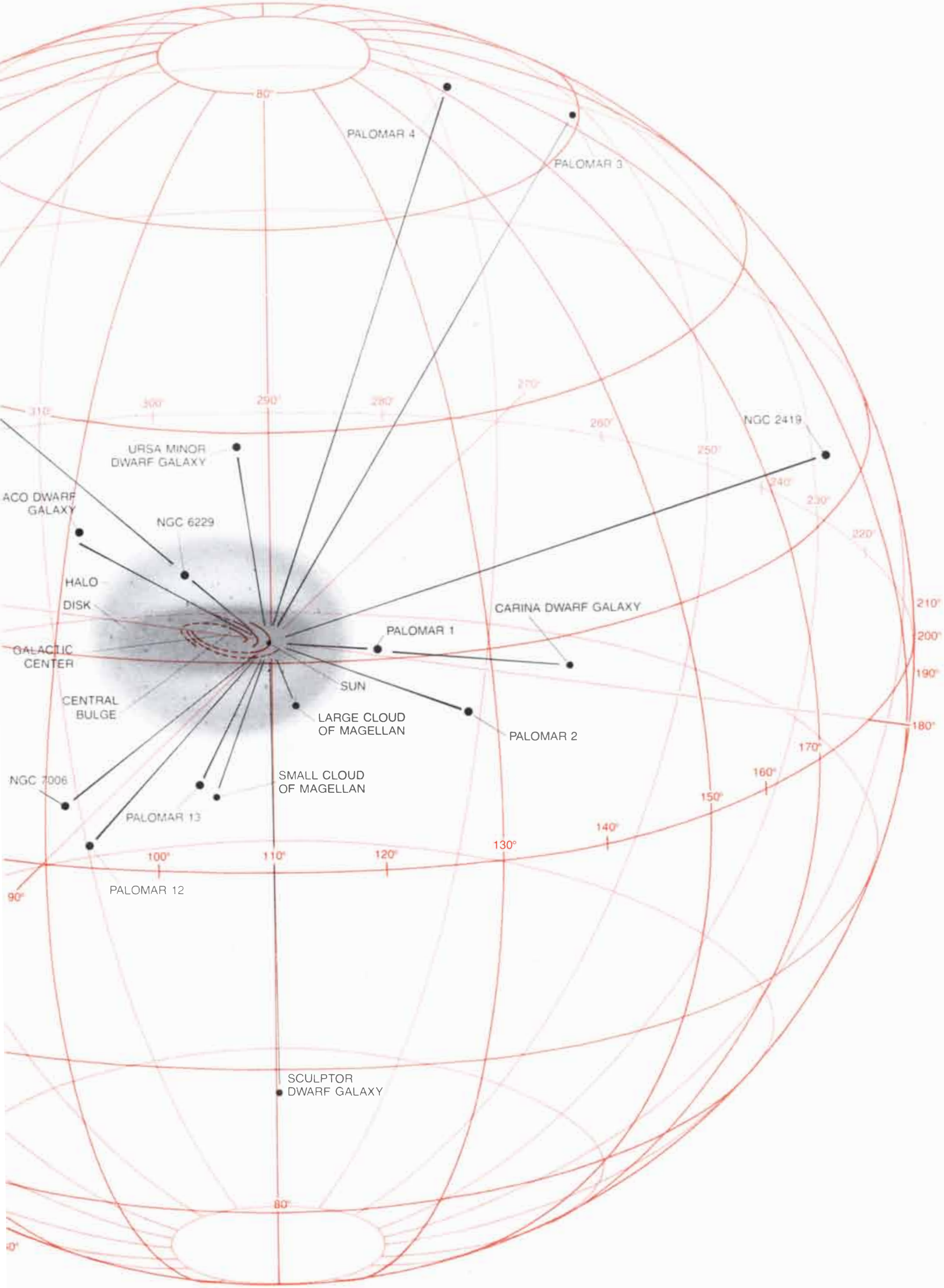
On a clear and moonless night free of the lights of civilization the most arresting thing in the sky is the luminous band of the Milky Way. Even without a telescope there is much one can say about it. One can see, for example, that stars become more numerous as one directs one's gaze across the sky and into the band. The fact that the band itself is made up mostly of stars then seems less astonishing. One can see that the band follows a great circle that bisects the celestial sphere. Hence the earth is embedded in the central plane of the band. One can see, moreover, that the band is widest and brightest in the direction of the constellation Sagittarius. Surely this is the direction to the center of the system. The system is the Milky Way galaxy.

A telescope will reveal that certain starlike points in the Milky Way are actually great aggregations of stars. They are plainly distant objects; they are known as globular clusters. Between 1918 and 1921 Harlow Shapley employed the telescopes of the Mount Wilson Observatory to demonstrate that the clusters, like the stars in the band, are commonest in Sagittarius. In an area of Sagittarius that makes up only 2 percent of the sky Shapley plotted a third of all the globular clusters then known. Evidently, therefore, the solar system is far from the center of the galaxy; the distance from the sun to the center is now estimated to be 8,500 parsecs. (A parsec is 3.26 light-years.) It has since been established that the solar system, along with the rest of the mass in the central plane of the galaxy, revolves about the center. The sun revolves at a rate now taken to be some 230 kilometers per second. It thus completes a revolution every 200 million years.

In 1930 Robert J. Trumpler of the Lick Observatory showed that an interstellar medium of gas and dust dims the light of the stars, particularly the stars in the central plane of the galaxy. In 1951 William W. Morgan of the Yerkes Observatory and his students Donald E. Osterbrock and Stewart L. Sharpless found evidence that the band is an edge-on view from the earth of the galaxy's spiral structure. Evidence for spiral features soon followed at radio wavelengths. In the 1960's Chia Chiao Lin of the Massachusetts Institute of Technology and Frank H. Shu, then at the Harvard College Observatory, suggested that waves of increased density in the interstellar medium precipitate the spiral features and are responsible for the formation of stars. Meanwhile the advent of infrared astronomy provided a means for exploring the interior of dark interstellar clouds. It now appears that the clouds are where most of the galaxy's new stars form. By 1970 radio astronomy had begun to reveal the composition of the dark clouds. They are



MAP OF THE MILKY WAY shows the galaxy in accord with the hypothesis that it is unexpectedly large and massive. The three long-recognized components of the galaxy lie well inside the coordinate grid. Their dimensions are given in parsecs. (A parsec is 3.26 light-years.) Among the three components the central bulge, with a radius of 4,000 to 5,000 parsecs, consists mostly of a dense packing of old stars. The galactic disk, with a radius of 15,000 parsecs, consists of younger stars and dust and gas. Its spiral features (*colored curves*) have been traced only in the sun's vicinity. The galactic halo, with a radius of 20,000 parsecs, consists mostly of a thin packing of old stars and of roughly half of the stellar aggregations called globular clusters. The hypothetical outermost component of the galaxy is called the corona; its presence is inferred from the velocities of the visible matter. Presumably the objects in the corona are not highly luminous. In this map the "galactic companions" visible in the corona to a distance of 100,000 parsecs are plotted by sun-centered coordinates. The direction from the sun to the center of the galaxy defines zero degrees galactic longitude. Angles above and below the galactic plane are measured in galactic latitude. The companions include 10 globular clusters, four dwarf spheroidal galaxies and the irregular galaxies the Large and Small Clouds of Magellan.



80°
PALOMAR 4
PALOMAR 3
310°
300°
290°
280°
270°
260°
250°
240°
230°
220°
NGC 2419
URSA MINOR
DWARF GALAXY
ACO DWARF
GALAXY
NGC 6229
HALO
DISK
GALACTIC
CENTER
CENTRAL
BULGE
CARINA DWARF GALAXY
PALOMAR 1
SUN
LARGE CLOUD
OF MAGELLAN
PALOMAR 2
210°
200°
190°
180°
170°
160°
150°
140°
130°
120°
110°
100°
90°
SMALL CLOUD
OF MAGELLAN
NGC 7006
PALOMAR 13
PALOMAR 12
SCULPTOR
DWARF GALAXY
80°

made up of dust and hydrogen, with an admixture of some surprisingly complex molecules, many of them organic.

For many years I have been a night watchman of the Milky Way galaxy. I remember the mid-1970's as a time when I and my fellow watchers were notably self-assured. The broad outlines of the galaxy seemed reasonably well established. The galaxy had two main components: a dense central bulge of stars with a boundary between 4,000 and 5,000 parsecs from the center, and a flat, much thinner disk of stars and interstellar gas and dust whose inner margin abutted the bulge and whose outer margin lay some 15,000 parsecs from the center. The bulge is in Sagittarius; the disk is flung across the sky.

The combined mass of the disk and the bulge was then calculated to be well under 200 billion times the mass of the sun. The disk and the bulge were surrounded, however, by a "halo" of matter on each side of the galaxy's central plane. In overall form the halo is a slightly flattened sphere. In the plane of the galaxy it has a radius on the order of 20,000 parsecs. It is notable for old stars, and also for a scattering of perhaps 100 globular clusters. (Another 100 globular clusters lie in or near the galactic disk.) The halo might add at most 100 billion solar masses to the total mass of the galaxy.

Much work was still to be done. We felt confident nonetheless that the basic findings would stand, and that astronomers would be undistracted as they looked into such matters as how stars form and how the Milky Way's spiral features evolve. We did not suspect it would soon be necessary to revise the radius of the Milky Way upward by a factor of three or more and to increase its mass by as much as a factor of 10. The revisions are emblematic of a number of recent upheavals. Here I shall take up several aspects of the current effort to understand the Milky Way. I include most of them because they are areas of notable ferment where progress in understanding may be imminent. I include one, in contrast, because I fear it is reaching a dead end.

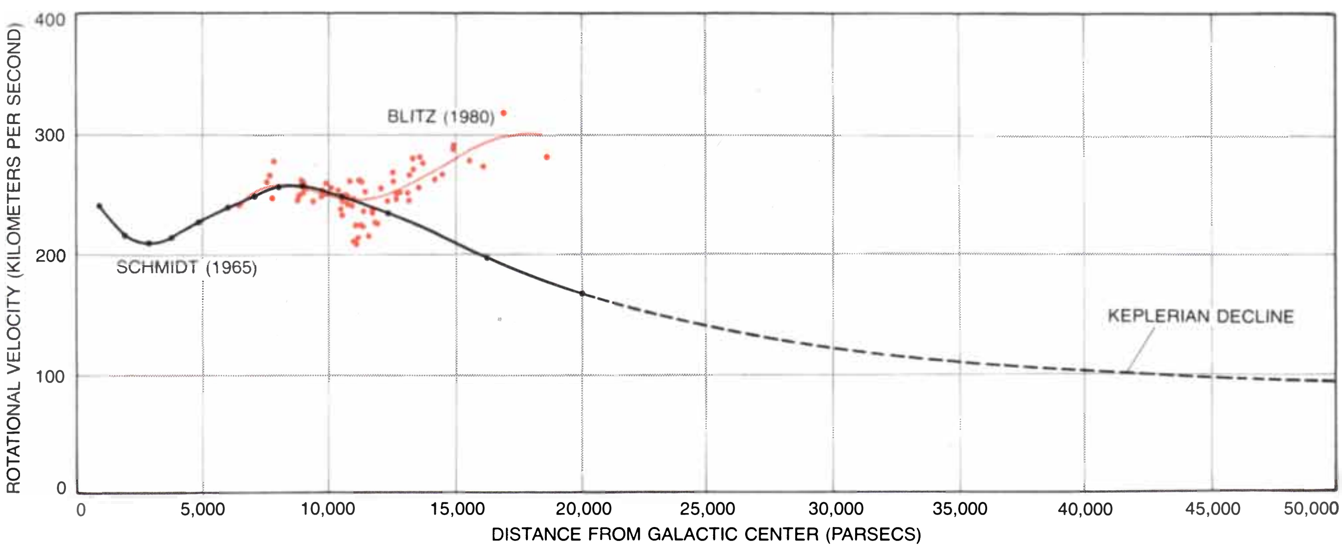
Galactic Coronas

Suggestions that the Milky Way galaxy is unexpectedly large and massive were offered as early as 1974. On the theoretical side they came principally from Donald Lynden-Bell of the University of Cambridge and from Jeremiah P. Ostriker, P. J. E. Peebles and Amos Yahil at Princeton University. The basic argument was that the dynamical stability and permanence of the galaxy cannot be guaranteed unless the galactic disk is surrounded, and thereby stabilized gravitationally in spite of its thinness and its delicate spiral structure, by an extended and massive halo.

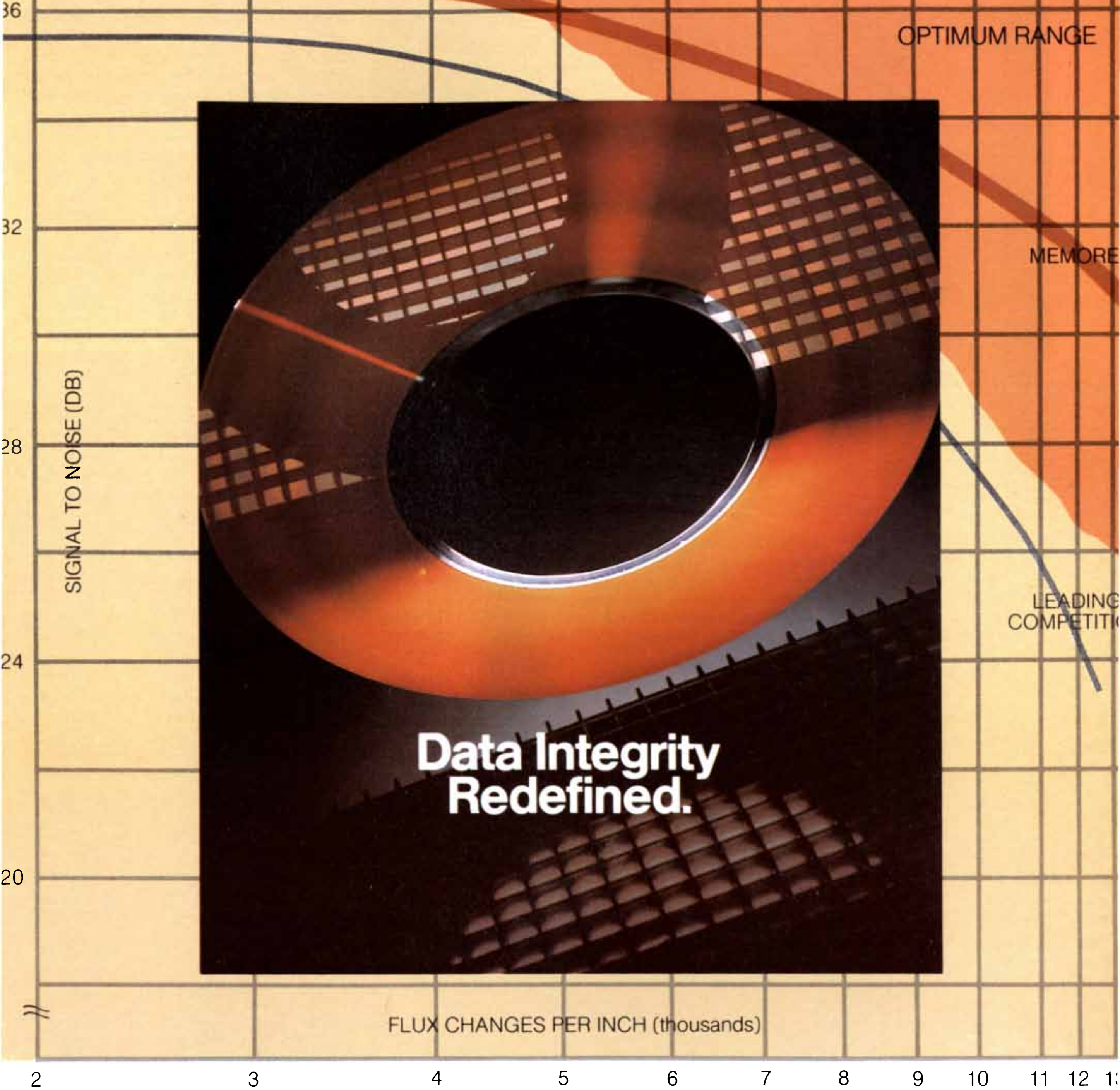
J. Einasto and his associates at the Tartu Observatory in Estonia came forward with a different chain of reasoning. For several years investigators looking into the dynamics of the Milky Way had argued that the velocity of the sun with respect to the globular clusters in the halo of the galaxy is only some 180 kilometers per second. Since the clusters are scattered throughout a large spherical volume, it seems plain that on the whole they do not participate in whatever rotation the galactic disk may have. The distribution of the clusters thus should be more or less stationary with respect to the center of the galaxy. Specifically, it is thought the rotational velocity of the system of globular clusters about the center of the Milky Way can be no more than 50 kilometers per second. Hence the rotational velocity of the sun about the center of the galaxy should be no more than 230 kilometers per second, and certainly no more than 250.

Meanwhile other investigators had determined the velocity of the sun with respect to the average motion of nearby galaxies, namely those whose distance from the sun is no greater than a million parsecs. Their result for the rotational velocity of the sun about the center of the Milky Way was 300 kilometers per second. The difference between the two results could best be interpreted as indicating that the center of the Milky Way galaxy is moving at a velocity of 50 to 80 kilometers per second with respect to the nearby galaxies. That velocity seemed surprisingly large.

Einasto argued that the result might nonetheless be correct and that the unsuspected massiveness of the Milky Way system might be the cause of it. To test his hypothesis he examined the motion of the Milky Way within the group of galaxies of which it is a member. Einasto proposed that there are subgroups within the group and that in one of the subgroups the Milky Way is gravitationally dominant over several other aggregations of stars, such as the two small nearby galaxies called the Large and Small Clouds of Magellan and a



ROTATION CURVE graphs the circular velocity of matter in rotation about the center of the Milky Way. Here two such curves are drawn. A rotation curve plotted in 1965 by Maarten Schmidt of the Hale Observatories (*solid black line*) shows a circular velocity that declines toward the limit of the visible galaxy at 20,000 parsecs. If all the mass in the galaxy lay inside that limit, a test mass placed farther out would rotate at lower speed, in approximate obedience to a law first formulated (for the motion of the planets) by Johannes Kepler (*broken line*). Data analyzed by Leo Blitz and his colleagues at the University of California at Berkeley now yield a rotation curve (*colored line*) that rises toward a value of 300 kilometers per second at 20,000 parsecs. The rise of the newer curve implies unseen mass in great quantity outside the visible limit of the galaxy. Each point in the newer data is the circular velocity of a cloud of hydrogen atoms.



The new Memorex® M Formula magnetic coating is redefining performance standards for all new generation rigid disc drives. Containing our exclusive Hi/Oxide™ formulation, M Formula offers the highest signal to noise ratio—>30 dB at 10,000 flux changes per inch—of any rigid media in the industry. With higher performance margins and protection against random errors, you're assured greater data integrity. And that means more flexibility for the advanced drive designer, too.

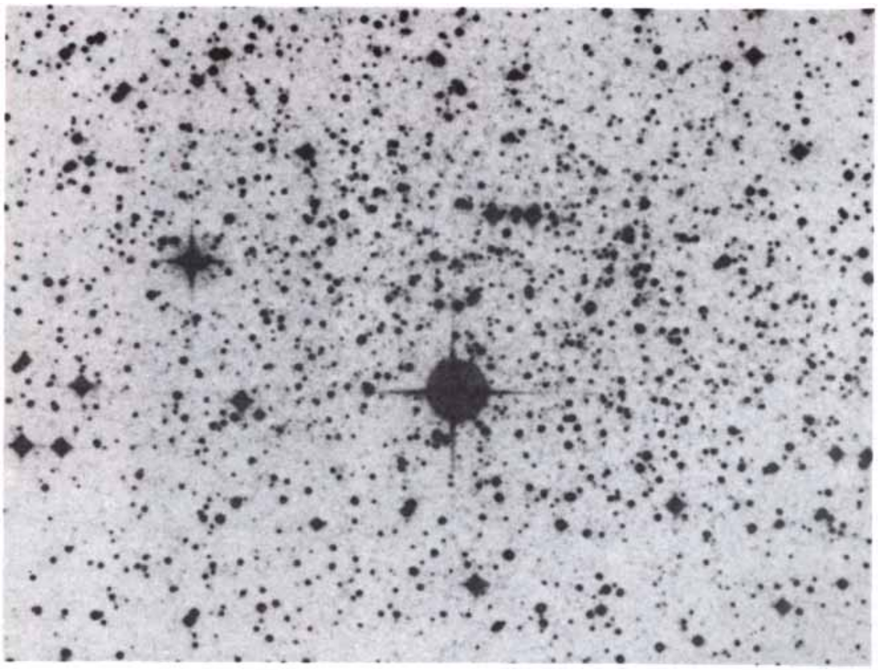
The unique M Formula coating has created a new benchmark for surface smoothness. CLA has been reduced to $<.5\mu''$ and the deviation has been reduced to within $5\mu''$ This represents a 100% improvement over existing industry averages. The smoother surface significantly reduces noise and improves head flying uniformity.

Our revolutionary Hi/Oxide™ coating forms a hard surface that resists wear and has set new standards for durability. M Formula's extraordinary thermoset properties also allow it to withstand a wide range of temperature changes.

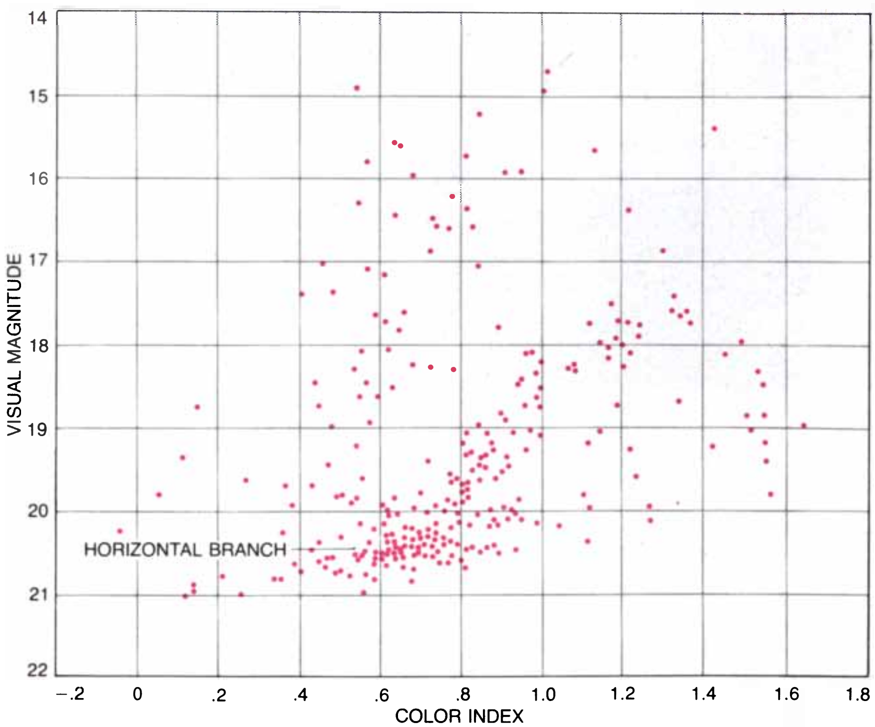
It all adds up to trouble-free data handling. But most importantly, the new M Formula ensures far greater data integrity than you'll find in any other rigid media available.

To learn more about how Memorex is redefining data integrity, contact Memorex Rigid Media and Components Division, 1500 Memorex Dr., Santa Clara, CA 95052 or call N. "Vis" Viswanathan at (408) 987-1928.

MEMOREX



CARINA DWARF GALAXY is the latest addition to the catalogue of galactic companions; its estimated distance of some 100,000 parsecs from the center of the Milky Way apparently makes it an outer member of the galactic corona. In this negative print of a photograph the dwarf galaxy is the loose, inconspicuous aggregation of small dark points near the center. The larger points are stars much closer to the solar system along the same line of sight. The photograph was made as part of the United Kingdom–European Southern Observatory Sky Survey.



HERTZSPRUNG-RUSSELL DIAGRAM for stars in the central part of the Carina dwarf galaxy makes it possible to determine the distance of that galactic companion. The diagram plots the apparent luminosity of each such star (*vertical axis*) against a measure of its color (*horizontal axis*). The stars in what is called the horizontal branch of the diagram turn out to have an apparent luminosity of approximately 20.4. Their intrinsic luminosity would place them 20 magnitudes higher. From the dimming of their luminosity the distance of 100,000 parsecs is inferred. At a galactic latitude of −22 degrees, the Carina dwarf galaxy is well outside the central plane of the Milky Way; hence the correction of the calculation for the absorption of light by intervening interstellar matter should be small. The data were gathered by Russell Cannon and his associates at the Royal Observatory in Edinburgh and at the Anglo-Australian Observatory. The limit of sensitivity of the workers' apparatus is just below the horizontal branch.

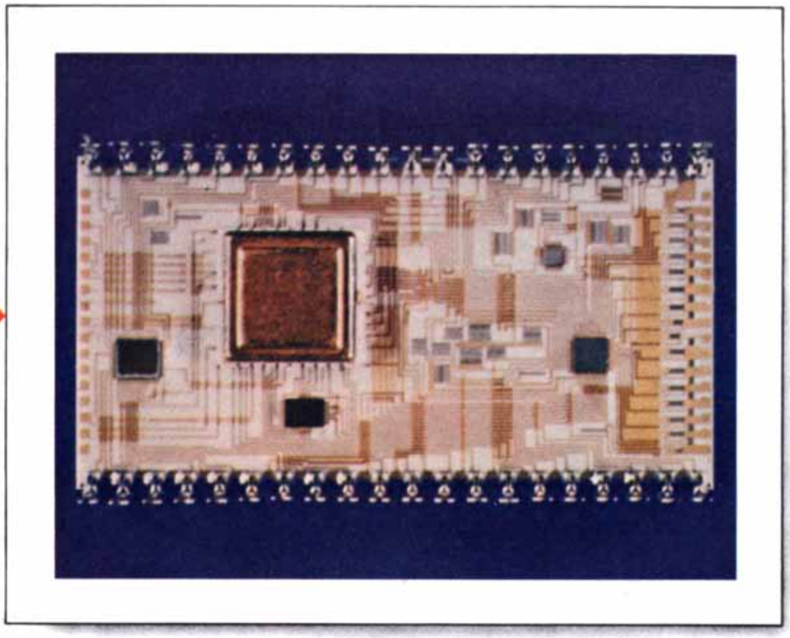
number of dwarf spheroidal galaxies, of which seven are now known. One of the dwarf spheroidal galaxies lies some 150,000 parsecs from the center of the Milky Way. The velocity of the sun with respect to the average motion of these galactic companions proved to be almost 300 kilometers per second. It too was surprisingly high. Einasto interpreted the velocity as an effect of the great mass of our galaxy on its nearest neighbor galaxies.

Einasto was therefore strengthened in his suspicion that the Milky Way is more extended and more massive than had been supposed. In 1976 he offered a model of the Milky Way system in which the mass of the central bulge, the disk and an extended halo is 900 billion solar masses. Even that mass is insufficient to account for the great velocities observed among the galaxy and its companions. Einasto therefore proposed that the bulge, the disk and the halo are embedded in a still larger but nonetheless unseen component of the galaxy, the corona (as he called it), which extends out to at least 100,000 parsecs from the center and has a mass of 1.2 trillion solar masses. The total mass of the Milky Way would then be 2.1 trillion solar masses, or at least seven times the value accepted in 1975.

Supporting evidence was forthcoming from several investigations. First, Vera C. Rubin, W. Kent Ford, Jr., and Norbert Thonnard of the Department of Terrestrial Magnetism of the Carnegie Institution of Washington examined the Doppler shifts of spectral lines in the light emitted by matter in the outer part of 17 galaxies. Each such shift is a displacement of a spectral line to a wavelength different from the one it would have if the source of the radiation were motionless with respect to the instrument that receives it. The investigators concluded that the outlying matter in each galaxy circles the center of the galaxy just as fast as the matter nearer the center. To put it another way, a rotation curve—a graph of rotational velocity as a function of distance from the center of the galaxy—was essentially a horizontal line for the outer part of each galaxy they studied.

The discovery of such flat rotation curves is remarkable. After all, the distribution of brightness in the optical image of a typical spiral galaxy leads one to infer that the galaxy's visible matter is concentrated toward the center and gets sparse at the periphery. From this concentration one can deduce that the outermost visible matter ought to be moving in response to forces analogous to those acting on the outermost planets of the solar system. The peripheral matter ought to be rotating about the center of the galaxy at a lower circular velocity (expressed in kilometers per second) than the matter nearer the center. Evi-

into 3¼ x 4¼" prints. In seconds.



Now you can produce your own ›rints from your slides. On the spot. In ;econds. With Polaroid's remarkable ıew Polaprinter slide copier there's ıo more waiting for days to get your ›rints back.

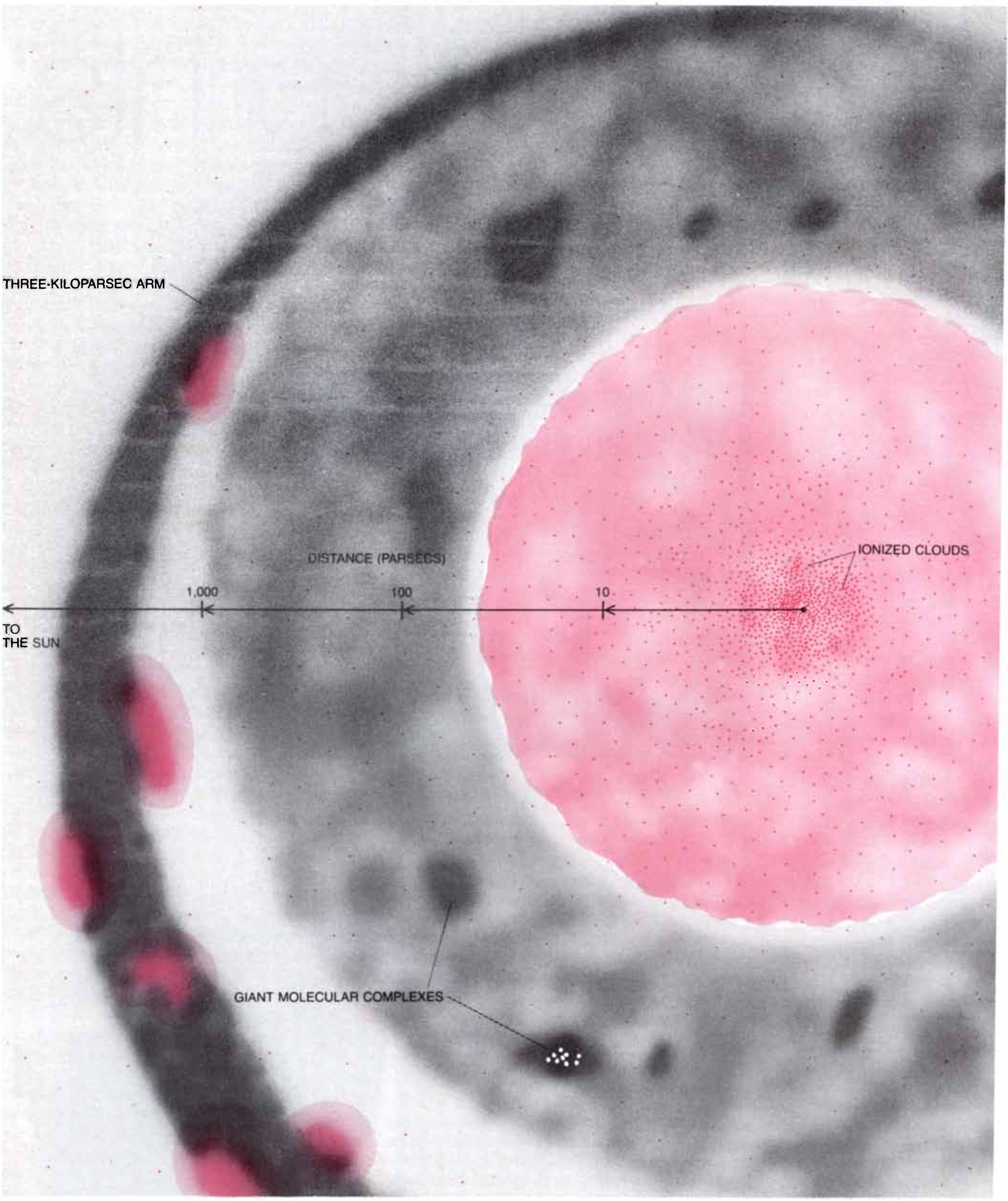
How does it work? Frame your ;5mm slide on the illuminated preview creen. Then insert it in the copier and ›ush a button. Seconds later you have . beautiful, fully-developed color or ›lack-and-white 3¼ x 4¼" print. Every 0 seconds you can start producing nother picture from a slide. It's a ıo-fuss, push-button job.

The Polaprinter uses our Type 668 ʼolacolor film to give you sharp, vivid olor prints in 60 seconds. With our ype 665 positive/negative film, you ıet a black-and-white print plus high-resolution negative in just 30 econds.

The Polaprinter itself is small and ghtweight and takes up less than a quare foot of space. It has fully auto-ıatic exposure control, a specially esigned lens for optimum sharpness nd an independent control for adjusting contrast to your liking. A cropping selector is included to adjust for the slight disproportion between the slide and print formats.

You'll find that the Polaprinter is the quickest, easiest way to prepare prints from your horizontal or vertical slides to illustrate reports and articles. Or to keep records of patients or projects. Or to supplement correspondence and proposals. If you have a large collection of slides, you can make yourself a handy print file for quick reference. And, of course, you always have the security of doing your own copy work at your desk or in your lab.

Bring one of your 35mm slides to a Polaroid Polaprinter dealer for an on-the-spot demonstration. See for yourself how easily and quickly you can turn important data or pictures on your slides into superb prints. For the name of your nearest Polaprinter dealer, call us toll-free from the continental U.S.: 800-225-1618. In Mass., call collect: 617-547-5177.



CENTRAL PART of the Milky Way is shown schematically in an illustration devised by Thomas R. Geballe of the Hale Observatories. The view is from above. The scale decreases logarithmically with distance from the center, so that the innermost parsecs are magnified. The three most central parsecs include the densest congestion of stars (*colored dots*) in the galaxy. (Throughout the illustration the density of stars has been reduced by a factor of 2,000.) The region also harbors compact clouds of ionized gas (*dark color*) that are made up mostly of hydrogen. The velocity of the clouds suggests they are circling a supermassive object, perhaps a black hole, that lies precisely at the center. A thinner distribution of ionized gas (*light color*) permeates the central 100 parsecs. It is surrounded in turn by a ring of cooler, unionized gas (*gray*) whose hydrogen consists of both atoms and molecules. The ring includes giant complexes of dust and molecules (*dark

gray). In some of them young stars (*white dots*) have formed. A dense band made up mostly of un-ionized hydrogen appears at the left. It is the innermost part of an expanding feature called the three-kiloparsec arm. On one hypothesis the arm was created by an explosion at the galactic center 30 million years ago.

dently it does not. The rotation curves calculated by Rubin, Ford and Thonnard thus imply the presence of unseen matter in large quantity beyond the apparent periphery of each of the galaxies.

The Corona of the Milky Way

Studies similar to those of Rubin, Ford and Thonnard bear directly on the distribution of mass in the Milky Way. In one of them F. D. A. Hartwick of the University of Victoria and Wallace L. W. Sargent of the California Institute of Technology calculated the velocities of globular clusters at distances greater than 20,000 parsecs from the center of the galaxy. In other studies James E. Gunn, Gillian R. Knapp and Scott D. Tremaine employed data gathered at the Owens Valley Radio Observatory to determine the velocities of clouds of interstellar hydrogen atoms. Maurice P. Fitzgerald of the University of Waterloo, together with Peter D. Jackson of the University of Maryland and Anthony Moffat of the University of Montreal, determined the velocities of stars and star clusters as much as 17,000 parsecs from the center of the galaxy. William L. H. Shuter of the University of British Columbia determined the velocities of clouds of hydrogen and carbon monoxide that lie in the galactic anticenter: the direction opposite to the line of sight from the earth to the center of the galaxy. Recently Leo Blitz and his colleagues at the University of California at Berkeley have measured the Doppler shifts in both the optical and the radio parts of the electromagnetic spectrum for lines from 184 nebulas and large clouds of interstellar hydrogen and carbon monoxide in the anticenter. All the results are consistent with a rotation curve that does not fall.

From the evidence available today it seems fair to conclude that the rotation curve for the Milky Way attains a value of 230 kilometers per second at 8,500 parsecs, the distance from the galactic center that marks the position of the sun. From there the rotational velocity continues to increase. It reaches 300 kilometers per second at a distance of 20,000 parsecs.

Several objects more distant than 20,000 parsecs are visible; a list of them was published by the International Astronomical Union in 1979. Four globular clusters lie between 20,000 and 40,000 parsecs from the galactic center; the Large Cloud of Magellan and two globular clusters lie between 40,000 and 60,000 parsecs; two dwarf spheroidal galaxies and the Small Cloud of Magellan lie between 60,000 and 80,000 parsecs, and one dwarf spheroidal galaxy and three globular clusters lie between 80,000 and 100,000 parsecs. Four more dwarf spheroidal galaxies and two globular clusters lie between 100,000 and 220,000 parsecs, but their claims to membership in the corona of the Milky Way are more dubious.

Plainly our galaxy is far more extended and of much greater mass than was hitherto thought; the Milky Way has been elevated to the rank of a major spiral galaxy. Taken together, however, the visible constituents of the corona have only a tiny fraction of the hundreds of billions of solar masses that ought to be there. Apparently the Milky Way shares with the galaxies studied by Rubin, Ford and Thonnard the property that much of its outlying matter is dark. Indeed, John N. Bahcall and Raymond M. Soneira of the Institute for Advanced Study infer from the invisibility of the hypothetical matter that if there are stars in the corona that are not bound into clusters or into dwarf galaxies, their intrinsic brightness should be less than a thousandth that of the sun.

What, then, is the unseen mass? Three facts are worth noting. First, dwarf galaxies and globular clusters consist mainly of old stars. Second, old stars are not highly luminous. Third, no one has detected from the corona the spectral lines that characterize clouds of gaseous matter such as hydrogen and carbon monoxide in more central parts of the galaxy. At present, therefore, the best suggestion is that the corona of the Milky Way is composed mainly of old, burned-out stars. On the other hand, the unseen mass of the galaxy's corona may not fit any of the categories based on what can be seen in more accessible regions. We do not know yet what is out there.

The Central Bulge

It may come as a surprise that the center of the Milky Way is no less mysterious than the galactic corona. In fact it is only slightly more visible to observers on the earth. Twenty-five years ago, when observations could be made only at visible and radio wavelengths, three kinds of object were known to be plentiful in the galactic center. Each of them is old. The first is globular clusters. The second is RR Lyrae variables. These are old stars that alternately brighten and dim with a period on the order of a day. The third is planetary nebulas. Most of them are the old, collapsed stars called white dwarfs, each one surrounded by a cloud of gas that is thought to be the shed atmosphere of the star. The galactic center itself was hidden from optical view by layers and shells of dust that are estimated to dim the light from the center by as much as 30 astronomical magnitudes, or a factor of 10^{12}. The dust is particularly effective at blotting out the blue light to which the photographic emulsions employed in earlier days of astronomical photography were often made most sensitive. Some of the dust lies in a lane at the margin of the central


THE MOMENT WAS MARTELL.
You remember how very special the moment was, how very relaxed, how totally perfect.
And, of course, you remember why. Martell Cognac.
MARTELL
COGNAC
SINCE 1715
MARTELL
S.P., V.S.O.P. and Cordon Bleu.
nported by the Jos. Garneau Co., New York, N.Y. © 1980. 80 Proof.

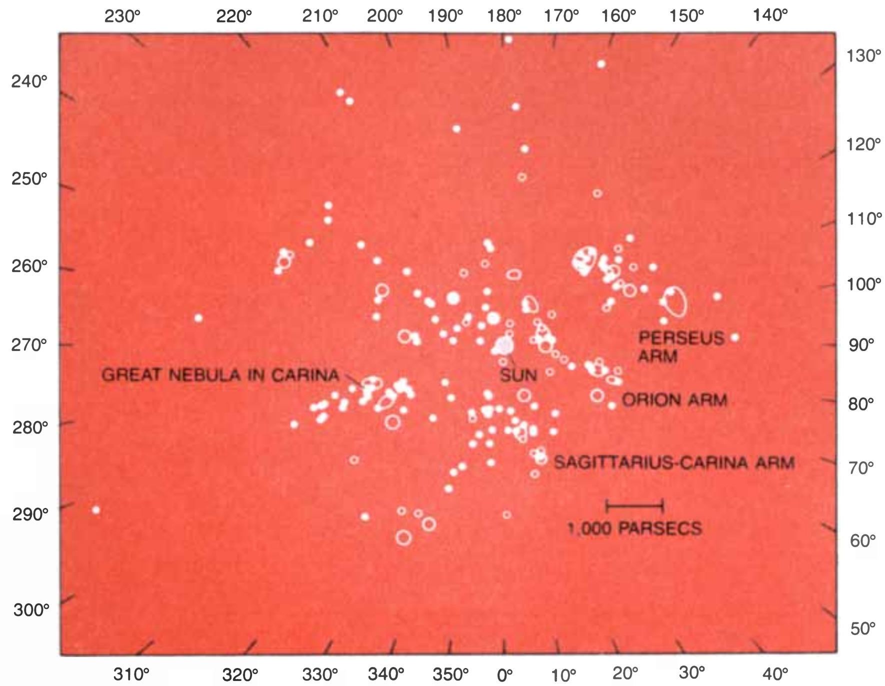
bulge, a few thousand parsecs from the solar system. Some of it lies only 300 parsecs from the solar system, in the constellation Ophiuchus, along the line of sight from the earth to the center. The only thing then known about the center was that it harbored a strong radio source. Electromagnetic radiation in the radio part of the spectrum is not absorbed by the dust.

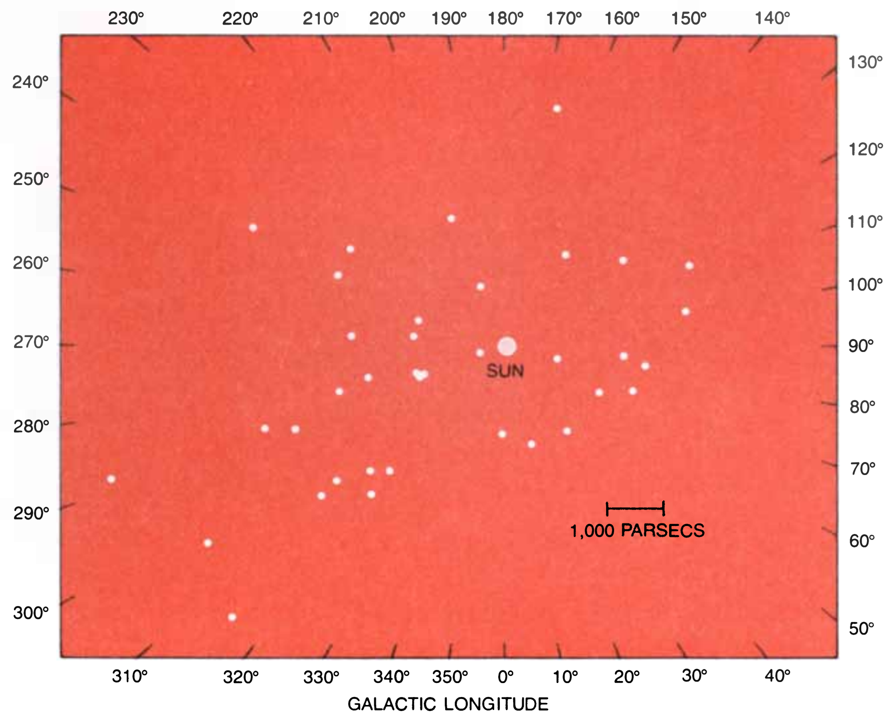
Some 15 years ago the prospects began to brighten. The detection of radio emissions at a wavelength of 21 centimeters served to delineate the sources of such radiation: clouds of neutral (unionized) hydrogen atoms. The detection of emissions at other wavelengths revealed the presence of dark interstellar clouds consisting mostly of molecules. Above all, the advent of infrared astronomy opened a window to the center of the galaxy at wavelengths of from a micrometer to a millimeter. Infrared radiation constitutes a second band of wavelengths that the interstellar dust does not heavily absorb. One of the most useful new techniques is the detection of the infrared radiation emitted at a wavelength of 12.8 micrometers by ionized neon atoms. Because neon is a by-product of energetic events such as stellar explosions it becomes a ubiquitous constituent of interstellar clouds. The neon therefore serves as a tracer.

The observations at radio and infrared wavelengths yield four kinds of data. First, they reveal the presence of local maximums in the rain of radiation from certain directions in the sky. Second, the Doppler shift of spectral lines in the radiation reveals the radial velocity of such a source: the speed of the source either toward the solar system or away from it. Third, the broadening of a spectral line can suggest that the source is expanding or contracting. Fourth, the relative intensities of certain lines in the spectrum suggest the temperature of the source. Even when all possible information has been extracted from the radiation, however, much remains uncertain. It cannot be ascertained, for example, whether a source in approximately the direction of the center of the galaxy lies in front of the center or behind it, or whether the radial velocity of a source actually means it is rotating about the center.

Let me summarize the present state of knowledge about the central bulge of the Milky Way. In its overall shape the central bulge is a slightly flattened sphere. Its outer boundary, which lies 5,000 parsecs from the center, is marked by a ring of what are now called giant molecular complexes. They are large, dark, clumpy interstellar clouds made up mostly of hydrogen molecules. I shall be returning to them. The bulge itself consists in general of a dense clustering of old stars in a rather thin matrix of interstellar gas and dust. The stars are known from their infrared radiation, which can be distinguished from that of gas or dust. One way to account for the relative scarcity of gas and dust in the bulge is to assume that much of it placidly condensed long ago to form the stars of the bulge. On the other hand, several interstellar features in the bulge suggest that the center of the galaxy has had a complex and violent history.

The outermost detectable feature of the bulge is a ring of neutral hydrogen at a distance of 3,000 parsecs from the center. The ring was discovered in 1964 by





SPIRAL STRUCTURE of the disk of the Milky Way in the vicinity of the sun has been traced by observations at optical wavelengths with varying success, depending on what classes of astronomical objects are employed in the attempt. The upper chart shows the positions of loose groups of young stars (*open circles*) and of clusters in which at least some of the stars are young (*solid circles*); the distribution of such stars confirms the presence of the spiral features called the Perseus arm, the Orion arm and the Sagittarius-Carina arm. The lower chart shows the positions of long-period Cepheid variables. These young giant stars are thought to form inside spiral segments. Although the same part of the galaxy is mapped, no spiral structure appears. Data for the charts were collected by Roberta M. Humphreys of the University of Minnesota.

BRIGHTEST SPIRAL FEATURE in the vicinity of the sun is the Great Nebula in Carina, which marks the place where the Sagittarius arm and the Carina arm meet. The nebula is a cloud of hydrogen that has been ionized, and hence rendered luminous, by the ultraviolet radiation of newly formed blue-white supergiant stars in its midst. The nebula cannot be seen from the Northern Hemisphere. The distance to the nebula is 2,700 parsecs. The photograph was made at Mount Stromlo Observatory of the Australian National University.

Jan H. Oort and G. W. Rougoor of the Leiden Observatory. The Doppler shifts of the radiation it emits show it is rotating and, more important, expanding, with velocities away from the center that range from 50 to 135 kilometers per second. Perhaps the ring is a new spiral arm unfurling. One is equally tempted, however, to speculate that the center of the galaxy expelled a kind of smoke ring some 30 million years ago. It is as if there had been a titanic explosion there. Perhaps the explosion swept away much of the gas and dust in the bulge.

The next feature inward lies at a distance of about 1,500 parsecs from the center. It is construed by its discoverers, Butler Burton of the University of Minnesota and Harvey S. Liszt of the National Radio Astronomy Observatory, to be a disk of both atomic and molecular hydrogen. It too is both rotating and expanding. Surprisingly, the best way to interpret the data is to assume that the disk is tilted at an angle of 15 to 20 degrees to the plane of the galaxy.

One might hope that the composition of the central bulge would be more or less homogeneous from the Burton-Liszt ring to the center, if only to simplify the task facing those who attempt to understand the structure of the bulge. More surprises, however, await. Another smoke ring is evident some 300 parsecs from the center. This one too is a mixture of molecular complexes, dust clouds and regions of atomic and molecular hydrogen. The atomic hydrogen is ionized in places, which means it is quite hot: well over 10,000 degrees Kelvin. Associated with these hot spots are clusters of newly formed blue-white supergiant stars. Why should these realms of high temperature and star formation lie precisely in this ring? One is particularly puzzled because a cooler ring of only mildly ionized atoms at a temperature of 5,000 degrees lies a mere 10 parsecs from the center. The 10-parsec ring is rather dense, and it is rotating.

The central three parsecs of the galaxy evidently includes several million stars; they give the center the densest packing of stars in the galaxy. The core region also includes a number of compact clouds of ionized gas; a group of workers led by John H. Lacy and Charles H. Townes of the University of California at Berkeley has detected 14 of them. A typical cloud has about the mass of the sun and a diameter of a fraction of a parsec, and it is speeding around the center: it completes an orbit in some 10,000 years (compared with the sun's orbital period of some 200 million years). Luis Rodriguez and Eric J. Chaisson of the Harvard College Observatory have shown that the velocity of the ionized gas increases with proximity to the center. All of this suggests the clouds are satellites of a supermassive innermost object.

Whatever it is at the very center appears as a bright infrared source in maps made by Eric E. Becklin and Gerry Neugebauer of the California Institute of Technology. According to Bruce Balick of the University of Washington and Robert L. Brown of the National Radio Astronomy Observatory, who have studied the radio emissions of the central object, it has a diameter no greater than 10 times the distance from the earth to the sun. Its mass may be as great as 50 million solar masses. The most likely conjecture is that the very center of the Milky Way harbors a black hole created by the infalling of hundreds of thousands of stars. The center would then be in essence a stellar graveyard.

Optical Spiral Structure

When the periphery and the center of the Milky Way have been considered, there remains a middle region, which has the solar system in its midst. It is the part of the galaxy in which spiral structure prevails.

The tracing of the spiral arms of the Milky Way began in earnest three decades ago, when Morgan, Osterbrock and Sharpless distinguished three spiral-arm segments. To trace them they plotted the positions of the blue-white supergiant stars classified on the basis of their pattern of spectral lines as *O* and *B* stars, together with the bright clouds of ionized hydrogen atoms that often surround such stars. Their diagrams show an Orion arm, of which the sun is a member; a Perseus arm, 2,000 parsecs farther from the center of the galaxy, and a Sagittarius arm, 2,000 parsecs closer to the center of the galaxy. The naming of the arms reflects the general practice in astronomy of employing the constellations to signify directions in the sky. To the three arms distinguished by Morgan and his colleagues has since been added the Carina arm, which may be a continuation of the Sagittarius arm; the concatenation is called the Sagittarius-Carina arm. The two segments that compose it meet at the Great Nebula in Carina, which enmeshes a large number of *O* and *B* stars.

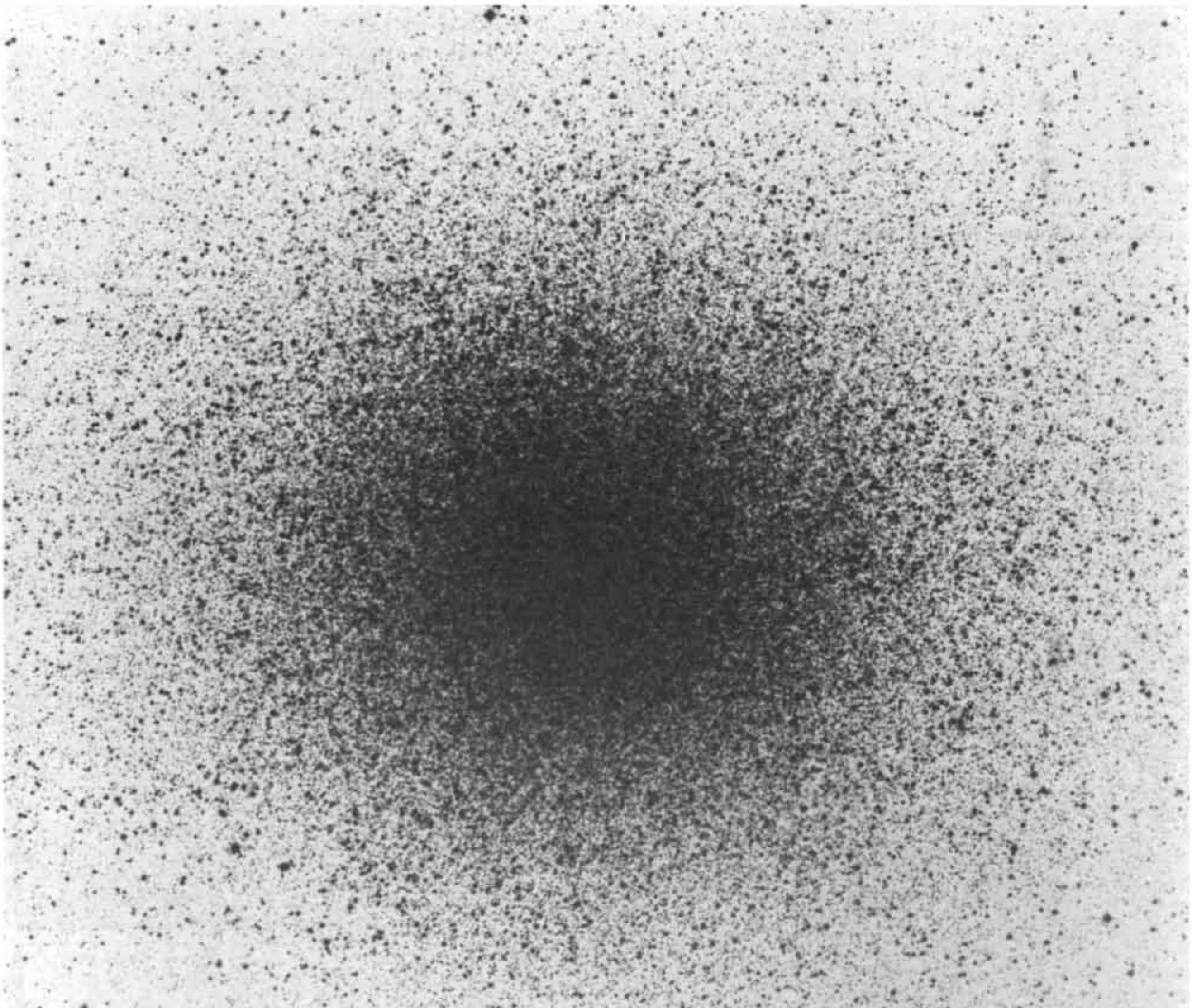
Recent work by Roberta M. Humphreys of the University of Minnesota confirms that *O* and *B* stars are abundant in the principal arms that were recognized three decades ago. This is cheerful news. It means the spiral arms are indeed delineated by very hot, blue-white supergiant stars, by the clusters made up of such stars and by the bright clouds of gas in which such stars and clusters are found. *O* and *B* stars are quite young; the ones seen now in the spiral-arm segments of the Milky Way were formed no more than 10 million years ago. There is no question the spiral arms are regions of star formation.

The inner margins of most spiral arms

are marked, it seems, by dark nebulas: cold clouds of atoms, molecules and dust. According to the density-wave theory of Lin and Shu, the dark nebulas signal the compression of the interstellar matter by a wave of pressure that advances through the galactic disk. The compression precipitates the condensation of the matter into the stars of the spiral segment. The young stars might be expected to lie behind the advancing density wave and dark nebulas to lie inside it. According to the theory, however, the wave moves at only two-thirds the speed of the galactic disk's rotation. Hence the wave is overtaken by the stars that formed inside it, and the loci of dark nebulas come to lie at the trailing edge of the stars.

Certain of Humphreys' results, it should be said, are less than cheerful. Some of the stars called Cepheid variables are as young as *O* and *B* stars; in particular the Cepheid variables that alternately dim and brighten with a period longer than 15 days have ages of no more than 10 million years. Moreover, they too are large and bright, and so they can be seen at great distances. Since they are young, they should not have moved far from their birthplace: presumably a dark nebula at the inner edge of a spiral arm. All things considered, the long-period Cepheid variables should be excellent delineators of spiral structure. They are not. The Cepheid variables plotted by Humphreys form an essentially random distribution of points in the galactic plane.

One thing must always be kept in mind by those who search for spiral structure in optical observations of the Milky Way. The usual method of calculating the distance to a star is to compare its observed luminosity with what its in-



RICHEST GLOBULAR CLUSTER in the vicinity of the sun is Omega Centauri, shown here in a negative print of a photograph made by Gary S. Da Costa with the 2.5-meter telescope of the Las Campanas Observatory in Chile. The cluster includes several hundred thousand stars within a diameter of approximately 30 parsecs. Omega Centauri is slightly more than 5,000 parsecs from the solar system.

The all-new Toyota Cressida. Lavish. Extravagant. Uniquely European in look. This is a sedan different from any Toyota you have seen before.

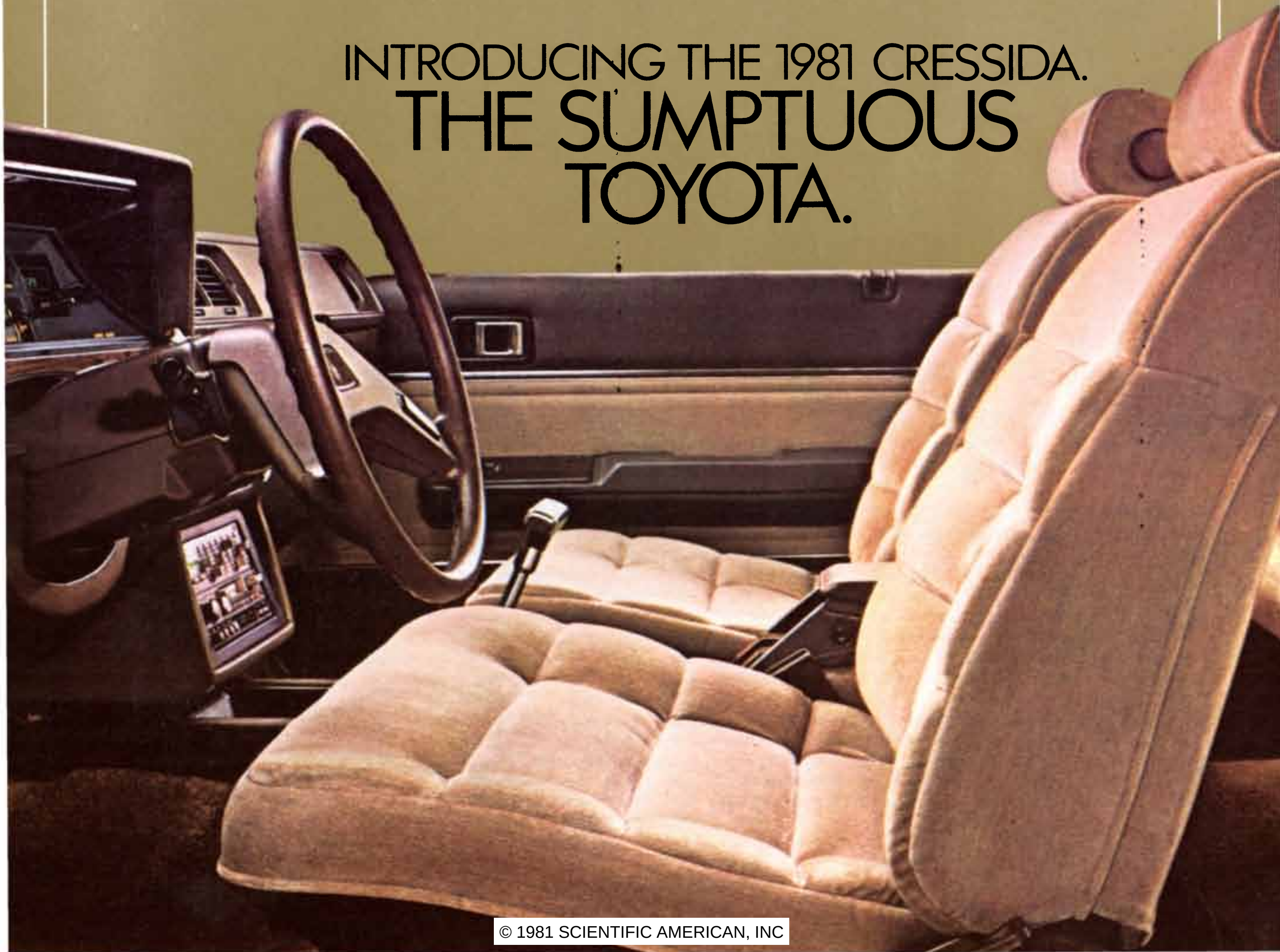
Inside, the new Cressida offers every comfort and convenience feature you could expect in the world's highest-priced automobiles—and then some. Power windows, cruise control, climate control and AM/FM MPX 4-speaker stereo are standard equipment, of course.

But Cressida's state-of-the-art engineering includes much more. Like a unique automatic shoulder-seat belt system. The belt is guided around you by an electric motor! And an optional cassette player with amplifier and graphic equalizer, for the ultimate in automotive sound.

Performance of the Cressida is also impressive. Both the Cressida Sedan and 5-Door Wagon feature electronic fuel injection and an innovative 4-speed overdrive automatic transmission, to enhance the new 2.8 liter engine's efficiency. MacPherson strut front suspension assures a soft, controlled ride.

The Toyota Cressida. If you're wondering why you should buy a luxury car from a company famous for economy cars, you need do only one thing. Drive a Cressida, and feel how sumptuous a Toyota can be.

INTRODUCING THE 1981 CRESSIDA.
THE SUMPTUOUS TOYOTA.



trinsic luminosity is taken to be. The intrinsic luminosity is assigned on the basis of the pattern of lines in the spectrum of the star's radiation. The diminution of the brightness is a correlate of the distance. Even with the best techniques, however, the distance of an *O* star or a *B* star is uncertain by plus or minus 10 percent. For example, a star that is calculated to be 2,700 parsecs from the sun—say a star in the Carina nebula—could be as little as 2,400 or as much as 3,000 parsecs away. The result is a purely observational blurring of several hundred parsecs in the plotting of what might actually be a well-defined spiral feature.

With the most modern telescopes, spectrographs and photographic equipment, *O* and *B* stars can be detected, and their spectra can be examined, at distances calculated to be as great as 8,000 parsecs from the sun. At that distance, however, the uncertainty in the calculation is plus or minus 800 parsecs. In such a case a spiral feature may well be unrecognizable. There seems little prospect of tracing the spiral structure of the Milky Way by optical means to distances beyond 8,000 parsecs.

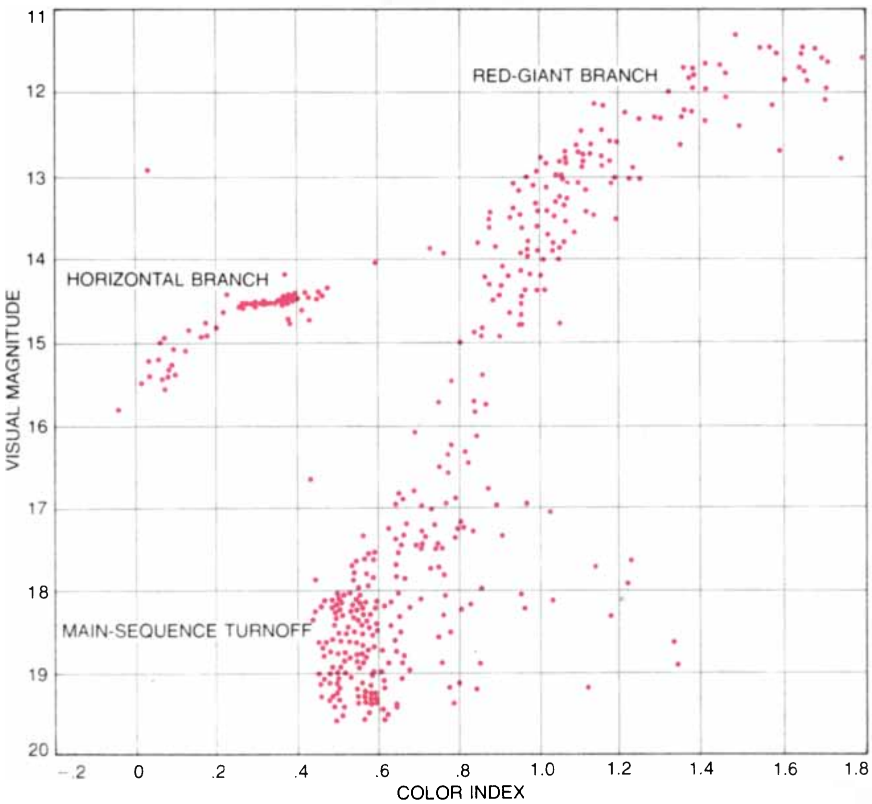
Radio Spiral Structure

How do the efforts to trace spiral structure fare at radio wavelengths? Again the prospects are far from encouraging. By the early 1950's, when Morgan, Osterbrock and Sharpless presented the key optical features of the spiral structure, Harold I. Ewen and Edward M. Purcell of Harvard University had detected radiation from interstellar clouds of neutral atomic hydrogen at a wavelength of 21 centimeters. Within a few years the first radio maps of the galaxy were available. They suggested that at least in the outer parts of the galactic disk the clouds of neutral atomic hydrogen are arrayed in nearly circular spiral features.

Some details of the way in which such data are amassed and interpreted are worth examining. When a radio telescope detects radiation at and near a wavelength of 21 centimeters from a small area of the sky, it is actually receiving radiation from a number of clouds of neutral atomic hydrogen along a single line of sight. Each cloud has its own velocity of approach or recession with respect to the telescope, and hence its radiation has a distinctive Doppler shift. As a result the telescope receives from a single direction a profile of peaks and valleys in a graph of intensity v. wavelength. The telltale 21-centimeter line is actually a set of closely spaced lines with various intensities, various degrees of broadening and various Doppler shifts. Since neutral atomic hydrogen is assumed to have its greatest density at the trailing edge of spiral features, it seems reasonable to assume that a peak of great strength in the profile should lie at the wavelength corresponding to the velocity of approach or recession of hydrogen where the line of sight crosses a spiral feature. With the aid of a rotation curve for the Milky Way one should then be able to determine the distance to the feature.

This chain of reasoning is simple and lovely, but it does not stand up under scrutiny. The first complication is that the clouds of neutral atomic hydrogen have motions of their own, quite apart from the motion of the spiral features. These independent motions can easily alter the total velocity of a cloud by as much as six kilometers per second. An additional complication is that the matter of the galactic disk apparently exhibits large-scale streamings. Burton and his associates at the National Radio Astronomy Observatory have demonstrated that even slight motions of this kind can give rise to peaks of intensity in the 21-centimeter radiation for directions in which the line of sight does not cross a spiral feature. Further still, there are directions in the sky for which the velocity of approach or recession of a cloud of gas with respect to the solar system may change only slowly with distance. The 21-centimeter profile may then show a peak caused by contributions from the width of a single great expanse of gas of uniform density.

Several other approaches can be made to the study of spiral structure. One is to observe the spectral lines emitted by molecules of carbon monoxide at radio wavelengths close to 2.6 millimeters. The carbon monoxide lines give access to a class of clouds that are cooler than the ones composed of neutral atomic hydrogen. The cooler clouds are composed mostly of molecular hydrogen, with some admixture of carbon monoxide and other substances. (The



CURIOUS STELLAR COMPOSITION of the Omega Centauri globular cluster is suggested by a Hertzsprung-Russell diagram for the stars that make up the cluster. The branch of the diagram consisting of red-giant stars (*upper right*) is unusually long. That signifies an unusual degree of variation in the giant stars' content of "metals," or atoms heavier than carbon. Specifically, the metal atoms in the star increase its opacity to the radiation escaping from inside. They thereby change its color: the metal-rich red giants are redder, and on a Hertzsprung-Russell diagram their position is well toward the right. The variation in metal content suggests in turn that the stars in the Omega Centauri cluster formed in more than a single episode. Further deductions are problematical. In general the age of a globular cluster is determined by comparing the luminosity and the color of the stars at what is called the main-sequence turnoff of the Hertzsprung-Russell diagram (*bottom*) with that of the stars at the corresponding position in diagrams generated in computer simulations of a cluster's evolution. In the case of Omega Centauri the data do not reveal the full extent of the main sequence, and so it is difficult to place the turnoff precisely. That makes it difficult to establish the age (or ages) of the cluster. The data were collected by Cannon and N. J. Stewart of the Royal Observatory in Edinburgh.

general principle is that increasing temperature leads first to the dissociation of molecules into neutral atoms and then to the ionization of the atoms.) William Herbst of Wesleyan University has shown that the presence of carbon monoxide in an interstellar cloud is correlated with the presence of dust. The carbon monoxide lines therefore aid in the detection of distant dust clouds and dust complexes that might otherwise escape notice.

A low-resolution survey of carbon monoxide clouds by Richard S. Cohen and Patrick Thaddeus of the Goddard Institute for Space Studies and Thomas M. Dame of Columbia University shows evidence of the existence of spiral features. Indeed, in some instances the carbon monoxide clouds seem to fill gaps between clouds of atomic hydrogen in some of the recognized spiral features. Still, I doubt the spiral structure of the Milky Way would have been discovered if the only data available had been the radio observations at a wavelength of 2.6 millimeters.

I thoroughly dislike having to be pessimistic about the prospects for mapping the spiral structure of the Milky Way beyond a distance of about 8,000 parsecs from the sun, particularly since the structure seems firmly established up to that distance, but I see no chance for improvement in the next decade or two. This is not to say there are now no possibilities whatsoever for the study of spiral structure. Among optical telescopes the great reflectors already in place and the Space Telescope, an instrument that will orbit the earth, promise to reveal the fine details of spiral structure and the motions of small-scale features. Among radio telescopes the Westerbork Array in the Netherlands and the Very Large Array in New Mexico, which synthesize images from a set of detectors, will advance the investigation of the cold, dark clouds in spiral features and no doubt will suggest much about the causes of spiral structure. These promising developments, however, apply mainly to spiral galaxies other than our own. The spiral galaxies Messier 31, 33, 51, 81 and 101 are a few fairly nearby examples.

The Ages of Globular Clusters

One realm of research that is flourishing in both galactic and extragalactic astronomy is the examination of globular clusters. The clusters are important in part because they seem to be the oldest objects in the Milky Way. They therefore hint at the birth and evolution of the galaxy and indeed of the early universe. The simplest hypothesis about the clusters is that they all formed within a short time (say a billion years) of the big bang: the instant when all the matter in the present universe emerged explosively, it is thought, from a single point. The clusters would then have been among the first objects to condense as the gal-

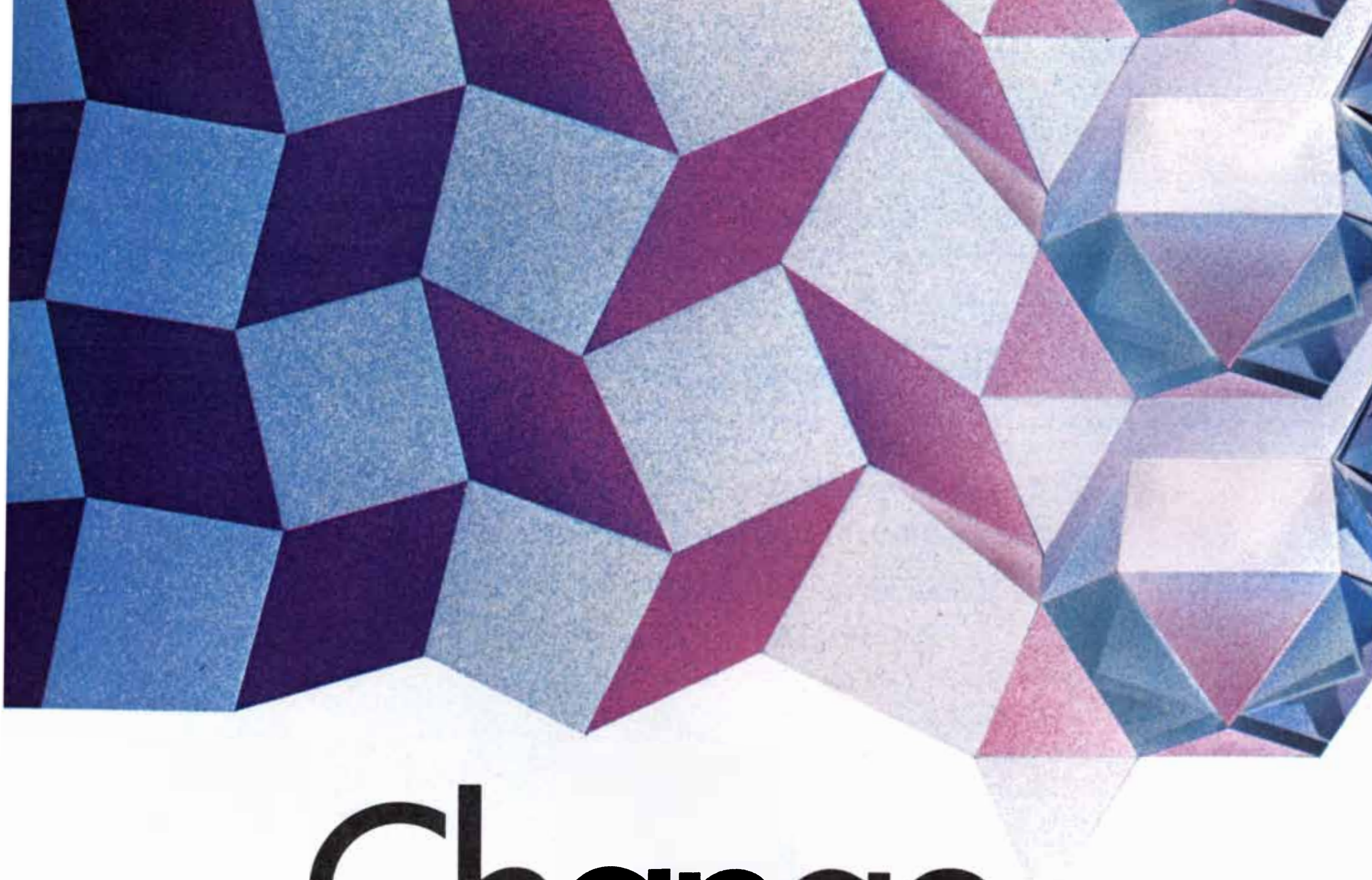
BIZARRE BIRTH OF STARS is visible in the small dark cloud at the left in this photograph. The cloud is one of those called a globule. It is half a parsec in diameter and probably has somewhat less than 100 times the mass of the sun. The mass of the globule is composed of molecules (mostly hydrogen) and dust. From this particular globule a pair of incipient stars connected by a luminous strand have evidently been expelled. The strand crosses the upper edge of the cloud. The expulsion defies the hypothesis that a globule condenses in about a million years to form a single new star. The contour of brightness extending diagonally from the bottom toward the right of the field is an edge of the Gum nebula, which is thought to be in part the debris of a supernova explosion. Both the globule and the nebula are calculated to be approximately 300 parsecs from the solar system. They lie in the part of the southern sky marked by the constellation Vela. The photograph was made by the author in 1978 with the four-meter telescope at the Cerro Tololo Inter-American Observatory in Chile.

axies took shape, each cluster evolving from a large blob of gas. It counts in favor of this hypothesis that half of the roughly 200 globular clusters in the Milky Way lie scattered throughout the almost spherical volume of the galactic halo. Presumably they formed there well before the galactic disk took shape. The orbit of such a cluster is typically a rather eccentric ellipse and not the more nearly circular orbit characteristic of matter in the disk. The major axis of the orbit is sometimes several tens of thousands of parsecs, and once every billion years the orbit sends the cluster rushing through the thickness of the disk, which is only a few hundred parsecs.

On the simplest hypothesis the stars of the globular clusters would have formed at a time when the available matter in the galaxy was mainly hydrogen and helium, the two chemical elements presumed to have been created in the immediate aftermath of the big bang. In contrast, a later-born star would condense from interstellar gas part of which had been cycled through the interior of the first-born stars. Some of the gas, for example, would have been propelled into space by supernova explosions. It would be matter in which heavier atoms had been created by thermonuclear fusion. The later-born star would therefore have higher concentrations of the heavy chemical elements; in the shorthand of astrophysicists all such elements are known as metals.

It is the concentration of metals in the various globular clusters that imperils the simplest hypothesis. To be sure, the ratio of metals to hydrogen and helium is 100 times greater in the sun than it is in the stars of metal-poor globular clusters such as M3 (for Messier 3). Moreover, the metal-poor clusters tend to be the ones that are outermost in the galaxy. An age of 15 billion years is now assigned to them. On the other hand, the globular cluster 47 Tucanae is relatively metal-rich. Its age is thought to be 10 billion years, which is twice the age of the sun. Omega Centauri, the most impressive globular cluster in the Milky Way, shows a range of metal concentrations. Evidently it is an idiosyncratic agglomeration of stars that were born at different times. The spread of ages for the globular clusters conflicts with current models of how the galaxy evolved. No model allows as much as several billion years for the galactic disk to have condensed.

With respect to the question of age it is notable that the maximum age of the universe can be inferred from the velocities at which galaxies are receding from one another. The current value for the maximum is close to 15 billion years. Recent investigations of the recession rate have tended to reduce the maximum age. If the trend continues, the age could conceivably fall to as little as 10



Change the shape of change.

The history of business is the history of change.
Now, from the Bell System, a tool designed to alter the nature of business by letting you change the shape of change.

billion years. Then one would have to explain the finding that certain globular clusters seem to exceed the age assigned to the universe. It is clearly important that trustworthy values be calculated for the ages of the globular clusters. It would be particularly useful if such ages were available by the end of this decade, because by then the Space Telescope should have yielded far more reliable values for the recession velocities and the distances of the galaxies. It will thus have yielded a far more reliable value for the maximum age of the universe.

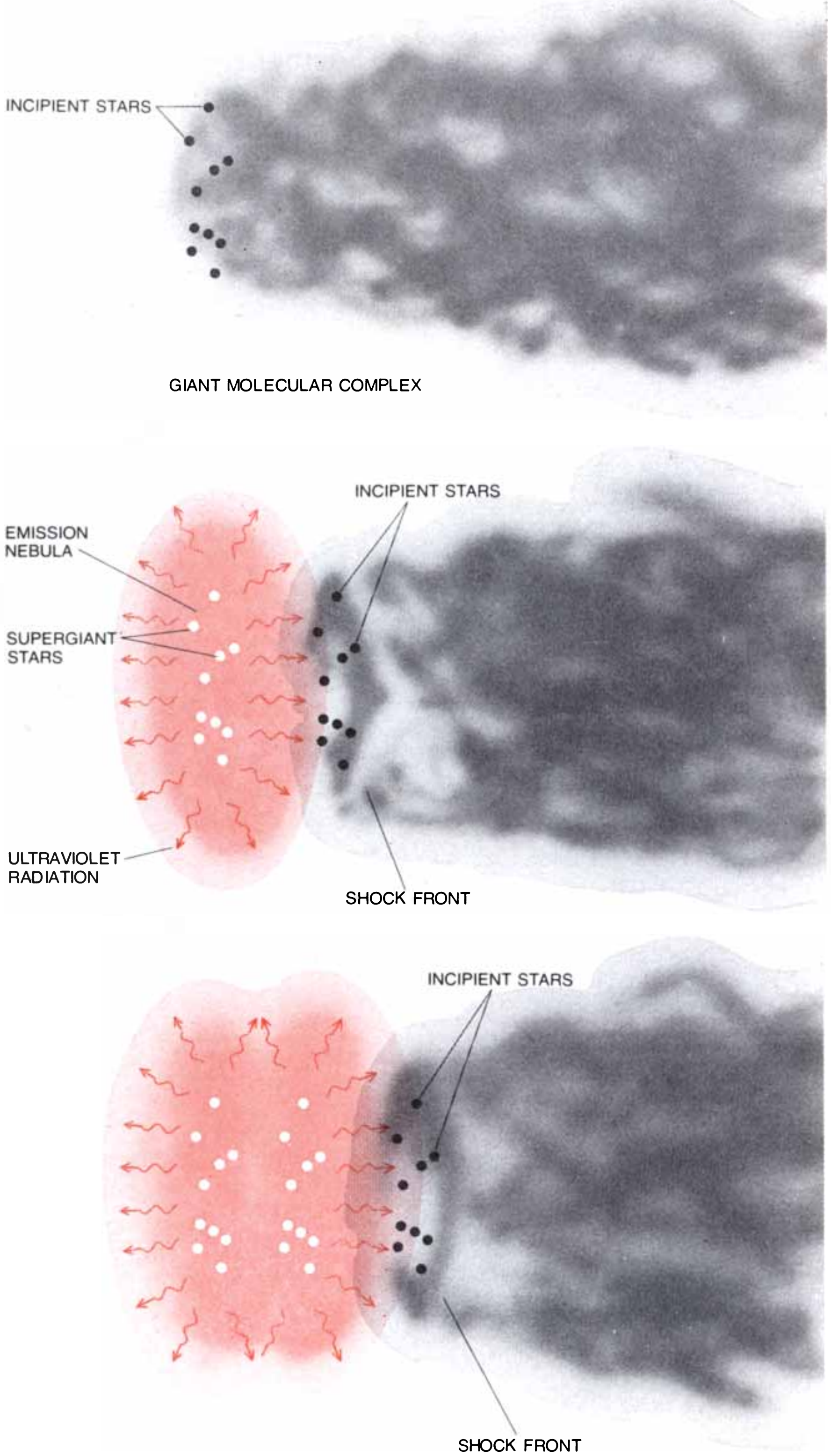
The Evolution of Globular Clusters

The internal dynamics of the globular clusters are well described by three characteristic times. The first is the crossing time: the time it takes a star to move across the cluster under the gravitational attraction of the cluster as a whole. The second is the relaxation time: the time in which the star settles down and becomes a stable member of the cluster under the influence of its gravitational interactions with nearby stars. The third is the evolution time: the time in which a stable cluster changes its form and its stellar composition significantly. For a globular cluster rich in stars the crossing time is much shorter than the relaxation time, which in turn is much shorter than the evolution time.

The work of Ivan R. King of the University of California at Berkeley has suggested some further quantifications. When a star-rich globular cluster has been in existence for roughly 50 crossing times, its stars will have settled down; the statistics of their velocities will be much like those for the velocities of the molecules in a cloud of gas. This state of equilibrium might persist indefinitely if the stars never changed internally and if no stars left the cluster.

As early as the 1930's, however, Lyman Spitzer, Jr., of Princeton and Victor A. Ambartsumian of the Byurakan Astrophysical Observatory in Armenia had concluded that the stars of lowest mass in a cluster were most likely to have the greatest velocity, and that in many cases the stars with the greatest velocity would escape from the cluster. Because of their great velocity the escaping stars would take with them more than the average share per star of the total energy of the cluster. As a result the cluster would contract. Over a period of time on the order of the evolution time the cluster would lose appreciably in average energy per star, and the most massive stars would settle close to the center.

Internal changes among the stars that remain can only quicken this trend. To begin with, the loss of mass by stars is now recognized to be an astrophysical commonplace. Early in its evolution a typical star is embedded in a shell of gas

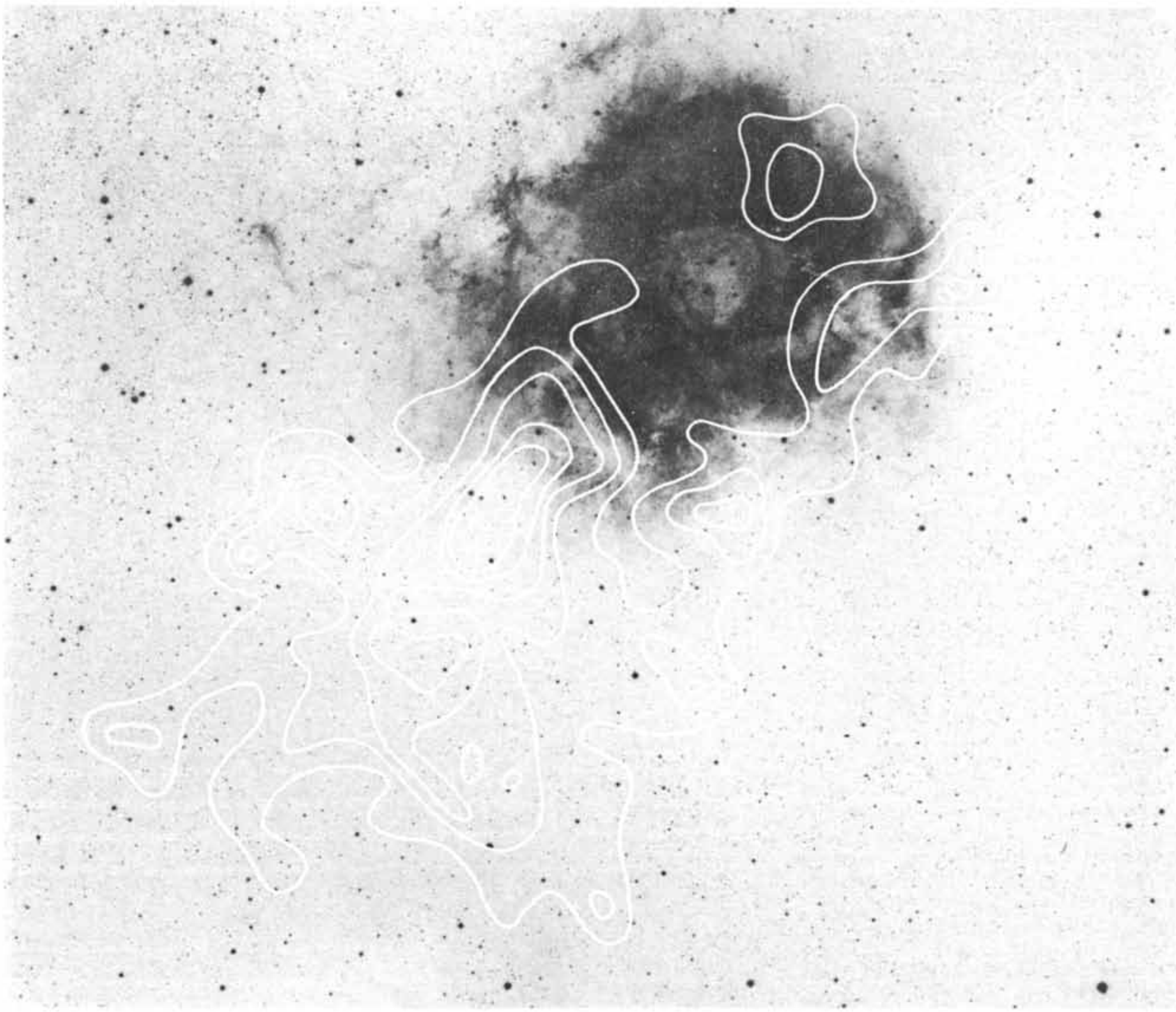


BIRTH OF SUPERGIANT STARS is thought to take place in repetitive stages. The sequence begins (*top drawing*) with the formation of a cluster of supergiant stars near the periphery of a giant molecular complex, a cold, dark region of molecules (mostly hydrogen) and dust. After a few million years the ultraviolet radiation of the new stars has ionized the nearby hydrogen, so that the stars are enmeshed in an emission nebula: a cloud of luminous gas. In addition the pressure of the radiation has compressed the matter in the giant molecular complex. (Each quantum of the radiation has momentum, and so it pushes the molecules and the dust.) The result is the formation of a second group of supergiant stars (*middle drawing*). In another few million years a similar sequence may form still another group of stars (*bottom drawing*). The diagram is based on a hypothesis put forward by Bruce G. Elmegreen of Columbia University and Charles J. Lada of the University of Arizona. The hypothesis accounts for clusters of supergiant stars in nebulas, such as the Great Nebula in Orion, that lie next to a molecular complex.

that is gradually expelled by the outward-directed pressure of the star's electromagnetic radiation. (Each photon, or quantum of the radiation, has momentum, and so it pushes whatever absorbs it.) Later the star begins to shed a gentle stellar wind of particles. Still later certain kinds of star explode. In a nova the exploding star ejects into space the equivalent of the mass of a planet such as Jupiter. A supernova is an explosion in which an entire star is destroyed. Because mass is lost through mechanisms such as these the stars in a globular cluster move systematically to classes with lower mass (and also fainter intrinsic brightness). Over a period of time on the order of the relaxation time they acquire the higher velocity of a less massive star. Hence their chance for escape from the cluster improves.

Throughout the hundreds of billions of years a cluster may spend in the thin outer halo (or even the corona) of the Milky Way the evolution of the cluster is affected only by the internal events I have just described. There comes a time, however, when the cluster traverses the galactic disk. At such a time the cluster is at risk not so much of collision as of gravitational interaction with the matter in and near the galactic plane. The force of the interaction may tear the cluster apart.

As the orbit of a cluster brings it close to the center of the galaxy a further threat arises. The threat exists because the central mass of the galaxy exerts a greater attractive force on the side of the cluster that passes nearest the center than it does on the side farthest away. The difference between the inner and the



BRIGHT NEBULA AND DARK NEBULA are neighbors in the constellation Monoceros. Together they form an example of the star formation shown on the preceding page. The bright (or emission) nebula, called the Rosette, is the dark cloud at the upper right of this negative. Its hydrogen atoms have been ionized by the ultraviolet radiation of a group of young blue-white supergiant stars. Several such stars lie in what appears to be a hole at the center of the nebula. The radiation from the star precisely at the center is particularly intense and may have blown out the hole. The dark nebula is a giant molecular complex. It consists almost entirely of cold molecular hydrogen. Its full extent is best revealed, however, by the radiation of carbon monoxide molecules at a radio wavelength of 2.6 millimeters. The contour lines in the illustration map antenna temperature, which indicates the strength of the carbon monoxide radiation. The pattern of the contours suggests that the ultraviolet radiation of the supergiant stars is eating into the complex. A shock front along which the matter in the complex is compressed by the radiation apparently lies at the lower left edge of the Rosette. The front is presumably a place where new stars will condense. One peak of the carbon monoxide radiation corresponds to a source of infrared radiation (*IR*) now identified as a newborn star embedded in gas and dust. Both the Rosette nebula and the giant molecular complex are 1,600 parsecs from the solar system. The illustration displays the results of a study of the interaction between the nebulas by Leo Blitz of the University of California at Berkeley and Patrick Thaddeus of the Goddard Institute for Space Studies. The photograph is from the Palomar Sky Survey.

ANOTHER TECHNOLOGICAL ADVANCE FROM SHARP.

MURPHY'S LAW DEFIED.

"If anything can go wrong, it will" and usually does, especially when you're entering an equation into a calculator.

Until now, you may have been using the error-prone procedure of encoding and decoding the steps of an assembly code. The Sharp 5100 Scientific Calculator eliminates this needless procedure. It's the first programmable calculator that allows you to enter directly, using algebraic, logarithmic, exponential and other functions in their familiar form.

That's Murphy's Law defied.

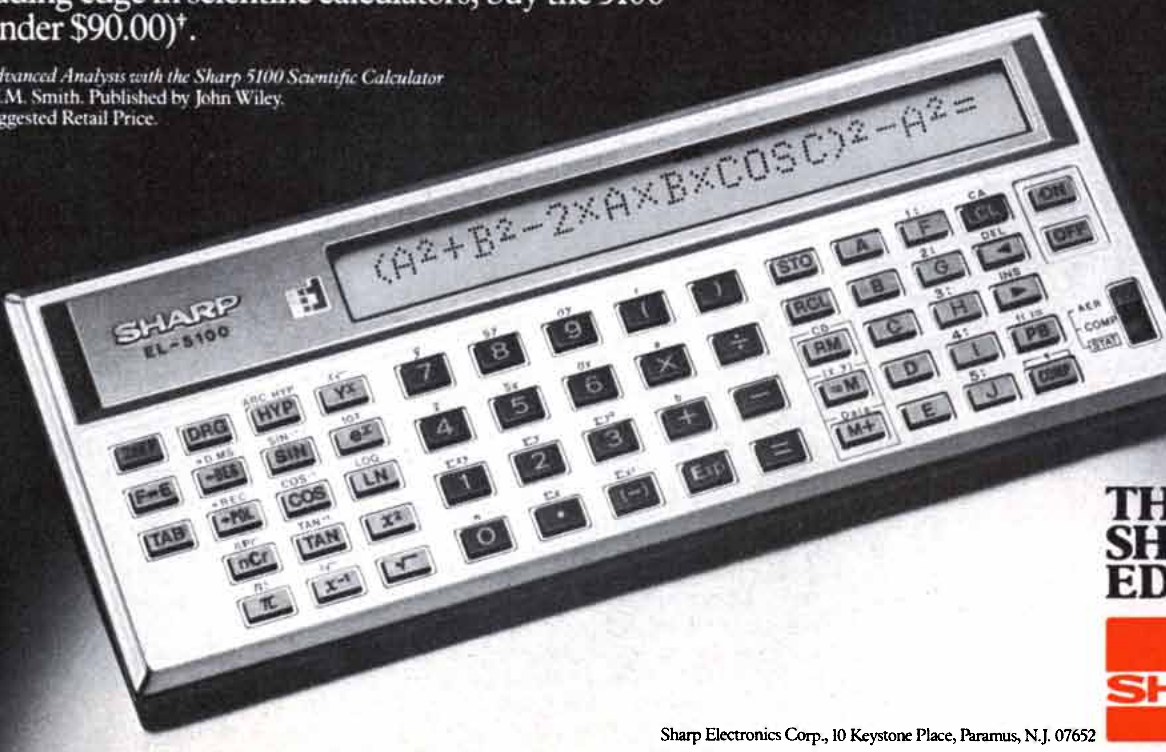
The 5100 is also the first hand-held to give you a complete display of the equation: you see 24 characters at a time in a dot-matrix display that rolls right or left to accommodate one or more formulas of up to 80 steps. So that while other programmable calculators display only one *program step* at a time, the 5100 displays the actual equation in its entirety, and in the way you're accustomed to seeing it.

If, anywhere along the line, you've entered an incorrect character (Murphy's Law at work), the display makes it easy for you to spot. You can then move the 5100's computer-like cursor to any part of the display to correct it.

In addition to the fast solution of frequently used formulas, the 5100 is extremely convenient for deriving difference tables, performing data analysis, numerical integration, simulations of linear systems, statistical and probability calculations, and many other procedures. It has so many uses, it would take a book to fully describe them, and someone has even written one.*

Sharp offers you a complete line of highly advanced scientifics, and if you want the leading edge in scientific calculators, buy the 5100 (under $90.00)†.

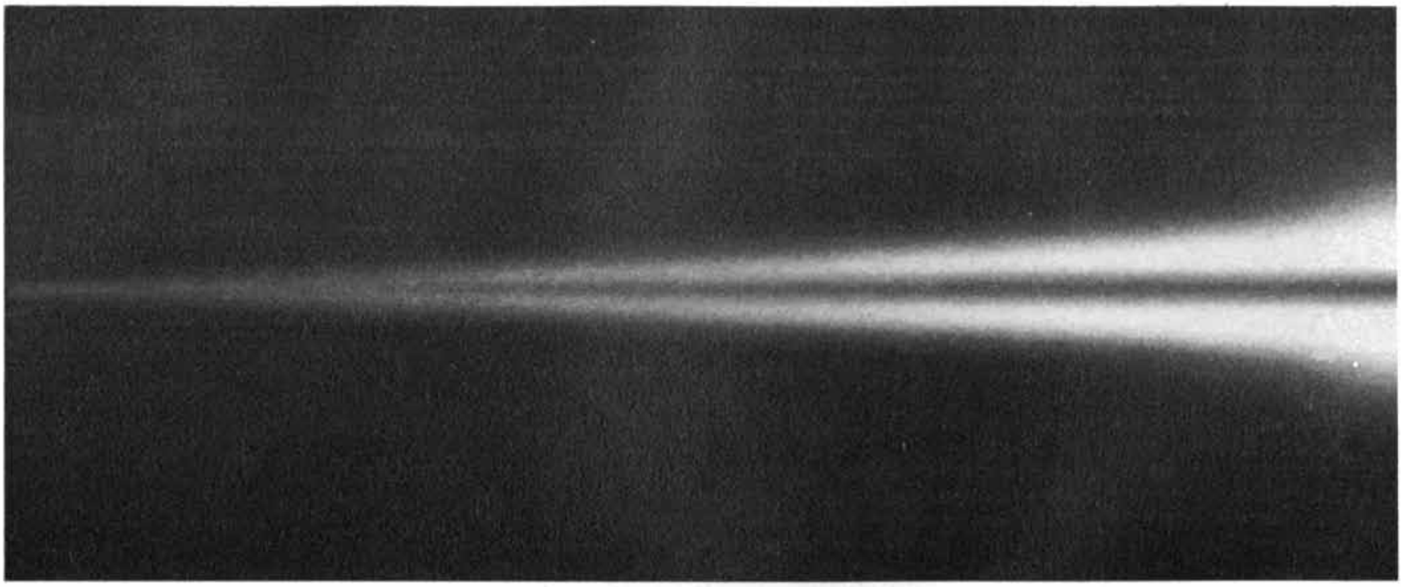
**Advanced Analysis with the Sharp 5100 Scientific Calculator* by J.M. Smith. Published by John Wiley.
†Suggested Retail Price.



THE SHARP EDGE

SHARP

Sharp Electronics Corp., 10 Keystone Place, Paramus, N.J. 07652



OUTSIDE VIEW of the Milky Way shows the galaxy edge on as it might be seen through a telescope by an astronomer in another galaxy. The view was generated with the aid of a computer by John N. Bahcall and Raymond M. Soneira of the Institute for Advanced Study.

outer force can deform the cluster. Indeed, for any globular cluster there is a critical radius: a distance from the galactic center within which some of the stars in the cluster will be sheared off as the cluster makes its closest approach to the center. The critical distance is called the tidal radius because the mechanism that deforms the cluster is similar to the one that raises tides in the oceans. It is not surprising that the globular clusters inside the galaxy's central bulge are smaller on the average than the ones in either the galactic halo or the galactic corona.

One final factor may affect the globular clusters, but I mention it with some misgivings. As recently as 15 years ago few astronomers would have thought the formation of double star systems would be important in star clusters. Then Sebastian von Hoerner of the National Radio Astronomy Observatory made the first computer simulations of open clusters: aggregations of stars that differ from globular clusters in that the stars are young, fewer and more widely dispersed. Von Hoerner based the simulations on equations that represent the gravitational interactions among small numbers of stars. The simulations consistently showed that the stars in a cluster tend to unite by pairs into binary systems.

In more recent computer simulations incorporating first 100 and later 500 and even 1,000 interacting stars, the same effect was noted. In 1975 Douglas C. Heggie of the University of Edinburgh showed that a newly formed binary system is a formidable sink of the energy available in the cluster. In essence the formation of each binary system takes kinetic energy from the motion of the stars and converts it into energy that binds the two stars into orbit, each one around the other. The pair accordingly sink toward the center of the cluster. Still more recently, Spitzer and his associates at Princeton have devised models suggesting that in the crowded central region of a globular cluster, where the most massive stars are huddled, each star is likely to capture one or even two or more companions. As a result the central region may become even more densely packed as the globular cluster evolves.

The reason for my misgivings in the face of all this circumstantial evidence is simple. Massive binary star systems do indeed prevail in open clusters. Mizar is a prime example. It is a star in the Ursa Major open cluster. More recognizably, it is the star at the bend of the handle of the Big Dipper. Mizar has a companion called Alcor, which is faintly visible to someone with excellent eyesight. Evidently, then, Mizar is a binary system. Actually both Mizar and Alcor have companion stars. Each one is a binary system. That makes it all the more exasperating that the search for instances of massive binary star systems near the center of globular clusters has not been equally successful. In fact, it has had no successes at all.

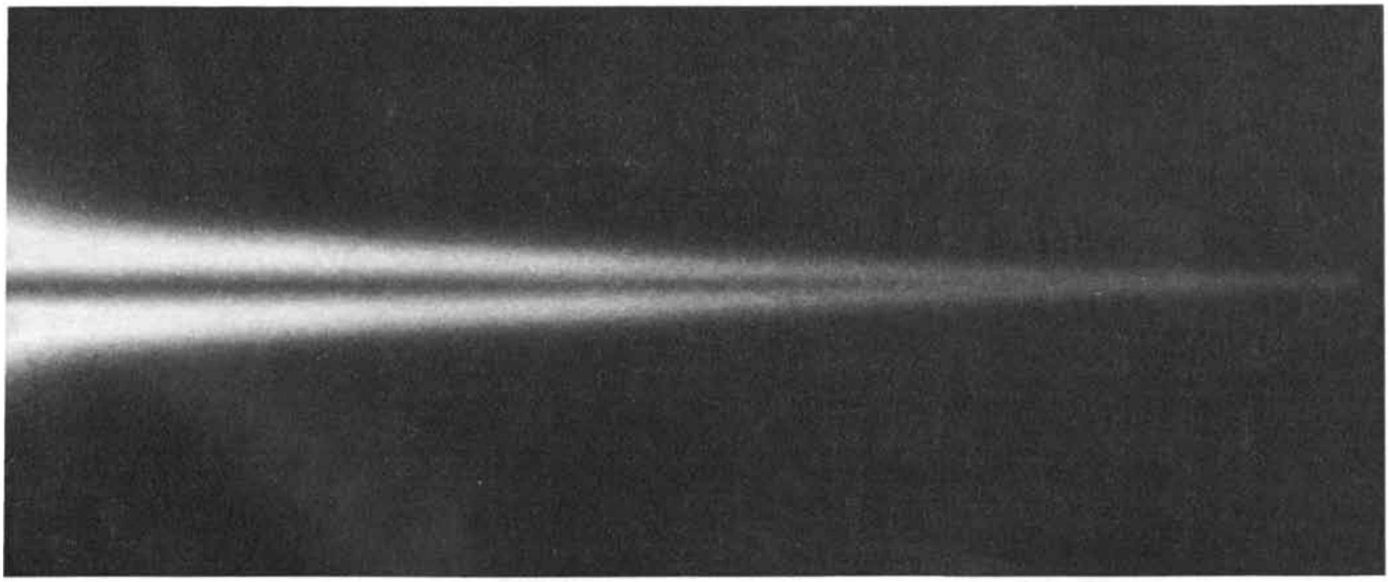
Something else has been found, however. In 1976 Jonathan E. Grindlay and Herbert Gursky of the Center for Astrophysics of the Harvard College Observatory and the Smithsonian Astrophysical Observatory reported that an X-ray burst of terrific strength had reached the solar system from a direction close to that of the center of the globular cluster NGC 6624. The burst lasted for only eight to 10 seconds. Similar bursts have since been detected repeatedly from six globular clusters. According to one hypothesis, there is a black hole with a mass not less than 100 solar masses at the center of each such cluster. The bursts would then result from the intermittent collision of interstellar gas with a hot accretion disk of gas that surrounds the black hole. It comes to mind that the hypothetical black hole might be the ultimate consequence of the gradual collapse of stars toward the center of the cluster, a collapse to which the formation of binary star systems may have made a contribution, and perhaps even given the final pushes.

Star Birth in Molecular Complexes

It was apparent even 35 years ago that the interstellar medium in the disk of the Milky Way includes clouds of gas and dust in which newborn stars are condensing. After all, some of the most striking objects in the galaxy are emission nebulas: bright clouds of ionized hydrogen atoms that are energized by the groups of giant stars inside them. The stars emit radiation at such a prodigious rate that they cannot have been doing so for more than a few tens of millions of years. The stars are therefore quite young. With only optical observations of the clouds, however, it was difficult to advance our understanding of how the stars actually formed.

All of this has now changed. Radio astronomy has revealed the presence in the clouds of more than 50 kinds of interstellar molecules, from hydrogen molecules, the lightest and by a factor of 1,000 the commonest, to a nine-carbon chain, the heaviest. Each species of molecule emits electromagnetic radiation at characteristic radio wavelengths. Moreover, infrared astronomy is now equipped to reveal the incipient stars themselves within their dense, obscuring clouds.

In almost all the processes in which a



The beginning of the simulation was a mathematical description of the number of stars of a given brightness in a given direction from the solar system. The computer displayed the resulting pattern of luminosity (and a central dust lane) as it would look to an outsider.

new star is thought to form, the initial step is the development of a concentration of matter—I shall call it a nodule—inside a cloud of interstellar atoms and molecules (mostly hydrogen molecules), with a small admixture of dust. Some of the clouds are actually enormous clumpy distributions of matter. A giant molecular complex, for example, can have a mass of several hundred thousand solar masses. Other clouds are much smaller. Some of the clouds called globules have masses of only 20 solar masses. What all the nodules have in common is that they tend to collapse, mostly under the influence of their own gravitation, with perhaps occasional pressure from outside. An outside pressure that may be quite important is the one exerted on a cloud and the incipient nodules inside it when a density wave passes through and a spiral arm begins to form. The wave may well accelerate each of the processes I shall now describe.

The conditions in the roughly 4,000 giant molecular complexes that lie within 13,000 parsecs of the center of the galaxy seem ready-made for the formation of new stars. For one thing, each complex has plenty of matter. The mass of a typical complex is as great as several hundred thousand solar masses. Its diameter is about 50 parsecs. It is the most massive object in the galaxy (unless the galactic center has a supermassive black hole). The mass of a giant molecular complex is thought to be almost entirely hydrogen molecules; at a temperature of 20 degrees K., the cloud is too cold for the molecules to dissociate into atoms. The cloud is also too cold in most places for the hydrogen molecules to emit detectable amounts of radiation at their characteristic wavelengths. The presence of molecular hydrogen must therefore be inferred. The cloud is detected best by the emission (at 2.6 millimeters) of the next-commonest molecule, carbon monoxide, and by the emissions of still less common molecules such as formaldehyde.

One other component of a giant molecular complex is important for star formation. That component is the dust. The dust particles are sites on which the surrounding gas can collect. The dust also shields the incipient stars from ultraviolet radiation, which would disrupt the condensation. It is thought that for every 100 to 200 grams of molecular hydrogen in the giant molecular complex there is a gram of dust.

Two nearby complexes of carbon monoxide and dust (and presumably hydrogen in abundance) have been particularly well studied. They are the Orion complex (which is centered on the bright region of ionized hydrogen and newborn giant stars called the Great Nebula in Orion) and the Ophiuchus complex (which blocks the light from the center of the galaxy). In the Orion complex Becklin, Neugebauer and their associates at Cal Tech, together with Frank J. Low of the University of Arizona, have found evidence at infrared wavelengths for the presence of condensed objects that seem to be intrinsically very red. Each one may be a newborn star still embedded in a thick cocoon of dust that the star's ultraviolet radiation has not yet blown away completely. The radiation heats the dust, which then radiates in the infrared. The star may have formed in the first place when the cooling of a small part of the complex reduced the pressure that results from the heat of a gas. The cooled region would thus have begun to collapse under the influence of its own gravitation.

In the Ophiuchus complex a group of 30 stars has been found by Gary L. Grasdalen of the University of Wyoming, Stephen E. and Karen M. Strom of the University of Arizona and Frederick J. Vrba of the U.S. Naval Observatory. The stars were not detected earlier because they are hidden by at least 30 magnitudes of optical absorption; it took an infrared search to find them. More recent observations by Charles J. Lada and Bruce A. Wilking of the University of Arizona in an extremely dense dust cloud near the star Rho Ophiuchi have revealed the presence of 20 similar stars whose optical radiation is dimmed by as much as 100 magnitudes. The stars lie only half a parsec from the ones discovered earlier.

The evidence to date suggests that giant molecular complexes give rise spontaneously to stars that have masses no greater than a few times the mass of the sun. In particular it seems they give rise to stars of the spectral classes *B, A, F* and *G*. (The sun is a star of the class *G*.) George H. Herbig and his associates at the Lick Observatory have found near the borders of the well-studied giant molecular complexes some groups of small, rather dim and nebulous stars that are said to be young. Almost all of them have a brightness that varies irregularly. They are called T Tauri stars. Perhaps they are products of the interrupted condensation of nodules. They may have wandered out of the complex. Often they are seen in places where the ultraviolet radiation of newborn massive stars has blown away the dark nebula's gas and dust.

Star formation of another kind is typical, it appears, in places where an emission nebula with supergiant *O* and *B* stars embedded in it lies next to a giant molecular complex. The clearest exposition of what happens in such cases has been given by Bruce G. Elmegreen of Columbia and by Lada. According to Elmegreen and Lada, the *O* and *B* stars emit ultraviolet radiation whose pressure piles up cold gas and dust at the outer edge of the complex. The result is the condensation of protostars there. In places such as the Orion nebula the process appears to be advancing sequentially. In the Orion nebula a group of *O* and *B* supergiants is fading after a lifetime of a few tens of millions of years. The radiation from these stars has triggered the formation of a younger generation of *O* and *B* supergiants, and the younger stars in turn are now emitting radiation that eats its way slowly but persistently into the giant molecular complex, where a third set of *O* and *B* supergiants will presumably form. Why the process should give rise to *O* and *B* giants and supergiants rather than the smaller *B, A, F* and *G* stars that condense spontaneously in the complex is not yet understood.

Star Formation in Globules

I turn now to the class of dark clouds known as globules. Roughly 200 of these objects have been found within 500 parsecs of the sun. They have remarkably similar properties. Each one is dark and distinct and on a photographic plate is almost circular. No doubt they are nearly spherical. Their radius varies from .2 to .6 parsec, their mass from 20 to 200 solar masses and their internal temperature from five to 15 degrees K. They are impenetrable to visible light. On the other hand, some images recorded at near-infrared wavelengths, either photographically or by electronic imaging techniques, show the stars behind the globule. The dimming of the infrared radiation from such stars allows an estimate of the globule's content of dust.

Radio observations at 2.6 millimeters show that the globules are rich in carbon monoxide. Other molecules, notably formaldehyde and ammonia, have now been found as well. Evidently, then, a globule is a small and often isolated spherule of darkness quite similar in composition to a giant molecular complex. Presumably the globule is composed predominantly of molecular hydrogen too cold to emit detectable radiation.

The radio data also show that several of the globules are collapsing under the influence of self-gravitation. The rate of collapse is roughly half a kilometer per second, which corresponds to half a parsec per million years. Since the radius of a typical globule is half a parsec, a million years is roughly the time it takes for the collapse to be completed. The collapse again is crucial to models of star formation. In almost every model the center of the globule collapses faster than the periphery, and so a nodule forms. The collapse converts into kinetic energy the gravitational potential of the infalling matter. Eventually the energy at the center raises the temperature of the matter enough for thermonuclear fusion to begin. This signals the birth of a star. If the star is large, it emits enough radiation to blow away the gas and dust that surround it.

In short, a globule ought to give rise to a single star in about a million years. Considering the time scale of the process and estimating the number of globules in the Milky Way, one concludes that the globules might account for the formation of 25,000 stars per million years, or about a sixth of the overall rate at which the stars in the galaxy form. (The formation of stars in giant molecular complexes is perhaps a more fecund process.) In 1977, however, Richard Schwartz of the University of Missouri observed what appears to be a pair of incipient, nebulous stars that are being expelled from a very dark globule some 300 parsecs from the solar system in the part of the southern sky marked by the constellation Vela. The creation of two stars rather than one is surprising, and so is their expulsion from the globule. One wonders where the energy to expel them came from. The two stars are connected by a luminous strand that might be described somewhat fancifully as a stellar umbilical cord.

I should like to give one final example to suggest how much remains to be learned. Investigators attempting to model the processes by which stars form have tacitly assumed that the protection of a cloud of dust is required. Inside such a cloud the temperature is low and the accretion of matter is undisturbed. In particular the dust shields the interior of the cloud from disruption by ultraviolet radiation from outside sources. Consider, however, the Clouds of Magellan, which lie in the galactic corona but defy the generalization that all the content of the corona is old.

In the Large Cloud of Magellan there is a grouping of roughly 50 luminous *O* and *B* stars known as Shapley's Constellation I. It is likely that their age is no greater than 20 million years. Their velocities are on the order of only 10 kilometers per second. Hence each star has moved no more than 200 parsecs from where it was born. Within a radius of from 200 to 300 parsecs of where the stars are now, measurements at 21 centimeters have revealed five million solar masses of neutral atomic hydrogen and measurements at optical wavelengths have revealed 60,000 solar masses of ionized hydrogen. But the sky in that area is transparent. It probably harbors no giant molecular complex; it has little or no molecular hydrogen and it has no cosmic dust. How did the stars form there?

Advanced Technology Center
Santa Clara Valley, California

This is one of a series of special reports on centers of high technology throughout the United States by Peter J. Brennan in association with Development Counsellors International, Ltd.

California's sunny Santa Clara Valley is a bubbling pot of advanced technologies. It is a veritable thicket in which each new company founded on a better idea quickly sends out shoots and runners, which in turn soon mutate. It is a land where technologists ride the wavecrests of the future.

There is in the Valley a tremendous intellectual and technological ferment that constantly stimulates the creation of new products and new companies. There are the venture capitalists (most of whose venture capital comes from the East). The success rate is extraordinarily high—better than four of five new ventures healthily survive. And the advanced technology sector is far more diverse than basic semiconductors. It is a diversity that stems from historical trends, the enormous influence of a very few men, the climate, and accident.

An Endowment

But for Stanford University, the communities of Menlo Park, Palo Alto and the like would be no more than bedroom suburbs of San Francisco.

Leland Stanford, railroad baron, Governor, Senator, San Francisco resident, immensely wealthy, had an 8,800 acre summer ranch in Palo Alto. When Leland Stanford, Jr., died at age 16, Stanford founded the University in his memory and gave it the Palo Alto property, which the deed of trust forbade the school ever to sell. That provision came to play no small part in the Valley's industrial preeminence.

Opened in 1891, the new school soon reached a level of excellence equalled by few and surpassed by none. Faculty members have won more than their share of Nobel prizes, some 9 at last count. The most recent was awarded in 1980 for work that underlies the new technology of genetic engineering.

Retired Professor of Electrical Engineering and Provost Emeritus Fred Terman is one of the remarkable men whose influence so shaped the Valley, though he modestly demurs when others credit him with much of the region's economic development. "I didn't do it," he says. "I was a focus and got the publicity, but it was the University and the community."

In the late 1930s, there was a surge of inventiveness at Stanford where the elements of linear accelerators, nuclear magnetic resonance and microwave technology first appeared in hardware. Professor Terman, in the thick of it all, encouraged students to establish businesses locally, based on the new high-technology of electronics. These students had names like David Packard, William R. Hewlett and Russel and Sigurd Varian.

After World War II, Stanford had land and a shortage of money. Professor Terman and the administration encouraged the trustees to establish an industrial park, the land to be leased to suitable firms. Recalls Prof. Terman, "At first we didn't know if American companies would agree to build facilities on leased land. But we found that companies would actually pay more to lease close to the University than to own land on the flats."

Thus began the Stanford Industrial Park, the real nucleus of commercial advanced technology in the area. The first tenants were Varian Associates and Hewlett-Packard Co. The 660-acre park now has some seventy advanced-technology companies.

Some large national firms also took an increasing interest in the region. In the Fifties, Lockheed, General Electric, Ford, GTE and Raytheon among others moved into the Valley.

And there were major Government facilities, at NASA's Ames Research Center, Moffet Naval Air Station (whose enormous air-ship hangars memorialize a bygone technology) and the Lawrence Laboratories at Berkeley and Livermore not far away.

The Aerospace Factor

One of the earlier arrivals was Lockheed Missiles & Space Company, Inc., (LMSC) at Sunnyvale next to Moffet NAS. The region's largest employer with some 22,800 people, LMSC was founded in 1954 and moved to Sunnyvale in 1956.

The company was split off from Lockheed's main Burbank, Calif., site specifically to handle advanced programs. "We made a survey in a reasonably businesslike way to find a suitable site," recalls Robert A. Fuhrman, President, who is also chairman of the Santa Clara County Manufacturers Group. "We needed to be near a university and in an area where the science is competent.

"Our products are large and complex and take thousands of people," says Fuhrman. Some of these products include the Polaris, Poseidon and Agena missiles, the Discoverer program, deep submergence rescue vehicles, all-terrain ground vehicles, deep-sea mining and Ocean Thermal Energy Conversion (OTEC). It is a broad base of applied and very advanced technology that requires many disciplines and lots of outside help. "We probably subcontract 50% of our work," says Mr. Fuhrman, "and have helped to develop a lot of companies in the area."

Semiconductors Arrive

"If Professor Shockley's mother had not lived in Palo Alto," remarked

Three good reasons why professionals pick Apples.

1 In research

Apple personal computer systems help you collect, store and analyze data as fast as you can load a disk and execute a program. Because more than 100 companies offer software for Apple, you have the largest program library for manipulating your data in the personal computing world. Need special programs? Use any of Apple's development languages — BASIC, FORTRAN, Pascal.

2 In engineering

Apple personal computer systems let you define models, make trade-offs and refine prototypes. Want to study cause and effect of several variables? Apple computes new results instantly and displays them in colorful, easy-to-read graphs, charts or plots on a video monitor.

3 In production management

Apple personal computer systems make it easy to gather data, analyze productivity, measure yields and facilitate all phases of production control. Want to speed up repetitive tasks? Rely on Apple's word processing capabilities to write, edit and print your reports.

Apples grow with you.

Whichever system you pick, Apple never locks you into a single configuration. You can use up to four or eight I/O accessory expansion slots to add an IEEE bus, Apple's Silentype™ printer, a modem or a graphics tablet. Add memory up to 64K bytes or 128K bytes. Add up to four or six 5¼″ disk drives without adding any overhead.

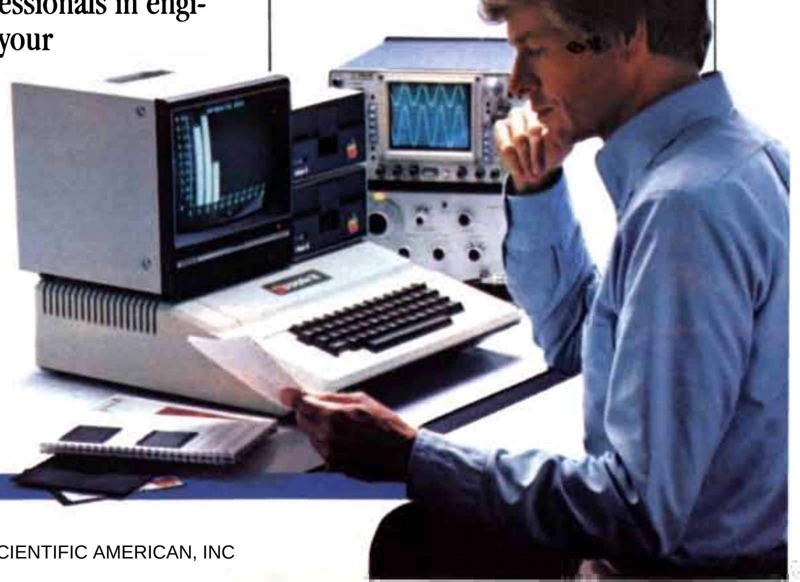
For support, service and the best extended warranty in the industry — Apple is the answer.

If you have any other questions about why Apple is the pick for professionals in engineering, see your nearest Apple computer dealer or call **800-538-9696**. In California, **800-662-9238**. Or write: Apple Computer, 10260 Bandley Drive, Cupertino, CA 95014.

	Apple II	Apple III
Maximum Memory Size	64K bytes	128K bytes
Screen Display	40 column (80 column with peripheral card) 24 Lines Upper Case	80 column 24 Lines Upper Case/Lower Case
Screen Resolution (B&W)	280 x 192	560 x 192
Screen Resolution (Color)	140 x 192 (6 colors)	280 x 192 (16 colors)
Keyboard	Fixed	Programmable
Numeric Key Pad	Accessory	Built-in
Input/Output	8 expansion slots	4 expansion slots plus built-in: disk interface RS-232 interface Silentype™ printer interface clock/calendar
Disk Drives	Add-on one to six drives	One drive built-in, plus interface to support three more drives
Languages	BASIC Fortran 77 Pascal Assembly Pilot	Enhanced BASIC Fortran 77 Pascal Assembly
Typical Configuration Pricing	CPU, 48K RAM, single disk drive, B&W Monitor (9″), Silentype™ printer, and BASIC. **$2875.00***	CPU, 96K RAM, integrated disk drive, B&W Monitor (12″), Silentype™ printer, SOS, Enhanced BASIC. **$4865.00***

*Suggested retail price.

apple computer inc.



a local executive, "there would be no Silicon Valley."

MIT-educated, Palo Alto-raised Shockley was co-inventor of the transistor while at Bell Laboratories in New Jersey. He won a Nobel Prize for it. In 1956, Dr. Shockley returned to Palo Alto and set up Shockley Transistor Co. The technology prospered but the company did not. Within two years, a number of Shockley's top associates departed and, backed by Fairchild Camera and Instrument Co., set up Fairchild Semiconductors Co. This was the true seed of the semiconductor industry and genesis of "Silicon Valley."

Shockley's associates, among them Dr. Robert Noyce, began a trend that has given the area its special flavor. Noyce himself left Fairchild in 1968 to start Intel. But before then, many more had departed to start their own firms: Rheem in 1959; Signetics and Amelco in 1961; Molectro in 1962; National Semiconductor in 1967; and in 1968, a bumper crop that, besides

Intel, included Precision Monolithics, Computer Micro-Technology, Qualidyne and Advanced Memory Systems. In 1969, Advanced Micro Devices and Four-Phase appeared.

Says former Fairchild president Dr. C. Lester Hogan: "It was a time full of turbulence."

A Community of Technology

Though the technology industries employ 630,000 people in the region, it almost seems as if everyone knows everyone else at every level. Meetings turn out freshly minted engineers and board chairmen at the same table. For the most part, the founders of even the biggest and most successful companies are still young men who thoroughly enjoy their success.

But size is not the criterion for success. Where four of five ventures succeed, many are small and remain so. Indeed, more than 80% of the high-technology companies in the Valley have less than 200 employees. One such is Zircon International Inc., a five year old electronics firm with only 46 employees. The privately-held electronics company has moved into consumer electronics. Small though Zircon is, it was large enough to buy Fairchild's electronic games business.

These smaller companies, of course, compete with their larger brethren for talent, but apparently with little difficulty. Says Peter Taylor, Director of Finance and Administration of Logical Machine Corporation: "We offer a challenge to the scientist or engineer who would rather work in a small company environment." Logical Machine itself was founded by John Peers, an English computer specialist who felt there was greater opportunity in Silicon Valley.

The Senior Citizen

It is odd perhaps to characterize a company that is barely forty years old as the old man of the industry. Yet that's the Hewlett-Packard Corporation, founded in Palo Alto in 1938. Older than most but still very young, Hewlett-Packard continues to typify the industry. The founders, William R. Hewlett and David Packard, gave H-P an openness and style of management that have become the hallmark of the industry in the Valley regardless of the industrial sector. It is a style more reminiscent of a community of scholars and equals than of hierarchical employees. Or perhaps that of people working together in a garage.

Packard, whom Professor Terman persuaded to return to the Valley from his job with General Electric in Schenectady, N.Y., started working with Hewlett in Packard's garage.

The company hired its first employees and moved out of the garage in 1940. In 1956, the company moved into the Stanford Industrial Park and in 1959, made its first expansion outside the Valley. "I was one of an H-P team which opened up the first shop outside Palo Alto," recalls German-born Fred Schroeder, Director of Corporate Development and himself a one-time student of Terman's classical work. "That was in Germany. We were quick to take advantage of geographic market conditions, the Common Market."

Hewlett-Packard, like virtually every other company in the area, is constantly expanding. It now has 57,000 employees, some 18,500 in the Bay area and 16,600 in Palo Alto alone.

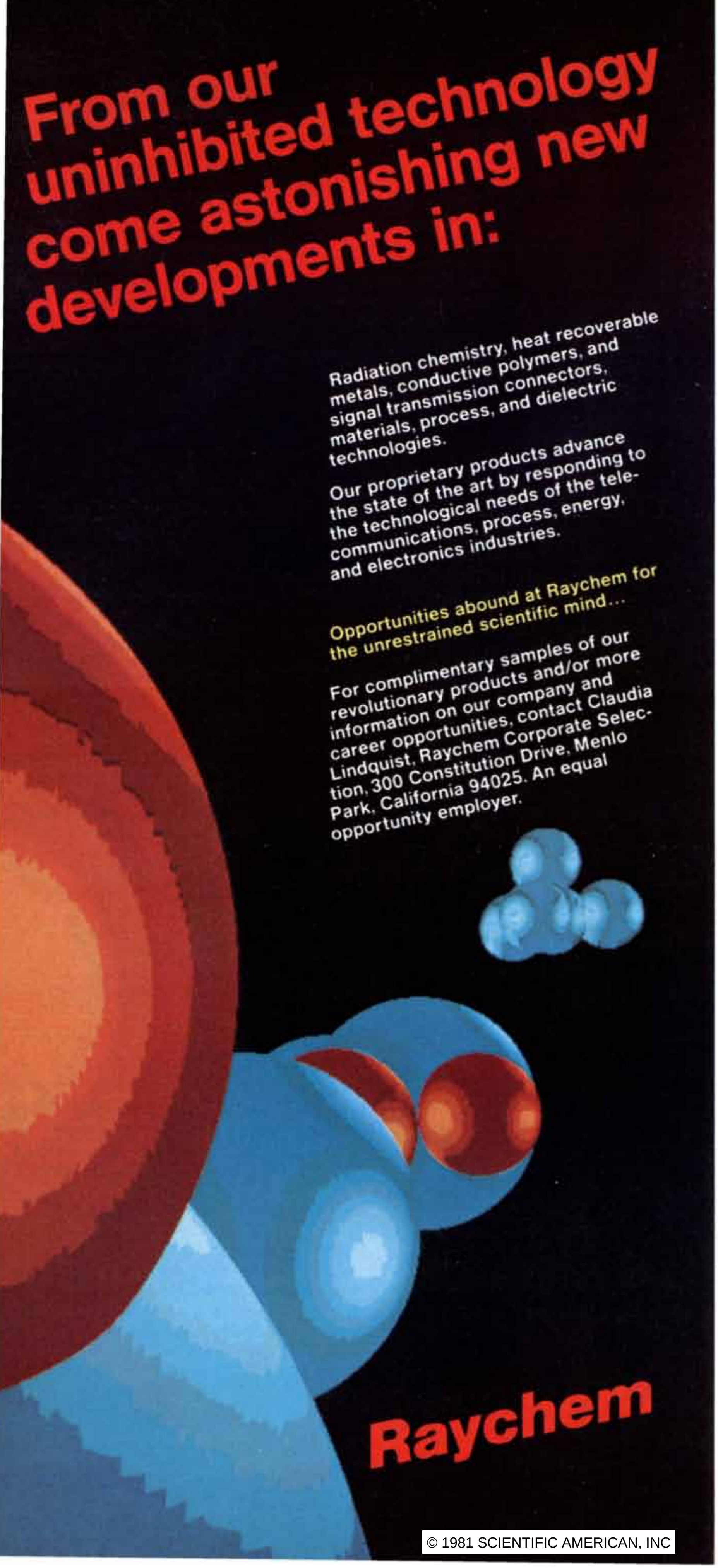
H-P values the small-company feel it started out with and has striven mightily to maintain it as it has grown. Management maintains not an open-door policy but a no-door policy.

The concept is very common throughout the Valley. At Intel, for example, Dr. Robert Noyce's only concession to his position as founder and Vice Chairman is that his ground-floor cubicle is on a windowed corner. In such ways has H-P's influence been pervasive.

Microwave Beginnings

The Varian brothers started about the same time as Hewlett and Packard. But their work on the Klystron tube, the first microwave device, was preempted by World War II. Not until 1948 did they return to Palo Alto and under the shadow of Stanford University establish Varian Associates with $22,000 and six employees. The first plant in San Carlos wasn't a garage but it wasn't much bigger.

The company was at first only a maker of microwave tubes. But its close association with Stanford enabled the firm to develop rapidly much of the new technology coming out of the University's physics laboratories. Varian soon was in nuclear



magnetic resonance (NMR), a range of analytical instruments, geophysical instruments and linear accelerators.

One thing led to another. "You have to pump a tube to get the vacuum," recalls Edward J. Barlow, Vice President for Research and Development, "so we invented a new type of electronic vacuum pump for our own use. We soon saw it had other uses and were in the vacuum pump business."

That process eventually led to a wide range of products in distinct lines, now manufactured and marketed through appropriate groups. "The common thread is physics and electronics," says Vice President Barlow.

"We have gone from a technology-driven industry to a market-driven one," says Mr. Barlow, countering the oft-quoted observation that much of the Valley's technology is "a solution in search of a problem." Those markets have brought the company to a $600-million per year sales level, supplied by 14,000 people in 27 plants in six states and 10 countries. Some 7,000 employees are in the Valley.

Though naturally there is great Government demand for many of its products, Varian's markets are preponderantly commercial, in common with most valley industry. "Some of our Groups are 100% commercial," says Barlow.

It is commercial at a very advanced level. There is an industrial sputter-coating machine to protect auto grills and headlamp bezels with microscopically thin and uniform coatings; reaction ion beam etching for semiconductor manufacture; ultrasonic scanning of the body that shows internal body organs functioning in real time.

"Stanford has been very important to us," acknowledges Barlow. "There is a tremendous advantage in being here."

Where The Silicon Is

The core of Valley industry is of course semiconductors. These many companies largely bred out of Fairchild Semiconductor bear strong family resemblances. Yet as in any large family, there are also significant variations.

Fairchild Camera and Instrument is now a unit of Schlumberger Limited. It is also one of the oldest corporate names in the Valley, having been founded in 1920 in New York as the Fairchild Aerial Camera Co.

The company entered the semiconductor business by funding Dr. Robert Noyce and his associates when they left Shockley Transistor and, in turn, fostered many more companies.

That is all past. Fairchild means to maintain its reputation as an innovator in the semiconductor field in its own right, not as a breeding ground for others. Under Schlumberger Vice President Dr. Thomas A. Longo, an old Fairchild hand and the company's chief technical officer, the company has embarked on an ambitious long-term research and development program. "Silicon was the base of the industry and will continue to be so," avers Dr. Longo, "so we know where we have to go. In three or four years, we will be spending 7 or 8 times as much as we have been on R & D" to bring into production the next generations of very large scale integrated circuits (VLSI).

All that R&D and new manufacturing will take lots of staff and a technical environment that will retain them. The company has introduced new personnel practices that Dr. Longo credits with already reducing turnover to 12%, half the Valley average of 23%. "Fairchild has trained much more than its share of people in the Valley," says Longo, who wants some of them back. "We know there are lots of good people out there. And once you've got a super job for a guy, it's not hard to get him to come."

Catch The Wave

A Fairchild descendant that is just as aggressively recruiting is Advanced Micro Devices (AMD) of Sunnyvale. Founded in 1969 by W. J. Sanders III and several associates, all from Fairchild, the company has grown prodigiously. In 1980 sales reached $225 million, up from $11.2 million in 1972, when the company first offered stock to the public.

AMD has positioned itself to be a major merchant supplier to the telecommunications industries. It specializes in new architecture, logic and circuit design techniques in VLSI technology, a wave that is just about to break over the electronics-using world. Indeed, AMD's recruiting campaign uses the phrase and image "Catch the Wave."

The company's products cover a wide range of devices requiring edge of the art technology to design, develop and manufacture. They include advanced microprocessors, memory chips and specialized telecommunications devices such as SLIC and SLAC (subscriber line interface circuit and subscriber line audio processor circuit). These will become increasingly important as telephone systems move to all digital switching.

AMD has used as a symbol the asparagus, a plant that takes several years from planting to harvesting but then yields abundantly and indefinitely. It has fostered an aura of technological excitement and personal reward and has striven to retain an entrepreneurial spirit through all employee levels. One factor that contributes to that spirit is substantial employee ownership through options

Signetics and American Microsystems are two more among many semiconductor companies that provide interesting contrasts.

Signetics was founded in 1961 as the first company specifically to manufacture and sell integrated circuits. The new company was soon bought out by Corning Glass Corp., but was then divested and, in 1975, bought by U.S. Philips Corp., an arm of N.V. Philips Gloelampenfabrieken of Holland.

With 5,000 local employees and 17,000 total, the company is the second largest in Sunnyvale and 11th in the Valley. It makes only semiconductors and integrated circuits but does not consider that to be limiting. "With a $6-billion current market and a projected $50 to $70-billion market by 1990," says Michael Hackworth, Senior Vice President of the MOS/ANALOG Products Group, "the sector offers a very high opportunity

Signetics expects the Philips connection to be beneficial for both companies. "Technology moves both ways," says Hackworth, "and picking and choosing the right technologies to pursue is a non-trivial problem. One requires an enormous investment just to decide which to use. We have that resource."

Signetics has its own Advanced Technology Center, headed by Dr. Kenneth A. Pickar, but also has access to Philips Research Laboratories in Sunnyvale headed by Dr. Else Van Doi. The emphasis here as elsewhere is on research.

At American Microsystems, Inc (AMI), the emphasis is also on research and development. The 14-year old company in Santa Clara is next to one of the few prune yards remaining north of San Jose. The employment level at the site has not changed significantly recently. But its makeup has as the company has established more manufacturing sites outside, reduced manufacturing here and increased its research and development staff.

"We are seeing as much circuit innovation as process innovation today," says President Glenn E. Penisten, who came to AMI from Texas Instruments four years ago. "And that is a software driven technology. It is becoming a software limited business, which means people limited.

AMI is today a $140-million company. "We would have been larger," says Penisten, "but we got off the track in the Sixties as many others did when we went into consumer goods.

The Supporting Technologies

While AMI sticks to what it knows best, MOS and CMOS (it was the first company to take MOS commercial) it has acquired a subsidiary, Millenium Systems, which makes systems to design or test microprocessor-containing systems. Millenium is part of a most important sector of Valley industry, and one just as technically advanced. That sector supplies the equipment, services and materials that the rest need to turn laboratory items into commercial products.

These firms, with names like Raychem, Calma, Millenium, Monsanto, Siemens, Speedfam and many more are often members of the Semiconductor Equipment and Materials Institute (SEMI), which is itself in Mountain View. Some 30% of SEMI's 600 members are in Silicon Valley.

Calma Company, a representative member and a subsidiary of General Electric Co., produces a range of sophisticated equipment for the semiconductor industry. The Sunnyvale concern makes automated insertion software, graphic data systems and, the heart of the semiconductor process, photo-mask making equipment and systems. It also makes computer-aided design (CAD) electronic equipment for this and other markets.

Another firm, not a member of SEMI, is Raychem Corporation of Menlo Park. Less exotic but equally necessary and technologically demanding, Raychem's products are in a sense the nuts and bolts of the electronics industry. Without Raychem's hookup wire and cable, coaxial cable, special-characteristic rubber and plastic tubing, termination devices and the like, all would be in isolation.

Data Storage

Representative of the huge peripherals industry in the region are such companies as Ampex and some of its descendants like Memorex and Shugart. Others include Dysan; Anderson-Jacobson, originator of the acoustic modem; and Verbatim, by the same Anderson.

Information storage is a difficult technology that works at the very edge of many disciplines—mechanical, chemical, electronic, electrical and physical. Tapes and disks must be made of materials that are flexible yet can withstand high forces in tension or rotation and remain dimensionally stable within extremely narrow limits. The equipment that drives the media must be built to even more exacting standards.

Electronics and mechanics must match. A disk carrying 965 tracks per radial inch must not wobble. Memorex, for example, had to design a new type of ball bearing so that it would not.

Current research aims at ever higher densities through semiconductor technology. Read-write heads now are made hundreds at a time on ceramic wafers by deposition methods, then diced and shaped before mounting. The design permits much narrower tracks and thus more of them. The limiting factor becomes the media. Thin film techniques use sputtering to deposit a uniform layer merely molecules thick. Such technology may by 1985 allow disk densities per square inch of 50-million bits, the equivalent of 4,000 pages of typescript.

Laser Pressure Testing

Test and design methods are equally exotic. The aerodynamics of the head are extremely important since it literally flies above the disk. A pressure measurement system uses lasers to produce interference patterns that can then be interpreted and plotted as pressure curves.

Memorex Corporation was started in Mountain View in 1961 specifically to exploit this difficult multi-disciplinary technology. The first product was computer tape, shortly followed by video tape. These soon led to complete tape drive systems and disk storage systems, a line of evolution the company has continued to follow at its campus-like laboratories in Santa Clara.

A high-technology manufacturing as well as research and development company, Memorex is people short in critical jobs, particularly engineers who can deal with millionths of an inch. It is high-precision work "We need engineers who can work in manufacturing as well as in the lab," says Gordon Smith, Vice President. "Software people are also hard to come by." Total employment is currently some 13,000 people worldwide with about 6,700 of those in the Valley.

At the other end of the Valley, in Redwood City, Ampex tills the same field. One of the early ones, founded in 1944 to make motors for radar systems, Ampex is also one of the bigger ones with sales exceeding half a billion dollars. Ampex was a pioneer in magnetic recording systems.

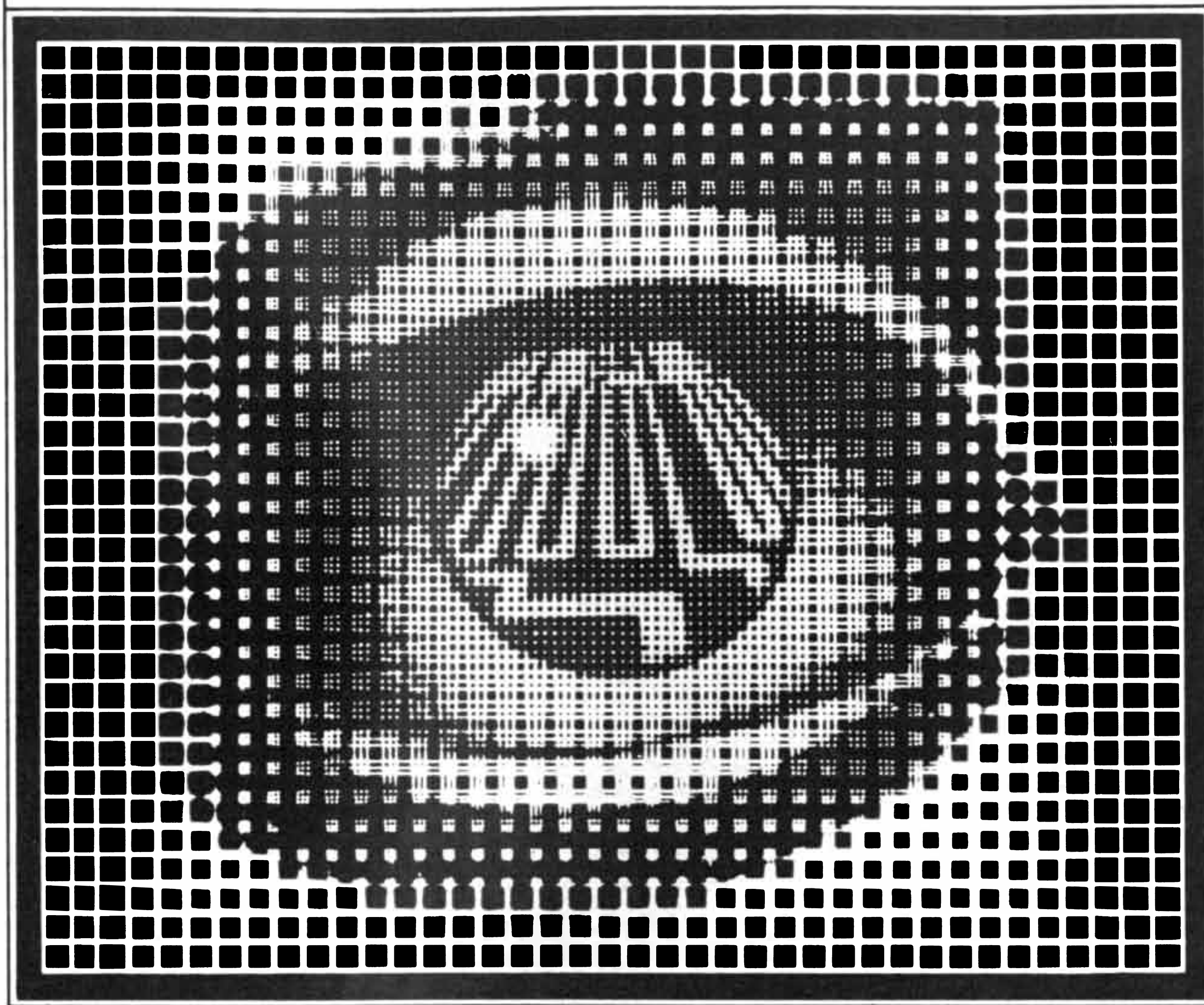
The company has 1,300 of its 13,000 worldwide employees at its Redwod City headquarters and laboratories and, unusual in that area, has plenty of land. "Somebody showed a lot of foresight," remarks C. Ridley Rhind, Vice President of Marketing.

Technologies Galore

Electronics is not the only technology in the Valley. Stanford University itself has schools for every discipline in the physical and life sciences as well as a renowned medical school. The nine Nobel prizes have come in many fields. And the inventive minds are readily available to the local industrial community for consultation

ENGINEERING &
COMPUTER SCIENCE
SPECIALISTS

We turn imagination into tomorrow's technology.



Whatever your area of specialization is, you'll find it to your advantage to look into the career opportunities at Lockheed.

Our challenging and diverse projects will provide you with a rewarding career that presents great opportunity for rapid growth and career development. You can get involved in such diverse areas as advanced spacecraft, unique deep sea vehicles, fleet ballistic missiles, ocean mining, communication satellites, electro-optical R&D, and computer science.

You can apply your experience in activities such as electronic or mechanical design, scientific or administrative programming, systems engineering, R&D, microprocessor applications, hybrid microelectronics, signal processing, remote sensing, test and integration, simulation, reliability, manufacturing, materials and processes, or in a number of other disciplines.

You'll find that in virtually all of our programs, scientists and engineers are working at the frontiers of their disciplines, advancing the state of the art.

You'll also find that our in-house training programs are among the best in industry. Experience, plus training. That's why at Lockheed, technical excellence is truly a way of life. Write now to:

Employment Department SA-381, Lockheed Missiles and Space Co., P.O. Box 504, Sunnyvale, CA 94086.

We are an equal opportunity affirmative action employer. Proof of U.S. Citizenship is required.

At Lockheed, Technical Excellence is a Way of Life.

LOCKHEED
MISSILES & SPACE COMPANY, INC.

and inspiration as are the school's facilities.

Another Nobel Prize factory, the University of California at Berkeley is just across the Bay. "Berkeley has a superb department in solid state, computer science and computer architecture," says Fairchild's Les Hogan.

In common with many another university, Stanford long ago established a separate institute to provide research facilities for local industry. The Stanford Research Institute (SRI) was spun off as a separate not-for-profit entity and is now housed in a motley collection of buildings in Menlo Park. It has additional offices and laboratories across the world.

The Institute did the original work on ink-jet printing, the high speed printing technology of the near future. It also pioneered magnetic coding on bank checks. An ultrasonic energy medical application was developed in the bioengineering department. SRI researchers also built a laser device for eye surgery. The Institute's process for biomass conversion is the only one known that makes oil. Its coal/air fuel cell may be more efficient than steam turbines. The list is long and varied, and much more than science.

The Life Sciences

Such diverse resources naturally attract similarly oriented companies. Prominent among these is the international pharmaceutical company, Syntex Corp., which moved its headquarters from Mexico to the Stanford Industrial Park. "Had Syntex not moved here, it would not have become the company it is," says Dr. Alejandro Zaffaroni, who managed Syntex when it moved to Palo Alto and who is now president of Alza Corp., another biological company just down the road.

Syntex started as a research organization interested in developing patents for new pharmaceuticals, mainly steroids, that it could then license. In 1960, the company opened a lab in the U.S. to be closer to that market. "In 1962," recalls Executive Vice President Richard G. Rogers, "we decided to move near a major medical institution with good medical research and good chemistry. We chose the Stanford Industrial Park because of Stanford University and because the climate and the area would help us attract good people."

The $580-million company has nearly 3,000 people in Palo Alto, 9,100 worldwide. Some 30% are professional. The emphasis is research in the life sciences and the development of new drugs. For years, the company was best known for steroid hormone derivatives, which yielded dermatological medicines and contraceptives. But the company now produces non-hormonal drugs, medical diagnostic systems, veterinary drugs and animal health products and even beauty care products.

The other companies are mostly descendants of Syntex. One is Alza Corporation, founded by Dr. Alejandro Zaffaroni, a Uruguayan who once headed Syntex's research operations in Mexico and Palo Alto. With his own money and backing from Syntex and other sources, Dr. Zaffaroni set up Alza Corp. in 1968 to develop new ways to administer medications in such a manner that the concentration of the medicine in the body remained at its optimum throughout the period of treatment. At its park-like laboratories in the Stanford Industrial Park, Alza's 350 people have developed an insert to treat glaucoma, an implantable contraceptive, a method of administering a motion sickness drug, a portable intravenous infuser and a different kind of timed release pill.

Advanced technology of many kinds goes into this research. The timed release pill, for example, has a hole measured in microns drilled into it with a laser. "The hole has to be large enough so pressure does not build up in the pill, yet small enough to allow membrane-control osmotic flow of the drug to proceed smoothly," says Dr. Zaffaroni. Alza worked closely with Coherent Inc., another Valley company specializing in laser technology, to develop the technique.

Alza itself has a spinoff, Dynapol Corp., which is working on new types of food additives that do not actually enter the blood stream but do enhance foods.

Still another Syntex descendant is Zoecon, which is developing genetic ways to control insects through their own hormones.

The technologies these companies are pursuing are very long range and costly—beyond their means, indeed. To continue the work, all three descendants have gradually given up equity to others to obtain financing. The pharmaceutical firm Ciba-Geigy owns 53% of Alza; DeKalb AgResearch owns 60% of Dynapol; Occidental Petroleum owns Zoecon.

Laser Technology

Lasers are now important manufacturing and research tools in the Valley where several companies are profitably pushing the edges of theoretical and applications technology. Among these are Spectra-Physics, Inc., and Coherent, Inc.

Spectra-Physics was started in 1961 by five people from Varian to exploit the very new laser technology, then confined to crude infrared ruby lasers. "Lasers were then a solution in search of a problem," remarks Jack M. Gill, Executive Vice President of the Instrument Group, "but widespread applications soon began to emerge."

Spectra-Physics, soon to move its headquarters from Mountain View to San Jose, is now a $130-million dollar company with eight operating divisions. Since 1961, it has found plenty of real problems to solve. "One of the earliest uses was in construction," says Gill, "where laser beams are far superior to any other levelling device." Present and expanding uses are in medicine, particularly for microsurgery, spectroscopy—"a laser gives forty times better resolution than a grating"—and manufacturing.

In Palo Alto, Coherent, Inc., founded in 1966 by Spectra-Physics alumni, also exploits laser technology. The company manufactures a wide range of products for industrial and medical uses.

Among the advanced designs is a laser annealing system for semiconductor manufacture. Using a laser, this device allows the operator to

Memorex Technicians looking into the vacuum chamber of a scanning Auger microprobe while sample is being bombarded by electrons to determine composition of different deposition layers heat precisely micron-wide bands across a wafer, thus manipulating the crystal structure without in any way affecting the composition. Standard annealing methods, which heat the entire wafer, can allow some migration of chemical elements during heating and cooling. One of the more exciting prospects of the technique is that it may make possible three-dimensional semiconductors.

The Computer Sector

Silicon Valley is becoming a leading center for computer manufacturers of every stripe. IBM was there first. Dr. Gene Amdahl was Director of that company's Advanced Computing Systems Laboratory at Menlo Park responsible for developing the largest and latest models. When IBM declined to undertake a new project that Dr. Amdahl urged, he resigned in 1970 and started his own company.

Six years and $40-million later, the Santa Clara company sold its first machine. The company was profitable in its first sales quarter and has been so ever since.

A more recent phenomenon is Tandem Computers, Inc., founded in 1974. many of the founder group and later senior staff were Hewlett-Packard alumni.

James G. Treybig, President and principal founder, was marketing manager for H-P's computer division. "Treybig left Hewlett-Packard to join Kleiner and Perkins, a venture capital firm. There he conceived the idea and developed the plan for what would become Tandem Computers," says Robert C. Marshall, Senior Vice President and employee number 33.

Rather than redundant standby computers, the usual way to assure systems reliability, Tandem builds multiple processor computer systems. All processors are on line all the time. If any part fails, the remaining units instantly take up the load, automatically doing the job of the failed part.

From 170 employees in 1978 when it shipped its first system, the Cupertino company has grown to 1,400, about half in Cupertino. Sales exceed $100 million. And in addition to two other manufacturing sites, the company is opening a plant in Fairfax County, Virginia "to be where our customer base is," says Bob Marshall.

Fun In Games

Electronic games started in the Valley. Pong burst upon the scene in 1972, an innovative application of microelectronics if ever there was one. Atari, Inc., the Pong maker, followed with a host of similar games and continues to do so.

A consumer-oriented high-technology company to whom marketing is everything, Atari is now headed by Chairman Raymond E. Kassar, who came to the technology company from the textile industry. "I am a most unorthodox, untypical chief executive for the Valley," where a surprisingly large percentage of top managers have technical backgrounds.

Atari is now a computer company. It makes and markets a line of personal and small business computers to consumer markets. The line, introduced in 1979, would seem a natural development from a company that pioneered microprocessors in computer products and "has built more computers than all the others put together," says Mr. Kassar.

Apple Computer in Cupertino is an offspring of Atari. As so often happens, Apple's founders had an idea their employer was not then prepared to implement, so they did it themselves. Gene Amdahl has much company.

Apple's founders, two engineers in their twenties, believed that the microprocessor revolution had brought prices within the reach of everyone and there was a market for real computers based on these devices that extended far beyond dedicated hobbyists. They designed and built a "personal computer," the Apple computer, based on much the same microprocessor chips then being used by Atari for games and by others for all sorts of intelligent applications. Interestingly, though the microprocessor had for years been widely hailed as the "computer on a chip," very few people other than dedicated hobbyists had actually based a real computer on such a chip.

There was indeed a market, though not quite the one anticipated. It is a professional person—writer, lawyer, doctor—and small business market rather than homeowner and hobbyist market. And it was not one Apple was to have to itself. Nonetheless, the company is now second largest in this fast-growing field and so well regarded by investors that it is valued in the stock market at an enormous multiple.

Klystron tybes for UHF television transmitters. Most UHF stations are powered by Varian klystrons.

Electronics are important to another Valley firm, Arrow Development Co. of Mountain View. The company's small engineering staff uses industrial-grade microprocessors and electronics to control the terrifying and thrilling amusement park rides it designs and builds.

Arrow's engineers, however, are more concerned with acceleration forces and strengths of material as well as the tolerance of the human mind and body to being seemingly on the edge of destruction. Its huge experimental rides loom in view of the Bayshore Freeway, leading many to think they have reached Great America, a theme park several miles south in Santa Clara.

From Computers to Phones

Messrs. Richeson, Oshman, Loewenstern and Maxfield, electrical engineers all, had an idea in 1969. They would make an off-the-shelf computer that would meet military specifications. It was a market until then served only by costly custom-built machines.

The engineers opened ROLM Corporation in a prune-drying shed and were at once successful.

In 1975, the company had another idea, to enter the telecommunications market that was now open to competition following a court decision and about to see great technological change. That idea was explosively successful. The company's growth rate since then has exceeded 50% per year.

ROLM still makes rugged computers, but telecommunications is now 80% of the business. "It was a matter of luck and judgment," says Richard M. Moley, Vice President of Marketing, Telecommunications Division, at the company's Santa Clara campus that is so sprawling, spectacular and opulent, one wonders how such a complex could be built in a mere five years. "We went ahead while others held back."

Equestrians Prohibited

There is perhaps a poignant note for a time past in a sign beside a bicycle path that parallels a Palo Alto road: "Pedestrians and Equestrians Prohibited."

Fortunately, the Santa Clara Valley technologist need not give all for science. Three universities, Stanford, the University of Santa Clara and San Jose State University, give ample opportunity for continuing education. The University of California at Berkeley is within easy reach. Nine community colleges yield a steady flow of well trained technicians and clerical people. And the 33 public school districts and over 100 private schools in the Valley's many communities provide quality education at lower levels.

Cultural and recreational opportunities abound. San Francisco with all the mystique it implies is barely an hour away.

Leisure life is outdoors where trees and flowers bloom all year and only the recreational areas of the Santa Cruz mountains bar one from the recreational areas of the Pacific beaches an hour away. Eastward are the endlessly diverse deserts beyond the Eastern ranges, in which the skiing season starts as early as late October. And there is the Bay, a sheltered boating ground that gives onto the Pacific through the Golden Gate.

Just A Nice Place

When asked why an engineer or scientist should want to come here all things considered, people eruditely discuss profession, career, education, technology, a sense of being part of something that will never be again, excitement, all worth whatever the cost. "It is a remarkable place for intellectual stimulation," says Fairchild's Lester Hogan. "In general, technological stimulation here is greater than in any university engineering department. Bright ideas are popping up all the time and there is a new process every week or two."

And truly, a scientist or engineer who is really serious about his future and about the technology almost cannot afford not to spend some time in the Valley. The area has become to virtually every sector of the electronics industry what Detroit is to automobiles, Akron to tires, New York to communications and the arts and Houston to petroleum.

But all at last falls back on location, the reason it's all here: "It's a very desirable place to live," says Intel's Robert Noyce.

An Engineering Tradition Challenging Today's Engineers. At Westinghouse.

We have always been identified by significant achievements based on solid technology. From marine vessel propulsion systems of the early 1900's

to precision missile launching systems of the '80s, excellence in engineering is a tradition at the Westinghouse Marine Division, Sunnyvale, California. And it's a tradition that stands today on a solid foundation of advanced engineering developent.

Whether you're ready to break new ground in the MX Missile Program or explore new research & development in Advanced Concepts Engineering or Product Development, we're challenging you to take a long, hard look into an opportunity to join one of the finest teams of professionals in the world.

If you're a Mechanical, Electronic or Aeronautical Engineer, consider a move to the company you can

count on for stability and continued growth. And the opportunity to pursue professional excellence in areas

such as Structural Dynamics & Stress Analysis, Gas Dynamics, Mechanical Design & Development, Systems Test, Instrumentation, Non-Metallics and Composites, Integrated Logistics, and more.

Capture the excitement of our past and create achievements of tomorrow at the Westinghouse Marine Division, Professional Employment, Department SA-2, Hendy Avenue, Sunnyvale, California 94088; (408) 735-2191. Call or send your résumé today. We're an equal opportunity employer, U.S. citizenship required.

Westinghouse
Marine Division
Sunnyvale, California

Have We Got A Job For You!

SMALLER. SMOOTHER. FASTER. FINER. STRONGER. THINNER. TOUGHER.

THOSE words are not mere product boasts. They describe, instead, the relentless realities of doing business in the marketplaces of high technology. Marketplaces with insatiable appetites for ever-smaller sizes, smoother surfaces, closer tolerances, faster accesses, greater capacities. Marketplaces with their own uncompromising process of natural selection.

Memorex® Corporation has been part of that environment since 1961. For 20 years, we have helped to shape the information recording, storage and retrieval industry. For 20 years, we have expanded our product lines to include everything from disc packs to disc drives, from word processing supplies to communications terminals. And for 20 years, in an environment where change is the constant—and an unsympathetic, ever-accelerating constant at that—Memorex Corporation has done more than simply manage to change so as to keep up. We have *managed change itself* and stayed ahead.

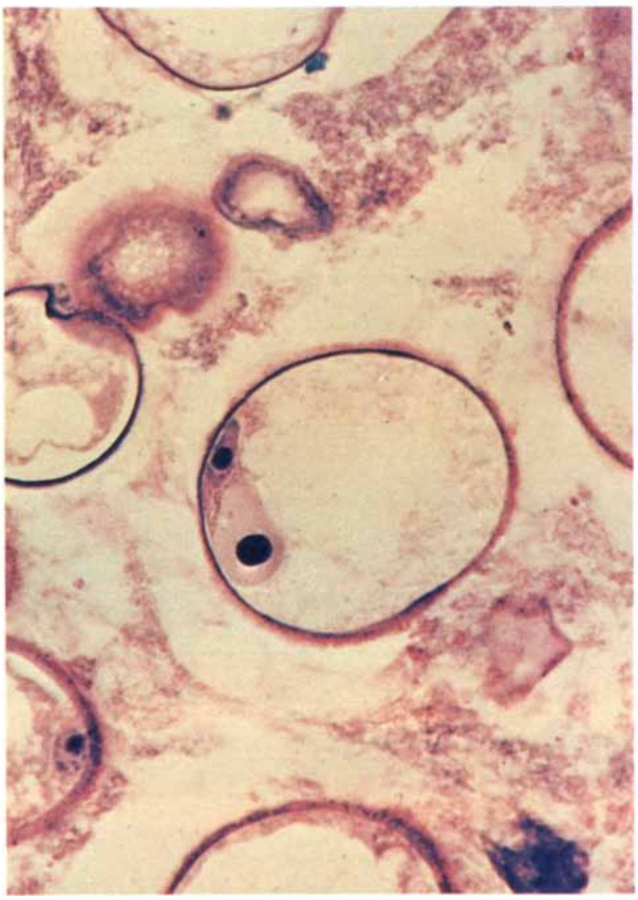
And since ahead is where we intend to stay, we are taking those mandates for smaller, smoother, faster, and the rest and making of them points of pride, "expressions of excellence." Certainly, the accomplishments our people have made possible these past 20 years are noteworthy. And if we are not dwelling on them, it is because excellence, like change itself, moves in only one direction.

MEMOREX

For twenty years, the 20 1961-1981 *expression of excellence.*

San Tomas at Central Expressway, Santa Clara, CA 95052 (408) 987-1000

COLONY OF TURTLE GRASS, one of some 50 species of sea grass that are pollinated underwater, is supported by a wave-cut platform on the coast of Kenya. The turtle grass is visible in this photograph because of low-tide conditions; at high tide it is well submerged.



TWO KINDS OF SEA-GRASS POLLEN appear in these photomicrographs. At the left are the immature grains of turtle-grass pollen, seen in a section of the anther, or male organ, of the plant. As is apparent in the central grain, these grains are at the two-cell stage of development. Wisps of the slime (*stained purple*) that will surround the pollen grains when they are ripe have begun to form. At the right are the linear strings of sea-wrack pollen; each string is a mucilaginous tube containing four pollen grains. The tubes are stained purple.

Submarine Pollination

The sea grasses, flowering below the ocean surface, must shed and capture pollen in an underwater habitat. What are the differences between their system of reproduction and that of land-based plants?

by John Pettitt, Sophie Ducker and Bruce Knox

In the life cycle of a flowering plant the events of pollination are the most important. Elaborate strategies have evolved for transferring pollen from one flower to another; some plants rely on the wind, some on bees, bats or other animals. Among certain groups of flowering plants, however, these devices do not operate. The reason is that the flowers bloom underwater. The sea grasses, which colonize coastal marine habitats in temperate and tropical regions around the world, are one such group.

How is pollination achieved underwater? On one level the answer to the question is very simple: the pollen is distributed by the movement of the water and in particular by the action of the waves. The method therefore seems to be somewhat analogous to that of windborne pollen, but on closer examination there is an intriguing complication. Among the flowering land plants a key element in successful reproduction is chemical "recognition" between molecules of protein or glycoprotein (which have sugar side chains) on the pollen grain or pollen tube and molecules of glycoprotein on the receptive surface of the stigma or stylar-cell zones of the pistil, the female organ of the flower. These molecules are water-soluble, a characteristic that is hard to reconcile with pollen recognition underwater. Do the sea grasses have a modified recognition system? It is this question that brought about our collaboration in the work summarized here.

The name sea grass is purely descriptive, as is the name seaweed with respect to marine algae; the sea grasses are not closely related to any of the land grasses. The sea grasses number 12 genera and a total of about 50 species. They are members of the order of flowering plants that includes the pondweed and frogbit families. Eelgrass and surfgrass are familiar examples in temperate regions. Sea wrack and turtle grass are widely distributed in the Tropics.

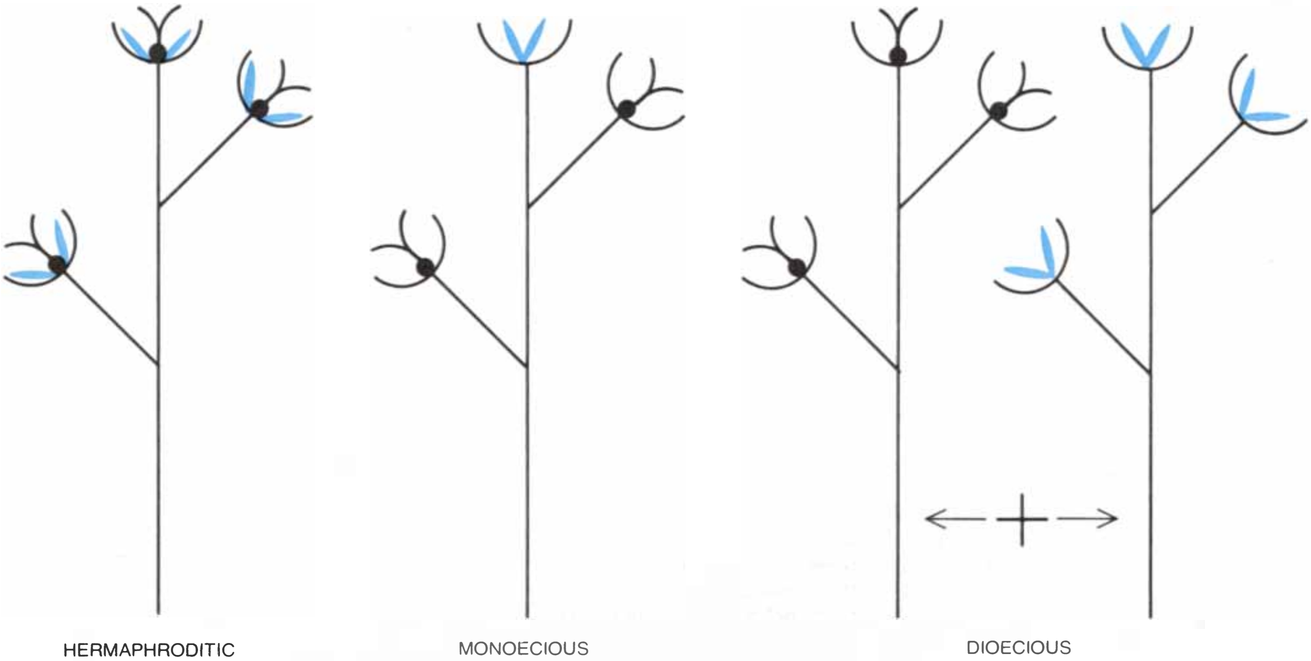
An extensive survey by Cornelis den Hartog of the University of Nijmegen has shown that the greatest concentration of sea grasses is in the Indo-Malaysian region. Eight of the 12 genera are found there. Another concentration is in the tropical New World, centered in the Caribbean. The distribution raises many questions about the plants' geographic origins, migrations and speciation. For example, turtle grass (the genus *Thalassia*) consists of two morphologically similar species that appear to be closely related and may derive from a common ancestral population. One of the species is found in the Indian Ocean and the western Pacific and the other in the Caribbean Sea and the Gulf of Mexico. The species do not interbreed. Is this because they are separated geographically or because they are not close genetically? If they were brought together, would they interbreed successfully or would they exhibit reproductive isolation? The question has not been answered for turtle grass, but it has been for another sea grass, the sea nymph (the genus *Amphibolis*). That genus also consists of two species, but there is no great distance separating them. They grow together along the southern and western coasts of Australia. A biological barrier to interspecific mating nonetheless seems to be present. Although the flowering periods of the two species can coincide or at least overlap to some extent, hybrid offspring are not observed.

With respect to reproduction, flowering plants generally fall into one or another of three categories. The vast majority are hermaphroditic, that is, each flower includes both the female pistils and the male, pollen-releasing anthers. Some plants, however, are monoecious (from the Greek for "single house"): a given flower has only pistils or only anthers but a single plant bears both kinds of flowers. Only 3 or 4 percent of all flowering plants are dioecious ("two houses"): all the flowers of some plants have only pistils and all the flowers of other plants have only anthers. The incidence of these reproductive states among the sea grasses differs markedly from that among flowering land plants. It is not a small fraction of the sea grasses that are dioecious but rather nine of the 12 genera.

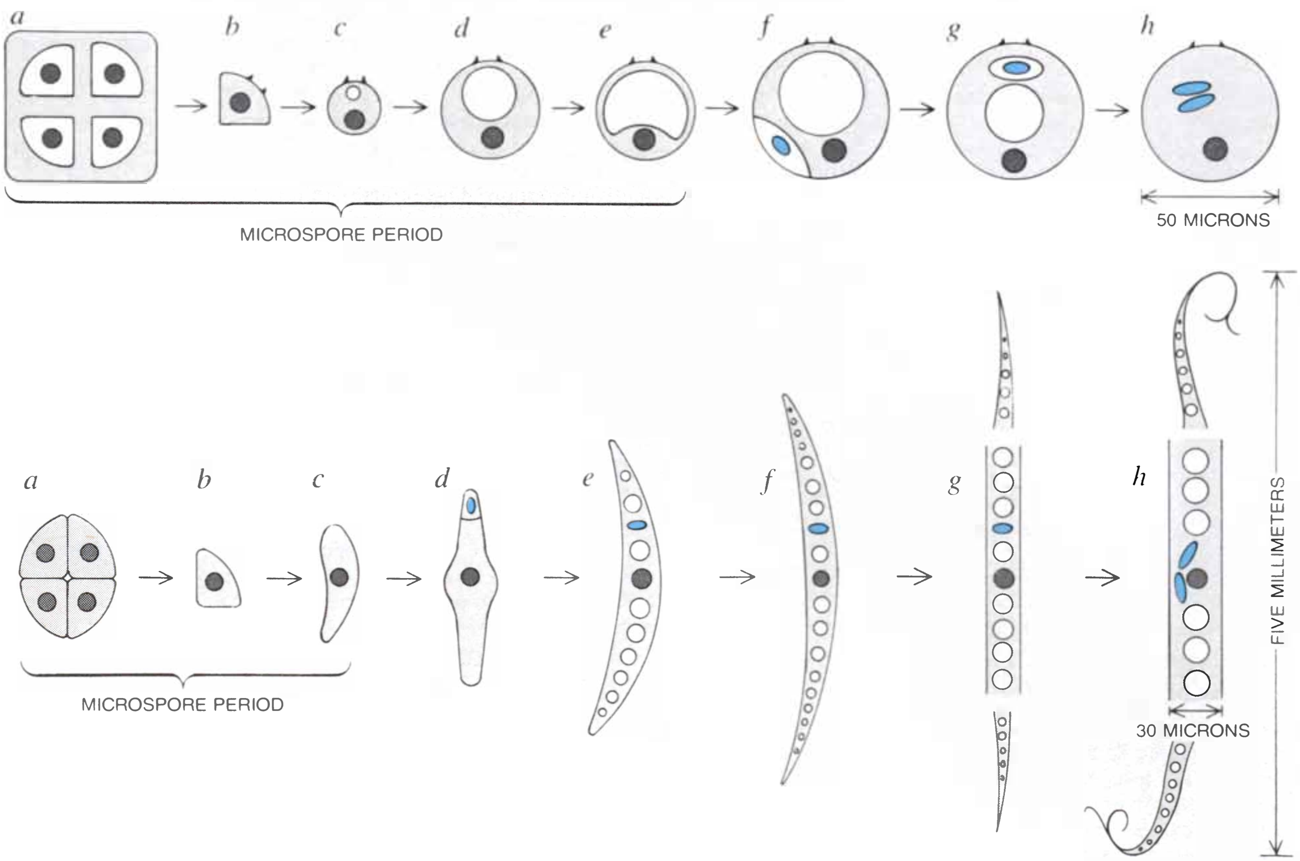
Regardless of the sexuality of the flower the sequence of events in the sexual reproduction of flowering plants is basically the same. When the flower matures, the anther opens to release pollen grains; each grain contains male gametes. The grains must then be conveyed to a mature pistil and adhere to the receptive stigma of the pistil. Recognition can take place at this time. If the grain is recognized as being compatible, it germinates and begins to produce a tubular structure. This pollen tube penetrates the cuticle of the stigma, grows through the pistil's style and eventually reaches the ovary, where the female gametes are stored. The male gametes pass through the pollen tube to the ovary, fertilization takes place and in due course seeds are set. If the pollen grain is not compatible, the pollen tube either fails to form, fails to penetrate the stigma or fails to grow through the style. In all three instances there is no union of the male and female gametes.

The flowering plants dominate land vegetation, and the sea grasses exhibit clear signs of descent from the flowering land plants. They seem to have migrated into the water in relatively recent times. They show many morphological and structural differences from terrestrial flowering plants, and some of the differences certainly have adaptive value in a submarine habitat. Nevertheless, the sea grasses still exhibit features of their terrestrial heritage, and the basic sequence of events in their reproduction is not discernibly different from that of land-based flowering plants.

The floral adaptations of flowering plants can take a variety of forms, but all of them must provide for efficient pollination, for the regulation of the level of outbreeding, for the nurture of the embryo and for the dispersal of seed in quantities sufficient for the maintenance of the species. The first of these provisions, efficient pollination, has led to the development of conspicuous flowers in those plants whose pollination depends on attracting animal visitors. Plants pol-



FLOWERING PLANTS generally bear their sexual organs in one of three arrangements. Most are hermaphroditic, as at the left; each flower has male anthers (*color*) that produce pollen and female pistils (*black*) that capture the released pollen. Some are monoecious: male (*color*) and female (*black*) flowers grow on the same plant, as in the middle. A few flowering plants, perhaps 4 percent of the total, are dioecious: each plant bears only female flowers or only male flowers, as at the right. Of the 12 sea-grass genera, 75 percent are dioecious.



POLLEN DEVELOPMENT is compared in a terrestrial flowering plant (*top*) and a submarine one (*bottom*). The production of the pollen grains of ryegrass begins with a tetrad of microspores. After release from the tetrad each microspore enlarges; the microspore period ends five stages later, and a division (*f*) forms a vegetative cell (*black*) and a sperm progenitor cell (*color*). A second division (*h*) produces the two sperm cells of the mature grain; maturation takes two weeks and the grain is then 50 micrometers in diameter. The production of the threadlike grain of sea-nymph pollen also begins with a tetrad of microspores. Three stages later, with the microspore period complete, the sperm progenitor cell (*color*) lies at one end of the elongated grain (*d*). Then it comes to occupy a central position before it divides into two sperm cells (*h*). Maturation takes 16 weeks; the grain is 30 micrometers in diameter and has reached a length of five millimeters.

linated by the wind, on the other hand, need not offer such attractions, and so their ancillary floral structures are completely or almost completely dispensed with.

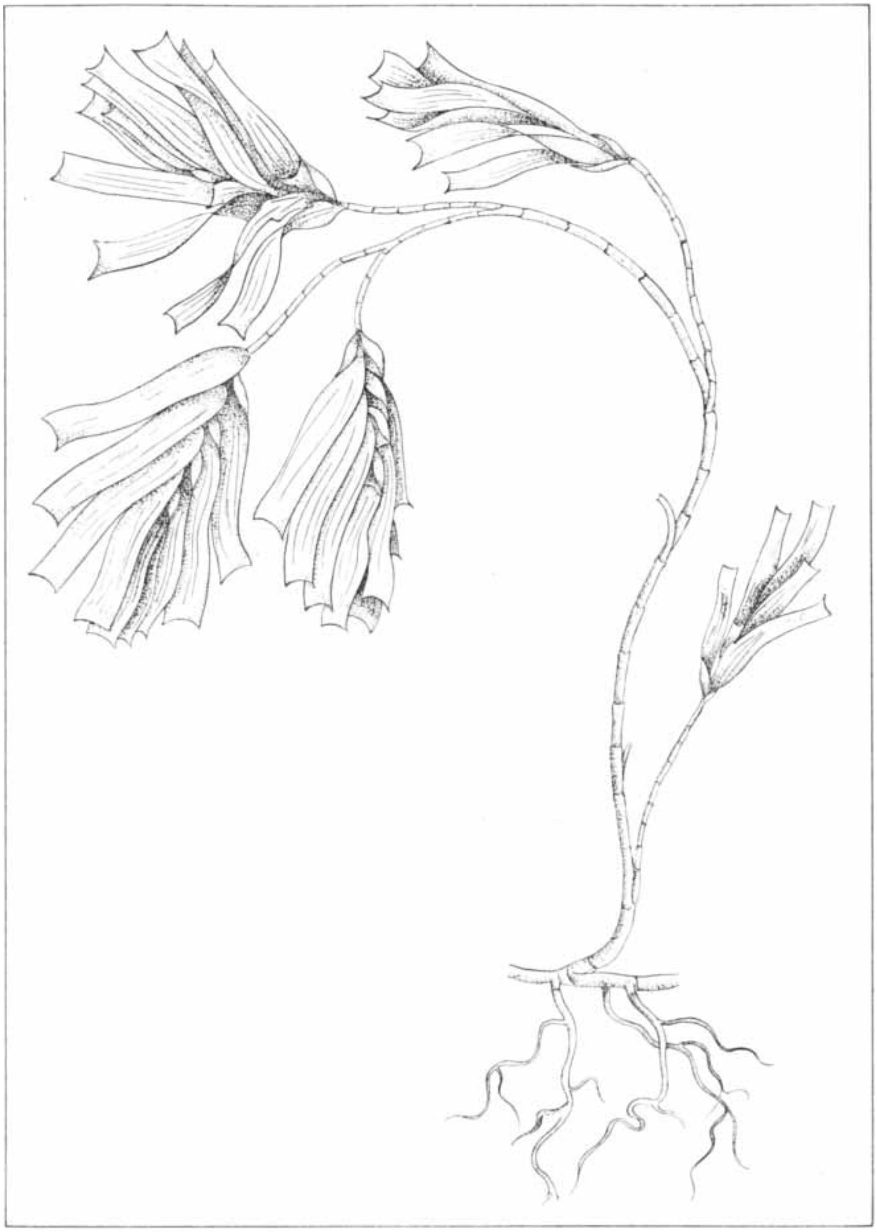
The same is true of the sea grasses. In these water-pollinated plants the attractive floral structures of an insect-pollinated plant (the colorful petals of the corolla or the sepals of the calyx) are few in number, if they are present at all, and they offer little or no display. Turtle grass is an exception: its flowers bear colored, petal-like structures that are a bright purple red when the flower opens. Whatever the reason for these showy structures, it cannot be to catch the attention of a potential animal pollinator.

Whether or not a flower is attractive, two of its vital tasks are the production and reception of pollen. In the flowers of the sea grasses the organs concerned with these functions are well developed. The mature stamens release buoyant pollen grains, and the pistils bear receptive stigmas to capture the floating grains.

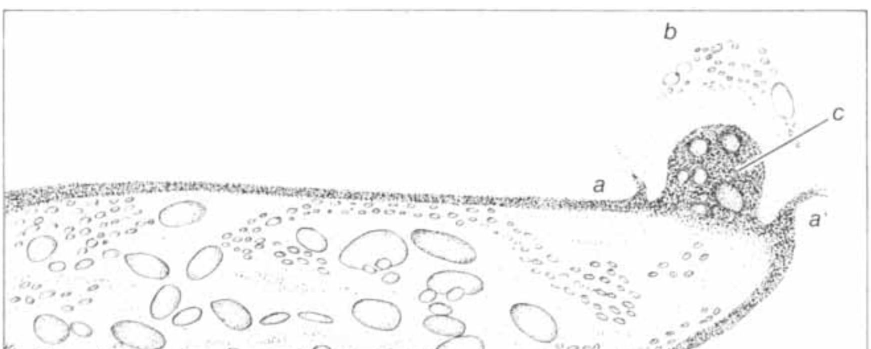
As for the second provision for reproductive success, the regulation of outbreeding, it is surely significant that most sea-grass genera are dioecious. Segregation of the sexual functions between different individuals is a most effective outbreeding strategy: it enforces the movement of pollen between plants and so yields the highest possible frequency of genetically variant progeny. At the same time dioecism is not a very efficient method of reproduction. The reason is that only half of the plants in any population, the females, contribute to seed production, a fact that bears on the last of the four reproductive provisions. The remarkably high incidence of dioecism among the sea grasses, one may presume, signifies that the gains in survival attributable to continuous outbreeding outweigh the disadvantages of diminished seed production.

Flowering in land-based plants takes place in response to changes in day length, temperature and other factors. What are the factors that stimulate the initiation and development of sea-grass flowers? At the turn of the century the Swedish botanist Nils Svedelius described the flowering of the flatweed (the genus *Enhalus*), an Indo-Pacific sea grass he had studied in Ceylon. He observed something that had apparently long been known to local fishermen: flatweed flowering is synchronized with the cycle of the tides. If the highest tides in the cycle—the spring tides that follow the new moon and the full moon—occurred in the daylight hours, flowering was coincident with them.

Four years ago we discovered a similar correlation in the reproductive behavior of turtle grass. The colonies we studied grow on a shallow platform along the coast of Kenya. During the



SEA NYMPH, a sea grass of the genus *Amphibolis*, was used in the authors' pollination experiments in Australia. The plant grows between 20 and 80 centimeters high, with upright stems.



APERTURE IN SEA-NYMPH POLLEN GRAIN, which appears after the grain is attached to a stigma, is flanked by shreds of the pollen wall (*a, a'*) dissolved by the action of an enzyme in this drawing based on a photomicrograph. Emerging through the aperture is a mucilaginous bubble (*b*); behind it a deeply staining zone of cytoplasm (*c*) presages the pollen tube.

neap tides, when the tidal variation in the water level is at a minimum, the male and female flowers begin to emerge from within the leaves that enclose them. By the start of the succeeding period of spring tides the plants are in full flower and pollen is released. Thus the plants flower when the circulation of water across the platform is at a maximum.

Each flower begins to wither a few hours after opening. The total period of flowering for all the plants in the population is only a few days. When we collected plants at the end of the spring tides, they were either producing fruit or beginning to develop flowers that would reach maturity at the next spring tides. The liberation of seed was also in synchrony with the cycle of the tides. For a population of plants growing in a fringe habitat this is an extremely effective means of promoting pollen distribution and seed dispersal.

In other sea grasses flowering appears to be regulated by the seasons rather than by the tides. For example, in the sea-nymph colonies along the coast of Victoria in Australia the first flower structures appear in early winter (May) and the flowers usually reach maturity in the period from early summer to midsummer (from mid-November to the end of December). The seeds of the sea nymph germinate while the fruit is still attached to the mother plant. On the Victoria coast the ripe fruits with their seedlings are shed during the winter months, from May to August. The flowering period of the sea-nymph colonies in Western Australia is somewhat later than it is in Victoria, which suggests that the initiation of the flower structures is regulated by day length.

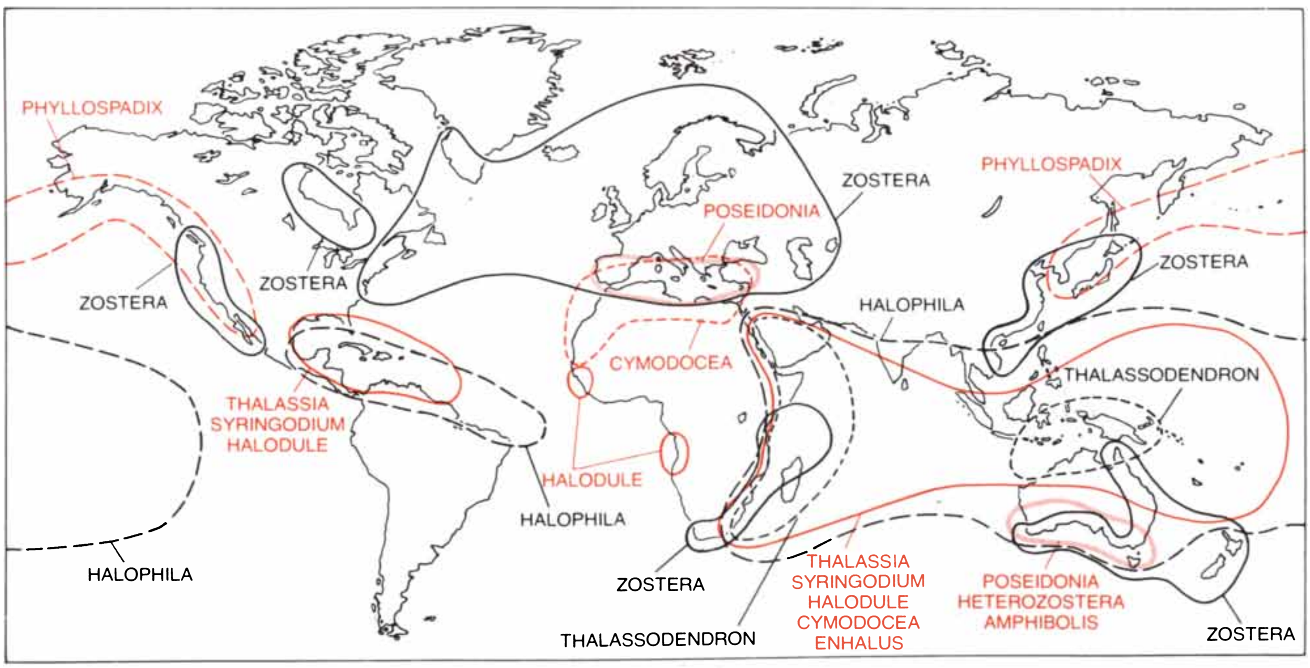
The pollen of the sea grasses has attracted the attention of naturalists at least since the end of the 18th century. It was then that Filippo Cavolini described and sketched at Naples the pollen of fiberball weed (the genus *Posidonia*). Cavolini observed that with the opening of the anthers a cloud of filaments appeared in the water. A few decades later, in 1826, the French botanist Charles Gaudichaud-Beaupré observed and sketched the similarly filiform pollen of the sea nymph in specimens collected from Shark Bay in Western Australia. A few decades after that, Edouard Bornet, another French botanist, described the development of pollen in colonies of the sea grass *Cymodocea* in Mediterranean waters.

At the beginning of the 20th century the Swedish botanist Otto Rosenberg fixed, sectioned and stained a series of anthers of eelgrass (the genus *Zostera*). He was able to trace the differentiation of the pollen from its origin at meiosis (the kind of cell division that halves the number of chromosomes) to maturity. In this genus too the pollen is filiform. The process of differentiation takes many weeks, and the mature pollen filament reaches a length of about three millimeters. The sperm cells, the male gametes carried by the pollen, are no more than five to 10 micrometers (thousandths of a millimeter) in diameter.

In the 80 years since Rosenberg's work with eelgrass pollen similar investigations have revealed the general features of pollen development in other sea grasses. Through procedures for the detection of specific cellular components it has been possible not only to resolve the structure and composition of sea-grass pollen grains in considerable detail but also to demonstrate the role of these components in reproduction.

Sea-grass pollen grains have certain special features that appear to be adaptive derivations from the pollen of terrestrial ancestors. To appreciate the extent of these adaptive changes it is useful first to consider pollination on land. The grain of pollen from a flowering land plant is usually spherical and is from 30 to 60 micrometers in diameter when it is released from the anther. Depending on the species of plant, the mature grain consists of either two or three cells. If there are only two cells, one is vegetative and the other is the gamete progenitor that divides to form two sperm cells after the pollen germinates. In a pollen grain with three cells one is again vegetative and the other two are sperm cells.

The wall of the pollen grain surrounding the sperm cells consists of two layers that differ in origin, structure and chemical composition. The outer layer is composed of the tough substance sporopollenin, which resists breakdown and is often elaborately sculptured. The inner layer is unsculptured and consists of fine fibrils of cellulose held in a matrix of polysaccharide. The material for the construction of the outer layer originates in the tapetum, the nutritive tissue of the anther; the inner layer is a product of the pollen-grain protoplasm.



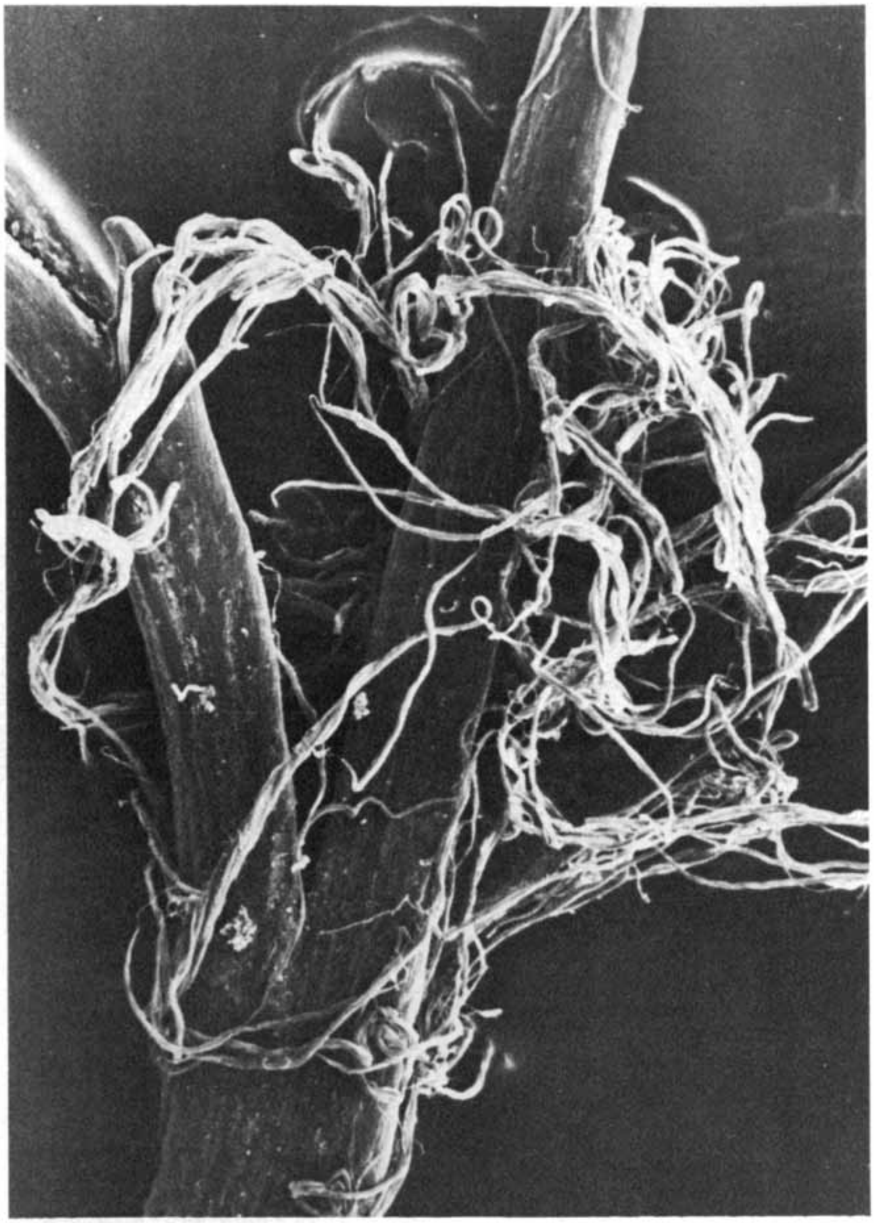
MAJOR CONCENTRATIONS OF SEA-GRASS GENERA are in the Tropics of Indo-Malaysia, where eight of 12 genera are present, and in the Caribbean, where four of the same eight grow. The cosmopolitan eelgrass, *Zostera*, has colonized as far inland as the Aral Sea.

Apertures or pores are usually present in the two-layer wall of the pollen grain; they vary widely in form, number and position according to the species. When the pollen germinates, the pollen tube emerges through one of the apertures. Certain enzymes, including acid phosphatase, are incorporated in the inner layer of the grain; their concentration is particularly high near the apertures. In pollen grains that have no apertures the pollen tube can emerge through the grain wall at any point; the enzymes are distributed evenly around the circumference of the grain in the inner layer. In either case the result is that enzymes are borne along at the tip of the pollen tube as it emerges. Enzymes are also transferred to the outer layer of the grain wall from the tapetal cells of the anther late in pollen development; such a transfer was first demonstrated in the early 1970's by John Heslop-Harrison, who was then at the Royal Botanic Gardens in Kew, and one of us (Knox). Among the transferred enzymes is esterase, which is now regarded as being characteristic of the proteins of the outer layer.

Sea-grass pollen grains are of three basic types: filiform, spherical and ellipsoidal. The filiform grains are characteristic of eelgrass, fiberball weed, sea nymph, grass wrack (the genus *Heterozostera*) and the genus *Thalassodendron*. These threadlike bodies, which are up to five millimeters long and only from 10 to 30 micrometers in diameter, are morphologically unlike the pollens of terrestrial plants. The grain lacks an outer layer of sporopollenin; the pollen wall consists only of cellulose microfibrils embedded in a polysaccharide matrix. When fresh, mature grains of filiform pollen are tested for acid phosphatase (the enzyme found in the inner layer of the pollen-grain walls of flowering terrestrial plants), the reaction is positive.

Covering the surface of the filiform pollen grains is a thin film of a material that includes carbohydrate and protein. Included in the film is the enzyme esterase (the tapetal-cell product found in the outer wall of pollen from terrestrial plants); it is deposited in the final phase of pollen maturation. Thus even though there are morphological differences between filiform sea-grass pollen and the pollen of land-based flowering plants there are also structural and biochemical similarities.

Two of the tropical and subtropical sea grasses, turtle grass and flatweed (the genus *Enhalus*), shed spherical pollen grains. Those of turtle grass are 100 micrometers in diameter and those of flatweed 150 micrometers; both are without apertures. The pollen of both plants closely resembles the pollen of terrestrial flowering plants in having distinct inner and outer layers. The enzymes in the inner layer include acid phosphatase; in the thin film that covers the outer layer esterase is present.



SEA-NYMPH STIGMAS, the branching structures seen in this scanning electron micrograph, have captured many grains of the threadlike pollen released by the male plant and directed at the female flowers by the experimenters. The structures are enlarged 20 diameters.

In the course of field studies in the Indian Ocean we noted that the pollen grains of turtle grass are released into the sea embedded in a viscid, insoluble slime; the mass of pollen resembles the spawn of amphibians, although the pollen is much smaller. Carl Julius Fritsche, a 19th-century German chemist, had drawn attention to the presence of this slime in aquatic plants. We were intrigued by the possibility that the insoluble substance might play a part in the pollination of turtle grass, perhaps in attaching pollen grains to the receptive stigmas. As we followed the development of pollen in turtle grass we found where the slime originates. The tapetal-cell protoplasm invades the anther cavity and is converted into slime in the final phase of pollen development.

In order to determine the composition of the slime we employed a centrifuge to separate the slime from the pollen. Analysis showed that the slime has a 5 percent content of carbohydrates. The sugars that make up the carbohydrate fraction were identified by chromatography. More than half of the sugar is mannose; the rest is xylose, galactose, glucose, arabinose and rhamnose. The technique of electrophoresis was employed to detect slime proteins and glycoproteins according to their molecular size. A single glycoprotein component was found.

Microscopic examination of the cen-

STIGMA AND POLLEN GRAIN of sea nymph are enlarged 15,000 diameters in this transmission electron micrograph. The dark bond between the surface film of the stigma, at the bottom, and that of the pollen grain, at the top, forms on contact and locks the two together.

trifuged pollen showed that after centrifugation a thin layer of slime still coated each grain. This finding suggests that even after wave action has broken up the pollen mass some slime still adheres to the individual pollen grains to serve as an adhesive at the surface of the stigma.

The third form of sea-grass pollen is produced by sea wrack (the genus *Halophila*). These ellipsoidal grains measure some 40 by 80 micrometers and are without apertures. The pollen is liberated from the anther in strings; each string is a mucilaginous tube that contains four pollen grains. The four grains are sister cells that stem from the same mother cell at meiosis. The grain wall has no outer layer and is of typical inner-layer construction, consisting of microfibrils in a polysaccharide matrix. Acid phosphatase is present. The mucilaginous tube is of similar construction and is coated with a thin surface film of proteins, including esterase, derived from the tapetum. The presence of the film at the surface suggests that it plays a role in the recognition and attachment of the pollen.

The capture and retention of the drifting pollen grains also depends on the receptive surface of the sea-grass stigma. In some genera the surface is smooth or ridged; in others it is covered with small projections. Both kinds of surface are covered by a thin cuticle that in turn bears a surface secretion.

The stigma secretions are directly comparable to those of certain terrestrial plants. One of us (Knox), working with John and Yolande Heslop-Harrison and Ole Mattsson at the Royal Botanic Gardens, determined several biochemical characteristics of the stigma surface layer in terrestrial flowering plants. These include the presence of esterase and an affinity for concanavalin *A*, a protein found in the jack bean (the genus *Canavalia*) that is a member of the class of substances called lectins. Concanavalin *A* is known to combine with several of the sugars found in polysaccharides and glycoproteins, among them mannose and glucose. Recently Adrienne Clarke and Paul Gleeson of the University of Melbourne have found that the animal lectin tridacnin, which is produced by the giant clam *Tridacna maxima,* also binds to the stigma surface of terrestrial flowering plants. Tridacnin too is known to have an affinity for certain sugars. The secretion from the stigma surface of terrestrial flowering plants includes a sugar component and is chemically related to plant gums. Its function may be to capture and retain pollen grains, thereby initiating the process of germination.

The surface secretions of all the sea-grass stigmas we have examined resemble those of terrestrial flowering plants in showing the presence of esterase. Moreover, the secretion of the turtle-grass stigma has an affinity for concanavalin *A*. There is, however, one significant difference. The terrestrial-plant stigma secretions are water-soluble, but the sea-grass secretion, once it is formed, does not disperse in water. The secretion retains its structural integrity and, as pollination experiments have shown, when the mating is compatible, the secretion is responsive to pollen contact.

The survey of sea-grass stigmas revealed a morphological correlation between stigma and pollen structures that might well be related to pollen capture, a process that depends on chance collision in a marine environment. Stigmas that are covered with small projections are captors either of spherical, slime-coated pollen grains, as in turtle grass and flatweed, or of strings of pollen with mucilaginous tubes, as in sea wrack. Stigmas that lack projections are invariably the captors of the threadlike filiform grains characteristic of eelgrass and sea nymph.

Among terrestrial flowering plants modification of the stigma surface drastically affects pollination. For example, one of us (Knox), working with Clarke at the University of Melbourne, found that when concanavalin *A* is applied to the stigma of the sword lily (*Gladiolus*), compatible pollination is disrupted. The stigma loses the ability to retain grains of *Gladiolus* pollen and there is no penetration of the stigma by the pollen tube.

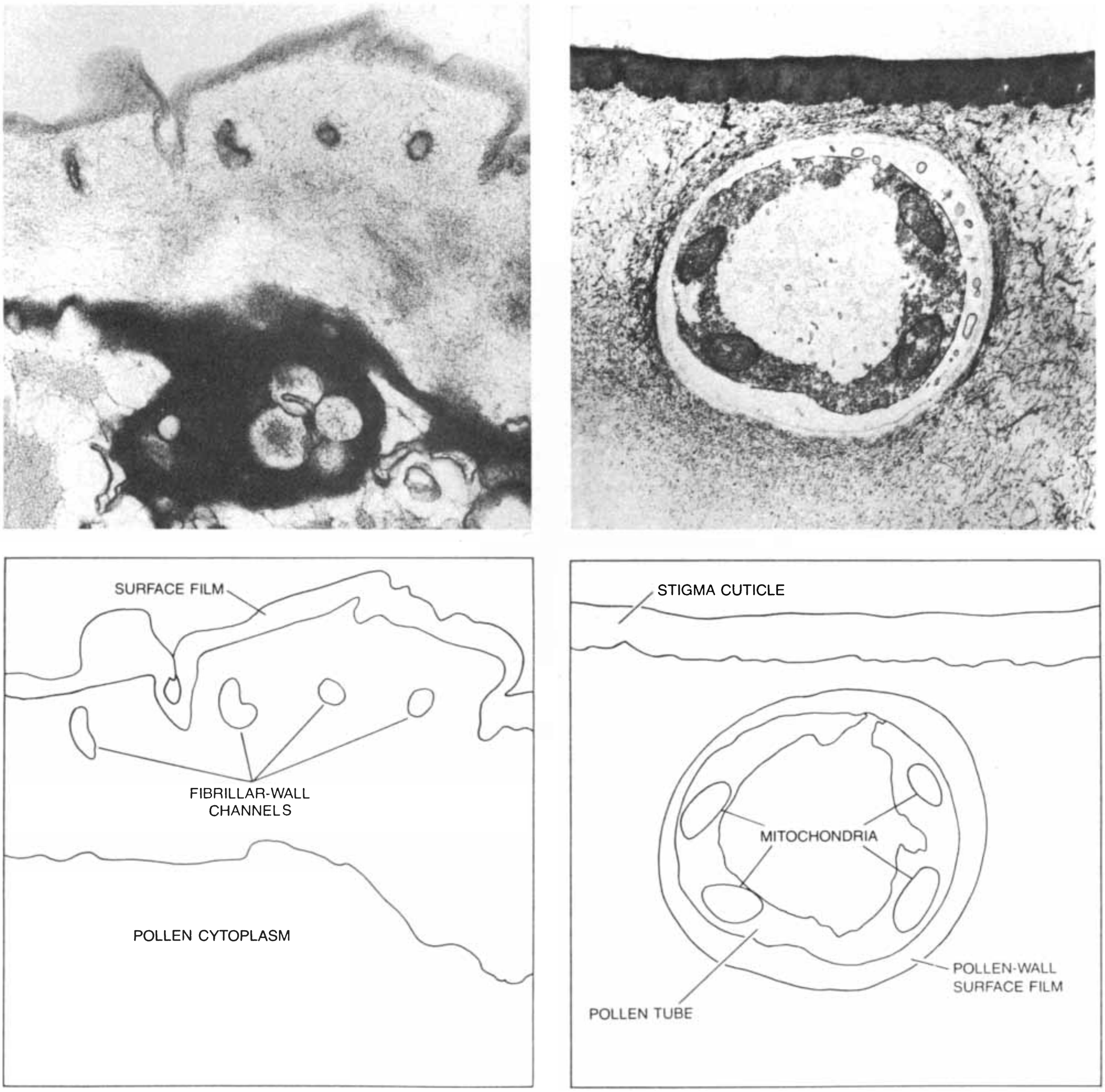
The finding suggests, first, that under normal circumstances the surface secretion of the stigma of terrestrial flowering plants is an effective adhesive for pollen. It also suggests that the surface sites on the stigma that bind to the lectin play an essential role in pollen-grain recognition. Additional evidence comes from experiments done by the Heslop-Harrisons and Hugh Dickinson with the corn cockle (*Agrostemma*) and cabbage (*Brassica*) in Dickinson's laboratory at the University of Reading. When they removed or modified the stigma secretion with enzymes, the normal pollination process was frustrated.

Thus the question arises: Is the capture, retention and germination of pollen different in the marine domain? Working together at the University of

Melbourne, we set up laboratory cultures of the sea nymph in order to study the pollination process, collecting male and female plants from colonies in Bass Strait and transferring them to seawater aquariums. Pollen released by the male plants was taken up and directed at the submerged stigmas of the female plants. Most of the pollen failed to encounter a stigma, but all the grains that did come in contact immediately adhered and were retained.

At intervals pollinated flowers were removed from the female plants and the stigmas were prepared for microscopic examination. Thin sections were cut through the pollen-stigma interface and examined by electron microscopy. They showed that the pollen grain was firmly attached to the stigma by a meniscus of adhesive. The attachment evidently formed at the moment of contact by the coalescence of the two surface materials; immediately on touching the stigma secretion, the film of carbohydrate and protein on the surface of the pollen grain combined with it to form a bond. The bonding action seems comparable to that produced by certain manmade adhesives, which consist of two components that set after mixing.

Such a two-component adhesive may provide the sea nymph with the means of discriminating between pollen of the same species of sea grass and foreign pollen. If adhesion depends on the interaction of specific large molecules on the surfaces in contact, then foreign pollen carrying a surface film with unsuitable



MICROSTRUCTURES OF GERMINATION appear in these transmission electron micrographs, which enlarge the structures respectively 50,000 and 18,000 diameters. The micrograph at the left shows small channels in the fibrillar wall of grass-wrack pollen. They may serve as conduits for the passage from the stigma to the pollen cytoplasm of substances that could stimulate germination. Also visible is the thin film on the surface of the pollen wall. The micrograph at the right shows a cross section of a pollen tube that is growing toward the style tissue of a pistil after penetrating the wall of a stigma receptive cell. Above the pollen tube is the dense stigma cuticle that the tube has penetrated at a point nearby. Key structures in both micrographs are identified in the drawings that appear below them.

molecular characteristics would simply not adhere to the stigma.

An earlier experiment had shown that viable sea-nymph pollen fails to germinate when it is incubated by itself in vessels of seawater. The finding is of particular relevance in interpreting pollination experiments because one can conclude that contact with the stigma is needed to stimulate pollen germination. What form does the stimulus take? In terrestrial flowering plants the stimulus is hydration: when the dehydrated pollen grain alights, it takes up water from the stigma. The moistened grain then changes from a dormant state to an active one and the pollen tube emerges.

The hypothesis that hydration similarly triggers germination in a submarine environment faces a number of difficulties. For example, it is doubtful that the degree of dehydration of the submerged pollen grain can even approach the degree found in terrestrial pollen. Of course the water content of a sea-grass pollen grain may still be less than that of the corresponding stigma.

A possible solution to these difficulties is provided by the work of A. W. A. M. de Cock at the University of Nijmegen. Investigating eelgrass pollen, de Cock found that the development of the pollen tube is enhanced if an extract from the crushed pistil of the plant is added together with sugars to a seawater medium. He has suggested that an organic molecule in the pistil extract (a molecule that in normal circumstances presumably diffuses from the stigma) activates pollen-tube development. The conjecture has intriguing implications. If such a molecule is instrumental in promoting growth in pollen of the same species, it could also play a role in species recognition in mating.

If molecules derived from the stigma pass to the pollen grain to trigger germination in sea grasses, presumably they must pass through the wall of the grain. How they do so would depend on their size. A substance of low molecular weight might be transmitted across the wall by diffusion, whereas special conduits might facilitate the passage of larger molecules. We have recently discovered such conduits. They traverse the wall of grass-wrack pollen grains and are some 100 nanometers (thousandths of a micrometer) in diameter.

When we observed the emergence of the sea-nymph pollen tube, we found something quite unexpected. In the pollen of terrestrial flowering plants the inner layer of the grain becomes extended to form the first part of the pollen tube. In the sea-nymph pollen grain, however, a region of the wall is dissolved away, apparently by enzymatic action. Under the microscope the incipient aperture can be detected as a small area of the grain wall that loses its affinity for certain stains. The area is near the point of attachment between the pollen grain and the stigma but never actually at the point. This indicates that the enzymes responsible for dissolving the wall originate in the pollen grain rather than in the stigma.

In sea-nymph pollen the emergence of a mucilaginous bubble through the aperture in the wall of the pollen grain indicates that germination has begun. Under the bubble is a zone of densely staining pollen cytoplasm that is delimited by the cytoplasm's outer membrane. The cytoplasm passes through the aperture to establish the pollen tube proper. It is not yet known how the materials for the tube wall are formed, but it appears that the polysaccharide components of the wall are synthesized at the front of the emerging cytoplasm.

The development of the pollen tube in the sea nymph differs from that in terrestrial flowering plants. In all known cases on land the developing tube is structurally continuous with the microfibrillar inner layer of the pollen-grain wall. An exception to this generalization has recently been discovered in the initial stages of pollen germination in the avocado (the genus *Persea*). Margaret Sedgley of the Australian Commonwealth Scientific and Industrial Research Organization in Adelaide has observed that the pollen tubes of this plant first consist only of the outer cytoplasmic membrane; synthesis of the tube-wall materials is delayed until sometime after the tube has made contact with the stigma. This apparent developmental similarity between a sea grass and the avocado could prove to be fortuitous; nevertheless, the existence of the same kind of phenomenon in such different plant species encourages attempts to find a functional explanation.

When our joint investigations began, the interaction of the pollen grain with the stigma had not been observed in any of the sea grasses. It was obvious that our studies might reveal some extreme structural and physiological departures from the terrestrial themes, but it was not unreasonable to expect that the general principles of the pollen-stigma interaction might be similar in the sea and on land. As has been shown here, that is the case.

What, then, are the special physiological provisions for submarine pollination? To answer the question one needs to know the precise function of the pollen mucilage. Does it serve merely as an adhesive? Does it regulate the water balance of the pollen grain? Or does it serve both of these purposes and also, as the English botanist Agnes Arber suggested more than half a century ago, prevent soluble substances (which would include recognition molecules) from diffusing into the water?

Educator, Entertainer, Accountant.

Your Challenger Personal Computer.

Through the miracle of modern technology, a complete computer as powerful as the multimillion dollar room-sized computers of a few years ago can be put in a package the size of a typewriter and sells for as little as a color television set!

Through its years of microcomputer experience, Ohio Scientific has effectively channeled this tremendous computer power into a "friendly" computer with hundreds of personal uses, via a huge software library of programs for a broad range of personal, home, educational and business use.

This available software allows you to use and enjoy your computer without becoming an expert. The Challenger, however, is a powerful, general purpose computer which can be programmed in several languages by those who choose to.

Here are just a few of the popular uses of an Ohio Scientific Challenger Computer:

Education

The personal computer is the ultimate educational aid because it can entertain while it educates. Software available ranges from enhancing your children's basic math, reading and spelling ability, through tutoring high school and college subjects, to teaching the fundamentals of computers and computer programming.

Entertainment

Many of the Challenger's games educate while they entertain, from cartoons for preschoolers to games which sharpen mathematical and logical abilities. But, entertainment doesn't stop here. The Challenger's graphics capabilities and fast operation allow it to display action games with much more detail than the best video games, providing spectacular action in games such as Invaders, Space Wars, Tiger Tank and more! All popular sports such as golf, baseball and bowling are available as simulated computer games as well as many conventional games such as chess where the computer plays the role of a formidable opponent.

Accounting

Your Challenger computer can keep track of your checkbook, savings account, loans, expenses, monitor your calorie intake and your biorythms.

If you are involved in a business, you can use it to do word processing; accounting, inventory control, order processing, customer lists, client records, mailing labels and planning.

And more:

This may seem like a lot of uses, but it's only the tip of the iceberg for a general purpose computer. For example, your Challenger can be expanded to control lights and appliances, manage your energy usage and monitor for fire and break-ins. Furthermore, it can communicate with you, with other computers and the new personal computer information services over the telephone.

In fact, the uses of general purpose, personalized computers are expanding daily as more and more people discover the tremendous capabilities of these new technological wonders.

Ohio Scientific offers you four personalized computer systems starting at just $479.

Anorthoscopic Perception

A figure revealed one section at a time through a narrow slit may be perceived in its entirety. Is this effect based on the "painting" of the image over the retina or on deeper psychological events?

by Irvin Rock

How one visually perceives the shape of things is an old topic in experimental psychology, but it is still not well understood. In some ways the eye is like a camera, in that an object of a certain shape gives rise to an image of the same shape on the surface of the retina. The simple analogy breaks down, however, in situations where the presence of such an image is neither sufficient nor necessary to account for the perceived shape. In order to learn more about the perceptual process investigators over the years have devised a variety of laboratory techniques. My colleagues and I at the Institute for Cognitive Studies of Rutgers University have recently concentrated on a method that calls for viewing a figure in successive sections through a narrow slit. In recognition of the unusual nature of this experimental arrangement we have adopted the 19th-century term anorthoscopic ("abnormally viewed") to characterize perception in such cases.

One of the first problems to arise with the eye-as-camera analogy is that it does not take account of the question of how the components of the retinal image are organized. Consider an array of three dots that are not in a line. They will be seen as the corner points of a triangle only if they are grouped into one unit by the perceptual system; in the presence of other dots their triangularity may be less evident. Even a well-known figure composed of lines may not be recognized if the perceptual system associates the lines with other lines to form different groupings. Then there is the matter of organizing a pattern into figure and ground: a familiar shape may not be perceived—even though an image of that shape is present on the retina—if some other part of the field of view is taken to be the figure.

Another problem is that of orientation. If one misperceives where the top, bottom and sides of a figure are, the mere presence of its image on the retina does not ensure that the figure will be recognized. The orientation of a figure in the third dimension can also be a factor in the perception of form. For example, a circle viewed obliquely projects an elliptical image on the retina, but it may nonetheless appear to be circular under the right conditions.

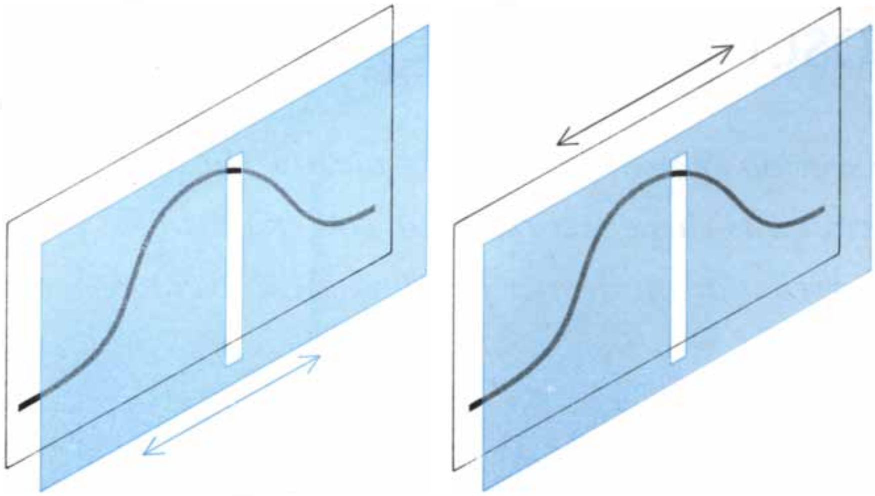
All these examples indicate that the existence of an accurate retinal image is not a sufficient explanation of the perception of a form. Other factors must be invoked. Indeed, an accurate retinal image is sometimes not a necessary factor. It can easily be shown that if only parts of a figure are present, other parts may be perceived as if they too were present. Consider a simple line drawing of two rectangles, one of which appears to overlap and occlude the other; the figure may be seen as two complete rectangles even though the image of one is incomplete. In other cases contours (called subjective or illusory contours) are seen in spite of the absence in the retinal image of the abrupt transition from light to dark that is usually considered the stimulus for a perceived contour [see "Subjective Contours," by Gaetano Kanizsa; Scientific American, April, 1976].

In the last two examples at least some of the contours of the figure are present in the retinal image. In other experiments one can perceive a form without any contour in the retinal image. With the aid of a stereoscopic viewer a random distribution of graphic elements can be presented to each eye separately in such a way that no figure is visible with either eye alone. If a cluster of the elements in the display presented to one eye is shifted horizontally with respect to the same cluster in the display presented to the other eye, however, the perceptual system may detect the similarity of the two clusters and infer that there is a shape either in front of the rest of the display or behind it. Here the perceived shape emerges not from retinal-image contours present in each eye but from the boundaries of a region the perceptual system infers is an entity in the scene [see "Experiments in the Visual Perception of Texture," by Bela Julesz; Scientific American, April, 1975].

Another example has a closer connection with my own recent work. Suppose a luminous point traverses a path in a dark room and an observer is instructed to track it with his eyes. The image of the point would always fall at about the same place on the retina (namely the fovea), but the viewer would nonetheless perceive the shape of the path. This impression would be very close to the perception of a form, and yet there would be no extended image on the retina corresponding to the path of the point. When Fred Halper and I tried this simple experiment in our laboratory at Rutgers, we found that observers do as well in apprehending the shape of such a path as they do when they keep their eyes still, so that the image of the point does move across the retina.

One might still argue that the perception of a path is not the same as the perception of a shape. After all, one sees a contour in the latter case, not just a moving point. With a small change in the experimental procedure, however, there is a way to achieve a strong impression of a shape. Instead of a dark room and a moving point, the observer is presented with a narrow slit in an opaque surface moving across a line drawing of a figure [*see illustration on next page*]. If the observer tracks the slit closely, the slit stimulates only a narrow strip on the retina and the image of the visible portion of the figure accordingly moves up and down within this retinal strip. Alternatively, the figure can be moved back and forth behind a stationary slit; if the eyes of the observer in this case remain stationary, fixating the slit, the slit again stimulates only a narrow strip on the retina. At no time in either case is there an image on the retina that resembles the shape of the figure. Nevertheless, under these conditions we find that figures are easily perceived.

In the 19th century several investigators, including Joseph Plateau, Fritz Zöllner and Hermann von Helmholtz,



ANORTHOSCOPIC PRESENTATION is an abnormal method of visual stimulation adopted by the author and his colleagues to test the retinal-painting hypothesis, which holds that an image is spread over the retina one area at a time. The technique calls for viewing successive sections of a figure through a narrow slit. Two experimental conditions are possible: either a slit moves in front of a stationary figure (*left*) or a figure moves behind a stationary slit (*right*).

employed a similar arrangement to study certain visual phenomena. They referred to their procedure as anorthoscopic presentation, presumably because it was a nonstandard method of visual stimulation. The experimental device called the anorthoscope consisted of two disks mounted axially one behind the other. The disks could rotate at different speeds and in opposite directions. The observer viewed figures on the rear disk through slits in the front disk, thereby perceiving certain illusory distortions of the figures.

This early work was apparently forgotten until 1965, when it was rediscovered by T. E. Parks of the University of California at Davis. Both in the 19th century and after the rediscovery a controversy arose over how to interpret the effect. Helmholtz and others thought an extended image of the figure behind the slit is always imprinted on the retina. This interpretation has since come to be known as the "retinal painting" hypothesis, since it is based on the assumption that an image is spread over the retina one area at a time. Where the slit moves in front of a stationary figure, retinal painting would require the observer to keep his eyes stationary and not track the slit. Where the figure moves behind a stationary slit retinal painting would require the observer to move his eyes back and forth. Helmholtz argued that the observer might not always know whether his eyes were still or moving.

If the retinal-painting hypothesis is correct, the perception of form under such conditions would hardly be surprising. An extended retinal image of a figure would then be present in both of the cases cited above, and the only difference between perception in the abnormal viewing situation and in the normal situation would be that in anorthoscopic viewing the image would be created in segments presented successively rather than simultaneously. Moreover, because the sensory cells of the retina tend to continue discharging for a short time after the stimulation has ended (a phenomenon known as neural persistence), the image presented in successive segments would not be significantly different from an image formed all at once, assuming that the successive presentations were fast enough. The observer would therefore be expected to perceive the entire anorthoscopic image as if it had been established normally.

To evaluate the retinal-painting hypothesis one needs to know exactly what the eyes of the observer are doing during an anorthoscopic presentation. In our experiments, therefore, we usually make videotape recordings of the observer's eyes during the presentation. In one test Joseph Di Vita and I had observers view curved-line figures that were approximately five and a half inches long through a slit an eighth of an inch wide from a distance of two and a half feet. Either the figure moved back and forth through five cycles or the slit did. The speed of the figure or the slit was one pass (that is, one half cycle) per second, which was lower than the speed adopted by other investigators of anorthoscopic perception. Eye recordings indicated at every moment the direction of gaze with respect to the display. The observer was told to press a switch only during the periods when a figure was perceived behind the slit. The signal was recorded with the videotaped record of eye movement.

When the slit was stationary, the data showed that by and large observers kept their eyes still whether or not they were told to do so and whether or not they were given a point on the slit to fixate. Of course, the eyes did not remain perfectly stationary, but the occasional small, rapid motions to one side or the other (known as saccadic eye movements) were not sufficient to account for the perception of form. The observers did not always perceive a figure, but there were many times when they did and when their eyes were more or less stationary.

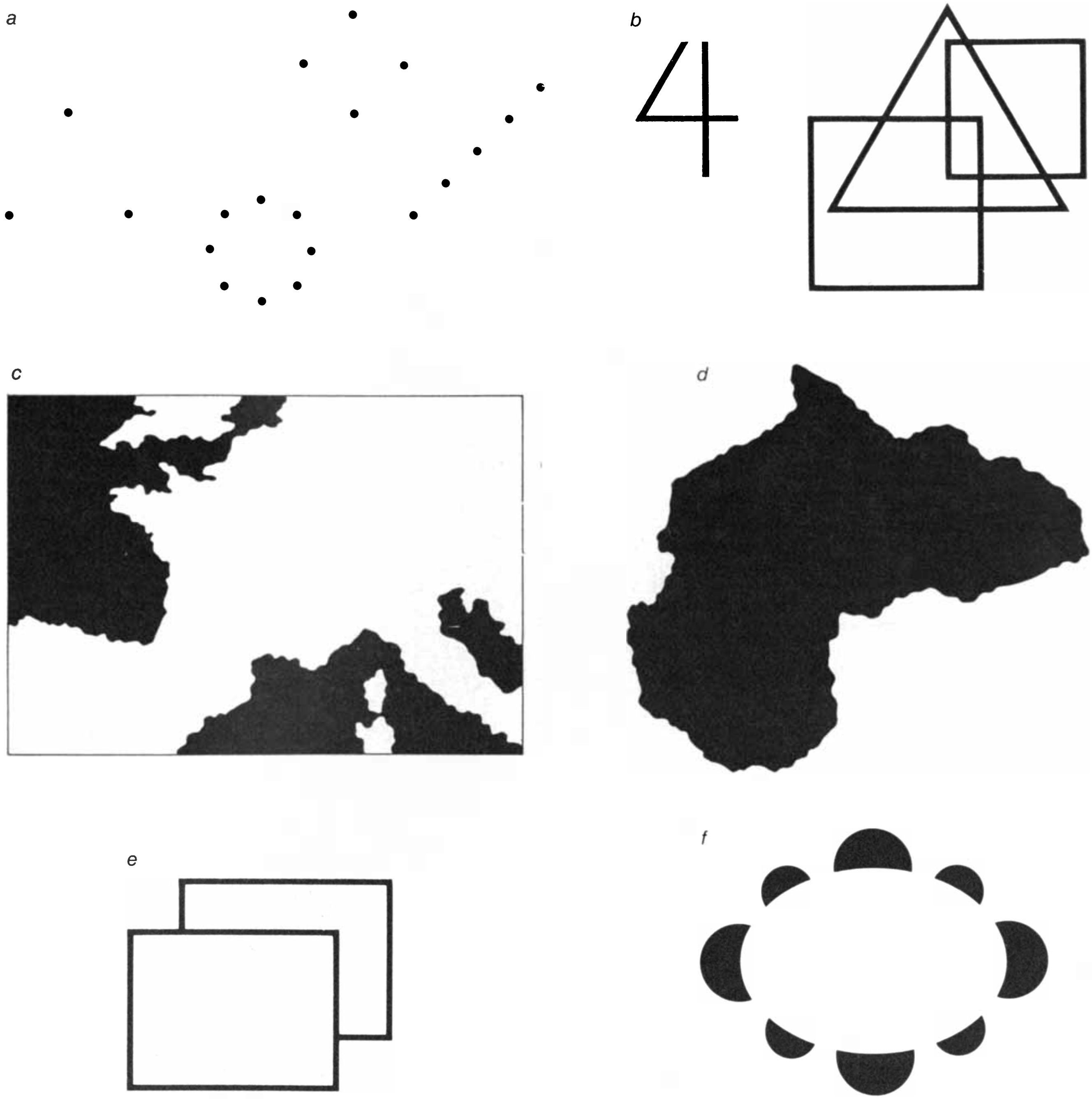
These findings were recently confirmed by Robert Fendrich and Arien Mack of the New School for Social Research. In their experiment the image of the display was stabilized in such a way that even if the eyes moved, the image of the slit would not be displaced across the retina. Nevertheless, their observers also perceived a figure moving behind the slit.

When the slit in our experiments moved over a stationary figure, the observers typically moved their eyes along with the slit. There was usually a lag, however, with the slit quickly getting a little ahead of the eyes and remaining ahead until it reached its terminal position, whereupon the eyes quickly caught up. It was nonetheless clear that the image of the slit did not move far enough over the retina to explain the perception of form; indeed, the retinal image did not even move consistently in the right direction. In this situation a figure was almost always perceived, and its size was perceived more or less correctly.

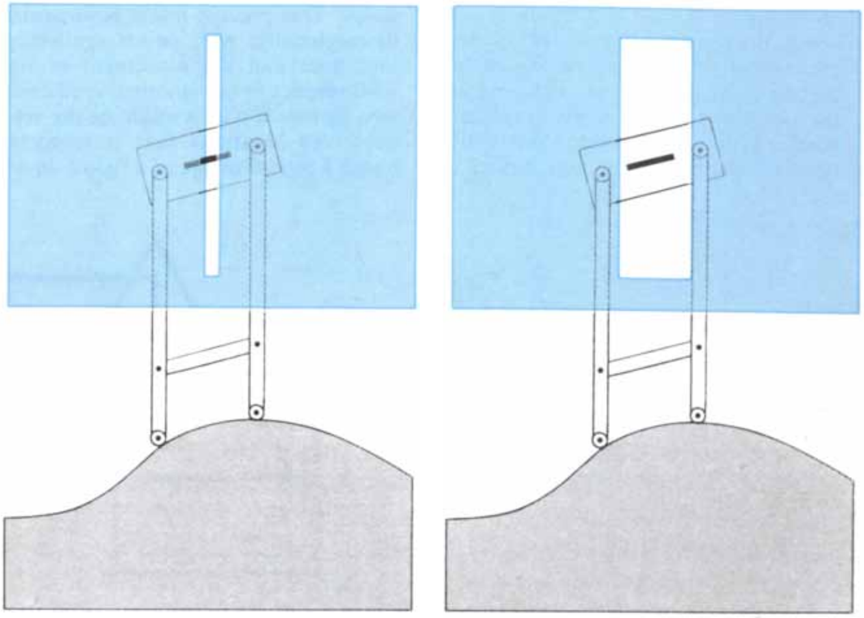
Other experiments done in our laboratory and elsewhere indicate that the only way the anorthoscopic effect could result from retinal painting would be if the image of the slit were to move over the retina very quickly. Only at a speed approaching five sweeps per second, or five times the speed we typically employ, could neural persistence account for the impression of a whole figure. Moreover, in the moving-figure condition retinal painting would have to be based on eye movements, and therefore the eyes would have to move back and forth at a rapid rate. Observers are unable to move their eyes that fast.

On the basis of this and other evidence we conclude that retinal painting is not a sufficiently general explanation of anorthoscopic perception. What then is the basis of form perception in experiments of this kind? Consider again the tracking of a luminous point in a dark room. Although the image of the point does not move over the retina, the perceptual system does receive information about the path of the point, because the eyes must move to track it. The perceptual system takes into account the position of the eyes in judging, from the retinal stimulus, the location of the point in space. Thus the perceived path of the point is derived from the set of perceived locations of the point as it moves.

Generalizing this argument to all form perception, one can say that when an extended image of a figure is present on the retina, the observer perceives each point of the figure as being in a certain place. Hence one might suppose the perceived shape is the result of a mental synthesis of the perceived locations of all the points constituting the shape. This process might be organized hierarchically, with points synthesized into lines and the placement of lines with respect to one another synthesized into figures. If it is, it might be the set of perceived locations that is important about a retinal image of a figure. In oth-



LIMITED ROLE OF THE RETINAL IMAGE in the visual perception of form is illustrated by means of six examples. In *a* the three dots at the left are seen as the corner points of a triangle only if they are organized into one unit by the perceptual system; in the presence of other points (*right*) the triangle may not be perceived. In *b* the numeral 4 (*left*) will not be seen as such if the perceptual system associates its constituent lines with other lines to form different groupings (*right*). In *c* the familiar shape of western Europe will not be perceived if the background (in this case the surrounding bodies of water) is taken to be the figure. Similarly in *d* the continent of Africa may not be recognized if one is not aware that the top of the figure (that is, north) is to the left. These four examples show that the mere presence of a particular retinal image is insufficient to account for what is perceived. In *e* two rectangular shapes, one shape overlapping and occluding the other, may be seen as two complete rectangles even though one is actually incomplete. In *f* a subjective (or illusory) contour is readily perceived even in the absence of the abrupt light-dark transition that is usually considered the stimulus for the perception of contours. The last two examples reveal that an accurate or complete retinal image is not necessary to explain what is perceived.



SIMULATION TECHNIQUE was devised by the author and one of his colleagues, Alan L. Gilchrist, to investigate the phenomenon of anorthoscopic perception. Here the technique is used to verify that for a figure to be perceived under anorthoscopic conditions the visible portion of the figure must be seen to extend across the slit from one side to the other. A small line segment drawn on a piece of transparent plastic is attached by thin rods to a set of rollers that ride on a cam not visible through the slit. The visible part of the line segment moves up and down and simultaneously changes its slope just as the visible part of an actual line figure would. When the slit is narrower than the length of the line segment, a figure is often perceived (*left*). When the slit is wider than the line segment, however, the figure is never perceived (*right*).

er words, perhaps the role of the retinal contours of a figure is to establish a set of locations that together yield a form.

The same argument can be applied to anorthoscopic perception. The perceptual system integrates the information about the locations of the parts of the figure as they appear successively through the slit. When the slit moves over a stationary figure, the visible element at one end of the figure is characterized as, say, "at the bottom of the slit and to the left," whereas a moment later the visible element is characterized as "higher in the slit and straight ahead." Integrating the directions leads to the perception of an oblique line sloping upward to the right. Therefore both anorthoscopic perception and normal form perception can be said to entail an integration of all given locations of parts of a figure with respect to one another into a description of the whole.

This explanation turns out to be incomplete for several reasons. First, direction with respect to the observer is given directly in the moving-slit condition but not in the moving-figure condition. In the latter case the slit remains stationary and straight ahead. Where then does the information come from that the successively revealed parts of the figure are in different places? Apparently it must first be hypothesized that something is moving behind the slit. Once that idea is entertained, the temporal succession of stimuli can be converted into a spatial configuration. Next, as in the moving-slit condition, the perceptual system can integrate the sequence of stimuli into a shape based on the set of directions of the parts with respect to one another.

Our analysis suggests another reason the location-synthesis hypothesis does not tell the whole story. Often observers do not perceive a figure in the anorthoscopic display. This result must be related in some way to the kind of figure we employ, because other investigators have not reported such failures. We have deliberately chosen figures with several distinctive features: they are single lines and are shaped so that only one figure element at a time is visible in the slit; they are fairly smooth curves rather than patterns with abrupt changes in direction; they are continuous lines rather than dotted or broken; they do not depict familiar objects. Others have employed figures such as squares, circles and ellipses or outline drawings of familiar objects such as animals.

We consider our kind of figure to be ambiguous in that the anorthoscopic display of it could logically represent either an element moving vertically in a slit or an extended figure being revealed through the slit. Therefore we reasoned that we could investigate what factors are important in figure perception, since a figure is not always perceived. Our guiding hypothesis has been that figure perception under anorthoscopic conditions is based on an unconscious process of problem solving. Even before any location-synthesis activity begins, the perceptual system must hypothesize that there is an extended figure behind the slit and that it is being revealed in successive sections.

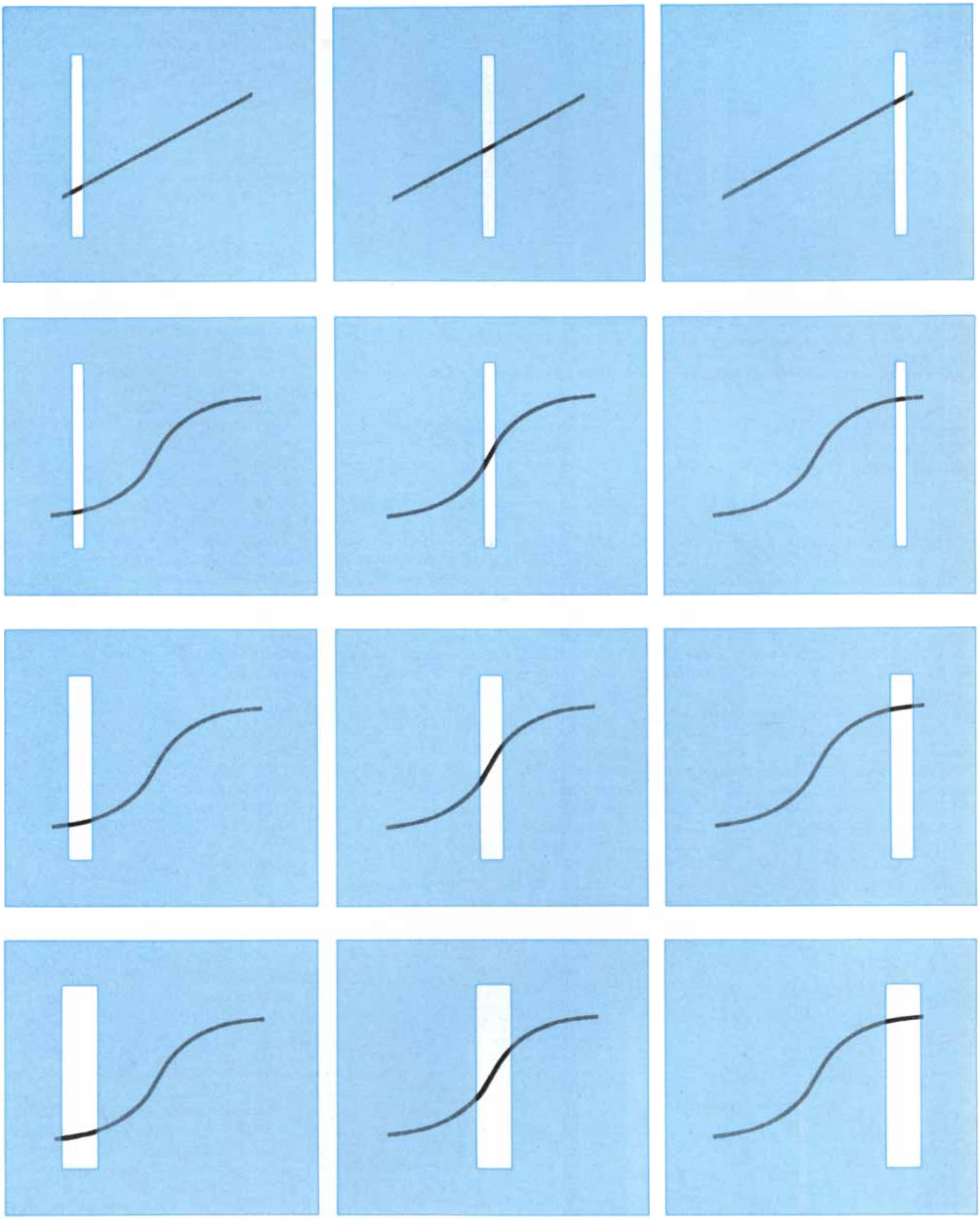
To clarify this hypothesis consider what would be seen if the figure were made luminous and were viewed in a dark room, so that the slit was not visible. In the stationary-slit condition all the observer would see would be a point of light moving up and down. In the moving-slit condition the observer would see a moving point of light traversing a path. In neither case would one describe the percept (the perceived entity) as an extended figure. Thus it seems the observer must realize that the momentary view is through a narrow slit and that the remainder of the figure is hidden by an opaque surface. Unlike problem solving as an act of thought, problem solving in perception requires support from the stimulus. To perceive a figure anorthoscopically it is not enough to know one is looking through a slit; the slit must be seen.

Under certain conditions the observer may even "invent" a slit when none is visible. Fendrich and Mack have found that if a circle is moved behind an invisible slit and the appropriate suggestion is made, observers who at first see the visible parts of the circle moving vertically later perceive a figure moving horizontally behind an illusory slit. Their experiment provides a good example of how the perceptual system tends to rationalize a solution adopted by the problem-solving faculty if the stimulus does not fully support it.

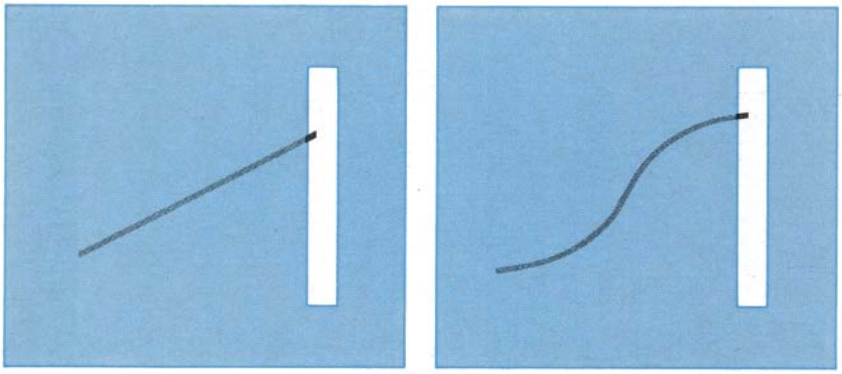
Not only must the slit be seen; the visible portion of the figure must also extend fully across it. To clarify this point Alan L. Gilchrist and I devised a technique for stimulating anorthoscopic presentation. Instead of a figure behind the slit, there was only a small line segment drawn on transparent plastic. The plastic was attached by thin rods to a set of rollers that rode on a cam not visible through the slit; the cam had the shape of the figure to be simulated [*see illustration on this page*].

Under some conditions simulation with this instrument leads to figure perception. In this instance, however, we widened the slit so that the visible line segment did not fill it but rather was isolated within it. Although the segment moved vertically and changed its slope just as it had when the slit was narrower, a figure was never perceived. The stimulus apparently violated the logic of the solution: that one was viewing a continuous figure being revealed through the slit.

Consider next the kind of figure dis-



STRONG TENDENCY to perceive an anorthoscopically displayed line figure as a line segment undergoing vertical displacement rather than as a horizontally revealed continuous line is demonstrated in four display sequences, each sequence representing three successive times. When an oblique straight line is displayed anorthoscopically (*top*), it never appears to be an extended, moving figure, presumably because there is no hint of any change in the segment's identity. When a curved-line figure is presented under the same conditions (*second from top*), the visible segment accelerates and decelerates as it moves vertically within the slit, but this additional information does not seem to change the perception. Even when the width of the slit is doubled, so that the slope and the change in the slope of the visible segment are detectable (*third from top*), most observers still do not perceive a figure. Only when the slit is widened still further, so that the curvature and the change in the curvature become detectable (*bottom*), is a figure perceived. All the figures were viewed from about three feet away.



APPEARANCE OF THE END POINTS of a line in the slit of an anorthoscopic display may aid in the perception of a shape, but this factor by itself does not seem to be adequate to shift the balance in favor of figure perception. An oblique straight line still tends to be perceived as a vertically displacing line segment even when its end points are visible (*left*). A curved line with its end points visible, on the other hand, is more likely to be seen as a figure (*right*).

played behind the stationary slit. If the figure is an oblique straight line and the ends do not pass across the slit (which is about a sixteenth of an inch wide and is viewed from a distance of about three feet), the stimulus consists of a line segment undergoing vertical displacement. This stimulus could result from either of two physical events: a segment moving up and down at a uniform speed or an oblique line moving back and forth horizontally at a uniform speed.

Under such conditions of ambiguity Hans Wallach of Swarthmore College long ago demonstrated that the shape of the aperture has an important influence on what is perceived. There is a strong tendency to perceive the line as moving parallel to the long axis of the aperture, probably because with a narrow slit there is no information to suggest any change in the identity of the segment. Accordingly we have found that under such conditions an extended, moving figure is never perceived.

Suppose the figure is a curved line but the conditions of presentation are otherwise the same. Here the only new factor is that the visible segment accelerates and decelerates as it moves vertically. Ann Corrigan and I have found that this information does not alter the outcome. All observers perceive a vertically moving segment, not a figure.

Next suppose the width of the slit is increased to an eighth of an inch. Under these conditions we found that the slope of the visible segment and the change in the slope can be detected through the slit. Even with this information, however, and the correlated vertical acceleration and deceleration of the segment, most observers still do not perceive a figure. Hence one might conclude that the tendency noted by Wallach for a segment to maintain its identity is so strong that it will do so even if the segment appears to tilt back and forth as it rises and falls.

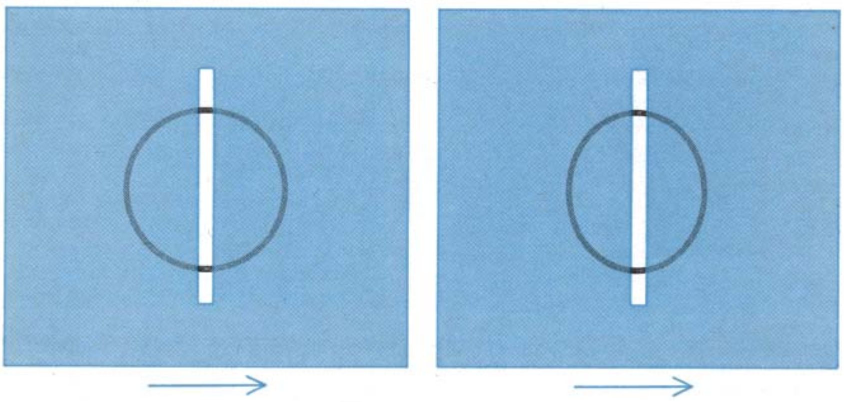
Only when the slit is widened still further, so that the curvature and the change in the curvature become detectable through the slit, does the balance shift in favor of figure perception. Why this is true is not yet clear, although several possibilities come to mind. One might speculate that the acceleration of a line segment upward just as its slope is tilting toward the vertical and just as its curvature is changing appropriately would represent too much of a coincidence if there were no extended curved line behind the slit. A moving-figure solution explains why these features vary together in this way, and the perceptual system tends to seek such solutions and to reject percepts that postulate mere coincidental variations.

Alternatively, one might say that the notion of a line segment moving vertically while simultaneously tilting and changing its curvature is a very complex perception. A moving figure is a simpler perception, and it has been argued that the perceptual system prefers the simplest description consistent with the stimulus. Moreover, one might speculate that the combination of changes, including changes of curvature, is recognizable to the perceptual system on the basis of past experience as the combination produced by an object passing behind a narrow aperture.

There are other kinds of information in the display that can affect the outcome. Allowing the ends of the figure to be seen through the slit may lead to the perception of a shape. Clearly when the ends are visible, there is a moment of unambiguous information to the effect that something is moving across the slit, not along it. One might think of this information as a cue that suggests the figure hypothesis, but in itself the information does not seem to be adequate; an oblique straight line still tends to be seen as a vertically moving segment except for the brief time when the end points are in view. When such end-of-line information is coupled with change-of-slope information, however, the balance is often shifted toward figure perception. Perhaps the reason is that a perceived change of slope tends to support the hypothesis triggered by seeing the ends of the line. The kind of unambiguous information given by the end points may be present throughout a figure if it has sharp discontinuities or changes in the direction of contours.

There is another factor we have found to be important. If the observer knows the display may represent a figure behind a slit, he is more likely to perceive it. Ordinarily merely knowing something (for example knowing that an illusion is an illusion) does not affect what one perceives, but in this case it does. Even with prior knowledge the conditions must be favorable for form perception. Such knowledge has no effect if the slope of the line segment never changes or if the figure is luminous and viewed in the dark so that the slit is not visible.

Our research on the characteristics



COMPRESSION EFFECT is often observed in anorthoscopic displays in which the figure moves behind a stationary slit. The perceived figure, in this case a circle (*left*), is typically compressed along its axis of motion, forming what appears to be a vertically elongated ellipse (*right*).

of the display that govern the perceptual outcome bears on the question of whether the anorthoscopic effect is genuinely perceptual when retinal painting is ruled out. Several investigators have contended that when only part of the figure is seen at any moment, the impression one has of the whole figure is more a matter of knowing it is present than a result of true form perception. Only if one sees the whole figure at once, they say, should the experience be described as form perception.

In my opinion there is a confusion here that can be traced to a theoretical preconception. If one believes form perception presupposes an extended retinal image all of whose parts are present simultaneously, then of course the anorthoscopic effect cannot be perceptual when retinal painting is ruled out. There is no reason, however, to equate form perception with simultaneity. Hearing a melody or a spoken sentence is certainly perceptual, even though what is perceived is extended over time. Only if the figure or the slit were moved very slowly might it be appropriate to say that one fails to perceive the figure although one knows it is there. Analogously, one might block the perception of a melody by separating the tones by extremely long intervals.

The fact is that those who have experienced the anorthoscopic effect in the laboratory are sure it is a perceptual phenomenon. One reason for the certainty is undoubtedly the clear differences that are observed under the different conditions. Hence the failures become very important. Observers may know that a figure is being presented, but they never perceive it unless the conditions allow. In other words, there are cases where it is appropriate to say the observer knows a figure is being presented but does not perceive it; these cases, in contrast, make it clear that the other cases are perceptual. Moreover, there are many instances where perception shifts from one possibility to the other during the presentation; such a reversal under ambiguous conditions is one of the hallmarks of the perceptual process.

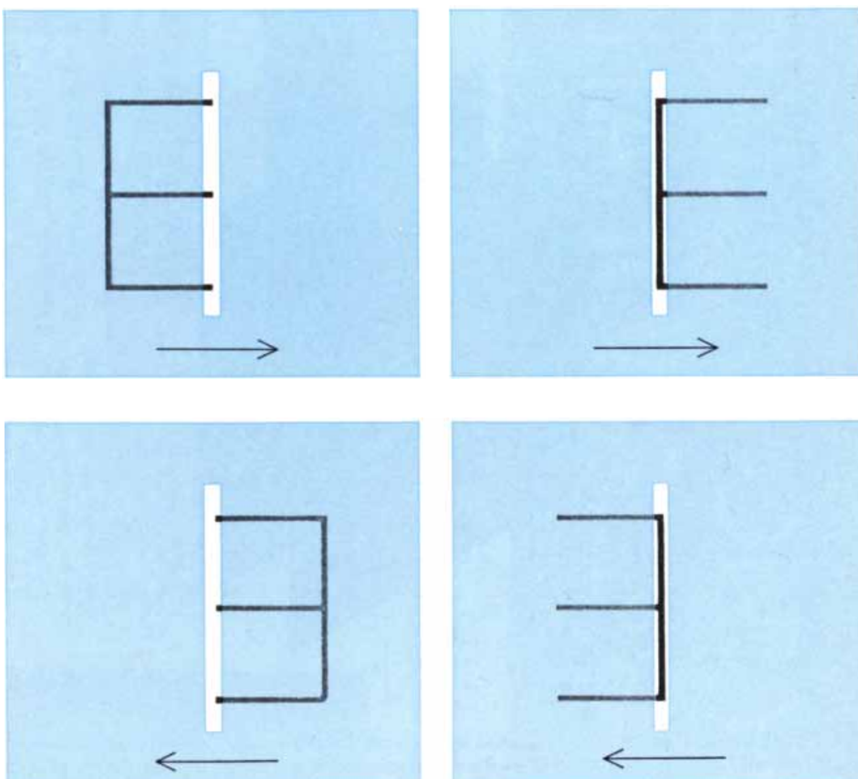
There is one peculiarity of anorthoscopic presentation I have not yet mentioned. When the slit is stationary and the figure is moving, the perceived figure is generally distorted; typically it is compressed along the axis of its motion. Thus a circle may look like an ellipse with its longer axis vertical. What is responsible for this effect? Advocates of the retinal-painting hypothesis have suggested that the distortion results from the failure of the observer to move his eyes in perfect synchrony with the figure. If the eye movement fell short of either the speed or the amplitude of the figure motion, the image painted on the retina would be compressed, and that could nicely account for the distortion.

An experiment carried out in 1967 by Stuart M. Anstis and Janette Atkinson of the University of Bristol tends to support this hypothesis. They introduced a moving target (a point of light) that the observer had to track back and forth. By varying the speed of the target point with respect to the speed of the figure, Anstis and Atkinson were able to "paint" a retinal image of the figure that varied in shape. For example, if the moving figure was a circle and the tracking target moved at half speed, the image spread over the retina would be an ellipse whose vertical axis was twice as long as its horizontal axis. If the figure moved to the right as the tracking target moved to the left, the image created would be the reverse of the one yielded by ordinary perception. The observers reported seeing figures whose shapes corresponded precisely with the retinal image established.

This finding puzzled my colleagues and me for some time because we had good reason to doubt that the retinal-painting hypothesis accounts for the anorthoscopic effect or for these perceptual distortions. Ultimately we devised an alternative explanation that seems to fit the facts better. To understand the explanation one must first consider more carefully the perception of figure motion when the slit is stationary and the figure is moving. According to the retinal-painting hypothesis, what should be seen is the whole figure within a large rectangular opening; the rectangle represents the image of the slit, which should also be painted on the retina. That is not what is seen. What is seen is a figure moving behind the narrow slit.

The question is: What determines the perceived speed of the figure? Given the narrowness of the slit, information about the speed of the figure can hardly be conveyed accurately by the brief passage of distinguishable contour components across the slit. Moreover, with the figures we employ there are no distinguishable components except at the ends. Therefore the speed of the figure is at best ambiguously represented. The perceived length of the figure depends entirely on its perceived speed, at least according to the problem-solving interpretation of anorthoscopic perception. Since the commonest outcome when there is no tracking target is one of perceived compression, we conclude that the speed is underestimated. We do not know why this is so, but it is important to keep in mind that there is no reason to expect accurate speed perception either.

When a tracking target is introduced, the perceptual system seems to assume



PERCEIVED FIGURE IS REVERSED if an observer misjudges the direction of its motion behind a stationary slit. In this experimental setup the letter *E* appears if the direction of motion of the figure is correctly perceived to be toward the right (*top*), but the reversed figure Ǝ is seen if the direction of motion is mistakenly interpreted to be toward the left (*bottom*).

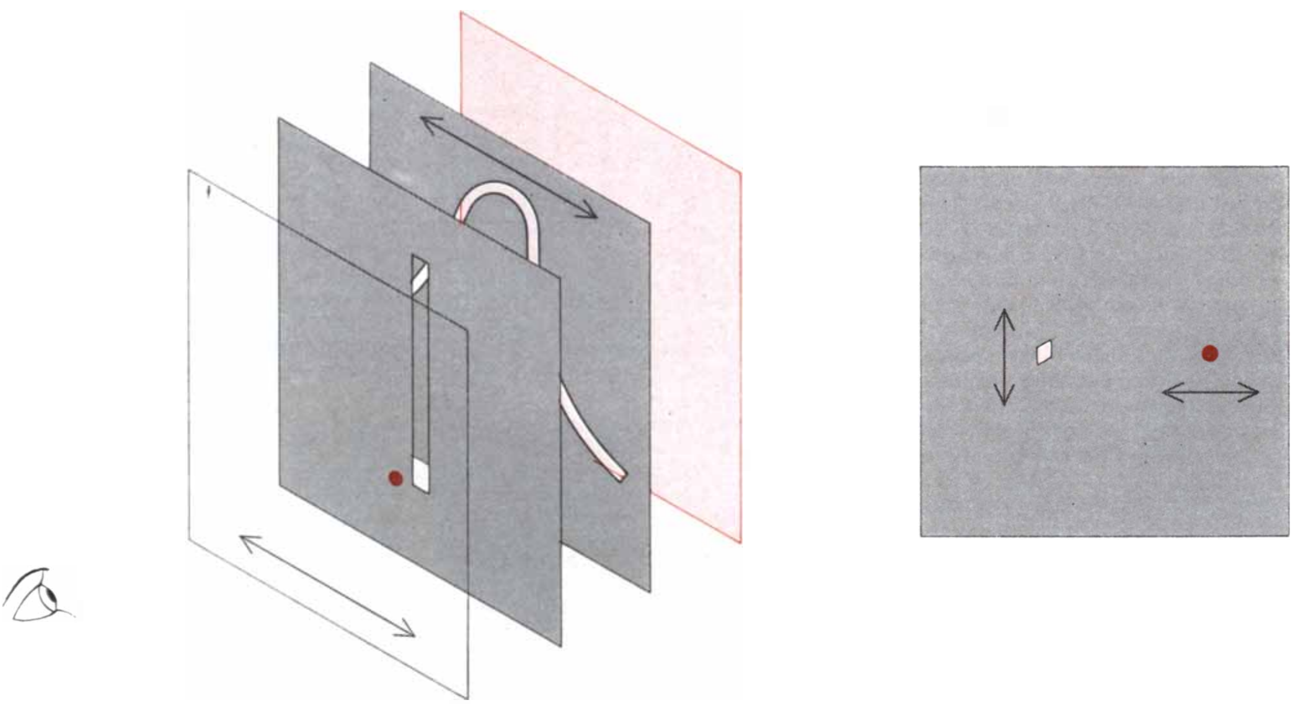
that the figure is moving at the speed of the target. One can formulate a general hypothesis that eye movement is a cue to figure motion under anorthoscopic conditions. Given this cue, the perceived speed of the figure is doubled when the target moves at twice the figure speed and is reduced by half when the target moves at half the figure speed. Since the apparent length of the figure depends on how far it seems to move behind the slit during the interval between the appearance of one end and the appearance of the other end, the distortion found by Anstis and Atkinson is explained. It results from a mental construction of length derived from the apparent speed of the figure and not directly from a distorted retinal image.

To provide evidence for our interpretation Di Vita, Halper, Deborah Wheeler and I did an experiment in which observers viewed a curved-line figure moving behind a slit at a certain speed and at the same time tracked a target dot. In effect we repeated the Anstis-Atkinson experiment but with our kind of figure. There was an important addition to the procedure as well. The observer not only indicated the perceived length of the figure (by adjusting a shadow-casting device that varied the length of a replica of the figure while holding its height constant) but also told us whether the speed of the figure appeared to be equal to, less than or greater than that of the moving target. Observations were made at several target speeds: equal to the speed of the figure, half the speed of the figure, twice the speed of the figure and the same speed as the figure but in the opposite direction. There was also a condition of free viewing, with no moving target present.

Whenever a target was tracked, all observers perceived a figure, whereas in the free-viewing condition none did. We deliberately made the slit very narrow in this experiment (a sixteenth of an inch), which we knew would eliminate the anorthoscopic effect under the more usual no-tracking condition.

The first point to be made about the results is that tracking is an important determinant of the anorthoscopic figure percept. We believe the major reason for this is that movement of the eyes provides an effective cue that a figure is moving back and forth behind the slit. Moreover, such a cue specifies the figure's speed, whereas without it the speed is indeterminate. It would be difficult to arrive at a figure percept of a definite length with the speed indeterminate. Virtually all the observers reported that the speed of the figure was the same or almost the same as that of the target regardless of the actual speed of each. As for the perceived length of the figure, if the actual length is assumed to be 1, the perceived length was .77 when the two speeds were the same, 1.65 when the target speed was twice that of the figure and .50 when the target speed was half that of the figure.

When the figure moved in one direction and the target in the opposite direction at the same speed, all observers perceived the figure as the mirror image of its true shape. For this test we employed an asymmetrical figure and had each observer make a rough sketch of what he perceived. The outcome was what one would predict on the basis of the reversed image painted over the retina, but it was also what we predicted from our eye-movement cuing hypothesis. Not only was the speed of the figure behind the slit ambiguous; its direction was too. For example, if the letter *E* was the figure and it moved to the right, the end points of its three horizontal prongs would be revealed first. If the figure seemed to be moving to the left but the end points of the prongs still appeared first and were followed by the vertical contour of the *E*, the figure would have to be a left-right reversed



MULTIPLE SURFACES were employed by the author and his colleagues in one of their control experiments on anorthoscopic perception. The first surface in front of the viewer is a transparent sheet of plastic on which a luminous tracking target (*colored dot*) is mounted; this sheet moves back and forth. The next surface is an opaque, stationary card with a vertical slit. The third surface is an opaque card with a cutout stencil of the test figure; this surface also moves back and forth. The back surface is a translucent sheet of plastic that diffuses the light from the battery of lamps behind it. The panel at the right shows what the viewer sees when the experiment is carried out in a dark room: a horizontally moving luminous tracking target and a line segment that appears to be moving vertically (or almost vertically). A moving figure is not seen under these experimental conditions because the slit is not perceived. When the same display is employed but the figure is instead drawn on the third surface and the room lights are on, a moving figure is always seen by the viewer.

figure: ᗱ. That the direction of motion was mistakenly perceived in this way is borne out by the incorrect response of all those observers who perceived the reversed figure when they were asked its direction of movement.

In this last experiment the tracking of a moving target has the effect of spreading the image of the figure over the retina. Although we have rejected the hypothesis that the image thus painted is the direct cause of anorthoscopic form perception and of the observed distortions, there is one aspect of the results that suggests an important role of the extended image. The figure percept is clearer and more similar to ordinary form perception under these conditions than it is when the eyes are stationary while viewing the stationary slit. We have found a similar result when, in the moving-slit condition, a stationary point is introduced and the observer fixes his gaze on it. Here too the image is spread over the retina, whereas in tracking the moving slit there is no such extended image. In both cases when such an image is present, an observer rarely fails to perceive a figure even if the conditions are otherwise poor for the anorthoscopic effect.

We call the improvement in anorthoscopic perception that can be attributed to an extended retinal image the facilitation effect. Apparently the perceptual system finds it easier to integrate the successive slices of the figure into a whole form when the slices are spread across the retina than when they all fall on one vertical strip. Although the extended image can thereby facilitate the anorthoscopic effect, it does not directly cause it. Reasons for this conclusion have been given above, and in addition it is supported by certain further control experiments.

In one of these experiments the figure was a horizontal straight line, and the observer tracked a target moving at the same speed as the figure. Although an image of the figure was spread over the retina, no observer perceived a figure. Instead the visible segment of the line seemed to be part of the slit and there was no impression of anything moving (except the target dot). We can only conclude that without some vertical displacement of the contour there is no reason for the perceptual system to infer that an extended figure is moving behind the slit; hence the extended image, although present, is not integrated into a mentally constructed form.

In another control experiment the figure consisted of a cutout stencil of a curved figure, which was illuminated from behind. A luminous target dot was also provided, and the display was viewed in an otherwise dark room. When the observer tracked the moving target, an image of the bright line was spread over the retina. Nevertheless, no figure was perceived. Rather, the visible segment of the figure, which was essentially a small point of light, appeared to move up and down, although its path seemed to be slightly tilted from the vertical.

Since tracking has the effect of spreading the image of the visible fragment over the retina, the failure to obtain a figure percept in this case shows dramatically how the anorthoscopic effect depends on the perception of the slit as an opening in an opaque surface. Only then is there the needed support for the mental construction of an extended figure.

In any event there seem to be two factors that combine to make the anorthoscopic perception of form particularly successful when a moving target is tracked. Eye movement serves as a cue that a figure is moving at right angles to the slit and imparts an unambiguous speed to the figure. The formation of an image spread over the retina facilitates the integration of the figure. Perhaps it is this facilitation that has led some investigators to subscribe to the retinal-painting hypothesis.

In conclusion, I have argued that the perception of form can be understood as a process of integrating information about the location of the parts of a figure with respect to one another (provided the parts are organized into one unit and are interpreted as figure rather than as ground). Physical contours or their representations in the retinal image are not necessary as long as some kind of information indicates where the boundaries of a figure are. Consequently anorthoscopic presentation of a figure one part at a time can yield form perception even though no extended image of the whole figure appears on the retina at any one time.

Because the anorthoscopic display is ambiguous and does not necessarily represent an occluded figure, however, the achievement of a form percept entails a process of problem solving. The presence of a region seen as a narrow aperture in a surrounding opaque surface is indispensable, and the stimulus must have certain other properties as well if perception of a figure is to be the preferred solution. Where a figure moves behind a stationary slit the perceptual system must also infer the figure's speed and direction in order to reconstruct its length and shape.

If this interpretation of the events that follow viewing an anorthoscopic display is correct, the perception of form is a process much closer to the cognitive level than has heretofore been recognized. It cannot be explained as a direct outcome of the physiological processing of contours stimulating the retina.

The Biochemistry of Resistance to Malaria

Genes for two lethal diseases, sickle-cell anemia and thalassemia, are favored by evolution because they protect against malaria. Now the mechanisms of that protection can be studied in the laboratory

by Milton J. Friedman and William Trager

Evolution results from natural selection, operating over a range of genetic diversity that arises from the mutation and recombination of genes. Variant genes that confer some selective advantage tend to increase in frequency, whereas deleterious variants tend to be eliminated. In human populations there are very few clear examples of selection for or against specific genes in response to specific forces. The best examples are inherited diseases. Selection acts against the genes that cause such diseases, and it acts most strongly against the severest conditions.

That being the case, lethal genetic diseases should be very rare. Yet certain inherited disorders of the red blood cells, notably sickle-cell anemia and thalassemia, are observed in some populations at surprisingly high frequencies. Does that argue against natural selection? On the contrary, the sickle-cell and thalassemia genes demonstrate the force of selection in evolution. The same variant genes that cause lethal blood-cell disease in homozygous individuals (who inherit two of the abnormal genes, one from each parent) protect heterozygous individuals (who inherit one abnormal gene and one normal gene) against the lethal effects of malaria, the agent of which is a parasite that infects red blood cells. That protection maintains the high frequencies of these otherwise deleterious genes.

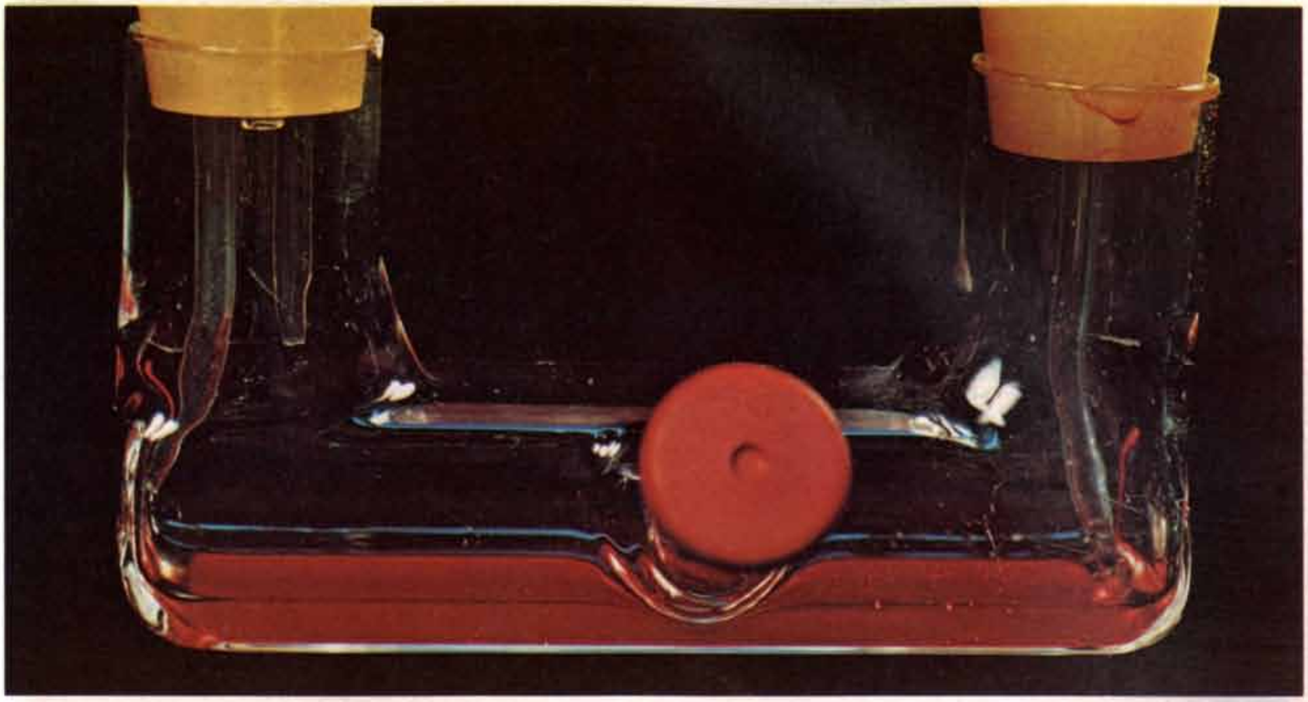
The strength of malaria as a selective force derives from the powerful effect of the parasitic disease on the health and reproductive capacity of human populations. Malaria has been a major cause of death throughout history. In Africa today malaria is endemic: it does not sweep through a population as an epidemic but rather is a constant affliction contributing to early-childhood mortality rates as high as 50 percent. It kills about 10 percent of its victims directly and contributes to the death of others by decreasing the immune system's ability to fight other infections. Because of malaria a significant number of children do not live to reproduce. Any genetic mutation that provides resistance to malaria must therefore have a high selective advantage.

It was the coincidence of the geographic range of sickle-cell disease with the range of malaria that first drew attention to the possibility that the sickle-cell gene might confer such resistance. Clinical evidence was harder to come by, but in 1954 Anthony C. Allison of the University of Oxford showed that children who were heterozygous for the sickle-cell gene had much less severe cases of the most lethal form of the disease than children who did not carry the gene. Because the parasite that causes malaria could not be maintained in a laboratory culture, however, the resistance could not be demonstrated at the cellular level, nor could its biochemical mechanism be established. Recently we have exploited a newly developed culture system to learn how the sickle-cell gene and some other variant genes that alter red-cell function confer resistance to malaria.

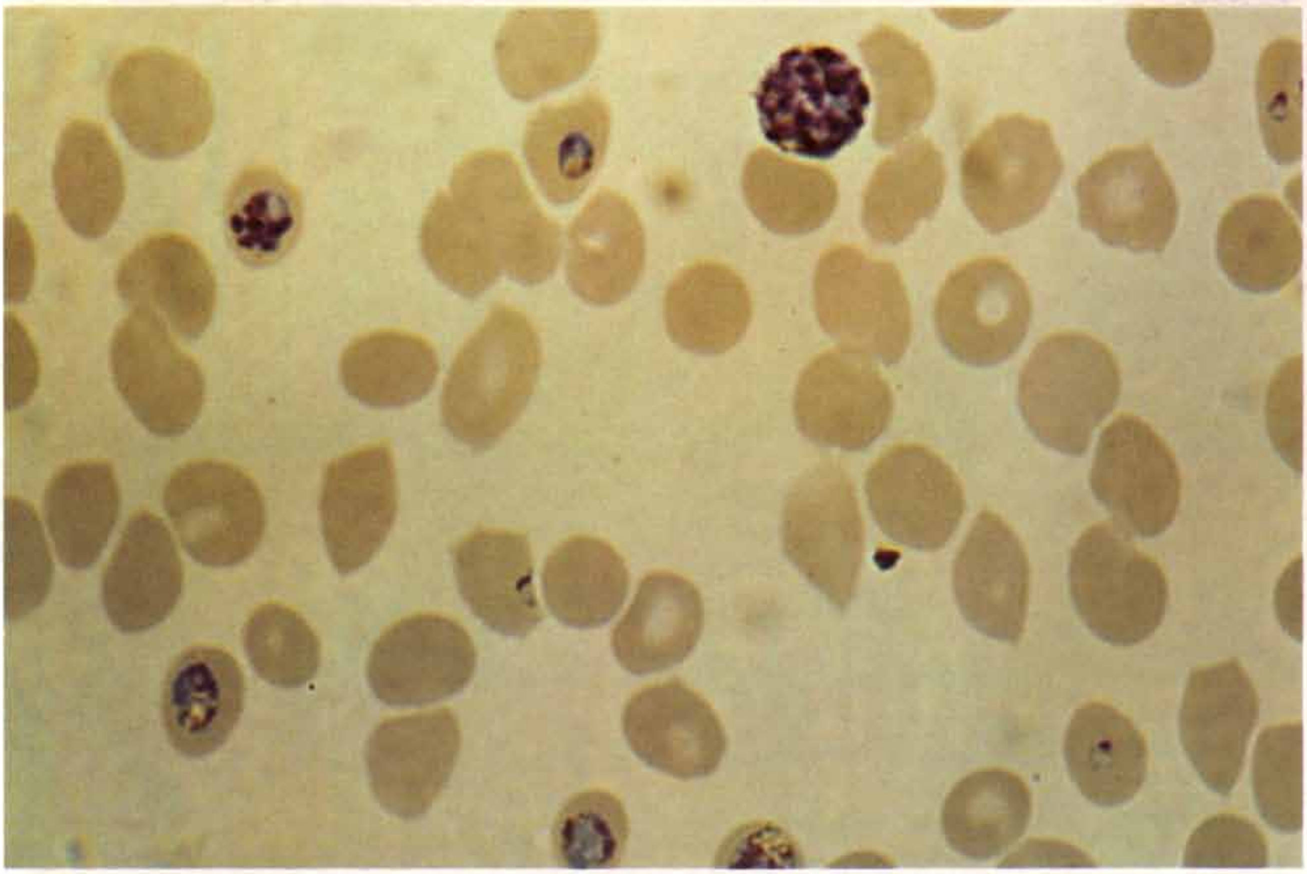
The red blood cell, where the malaria parasite encounters the altered cellular functions governed by these variant genes, is largely filled with hemoglobin: the protein that takes on oxygen in the lungs and carries it to the tissues. The other proteins of the red-cell cytoplasm are metabolic enzymes. Some of them catalyze glycolysis, whereby glucose is broken down step by step to form lactate, in the process synthesizing adenosine triphosphate (ATP). Others catalyze what is called the hexose monophosphate shunt, which maintains the coenzymes nicotinamide adenine dinucleotide phosphate (NADP) and glutathione in their reduced form. ATP is the all-purpose cellular energy carrier; reduced NADP (NADPH) and reduced glutathione are needed to prevent and repair oxidative damage. The cell membrane bounds the cell and controls its shape and deformability. It also controls the movement of ions into the cell and out of it; in particular it maintains—at the expense of ATP—a high-potassium interior against a tendency toward equilibrium with the low-potassium outside environment. On the outside of the membrane glycoproteins and glycolipids present a unique recognizable surface to the environment.

It is at this surface that the malaria parasite first interacts with the cell. The parasite is a small unicellular protozoon of the genus *Plasmodium*, four species of which cause malaria in man; the most lethal disease, responsible for a million deaths every year among African children, is caused by *Plasmodium falciparum*. A specialized form of the parasite is injected into the bloodstream by an *Anopheles* mosquito, migrates to the liver and there develops and divides to produce merozoites, the form that infects red cells. The merozoites reenter the bloodstream and recognize and bind to the red-cell membrane. A mechanism, as yet poorly understood, is activated that causes the merozoite to push in the cell membrane, which closes around it. Enclosed in a vacuole, the parasite grows, digesting hemoglobin to acquire the amino acids to make its own proteins and, we believe, utilizing red-cell glucose, ATP and reduced coenzymes in its metabolism. After a period of growth the parasite's nucleus divides several times, and then membranes enclose each nucleus and its surrounding cytoplasm. In this manner from 12 to 24 new merozoites are formed, which burst out of the cell and invade other cells.

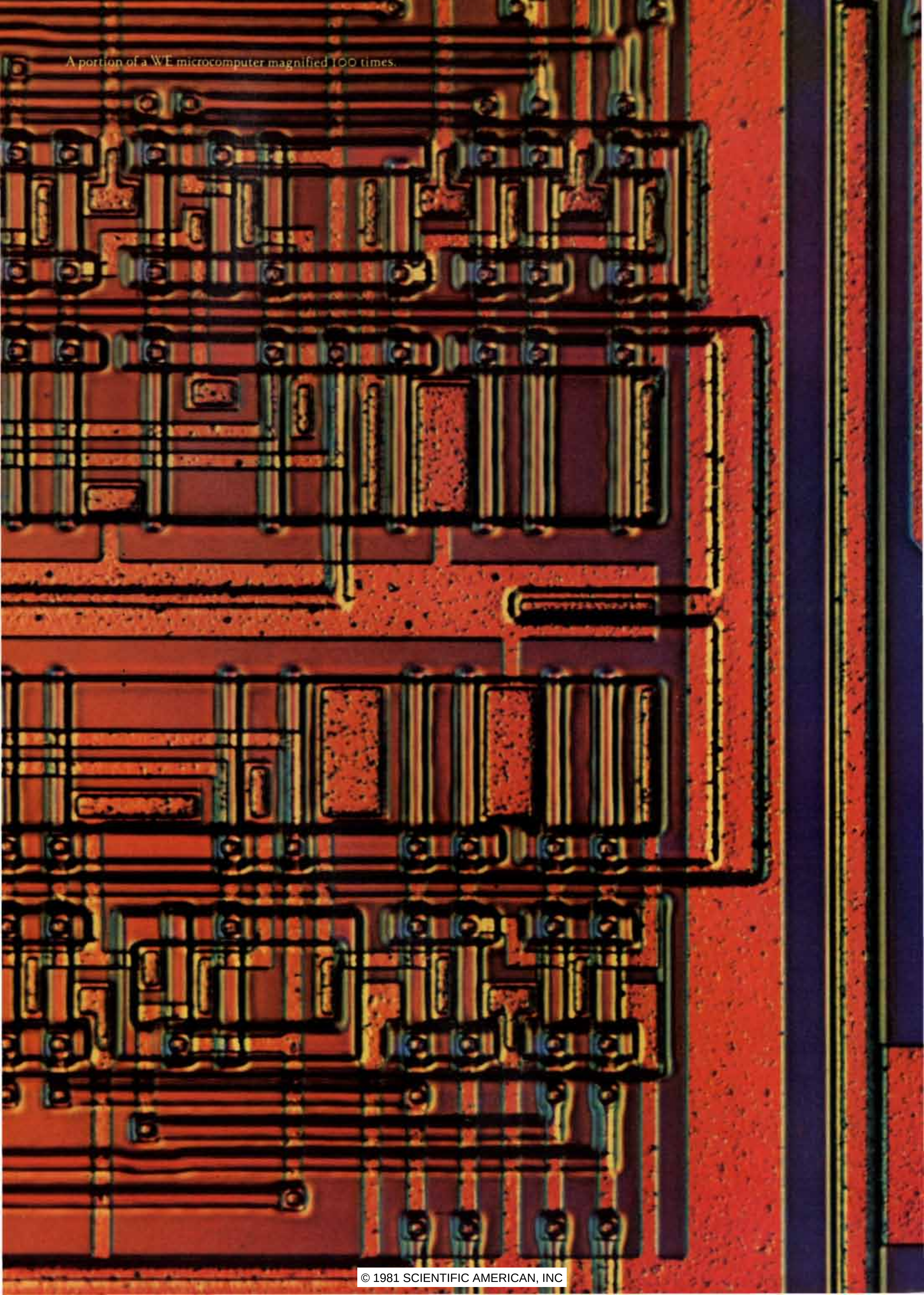
The characteristic periodic fever of malaria results from the synchronous release of merozoites, and of toxins produced by the parasite, throughout the body. In the case of *P. falciparum* this release takes place every 48 hours, the period of the parasite's developmental



CONTINUOUS CULTURE of the malaria parasite *Plasmodium falciparum* can be maintained in this apparatus. The parasites are grown in a thin layer of human red blood cells that coats the bottom of the horizontal tube and is covered with a nutritional medium. The oxygen content of the medium and other experimental conditions are manipulated by way of the vertical tubes. Cells are removed for microscopic analysis through the short tube at the front center. Until conditions for culturing the parasite were established in 1976, research on falciparum malaria, the most lethal form of the disease, had to be done with the blood of human volunteers or primate hosts.



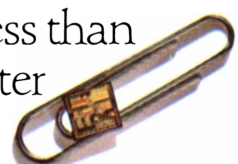
FATE OF MALARIA PARASITES is monitored in smears of the cultured cells on microscope slides. The stain (Giemsa's) colors the cell nuclei of the parasites dark purple and the parasite cytoplasm blue. The rate of multiplication is measured by counting the number of parasites per 100 red cells. This smear shows the parasites at various stages of their life cycle (*see bottom illustration on page 159*).


A portion of a WE microcomputer magnified 100 times.

Can you find the four mistakes in this microcomputer?

WE will. You're looking at a microcomputer smaller than your thumbnail. It has more than 30,000 parts, and performs complex tasks that once required a room-sized computer. Yet, if there's even the tiniest flaw, WE will find it.

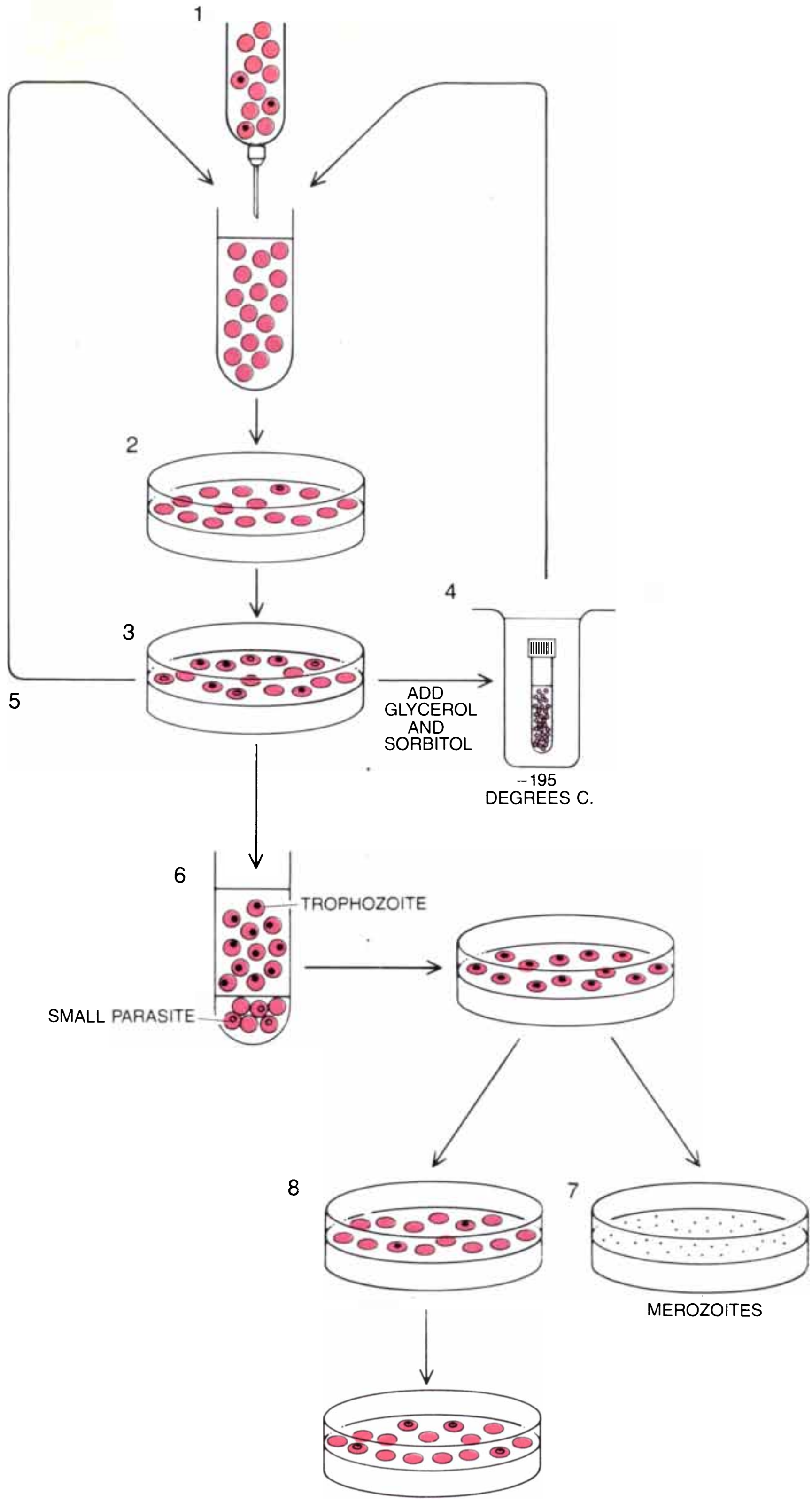
We're Western Electric. And in less than a minute, we can put a microcomputer through more than 100,000 separate checks. Components that don't measure up, don't make it into your phone system.



Microelectronics technology helps the Bell System keep down the cost of your phone service today. It will make dozens of new services possible in the future.

Western Electric products have helped to make your communications system the best in the world. And we're working hard to keep it that way.

Western Electric



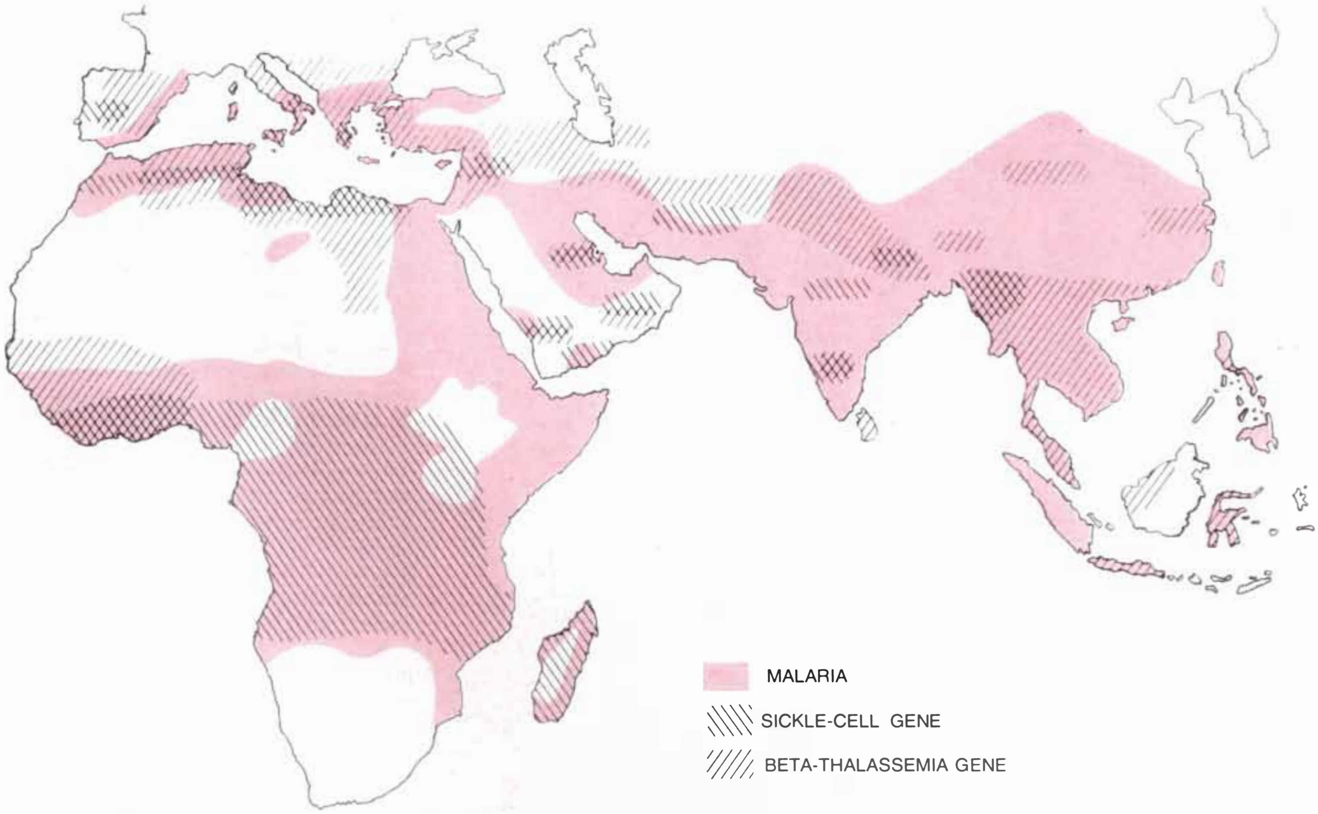
P. FALCIPARUM CULTURE is begun by inoculating fresh human red cells with parasitized cells from a patient or from a stock culture (*1*). The cells are grown at body temperature (in a culture dish, as is shown here, or in the U-tube depicted on page 155) as a thin layer covered by a nutrient medium (*2*). The culture can be sampled to measure parasite multiplication. After three or four days the cells contain a mixture of small and large parasites (*3*), which can be frozen for future use (*4*), inoculated into new cultures (*5*) or incubated in a gelatin solution to separate trophozoites (*6*); these large parasites can be grown alone to produce merozoites for further study (*7*) or can be mixed with fresh cells to produce a synchronous culture of parasites (*8*).

cycle in the red cell. The fever and the debilitation that accompany it are the major symptoms of malaria. In falciparum malaria, however, there is a more lethal effect. The infected red cell develops knobs on its surface that attach the cell to the walls of capillaries, where it lodges until the parasite is mature. When a large number of cells are thus sequestered in a vital organ such as the brain, death can result. The clearest indication of the protective effect of the sickle-cell gene is that very few carriers of the gene die from the cerebral complications of falciparum malaria.

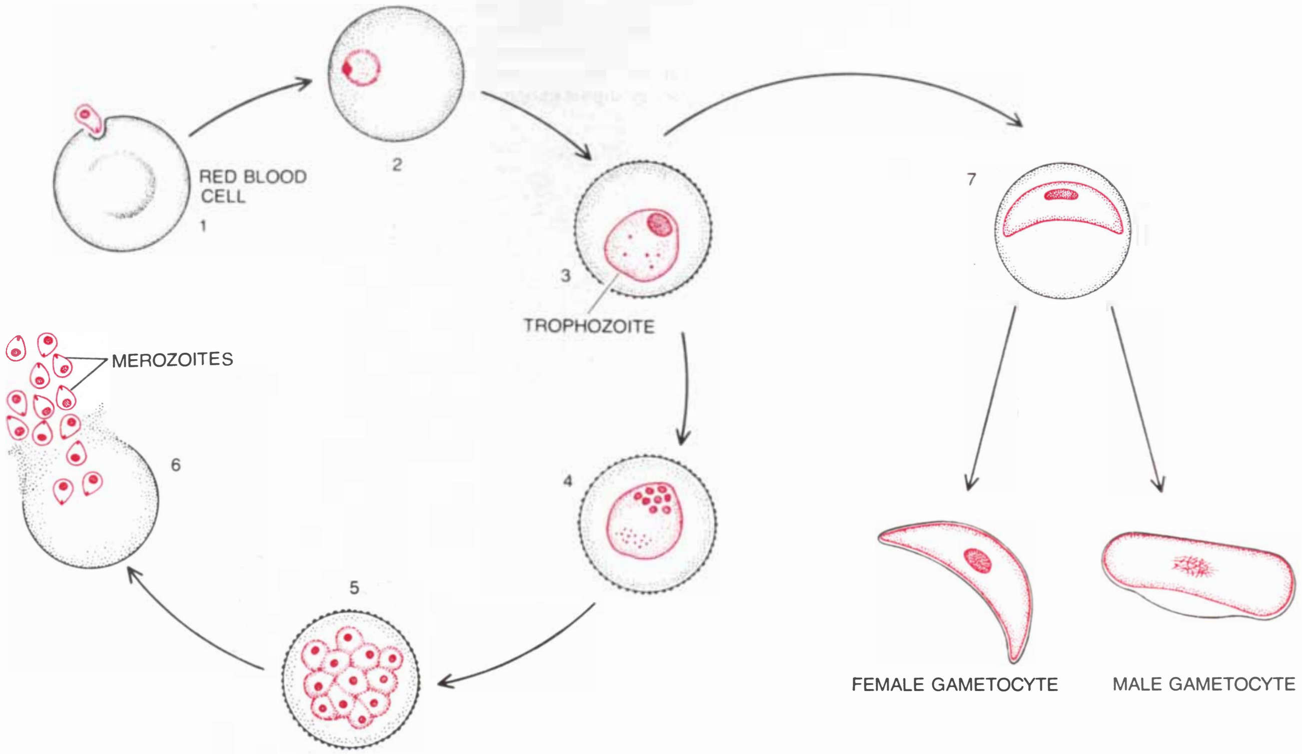
Linus Pauling and his colleagues defined a molecular disease for the first time when they demonstrated that in sickle-cell anemia the hemoglobin molecule is altered, and that people with sickle-cell disease have only the altered molecule, hemoglobin *S*, in their red cells, whereas some members of their families have about half hemoglobin *S* and half normal hemoglobin *A*. Family studies published at about the same time confirmed Pauling's findings, showing that the inheritance pattern of sickle-cell anemia could be ascribed to a single gene, with the disease appearing only in family members who are homozygous for that gene.

Adult hemoglobin is made up of two alpha chains and two beta chains, and Vernon M. Ingram of the University of Cambridge soon showed that in hemoglobin *S* only the beta chain is abnormal. The abnormality involves only one of the amino acids constituting the chain: a valine is substituted for a glutamic acid. Sickle-cell homozygotes (designated *SS*) have only hemoglobin *S* because they carry two of the mutant beta-chain genes. Heterozygotes (*AS*), who are said to have sickle-cell "trait," carry only one mutant beta-chain gene, and about 40 percent of their hemoglobin is hemoglobin *S*. If two people with sickle-cell trait have four children, the probability is, in accordance with Mendelian principles, that one child will have sickle-cell disease, one will be normal and two will have sickle-cell trait.

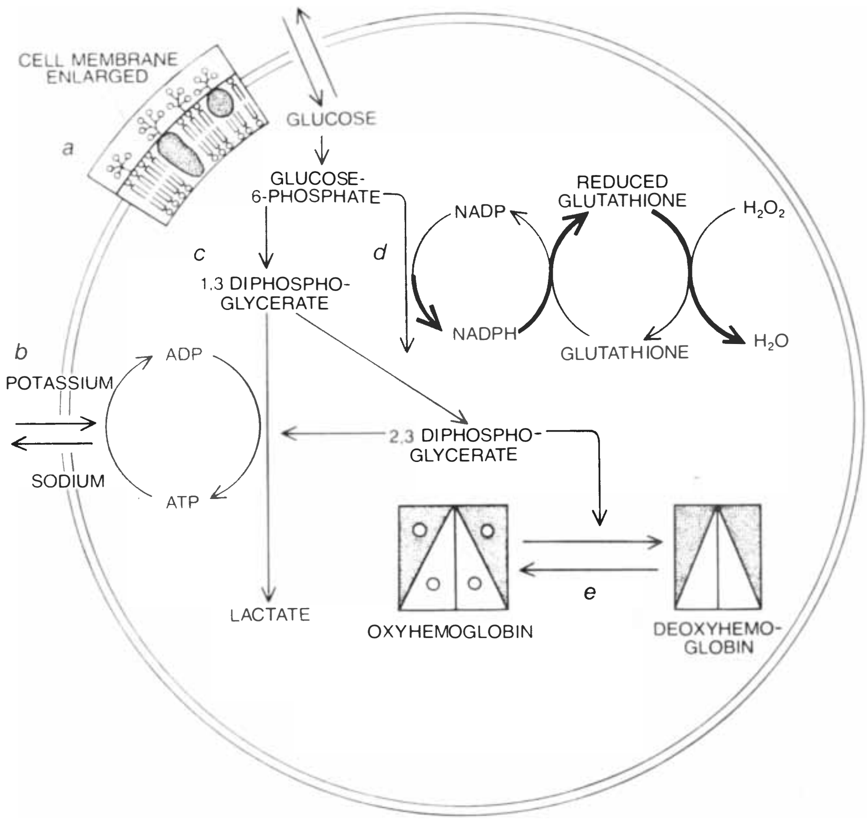
The symptoms of sickle-cell disease appear when the *SS* red blood cells lose oxygen as they circulate through the tissues. When hemoglobin *S* is deoxygenated, it tends to aggregate in long, thin fibers. The fibers distort the normally disk-shaped cells into angular forms, including the characteristic crescent shape. In a sickle-cell crisis some sickled cells block the local circulation of blood and impede the delivery of oxygen. As the oxygen level drops, more red cells sickle and the area of impaired circulation spreads, causing extensive tissue death. In the absence of advanced medical treatment the survival of hemoglobin-*S* homozygotes is very low. People with sickle-cell trait do not ordinarily suffer



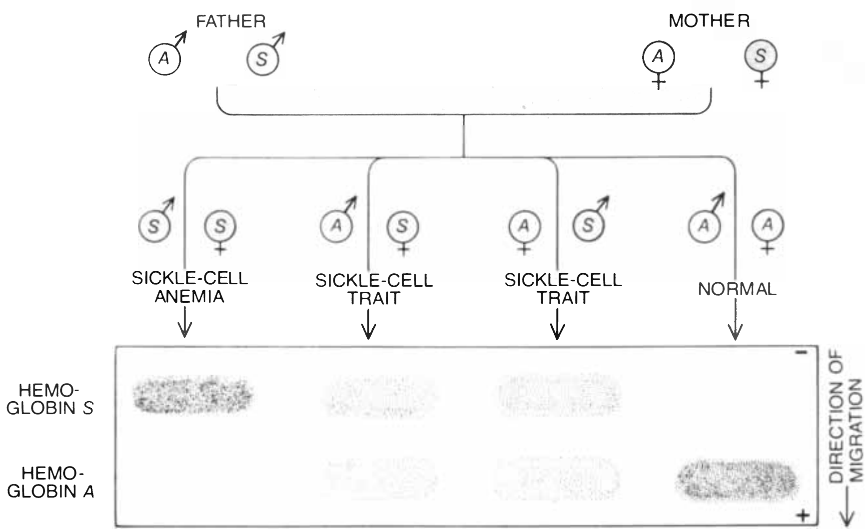
DISTRIBUTIONS of the sickle-cell gene and of the beta-thalassemia gene lie within the area where falciparum malaria was prevalent before 1930 (*color*). This geographic coincidence provided the first suggestion that resistance to malaria might be the evolutionary advantage that was tending to maintain the genes responsible for lethal blood diseases at high frequencies in certain human populations.



LIFE CYCLE of *P. falciparum* in the red cell begins with the invasion of the cell by a merozoite (*1*). The parasite engulfs a droplet of cytoplasm, so that in section it looks like a thin ring (*2*). The ring grows and fills in to become a trophozoite (*3*); knoblike structures develop on the cell surface and attach the cell to the blood-vessel wall. There the parasite's nucleus divides repeatedly (*4*); each daughter nucleus acquires a bit of cytoplasm (*5*), and the parasite divides into from 12 to 24 merozoites, which burst the cell and begin a new cycle (*6*). Some trophozoites develop (*7*) into male and female gametocytes, which are ingested by mosquitoes to initiate the sexual phase of the cycle.



RED BLOOD CELL, the simplest cell of the body, is no more than a membrane-bounded cytoplasm. Glycolipids and glycoproteins on the surface (*a*) determine the cell's interactions with its environment and with the malaria parasite. Proteins in the membrane control the transport of substances into and out of the cell; specifically, potassium is pumped in and sodium is pumped out (*b*). This process requires energy, which is provided by ATP generated through glycolysis, whereby glucose is broken down to form lactate (*c*). The hexose monophosphate shunt (*d*) produces reduced NADP and reduced glutathione, both of which prevent and repair oxidative damage to the cell membrane. The cytoplasm is largely filled with the four-chain protein hemoglobin (*e*), which binds oxygen in the lung and delivers it to the tissues, in the process undergoing structural changes that are to some extent controlled by diphosphoglycerate.



INHERITANCE of genes for abnormal hemoglobin *S* is diagrammed for parents who are heterozygous for the abnormal gene. One child is homozygous (*SS*) for the gene, that is, he inherits two hemoglobin-*S* genes and has sickle-cell anemia. Two children inherit one gene for hemoglobin *S* and one for normal hemoglobin *A*; they are heterozygous (*AS*) like their parents and have sickle-cell "trait." One child is normal (*AA*). The hemoglobins are identified by electrophoresis (*bottom*). Hemoglobin samples placed on cellulose acetate are subjected to an electric current. Hemoglobin *S* differs from hemoglobin *A* in only one amino acid; the substituted amino acid lacks a negative charge, and so hemoglobin *S* migrates toward the positive pole less rapidly than hemoglobin *A*. *AS* heterozygotes have a mixture of the two hemoglobins.

from the disease, however. Their *AS* red cells have enough normal hemoglobin so that they sickle only under extreme conditions, such as at high altitudes.

Four years ago, after half a century of attempts to grow the malaria parasite in the laboratory, one of us (Trager) was able to define the conditions that make it possible to maintain *P. falciparum* in a continuous culture of human red blood cells in an artificial bloodlike medium. The culture system is being exploited by some workers to develop experimental vaccines containing material from various stages of the parasite's life cycle, which are being tested in animals. Other investigators are trying to purify and analyze the biochemical agents that have particular effects in the course of an infection. The system has also made it possible for us to study in detail the interactions of the malaria parasite and variant host red cells.

To learn how sickle-cell hemoglobin protects a heterozygote carrier against malaria we cultured the malaria parasite in red blood cells taken from normal donors, from individuals with sickle-cell trait and from patients with sickle-cell anemia. We did so under our standard culture conditions, in an atmosphere with an oxygen concentration of 17 percent, which created an oxygen tension in the culture medium similar to that in the lungs; the hemoglobin was fully oxygenated and the variant cells did not sickle. Under these conditions the parasites grew equally well in all three kinds of cells. This showed that there is no major alteration of red-cell metabolism in the variant cells, and that hemoglobin *S*, like hemoglobin *A*, can be digested by the plasmodium.

To test the effect of sickling on parasite growth we added a small number of infected normal cells to cultures of *SS* and *AS* cells in a 17 percent oxygen atmosphere. During the next 48 hours, the period of one growth cycle, all the parasites left their normal host cells and invaded the variant cells. When we lowered the oxygen concentration to 3 percent, the *SS* cells sickled, as did some of the *AS* cells. We monitored the parasites daily. After one day in low oxygen almost no parasites were visible in the *SS*-cell cultures; they had lysed, or disintegrated, and so had their host cells. In the *AS*-cell cultures, on the other hand, it was only on the second day that the number of live parasites decreased significantly. And rather than disintegrating, the killed parasites were still visible as shriveled masses in the cells. They looked to us like parasites that had starved, perhaps as the result of some kind of metabolic inhibition that was an indirect result of sickling.

If that was the case, preventing sickling should protect the parasites. We treated *AS* cells with cyanate, which increases the affinity of hemoglobin *S* for oxygen, making it less likely to aggregate at a given oxygen tension. After cyanate treatment and washing, the *AS* cells remained competent as hosts for *P. falciparum,* but now they did not sickle as readily. When such cells were infected and then cultured in 3 percent oxygen, the parasites survived. The inhibition we had seen in untreated cells must therefore have been due to the sickling of the cells, not simply to low oxygen. By what mechanism might sickling inhibit parasite growth?

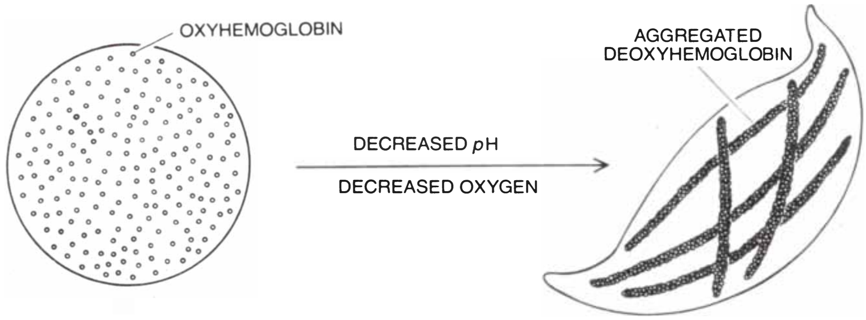
One of the things that happen when a red cell containing hemoglobin *S* sickles is that the cell membrane becomes more permeable to potassium, which leaks out; in the low-oxygen condition the potassium level in our *AS* host cells was decreased. It had earlier been shown that parasites maintained outside red cells require a high-potassium environment, and so we hypothesized that the loss of potassium on sickling might have inhibited parasite metabolism. To test the idea we again incubated infected *AS* cells in 3 percent oxygen, but in a medium with an elevated potassium content. The cells sickled as usual in low oxygen, but now the cellular potassium level stayed high—and the parasites survived. Preventing the loss of potassium, in other words, prevented the inhibition of parasite growth in sickled *AS* cells. (Under the same conditions plasmodia in sickled *SS* cells were not protected; they died by lysis. Electron micrographs showed why. After six hours of deoxygenation needlelike bundles of aggregated hemoglobin *S* could be seen penetrating some of the plasmodia; the membranes of other parasites had been disrupted, and they were partially lysed. In other words, in the *SS* cells the parasites were killed not by metabolic inhibition but by actual physical disruption.)

The sequence of events in *AS* cells, then, seemed to be as follows: Sickling lowered the potassium level, and the low potassium level killed the parasites. This finding could not, however, fully explain the heterozygote's resistance to malaria. Because an uninfected *AS* cell has less hemoglobin *S* than an *SS* cell, it does not normally sickle in nature; it circulates through regions of low oxygen tension too quickly for sickling to take place. The progress of a parasitized cell, however, is impeded by the knobs on its surface, and the cell remains in a low-oxygen environment for many hours. Even so, fewer than 5 percent of the cells would sickle were it not for still another effect of infection.

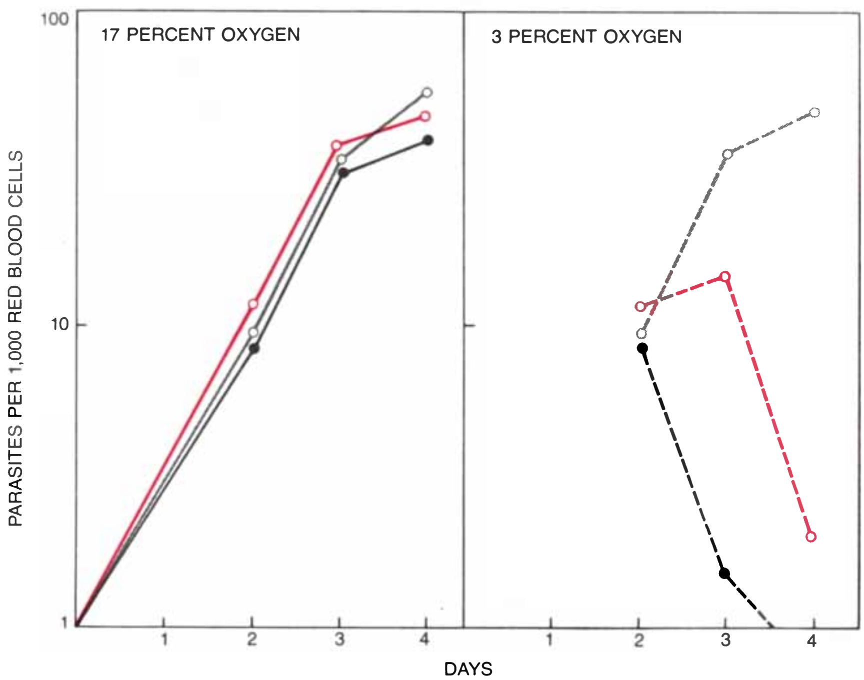
Lucio Luzzatto and his colleagues at the Istituto Internazionale di Genetica e Biofisica in Naples have shown that infected cells sickle much faster than uninfected ones. Why? We found that the intracellular environment of an infected cell is more acidic (.4 *p*H units lower) than that of an uninfected cell, so that the rate of sickling is significantly increased. We calculate, moreover, that the lower *p*H level of parasitized cells also increases the extent of sickling—up to about 40 percent.

Taken all together these observations suggest the following mechanism of protection against malaria in sickle-cell heterozygotes. The parasite in an infected *AS* cell develops normally until the cell is sequestered in the tissues. Then, given the low-oxygen environment and the low intracellular *p*H, the host cell sickles. The potassium level drops and the parasite dies. Such a process can protect against malaria even if not all the parasites are affected, because even a reduction in the rate of multiplication of the plasmodium can give the immune system the time it needs to mount a protective response of its own. (There is an alternative hypothesis. Infected cells might, for some reason we have not discerned, sickle while circulating rather than while being sequestered, and then they might be eliminated by the filtering action of the spleen. The first hypothesis is supported, however, by evidence that heterozygotes are not protected against types of malaria in which infected cells do not develop knobs and are not sequestered in the tissues.)

In addition to sickle-cell disease and sickle-cell trait there are other inherited disorders of the red blood cells



HEMOGLOBIN S in its oxygenated state (*left*) is dispersed through the red cell, which has a normal disk shape. Unlike normal hemoglobin *A*, however, hemoglobin *S* tends to aggregate when it becomes deoxygenated in the tissues, forming needlelike quasi-crystalline structures that distort the cell into a rigid, jagged shape (*right*). Sickled cells may block capillaries, decreasing blood flow, reducing the oxygen level and thus promoting the sickling of more cells.



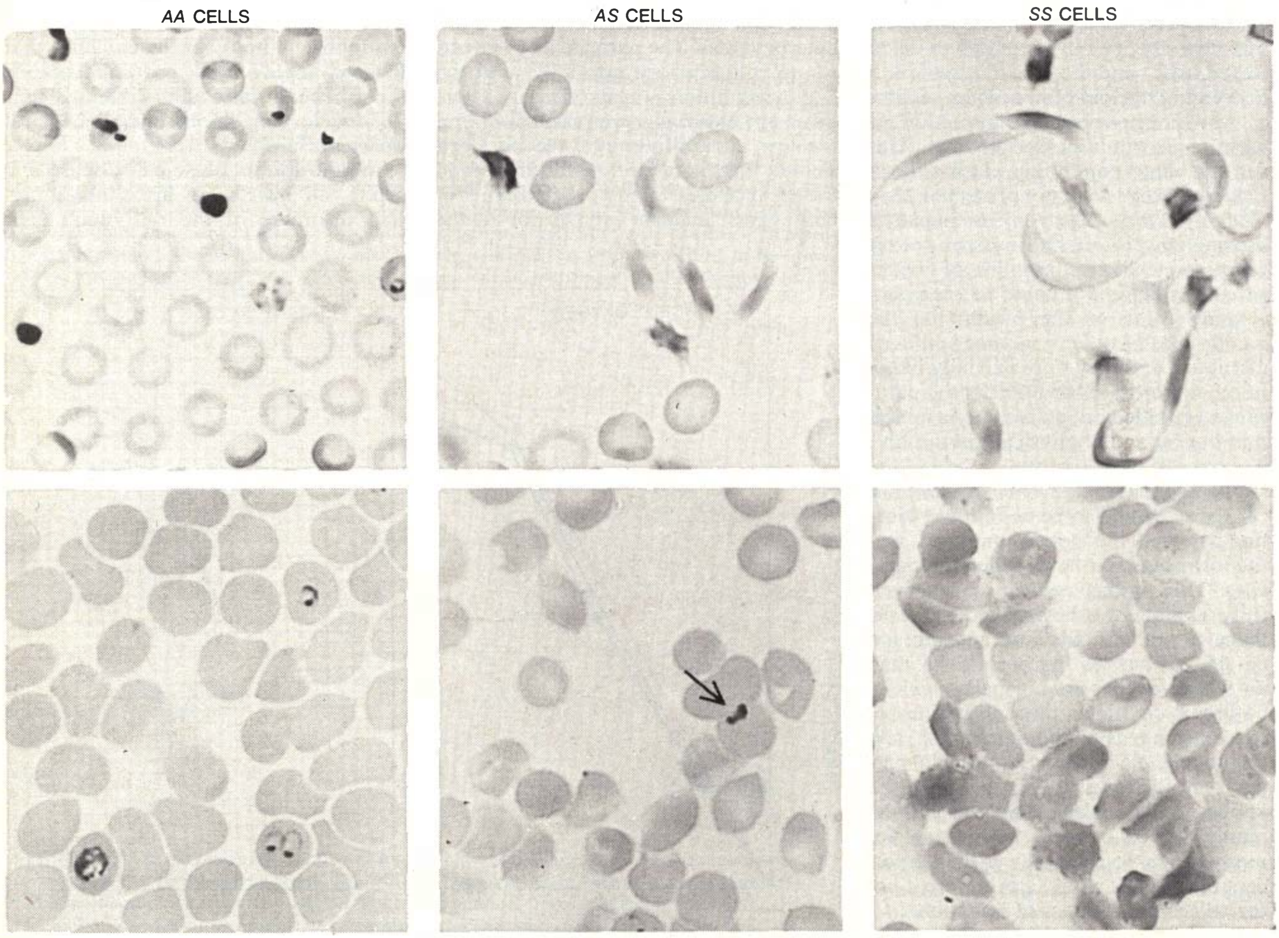
MULTIPLICATION OF P. FALCIPARUM is about the same in normal cells (*open circles*), *AS* cells (*color*) and *SS* cells (*black*) as long as the cells are in 17 percent oxygen (*left*). If after two days the oxygen level is lowered to 3 percent, however, only parasites in normal cells keep growing (*right*). The parasites in the *SS* cells die in a day, those in the *AS* cells in two days.

whose geographic incidence has been correlated with that of malaria, implying that the genes responsible for those disorders too may confer some resistance. Among those disorders are the thalassemias, which involve a deficiency in the manufacture of one or another hemoglobin chain. Beta thalassemia, for example, is a deficiency in beta-chain synthesis. Homozygous beta thalassemia, known as Cooley's anemia, is a severe disease in which little normal adult hemoglobin, if any, is synthesized; blood transfusion is usually the only means by which a patient's life can be prolonged. Yet throughout many malarial regions, and in particular around the rim of the Mediterranean, about 1 percent of all children born are homozygous for the beta-thalassemia gene and have Cooley's anemia; heterozygotes do not have the disease. Resistance to malaria has not been convincingly demonstrated for beta-thalassemia heterozygotes, but there is a very suggestive geographic correlation between the frequency of the gene and a regional history of malaria. To cite just one example, the gene frequency is much higher in the valleys of Sardinia, where malaria was for a long time endemic, than it is in the mountains, where malaria was rarer.

We set out to demonstrate a resistance effect in our culture system. We knew that one characteristic of any thalassemic cell is abnormal sensitivity of the cell membrane to damage by oxidation. When a molecule is oxidized, electrons are removed that would ordinarily have a role in the formation of chemical bonds. When lipid (fat) molecules, which are major constituents of the membrane, are oxidized, they fragment and disrupt the integrity of the membrane. The agents of oxidation in cells have not all been identified, but it is known that one such agent is hydrogen peroxide.

The malaria parasite generates hydrogen peroxide in its host cell (as has been demonstrated by N. Etkin and John W. Eaton of the University of Minnesota Medical School). Peroxides give rise in any cell to oxidative stress, challenging the cell's ability to preserve its integrity. In the more sensitive thalassemic red cell hydrogen peroxide might actually bring about damage to the membrane. We confirmed the likelihood of this effect by finding that parasites in heterozygous beta-thalassemia cells were more sensitive than parasites in normal cells to three experimental conditions. One condition was a high-oxygen environment (an oxygen concentration of from 25 to 30 percent). Another was the presence of certain chemicals that catalyze oxidation reactions. The third was the absence from the culture medium of one normal constituent: reduced glutathione, which is an intermediate in the metabolic pathway that reduces the cellular level of hydrogen peroxide.

Each of these three conditions was calculated to increase the oxidative stress on the cell. Eaton found that in mice infected with malaria oxidative stress and protection against it are finely balanced. Any additional sensitivity in a thalassemic cell, therefore, might well affect the course of the infection. (Although a thalassemic cell has a low hemoglobin content, there is apparently



PARASITE-INFECTED RED CELLS that have been cultured for one day in a 3 percent oxygen atmosphere are seen in these photomicrographs either fixed in formaldehyde to preserve their shape (*top*) or smeared on a slide for better observation of the parasites (*bottom*). The normal (*AA*) cells (*left*) are still disk-shaped; some of the *AS* cells (*middle*) and all the *SS* cells (*right*) have sickled. The parasites in the *AA* cells are alive. In *AS* cells parasites are dead (*arrow*) or will die within 24 hours. In *SS* cells no parasites are visible.

enough hemoglobin to maintain the parasite; the low hemoglobin level alone did not inhibit parasite multiplication in our system.)

We found that vitamin E, which protects cell-membrane lipids against oxidative damage, prevented the death of parasites in thalassemia-trait cells under all conditions. This was supporting evidence for the idea that the membrane of a heterozygous beta-thalassemia cell is damaged by oxidation in the course of malaria infection. As for the mechanism, again potassium appears to be implicated. In a high-potassium medium the parasites in the thalassemia-trait cells developed normally under conditions that led to their death in a medium with a normal potassium concentration.

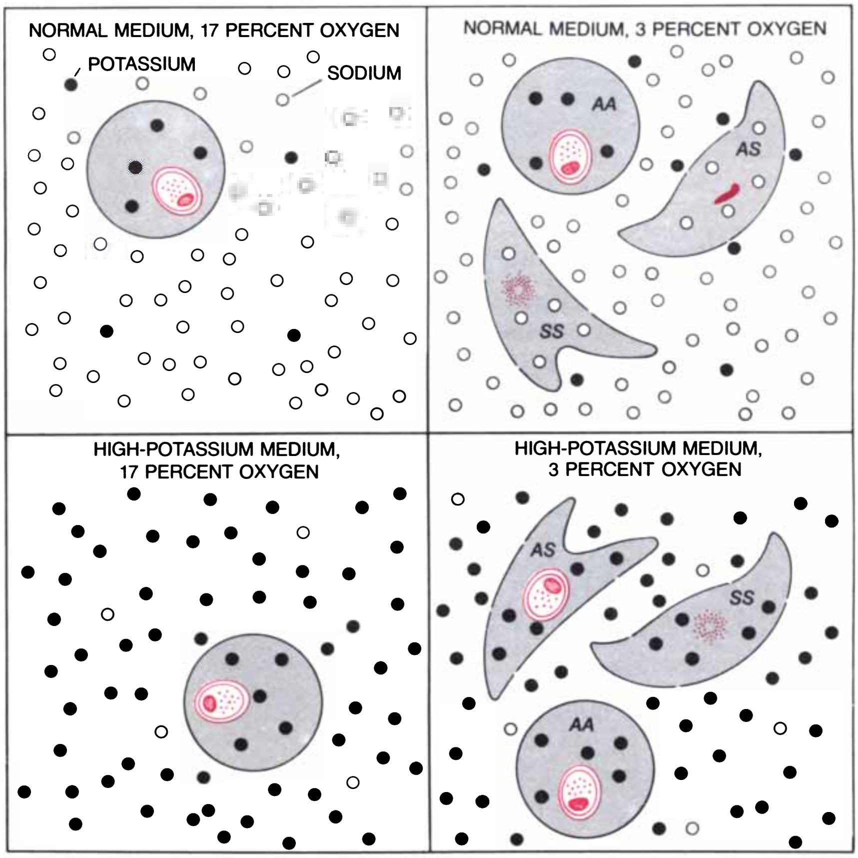
There were three interesting corollary results of this investigation. One was the reinforcement of a long-suspected link between malaria and an inherited deficiency in the red cell's supply of the enzyme glucose-6-phosphate dehydrogenase (G6PD). This deficiency too is prevalent in malarial regions; field studies have sometimes, but not always, demonstrated a correlation with resistance to malaria. G6PD is the first enzyme in the hexose monophosphate shunt, which regenerates NADPH, a coenzyme that is essential for protection against and repair of oxidative damage. It appeared that red cells deficient in G6PD, like thalassemia-trait cells, might be more sensitive to the hydrogen peroxide generated by the malaria parasite. We found that parasites in G6PD-deficient cells were indeed highly sensitive to stress by oxidants and were protected by antioxidant agents.

A second corollary implication of our results with thalassemic cells had to do with favism: a hemolytic anemia promoted by the ingestion of fava beans, which are consumed throughout the Mediterranean world. The fava bean contains a variety of substances that could increase the red cell's sensitivity to oxidants; some of the substances are related to the oxidation catalysts introduced in some of our experiments. Do those experiments mimic the consumption of fava beans by people with beta-thalassemic or G6PD-deficient red cells? If they do, the results would indicate that eating fava beans (and perhaps other foods as yet not identified) increases the level of protection against malaria in people who are heterozygous for these two red-cell disorders. Such a dietary effect could also explain the inconsistent results of studies of malaria resistance among such heterozygotes.

The third corollary result had to do with infants. During the first few months of life infants are almost completely protected against malaria. Geoffrey Pasvol, R. J. M. Wilson and D. J. Weatherall of the University of Oxford recently showed that fetal hemoglobin (which consists of alpha and gamma chains, persists for a time after birth and is found in some adults' red cells) may contribute to this protection; even adult red cells inhibited the growth of parasites if the cells contained fetal hemoglobin. We have found that malaria parasites in fetal red cells, like those in thalassemic cells, are highly sensitive to oxidative stress. It is not likely that malaria has been a selective force in the evolution of fetal hemoglobin, however. The fetal protein probably evolved under a different selective pressure. It has a higher affinity for oxygen than adult hemoglobin, and so it improves the delivery of oxygen to the developing fetus. The oxidant sensitivity of the fetal red cell and the resulting resistance to malaria are probably side effects of a developmental adaptation.

To sum up, the process of evolution has resulted in the selection of genetically variant red blood cells that function well enough under normal conditions but are susceptible to damage when they are infected by *P. falciparum*—damage that kills the parasite. In other words, these cells are so marginally viable that infection makes them unviable and unable to support the intracellular parasite. The genetic alteration in the case of *AS* cells is a borderline tendency to sickle, which is enhanced by the parasite's presence. The alteration in thalassemia-trait cells is increased sensitivity of the membrane lipids to hydrogen peroxide generated by the parasite. In both cases the end effect is a loss of potassium that inhibits the parasite's metabolism.

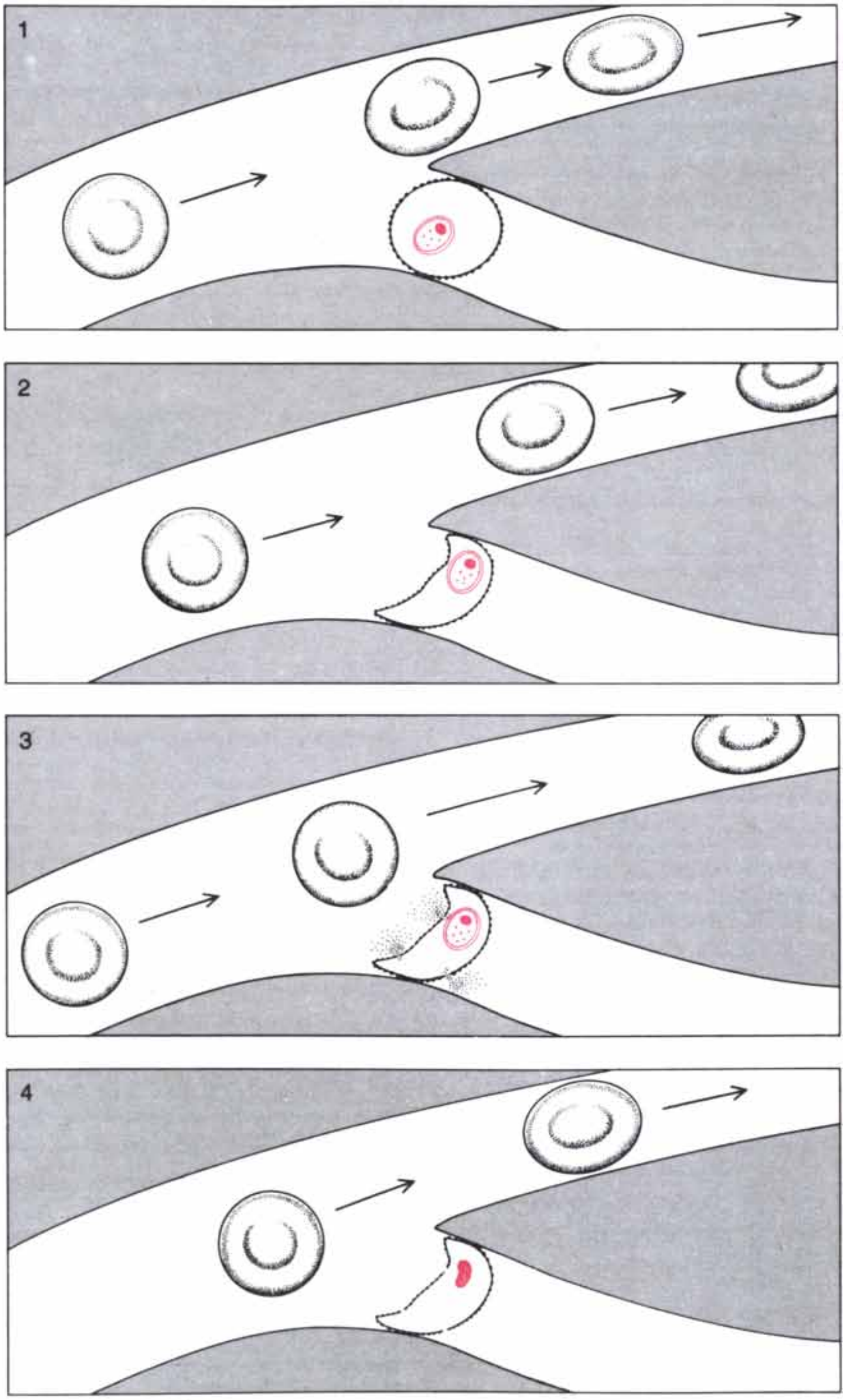
The evolutionary career of a gene can be quite complex. When a new random mutation arises, one cannot predict its effect or its potential. One could surely not have predicted that certain genes would cause lethal blood disease in a homozygote and protect the heterozygote against death from malaria. The delicately balanced contest between the selective effect of malaria on the one hand and of sickle-cell disease and beta thalassemia on the other has resulted in a balanced polymorphism: a situation



POTASSIUM LOSS is shown to be the cause of parasite death in *AS* cells. In 17 percent oxygen (*left*) the cell membrane remains intact and the potassium level is adequately maintained in either a normal physiological (low potassium) medium (*top*) or in elevated potassium (*bottom*). In a low-oxygen, low-potassium medium (*top right*) *AS* and *SS* cells sickle; their membranes are disrupted, they lose potassium and their parasites shrivel up (*AS*) or disappear (*SS*). Incubation in a high-potassium medium (*bottom right*) protects parasites in *AS* cells in spite of sickling and membrane disruption, but it does not protect parasites in *SS* cells. The cause of parasite death in *AS* cells, then, is deprivation of an essential metabolic factor, potassium. In *SS* cells, on the other hand, hemoglobin-*S* aggregates destroy parasites by physical penetration.

in which a heterozygote advantage coupled with a homozygote disadvantage maintains a variant gene at a low but consistent level in a population.

The mechanisms of life and death are rooted in chemistry and molecular biology. They can be explicated, to some extent, in the laboratory. Their ultimate effects on the human species are decided, however, not in the laboratory or even in the cells of individual human beings but slowly and unpredictably in evolutionary contests waged across continents and over millenniums.



CHAIN OF EVENTS depicted here may protect *AS* heterozygotes against malaria. A parasite-infected cell is characterized by knobs on its surface and by a low intracellular *p*H. An uninfected *AS* cell will pass through a short period of low oxygenation in a capillary without sickling (*1*); an infected cell, on the other hand, will be sequestered long enough so that the low oxygen level and the low *p*H cause it to sickle (*2*). Sickling causes the cell membrane to leak potassium (*3*). Deprived of potassium, the parasite dies (*4*). The death of some fraction of the infecting parasites may give the heterozygote's body time to develop its own immune response.

Learn how to use a computer for $249. And keep the computer.

Introducing the Computer Learning Lab, with the Sinclair ZX80 personal computer.

The Sinclair ZX80 is a powerful, personal computer already in use in tens of thousands of homes and businesses.

Now, for just $249*, the Computer Learning Lab will teach you and your children how to use this complete, expandable computer right in your own home. And when you're done, you get to keep the computer!

The Computer Learning Lab is a self-paced course that teaches the ins and outs of computers. It cuts away jargon and mystique, taking you straight into BASIC—the most common, easy-to-use computer language. And the computer itself does the teaching.

100 LESSONS MAKE LEARNING EASY.

You just take the ZX80 out of the box, connect it to your TV and an ordinary cassette recorder (connectors are provided), and slip in the pre-programmed Learning Lab cassettes. There's nothing extra to purchase.

You'll be working with the computer your very first day!

The cassettes take you through 100 experiments that teach you how to solve problems with the ZX80.

You learn by doing. By actually working with the computer.

And the lessons are designed to be fun and involving. You create your own programs for games, code breaking, interest calculation, and other topics. Then you can apply the principles you've learned to more complex problems.

In fact, you'll be a master of the ZX80. To use in your business, for home budgeting, or just for fun.

We'll also send you a catalog full of ready-to-use programs—from Loan Amortization to Lunar Lander—available for as little as $6.95.

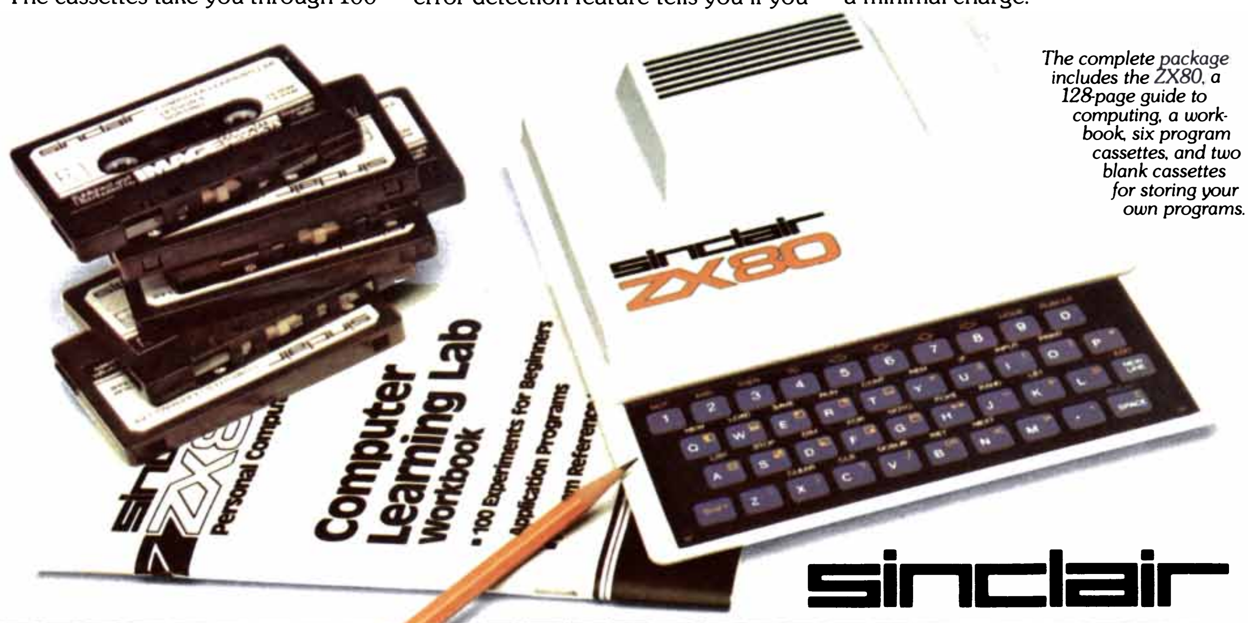
THE ADVANCED DESIGN OF THE ZX80.

The ZX80 is the world's first truly portable computer. It features a touch-sensitive keyboard and a 32-character by 24-line display.

And it performs like a much larger and more expensive computer. Single keystroke entries make typing programs fast and easy. An automatic error detection feature tells you if you make mistakes. And program editing helps you correct them. Yet the complete Computer Learning Lab, including computer, is still several hundred dollars less expensive than any comparable computer alone.

The Computer Learning Lab is a family learning aid. Children 10 and above will quickly understand the principles of computing—and have fun learning.

The ZX80 is backed by a 30-day money-back guarantee and a 90-day limited warranty with a national service-by-mail facility. Extended service contracts for the ZX80 are available for a minimal charge.



The complete package includes the ZX80, a 128-page guide to computing, a workbook, six program cassettes, and two blank cassettes for storing your own programs.

MASTERCHARGE OR VISA ORDERS CALL: (203) 265-9171. We'll refund the cost of your call.
INFORMATION: General and technical (617) 367-1898 or 367-1988. Phones open Mon.–Fri. 8 AM to 8 PM EST.

To: SINCLAIR RESEARCH LTD., 475 MAIN ST., P.O. BOX 3027, WALLINGFORD, CT 06492.

Please send me ________ Computer Learning Lab(s) with the ZX80 personal computer at $249 each (U.S. dollars).

Please send me ________ ZX80 computer(s) only, without the Learning Lab, at $199.95* each.

I enclose check/money order payable to Sinclair Research Ltd. for $______ (Add $5.00 for shipping.)

*For Conn. deliveries, add sales tax.

Name ____________________

Address ____________________

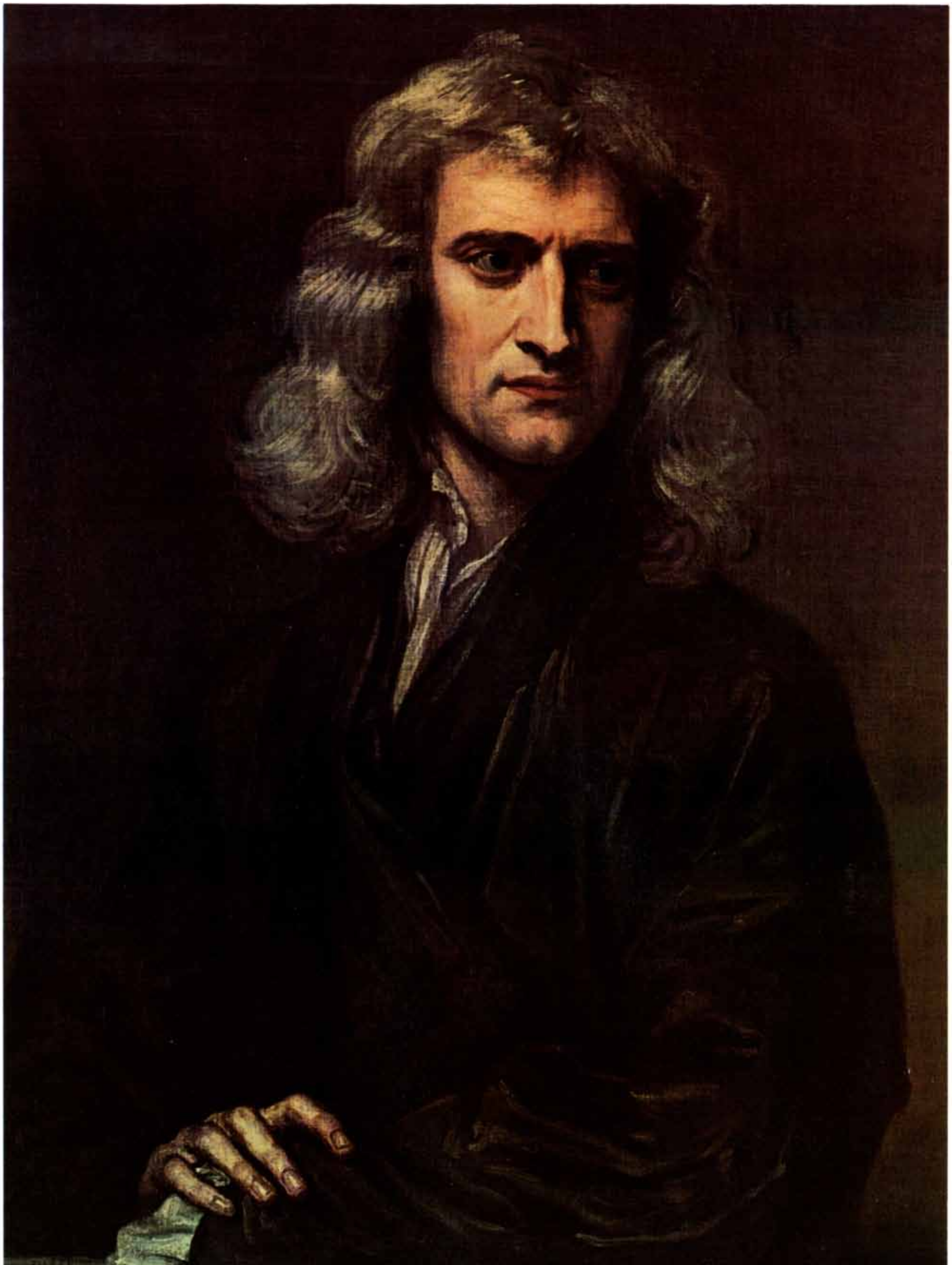
City ____________ State ________ Zip ________

Occupation ____________________ Age ______

Intended use of ZX80: ____________________

Have you ever used a computer? ☐ Yes ☐ No Do you own another personal computer? ☐ Yes ☐ No

SA-3-1



PORTRAIT OF ISAAC NEWTON was painted by Godfrey Kneller in 1689, when Newton was 46. Four years earlier Newton had developed the concept of universal gravitation. Newton's principal work *Philosophiae Naturalis Principia Mathematica* was published in 1687.

Newton's Discovery of Gravity

How did he come to develop the concept that marked the beginning of modern science? In essence he did so by repetitively comparing the real world with a simplified mathematical representation of it

by I. Bernard Cohen

The high point of the Scientific Revolution was Isaac Newton's discovery of the law of universal gravitation: All objects attract each other with a force directly proportional to the product of their masses and inversely proportional to the square of their separation. By subsuming under a single mathematical law the chief physical phenomena of the observable universe Newton demonstrated that terrestrial physics and celestial physics are one and the same. In one stroke the concept of universal gravitation revealed the physical significance of Johannes Kepler's three laws of planetary motion, solved the thorny problem of the origin of the tides and accounted for Galileo Galilei's curious and unexplained observation that the descent of a free-falling object is independent of its weight. Newton had achieved Kepler's goal of developing a physics based on causes.

The momentous discovery of universal gravitation, which became the paradigm of successful science, was not the result of an isolated flash of genius; it was the culmination of a series of exercises in problem solving. It was a product not of induction but of logical deductions and transformations of existing ideas. The discovery of universal gravity brings out what I believe is a fundamental characteristic of all great breakthroughs in science from the simplest innovations to the most dramatic revolutions: the creation of something new by the transformation of existing notions.

Newton developed the concept of universal gravity in the first few months of 1685, when he was 42. Physicists have usually made their greatest contributions at a much earlier age, but Newton was still in what he called "the prime years of my life for invention." The documents that have enabled me to date the discovery also make it possible to reconstruct the process that led to it.

A decisive step on the path to universal gravity came in late 1679 and early 1680, when Robert Hooke introduced Newton to a new way of analyzing motion along a curved trajectory. Hooke had cleverly seen that the motion of an orbiting body has two components, an inertial component and a centripetal, or center-seeking, one. The inertial component tends to propel the body in a straight line tangent to the curved path, whereas the centripetal component continuously draws the body away from the inertial straight-line trajectory. In a stable orbit such as that of the moon the two components are matched, so that the moon neither veers away on a tangential path nor spirals toward the earth.

The concept of a centripetal force replaced the older and misleading notion of a centrifugal, or center-fleeing, force. René Descartes and Christiaan Huygens had analyzed curved motion in terms of such a centrifugal force. Descartes, for example, had investigated the movement of a ball on the inner surface of a hollow cylinder and the movement of water in a bucket swung in a circle. The ball and the water seemed to flee the center of the system, and so Descartes attributed their motion to the influence of a centrifugal force. It is now clear there is no such force; a center-fleeing force cannot be traced to the interaction of physical objects. The illusion of a centrifugal force comes about when a moving object is viewed from a rotating frame of reference.

With the change in outlook from centrifugal to centripetal force came an appreciation of the fundamental role of the central body. The centrifugal analysis had focused on the revolving object, whose "endeavor to recede" from the center seems to be independent of the properties of the central body. The concept of a centripetal force, in contrast, depends fundamentally on the central body, toward which the revolving object is impelled or attracted. The interaction of the central, attracting body with the revolving, attracted object can obviously be expected to have a part in any theory of gravitation.

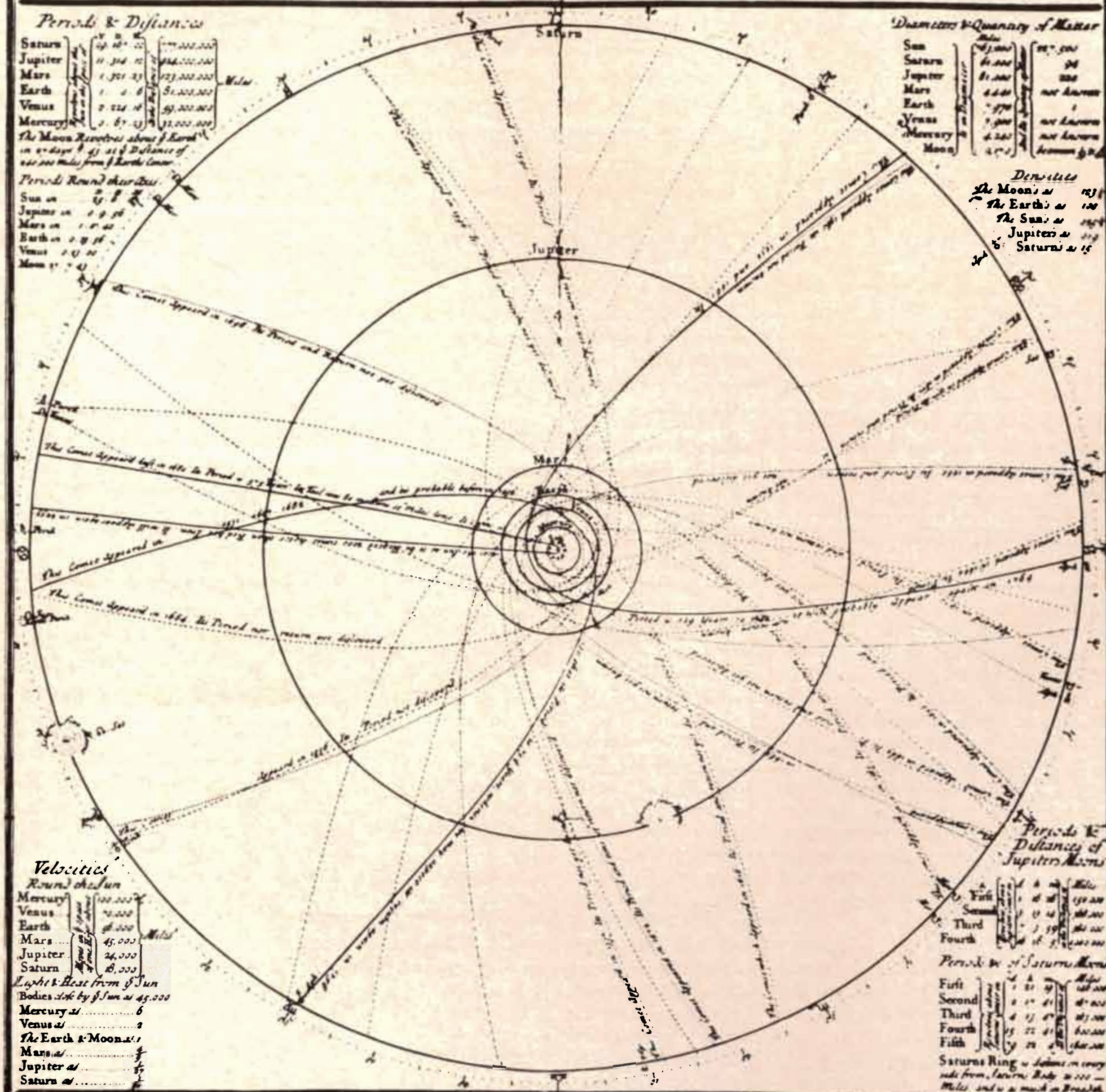
Hooke's analysis of curved motion may seem to be such an obvious and immediate consequence of the Cartesian principle of inertia that Newton would not have needed Hooke to instruct him on the subject as late as 1679. Newton had more or less accepted the inertial principle some 20 years earlier. Nevertheless, Newton, like Descartes and Huygens, was so mired in the concept of centrifugal endeavor that the full implications of inertial physics were far from obvious to him.

On November 24, 1679, Hooke wrote to Newton suggesting that they engage in a private "philosophical" correspondence on scientific topics of mutual interest. Six years earlier they had clashed publicly over Newton's experiments and theories on the prismatic dispersion of light and on the nature of color. Hooke was only one of several investigators who had rejected Newton's optical theories. Newton was so vexed at having to defend his work that he vowed to abandon "philosophy" (physical science) because she was "so litigious a lady" that a man who had anything to do with her would have to spend the rest of his life defending his opinions.

Hooke had since become secretary of the Royal Society of London. In spite of the earlier controversy his letter to Newton was friendly and gracious. The letter invited Newton to comment on any of Hooke's hypotheses or opinions, particularly on the notion of "compounding the celestiall motions of the planetts [out] of a direct motion by the tangent & an attractive motion towards the central body." This sentence was apparently Newton's introduction to the idea of decomposing curved motion into an inertial component and a centripetal one. There is no evidence that he had yet reached Hooke's level of understanding of circular motion. Indeed, Newton still often spoke of orbital motion in terms of centrifugal force.

In his letter Hooke ventured the suggestion that the centripetal force drawing a planet toward the sun varies inversely as the square of the separation. At this point Hooke was stuck. He could not see the dynamical consequences of his own deep insight and therefore could not make the leap from intuitive hunch and guesswork to exact science. He

M^r. WHISTON'S SCHEME of the SOLAR SYSTEM EPITOMIS'D. To w^ch is annex'd A Translation of part of y^e General *Scholium* at y^e end of y^e second Edition of S^r. *Isaac Newton's* Principia. *Concerning God.*



THIS most Elegant System of the Planets and Comets could not be produced but by and under the Contrivance and Dominion of an Intelligent and Powerful Being. And if the Fixed Stars are the Centers of such other Systems, all these being Framed by the like Council will be Subject to the Dominion of *One*, especially seeing the Light of the Fixed Stars is of the same Nature with that of the Sun, and the Light of all these Systems passes mutually from one to another. He governs all things, not as y^e Soul of the World, but as the Lord of the Universe, and because of his Dominion he is wont to be called Lord God παντοκράτωρ (i. e. Universal Emperor) for God is a Relative word, and has Relation to Servants: And the Deity is the Empire of God, not over his own Body (as is the Opinion of those, who make him the Soul of the World) but over his Servants. The *Supreme God* is a Being Eternal, Infinite Absolutely Perfect; but a Being however Perfect without Dominion, is not *Lord God*: For we say, *my God, your God, the God of Israel*, but we do not say, *my Eternal, your Eternal, the Eternal of Israel*; we do not say *my Infinite, your Infinite, the Infinite of Israel*; we do not say *my Perfect, your Perfect, the Perfect of Israel*. These Titles have no Relation to Servants. The word God frequently signifies Lord, but every Lord is not God. The Empire of a Spiritual being constitutes God, true Empire constitutes True God, Supreme the Supreme, Feigned the Feigned. And from his true Empire it follows that the true God is Living Intelligent & Powerful, from his other Perfections that he is the Supreme or Supremely Perfect. He is *Eternal & Infinite, Omnipotent* and *Omniscient* that is, he endures from Eternity to Eternity, and he is present from Infinity to Infinity, he Governs all Things, and Knows all Things which are or which can be known. He is not Eternity or Infinity, but he is Eternal and Infinite, he is not Duration or Space, but he Endures and is Present. He endures always and is present everywhere and by existing always and everywhere, he Constitutes Duration and Space, Eternity and Infinity. Whereas every Particle of Space is *always*, and every Individual Moment of Duration is *every where*, certainly the Framer and Lord of the Universe shall not be (*nunquam nusquam*) *never nowhere*. He is Omnipresent not Virtually only, but also Substantially, for Power without Substance cannot Subsist. In him are contain'd and moved all things (so the Antients thought ...) but without mutual Passion. God suffers nothing from the Motions of Bodies: Nor do they suffer any Resistance from the Omnipresence of God. It is confess'd that the Supreme God exists Necessarily, and by the same Necessity he is *always* and *every where*. Whence also he is wholy Similar, all Eye all Ear, all Brain, all Arm, all the Power of Perceiving Understanding and Acting. But after a manner not at all Corporeal, after a manner not like that of Men, after a manner wholly to us unknown. As a Blind Man has no notion of Colours, so neither have we any notion of the manner how the most Wise God perceives and understands all things. He is wholly destitute of all Body and of all Bodily shape and therefore cannot be seen, heard, nor touched; nor ought to be Worship ed under the Representation of any thing Corporeal. We have Ideas of his Attributes, but we know not at all what is the Substance of any thing whatever. We see only the Figures and Colours of Bodies, we hear only Sounds, we touch only the outward Surfaces, we smell only Odours and tast Tasts, but we know not by any sence or reflex Act the inward Substances and much less have we any Notion of the Substance of God. We know him only by his Properties and Attributes and by the most Wise and Excellent Structure of things, and by Final Causes, but we Adore and Worship him on account of his Dominion. For God without Dominion, Providence & Final Causes is nothing else but Fate and Nature

Engrav'd and Sold by John Senex at the Globe in Salisbury Court near Fleetstreet. Where may be had D^r. Halley's Scheme of the Total Eclipse of the Sun which will be in 1724 ... his Zodiack containing all the Stars in the ...

could go no further because he lacked both the mathematical genius of Newton and an appreciation of Kepler's law of areas, which figured prominently in Newton's subsequent approach to celestial dynamics. The law of areas states that the radius vector from the sun to a planet sweeps out equal areas in equal times.

On November 28 Newton wrote to Hooke that before reading Hooke's letter of the 24th he did not "so much as heare (that I remember) of your Hypotheses of compounding the celestial motions of the Planets of a direct motion by the tangent to the curve" and an "attractive" motion toward the sun. Having admitted that Hooke's analysis was new to him, Newton immediately changed the subject to a fancy of his own: the effect of the earth's rotation on a free-falling object. If a dropped object could pass through the rotating earth, what path would the object take? Newton had incorrectly concluded that it would follow a spiral trajectory.

In Hooke's next letter, dated December 9, he caught Newton's error and pointed out that the path "would resemble an Elleipse." Hooke was eager to get Newton going on the problem of planetary motion, and so he suggested that the correct description of an object falling through the earth and his own analysis of planetary motion were both cases of "Circular motions compounded by a Direct motion and an attractive one to a center."

On December 13, 1679, Newton responded guardedly to Hooke's correction but did not comment on his proposed analysis of circular motion. Hooke did not give up. In a letter written on January 6, 1680, he returned to his thesis about curved motion and repeated the quantitative supposition that the centripetal attraction is inversely proportional to the square of the distance. From this supposition Hooke concluded that the velocity of the revolving body is inversely proportional to the distance from the center. He then pointed out that his analysis "doth very Intelligibly and truly make out all the Appearances of the Heavens." Newton did not reply.

On January 17 Hooke sent a short supplementary letter in which he wrote: "It now remaines to know the proprietys of a curve Line (not circular nor concentricall) made by a centrall attractive power which makes the velocitys of Descent from the tangent Line or equall straight motion at all Distances in a Duplicate proportion reciprocally taken." In modern terminology Hooke's problem can be paraphrased as follows: If a central attractive force causes an object to fall away from its inertial path and move in a curve, what kind of curve results if the attractive force varies inversely as the square of the distance? He concluded: "I doubt not but that by your excellent method you will easily find out what that Curve must be, and its proprietys, and suggest a physicall Reason of this proportion."

Newton evidently did do almost that. He proved that an ellipse would satisfy the conditions outlined by Hooke. Nevertheless, he did not communicate the result of this proof to Hooke or to anyone else until August, 1684, when he was visited by Edmund Halley, the astronomer and mathematician. Halley came to see Newton in order to ask "what he thought the Curve would be that would be described by the Planets, supposing the force of attraction towards the Sun to be reciprocal to the square of their distance from it." The problem had been much discussed by the Royal Society. Halley and Christopher Wren were unable to solve it, and Hooke never produced a solution, although he maintained he had found one.

When Newton heard the question, he responded immediately: an ellipse. Halley asked him how he knew and Newton replied: "I have calculated it." Newton apparently could not find the calculations, but at Halley's urging he wrote them up for the Royal Society in the small tract *De Motu* (*Concerning Motion*). In *De Motu* Newton described his work on terrestrial and celestial dynamics, including his ideas on motion in free space and in a resistive medium. Newton must have finished *De Motu* by December 10, 1684, because Halley told the Royal Society then that Newton had recently shown him the curious treatise.

The exact progression of Newton's ideas in the time between his correspondence with Hooke and his completion of the first draft of *De Motu* is not documented. Nevertheless, I am certain it was Hooke's method of analyzing curved motion that set Newton on the right track. Although not all historians would agree with me, I believe the approach Newton takes to terrestrial and celestial dynamics in *De Motu*, which he further developed the following spring in the first book of the *Philosophiae Naturalis Principia Mathematica*, represents his thinking on planetary dynamics inspired by his correspondence with Hooke. In a few autobiographical manuscripts Newton said the correspondence either preceded or coincided with his demonstration published first in *De Motu* and then in the *Principia* that an object that has an inertial motion and is subject to an inverse-square centripetal force moves in an elliptical orbit.

It was this demonstration that brought out the physical significance of Kepler's law of elliptical orbits (the law stating that each planet moves in an elliptical path with the sun at one focus of the ellipse). The modern reader may be surprised that it was not Kepler but Newton who revealed the fundamental nature of Kepler's laws of planetary motion. Before the publication of the *Principia*, however, these laws (which were even called hypotheses) were not as highly respected as they came to be afterward.

Kepler's law of areas in particular had a diminished status in the 17th century. Most astronomical works did not even mention it. For example, Thomas Streete's *Astronomia Carolina*, from which Newton copied Kepler's third law (The cube of the average distance of a planet from the sun is proportional to the square of the orbital period), never discusses the law of areas or hints at its existence. Most 17th-century astronomers calculated planetary positions not by the law of areas but by a construction based on a uniformly rotating vector emanating from the empty focus of the planet's elliptical orbit [*see top illustration on page 174*]. Since astronomers rarely employed the law of areas, it required extraordinary perception for Newton to see its significance. Newton was the one who elevated Kepler's law of areas to the status it enjoys today.

The very first proposition of the *Principia* (and the discussion at the beginning of *De Motu*) develops the dynamical significance of the law of areas by proving that the curved motion described by the law is a consequence of centripetal force. The proof, which has three parts, shows how well Newton had learned Hooke's technique of decomposing curved motion into an inertial component and a centripetal one.

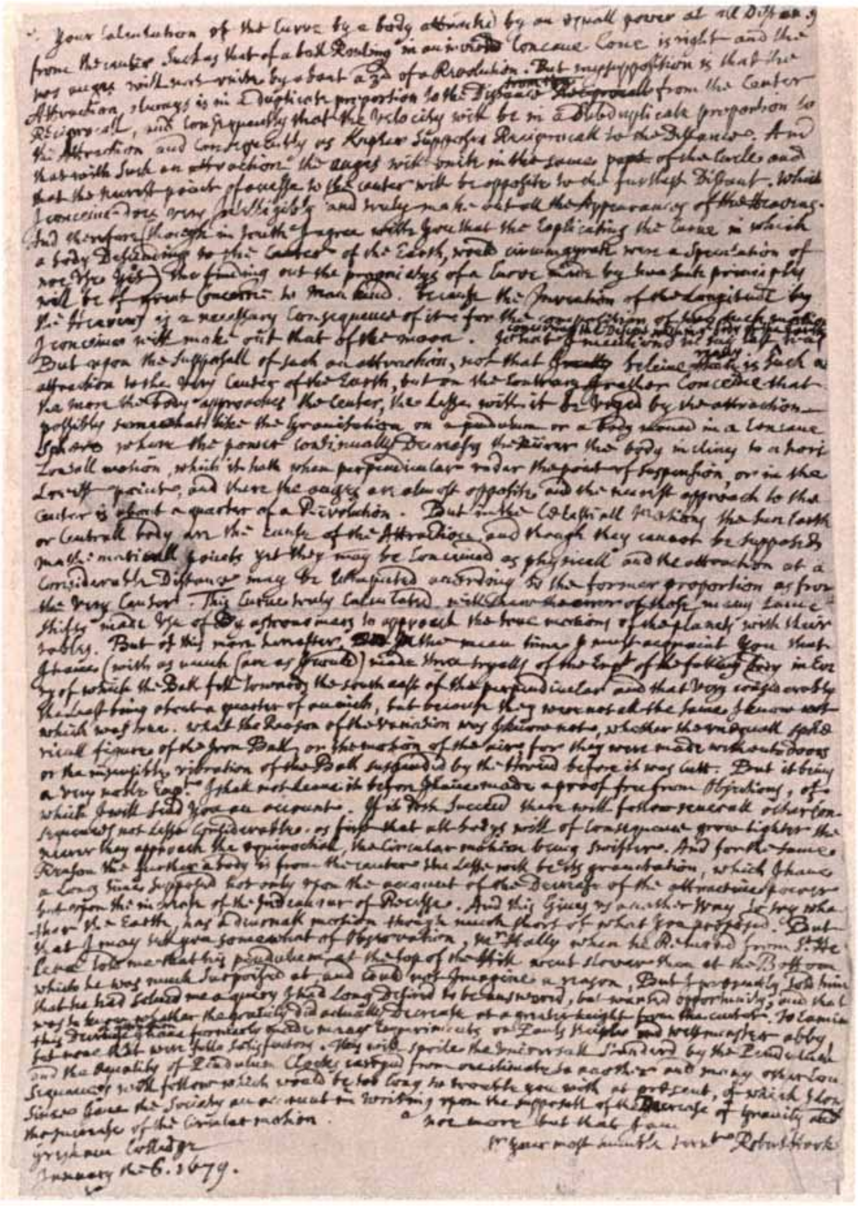
In the first part of the proof Newton considers a body moving along a straight line with a constant velocity. The line is divided into equal intervals to indicate that the body moves equal distances in equal times. A point P is chosen at a distance h above the line of motion. The triangles formed by connecting P to any of the equal intervals all have the same area because they have equal bases and the same altitude h. By this simple analysis Newton revealed

NEWTONIAN SYSTEM OF THE WORLD was diagrammed by William Whiston, who succeeded Newton as Lucasian Professor at the University of Cambridge. The diagram is from Whiston's broadside "Scheme of the Solar System Epitomis'd," published in 1724. The planets and the satellites of Jupiter and Saturn are shown orbiting the sun under the action of universal gravity. Remarkably, Whiston also included the orbits of comets. Newton had shown that orbits of comets are ellipses or parabolas in which a vector from the sun to the comet sweeps out equal areas in equal times. Below the diagram is Whiston's translation of part of the final General Scholium of the *Principia* (which is from the second edition, published in 1713). There Newton wrote that "This most Elegant System of the Planets and Comets could not be produced but by and under the Contrivance and Dominion of an Intelligent and Powerful Being."

an unexpected relation between inertial motion and the law of areas.

In the second part of the proof the body moves as before initially, but at the end of the second interval it receives an impulsive force—a blow—toward *P*. Therefore in the third interval the body no longer moves along the original straight line but rather along another straight line closer to *P*. Newton again showed by geometry that the triangle formed by connecting *P* to the ends of the trajectory traced in the second interval has the same area as the triangle formed by connecting *P* to the ends of the trajectory traced in the third interval.

In the third part the body is given a blow toward *P* at the end of each interval. As a result the body moves in a polygonal path around *P*. Again the area relation holds. In the limiting case where the interval between blows approaches zero the body is subject to a continuous force directed toward *P* and the polygonal path becomes a smooth curve or orbit. In this way Newton proved that a centripetal force generates a curve according to the law of areas.



LETTER TO NEWTON from Robert Hooke includes Hooke's views on the analysis of motion along a curved trajectory. (The letter is dated January 6, 1679, according to a version of the Julian calendrical system in which the year started in March; the modern calendrical system puts the date at January 6, 1680.) In the second sentence Hooke proposes that "the Attraction is always in a duplicate proportion to the Distance from the Center Reciprocall" (that is, the attraction is inversely proportional to the distance squared). As a result "the Velocity will be in a subduplicate proportion to the Attraction and consequently as Kepler supposes Reciprocall to the Distance." Hooke states that this analysis explains "all the Appearances of the Heavens." He stresses the importance of "finding out the proprietys" of curves because longitudes, which are "of great Concerne to Mankind," can be derived from the moon's curved motion.

The second proposition of the *Principia* proves the converse: Motion in a curve described by the law of areas implies a centripetal force. With these two propositions Newton demonstrated that the law of areas is a necessary and sufficient condition for inertial motion in a central-force field.

The two propositions are part of a sequence of demonstrations that begins with the law of areas and ends with a proof that an elliptical orbit requires an inverse-square centripetal force. This sequence of demonstrations, presented both in the *Principia* and in *De Motu*, marks a profound discontinuity in the history of the exact sciences. The demonstrations introduced a radically new celestial dynamics based on new concepts of force, momentum, mass and inertia and a wholly novel quantitative measure of dynamical force. The subtitle of Kepler's *Astronomia Nova* set the goal of creating a "celestial physics based on causes." Newton achieved this goal, of which Kepler had had only a visionary glimpse. Neither Galileo nor Descartes had conceived of such a celestial dynamics. And the Newtonian formulation left even the great physicist Huygens far behind.

From the early draft of *De Motu* that Newton probably wrote in November, 1684, it is clear he had not yet developed the concept of universal gravitation. The draft discusses centripetal force directed toward the focus of an ellipse and concludes with the scholium, "Therefore the major planets revolve in ellipses having a focus in the center of the sun, and radii drawn [from the planets] to the sun describe areas proportional to the times, entirely as Kepler supposed...."

Newton neither proved this scholium nor continued to believe it for long, and strictly speaking it is false. As he soon realized, the planets do not move according to the law of areas in simple Keplerian elliptical orbits with the sun at a focus. Instead the focus lies in the common center of mass because not only does the sun attract each planet but also each planet attracts the sun (and the planets attract one another). If Newton had already formulated his principle of universal gravitation, he would not have proposed the erroneous scholium.

Newton quickly realized he had not proved that the planets move precisely according to the law of elliptical orbits and the law of areas. He had only found that the laws hold for a one-body system: a single point mass moving with an initial component of inertial motion in a central-force field. He recognized that the one-body system corresponds not to the real world but to an artificial situation that is easier to investigate mathematically. The one-body system reduces the earth to a point mass and the sun to an immobile center of force.

WHAT'S BETTER THAN SPEED READING?

SPEED LEARNING

(SPEED PLUS COMPREHENSION)

Speed Learning is replacing speed reading. It's easy to learn...lasts a lifetime...applies to everything you read...and is the only accredited course with the option of college or continuing education credits.

Do you have too much to read and too little time to read it? Do you mentally pronounce each word as you read? Do you frequently have to go back and re-read words or whole paragraphs you just finished reading? Do you have trouble concentrating? Do you quickly forget most of what you read?

If you answer "yes" to any of these questions — then here at last is the practical help you've been waiting for. Whether you read for business or pleasure, school or college, you will build exceptional skills from this major breakthrough in effective reading, created by Dr. Russell Stauffer at the University of Delaware.

Not just "speed reading" — but speed reading-thinking-understanding-remembering-and-learning

The new *Speed Learning Program* shows you step-by-proven-step how to increase your reading skill and speed, so you understand more, remember more and use more of everything you read. The typical remark made by the 75,000 slow readers who completed the *Speed Learning Program* was: "Why didn't someone teach me this a long time ago?" They were no longer held back by the lack of skills and poor reading habits. They could read almost as fast as they could think.

What makes Speed Learning so successful?

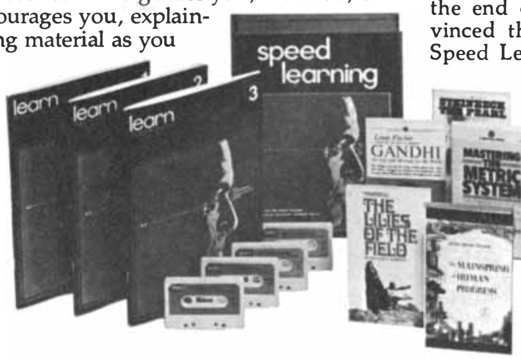
The new *Speed Learning Program* does not offer you a rehash of the usual eye-exercises, timing devices, costly gadgets you've probably heard about in connection with speed reading courses or even tried and found ineffective.

In just a few spare minutes a day of easy reading and exciting listening, you discover an entirely new way to read and think — a radical departure from anything you have ever seen or heard about. Research shows that reading is 95% *thinking* and only 5% eye movement. Yet most of today's speed reading programs spend their time teaching you rapid eye movement (5% of the problem) and ignore the most important part (95%) *thinking*. In brief, *Speed Learning* gives you what speed reading *can't*.

Imagine the new freedom you'll have when you learn how to dash through all types of reading material *at least* twice as fast as you do now, and with greater comprehension. Think of being able to get on top of the avalanche of newspapers, magazines and correspondence you have to read . . . finishing a stimulating book and retaining facts and details more clearly and with greater accuracy than ever before.

Listen-and-learn at your own pace

This is a practical, easy-to-learn program that will work for you — no matter how slow a reader you think you are now. The *Speed Learning Program* is scientifically planned to get you started quickly . . . to help you in spare minutes a day. It brings you a "teacher-on-cassettes" who guides you, instructs, encourages you, explaining material as you read. Interesting items taken from *Time* Magazine, *Business Week, Wall Street Journal, Family Circle, N.Y. Times* and many others, make the program stimulating, easy and fun . . . and so much more effective.



Executives, students, professional people, men and women in all walks of life from 15 to 70 have benefited from this program. Speed Learning is a fully accredited course . . . costing only 1/5 the price of less effective speed reading classroom courses. Now you can examine the same, easy, practical and proven methods at home . . . in spare time . . . without risking a penny.

Examine Speed Learning FREE for 15 days

You will be thrilled at how quickly this program will begin to develop new thinking and reading skills. After listening to just one cassette and reading the preface you will quickly see how you can achieve increases in both the speed at which you read and in the amount you understand and remember.

You must be delighted with what you see or you pay nothing. Examine this remarkable program for 15 days. If, at the end of that time you are not convinced that you would like to master Speed Learning, simply return the program and owe nothing. See the coupon for low price and convenient credit terms.

Note: Many companies and government agencies have tuition assistance plans for employees providing full or partial payment for college credit programs.

In most cases, the entire cost of your Speed Learning Program is Tax Deductible.

COLLEGE CREDITS

You may obtain *2 full semester hour credits* for course completion, wherever you reside. Credits offered through Whittier College (California). Details included in your program.

CONTINUING EDUCATION UNITS

National Management Association, the world's largest association of professional managers, awards 3.0 CEU's for course completion. CEU's can be applied toward the certificate in Management Studies.

PROFESSIONAL SOCIETIES

Speed Learning is offered internationally to members of professional associations such as: American Chemical Society, Foundation for Accounting Education, Institute of Electrical and Electronics Engineers and dozens more. Consult your Education Director for information.

BUSINESS, INDUSTRY, GOVERNMENT

Many companies and government agencies offer Speed Learning as a wholly-paid or tuition reimbursement program. Consult your Training or Personnel Director for details.

31SAD

learn INCORPORATED 113 Gaither Drive, Mt. Laurel, NJ 08054

YES! Please rush me the materials checked below:

☐ Please send the Speed Learning program @ $99.95 plus $4 postage and handling.
☐ Please send the Speed Learning Medical Edition @ $109.95 plus $4 postage and handling.
☐ Please send the Junior Speed Learning program (ages 11 to 16) @ $89.95 plus $4 postage and handling.

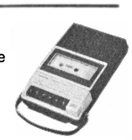
Check method of payment below: NJ residents add 5% sales tax.

☐ Check or money order enclosed (payable to learn incorporated)
☐ Charge my credit card under the regular payment terms
☐ Visa ☐ Master Card Interbank No.______ ☐ American Express
Card No.______ Exp. Date______

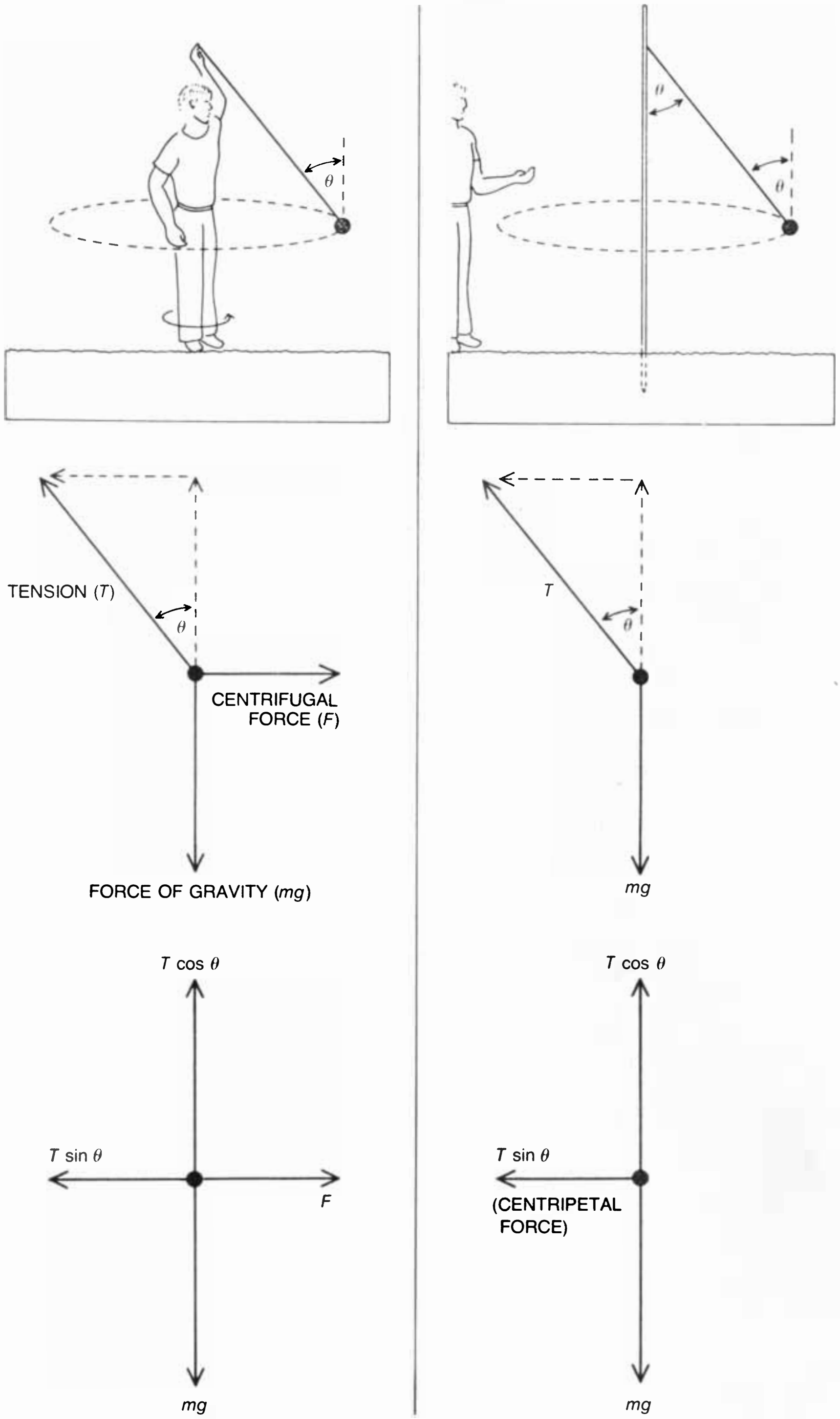
I understand that if after 15 days I am not delighted in every way, that I may return the materials in their original condition for a full refund. No questions asked.

Name ______
Address ______
City______ State______ Zip______
Signature______

If you don't already own a cassette player, you may order this Deluxe Cassette Recorder for only $49.95. (Includes handling and delivery.)
Check here to order ☐



Outside USA add $10 per item—Airmail extra



CENTRIFUGAL FORCE is a fictitious force. The illusion of a centrifugal, or center-fleeing, force can arise when a moving object is viewed from a rotating frame of reference (*left*), as when a ball is swung at the end of a string by an observer who rotates with the same angular speed as the ball. Two known forces act on the ball: the tension of the string and the force of gravity. The ball is not accelerating in the vertical direction, and so all vertical forces acting on it must be in balance; in particular the vertical component of the tension cancels the force of gravity. Since the observer and the ball are rotating together, the ball appears to be at rest and it seems that the horizontal forces should also be in balance. As a result the observer postulates a centrifugal force that cancels the horizontal component of the tension. No such force, however, can be traced to the interaction of physical objects. A different analysis of forces results (*right*) when the ball is rotating in the same way but the observer is at rest. In this stationary frame of reference the observer sees the same vertical forces on the ball as he saw in the rotating frame. In the horizontal direction, however, the ball is not at rest with respect to the observer but is moving in a circle. In other words, the ball accelerates continuously toward the center, so that the horizontal forces should not be expected to balance. The ball is subject to a centripetal, or center-seeking, force which is the horizontal component of the tension of the string. The centripetal force can be traced to the interaction of two physical objects: the string and the ball.

What enabled Newton to transcend the one-body system was his appreciation of the consequences of his third law of motion: the law of action and reaction. This law is perhaps the most original of his three laws of motion (the other two are the law of inertia and the force law). One testimonial to its novelty is that even today it is often employed incorrectly by those who relate it not to an impact situation or to the interaction of bodies but to a supposed condition of equilibrium.

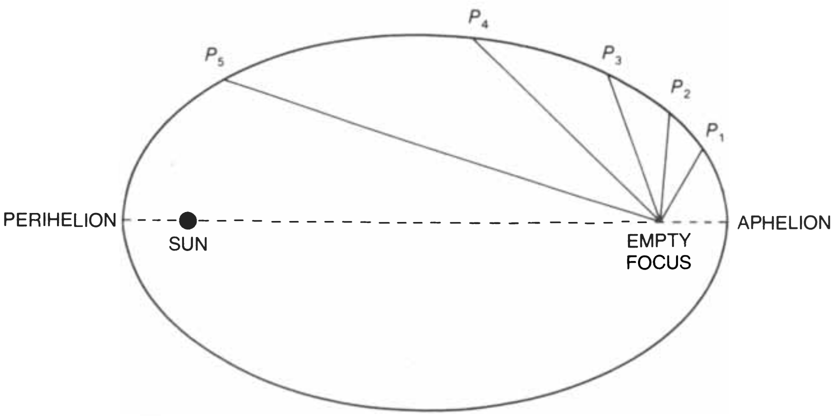
The development of Newton's thinking on action and reaction after he completed the first draft of *De Motu* is set out in the opening sections of the first book of the *Principia.* In the introduction to the 11th section Newton explains that he has confined himself so far to a situation that "hardly exists in the real world," namely the "motions of bodies attracted toward an unmoving center." The situation is artificial because "attractions customarily are directed toward bodies and—by the third law of motion—the actions of attracting and attracted bodies are always mutual and equal." As a result, "if there are two bodies, neither the attracting nor the attracted body can be at rest." Rather, "both bodies (by the fourth corollary of the laws) revolve about a common center, as if by a mutual attraction."

Newton had seen that if the sun pulls on the earth, the earth must also pull on the sun with a force of equal magnitude. In this two-body system the earth does not move in a simple orbit around the sun. Instead the sun and the earth each move about their mutual center of gravity. A further consequence of the third law of motion is that each planet is a center of attractive force as well as an attracted body; it follows that a planet not only attracts and is attracted by the sun but also attracts and is attracted by each of the other planets. Here Newton has taken the momentous step from an interactive two-body system to an interactive many-body system.

In December, 1684, Newton completed a revised draft of *De Motu* that describes planetary motion in the context of an interactive many-body system. Unlike the earlier draft the revised one concludes that "the planets neither move exactly in ellipses nor revolve twice in the same orbit." This conclusion led Newton to the following result: "There are as many orbits to a planet as it has revolutions, as in the motion of the Moon, and the orbit of any one planet depends on the combined motion of all the planets, not to mention the actions of all these on each other." He then wrote: "To consider simultaneously all these causes of motion and to define these motions by exact laws allowing of convenient calculation exceeds, unless I am mistaken, the force of the entire human intellect."

There are no documents that indicate


Suddenly, you're at the wheel of America's most popular sports car.
Suddenly, the moment is Mustang. Driving is a new sensation. Finely tuned excitement. Precision handling.
Suddenly, it's Mustang. Hot stuff.
FORD MUSTANG
FORD DIVISION
Ford



PLANETARY POSITIONS were often found in the 17th century not by Kepler's law of areas but by a construction based on a uniformly rotating radius vector that emanates from the empty focus of a planet's elliptical orbit. The position of a planet (P_1, P_2, P_3, P_4, P_5) at successive moments is the intersection of the ellipse and the vector. Kepler's law of areas states that the radius vector from the sun to a planet sweeps out equal areas in equal times. As a result the planet moves slower at aphelion than at perihelion. The diagrammed construction gives the same result. Correction factors were added to make the construction fit the data more accurately.

how, in the month or so between writing the first draft of *De Motu* and revising it, Newton came to perceive that the planets act gravitationally on one another. Nevertheless, the passage cited above expresses this perception in unambiguous language: "*eorum omnium actiones in se invicem*" ("the actions of all these on each other"). A consequence of this mutual gravitational attraction is that all three of Kepler's laws are not strictly true in the world of physics but are true only for a mathematical construct in which point masses that do not interact with one another orbit either a mathematical center of force or a stationary attracting body. The distinction Newton draws between the realm of mathematics, in which Kepler's laws are truly laws, and the realm of physics, in which they are only "hypotheses," or approximations, is one of the revolutionary features of Newtonian celestial dynamics.

I have assumed that the third law of motion was the key factor in the reasoning that led Newton to suggest mutual gravitational perturbations of planetary orbits. There is no direct evidence for my assumption because no documents exist in which there is an antecedent version of his statement "the actions of all these on each other." Nevertheless, there is strong indirect evidence. In the spring of 1685, a few months after revising *De Motu*, Newton was well on his way to finishing the first draft of the *Principia*. In the initial version of what was to become a second book, "The System of the World," he spelled out the steps that led him to the concept of planetary gravitational interactions. In these steps the third law of motion has the chief role, and I see no reason to believe they are not the same steps that led him to the same concept a few months earlier when he revised *De Motu*.

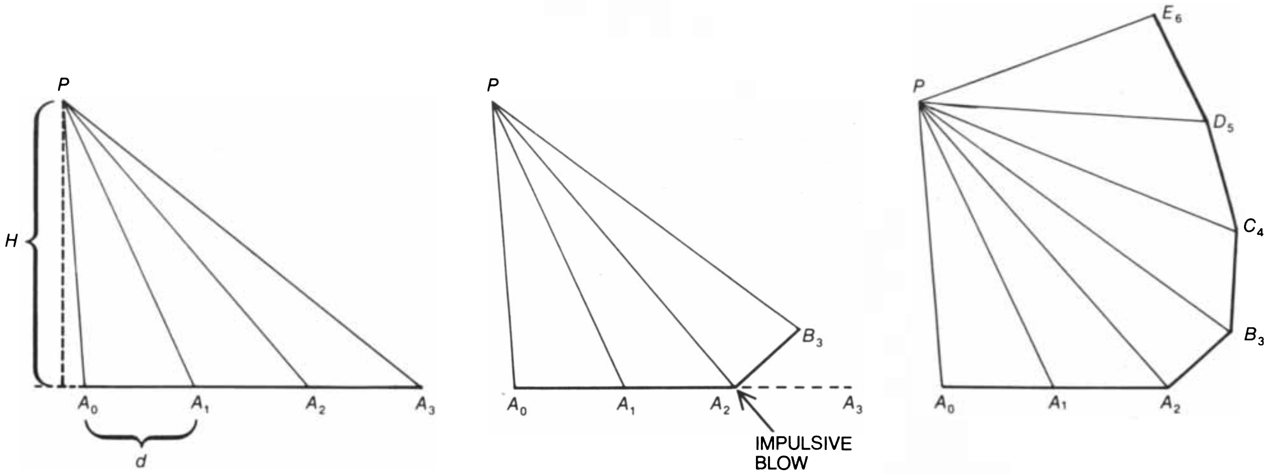
Here are two passages from the first draft of "The System of the World" (translated from the Latin by Anne Whitman and me) that bring out the crucial role of the third law of motion:

"20. *The agreement between the analogies.*

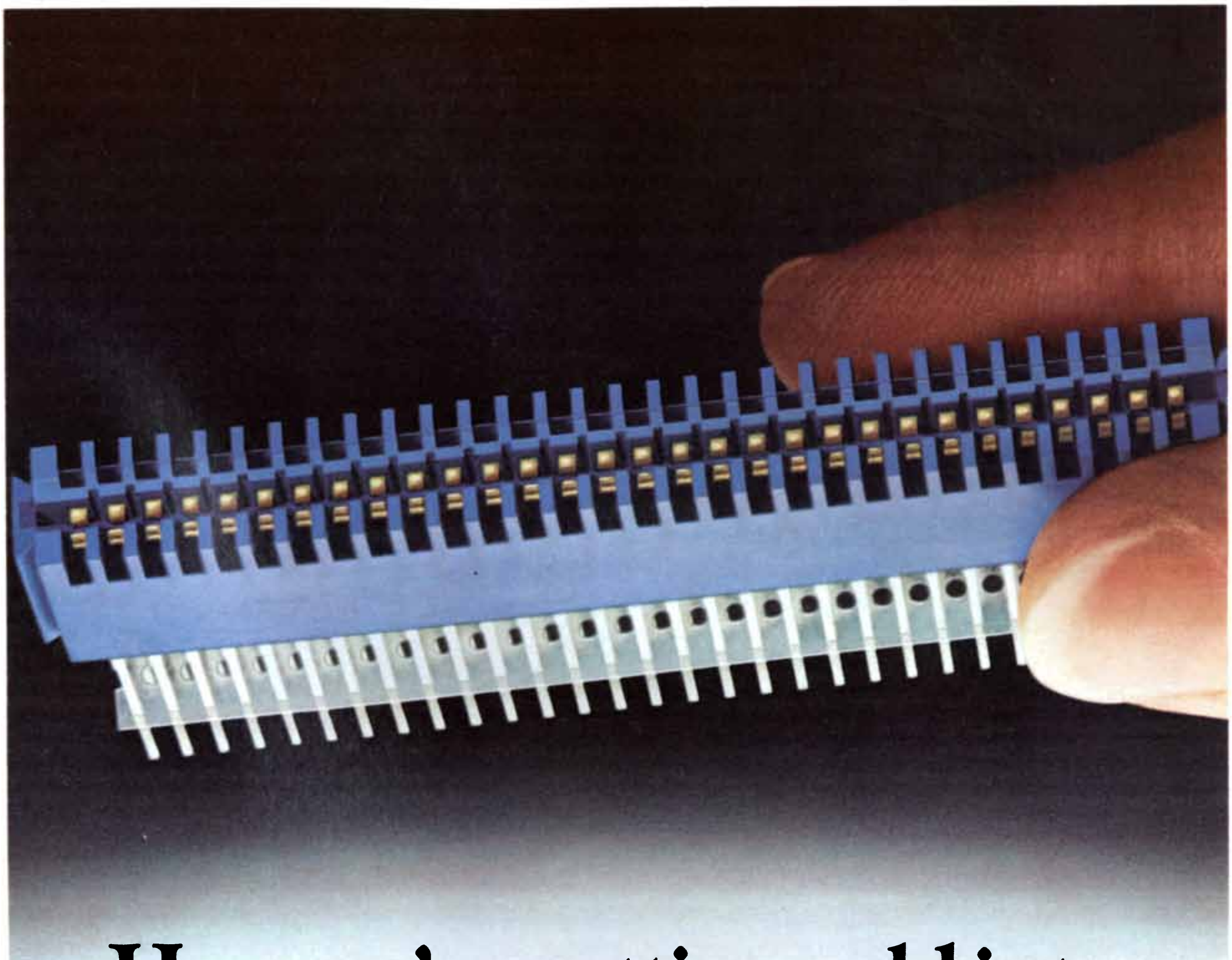
"And since the action of centripetal force upon the attracted body, at equal distances, is proportional to the matter in this body, it is reasonable, too, that it is also proportional to the matter in the attracting body. For the action is mutual, and causes the bodies by a mutual endeavor (by law 3) to approach each other, and accordingly it ought to be similar to itself in both bodies. One body can be considered as attracting and the other as attracted, but this distinction is more mathematical than natural. The attraction is really that of either of the two bodies toward the other, and thus is of the same kind in each of the bodies.

"21. *And their coincidence.*

"And hence it is that the attractive force is found in both bodies. The sun attracts Jupiter and the other planets,



CENTRIPETAL FORCE generates a curved trajectory consistent with the law of areas. This property of a centripetal force was demonstrated by Newton in the first proposition of the *Principia* and in the discussion at the beginning of the short tract *De Motu* (*Concerning Motion*). Newton began (*left*) by considering a body moving in a straight line at a constant speed. The body starts at A_0 and after successive equal intervals reaches first A_1, then A_2 and so on. A point P is chosen above the line of motion. The triangles A_0PA_1, A_1PA_2, A_2PA_3 and so forth all have the same area because they have equal bases and the same altitude. In a second stage of the analysis (*middle*) the body begins as before but at A_2 receives an impulsive blow toward P. Now the body moves along a straight line not to A_3 but to B_3. Newton showed by geometric methods that the triangles A_1PA_2 and A_2PB_3 have the same area. If the body receives a blow toward P at the end of each interval (*right*), it moves in a polygonal path around P. Again triangles can be formed that have the same area. In the limiting case where the time between blows approaches zero the body is subject to a continuous centripetal force directed toward P and the polygonal path becomes a smooth curve. Area is still conserved. This proof brought out the dynamical significance of the law of areas.



How we're putting gold into connectors you can afford.

A lot of companies that depend on electrical connectors are trying to reduce their usage of gold.

They can with AMP.

The reason? We've designed connectors with less gold that perform as well as ever.

One beauty of gold is its unique reliability in electrical contact areas. But the predictability of its price is something else. That's why we developed our own method of selective plating to put the gold only where it's needed. We call it ACCU-PLATE. It's precise. Its advantages work with other metals, as well. And it's saving our customers money.

Providing reliable platings that are cost-effective isn't new to AMP. Over the past decade, we've reduced the gold in our connectors by 33%.

What can we do for you in the '80s?

Whether you call us early with a problem, or simply call us for connectors, you'll get what the world has come to expect from AMP. Connectors and interconnection systems designed for performance. And as the industry leader in research and development, we'll continue to bring you products that reduce the cost of making yours.

AMP has a better way.

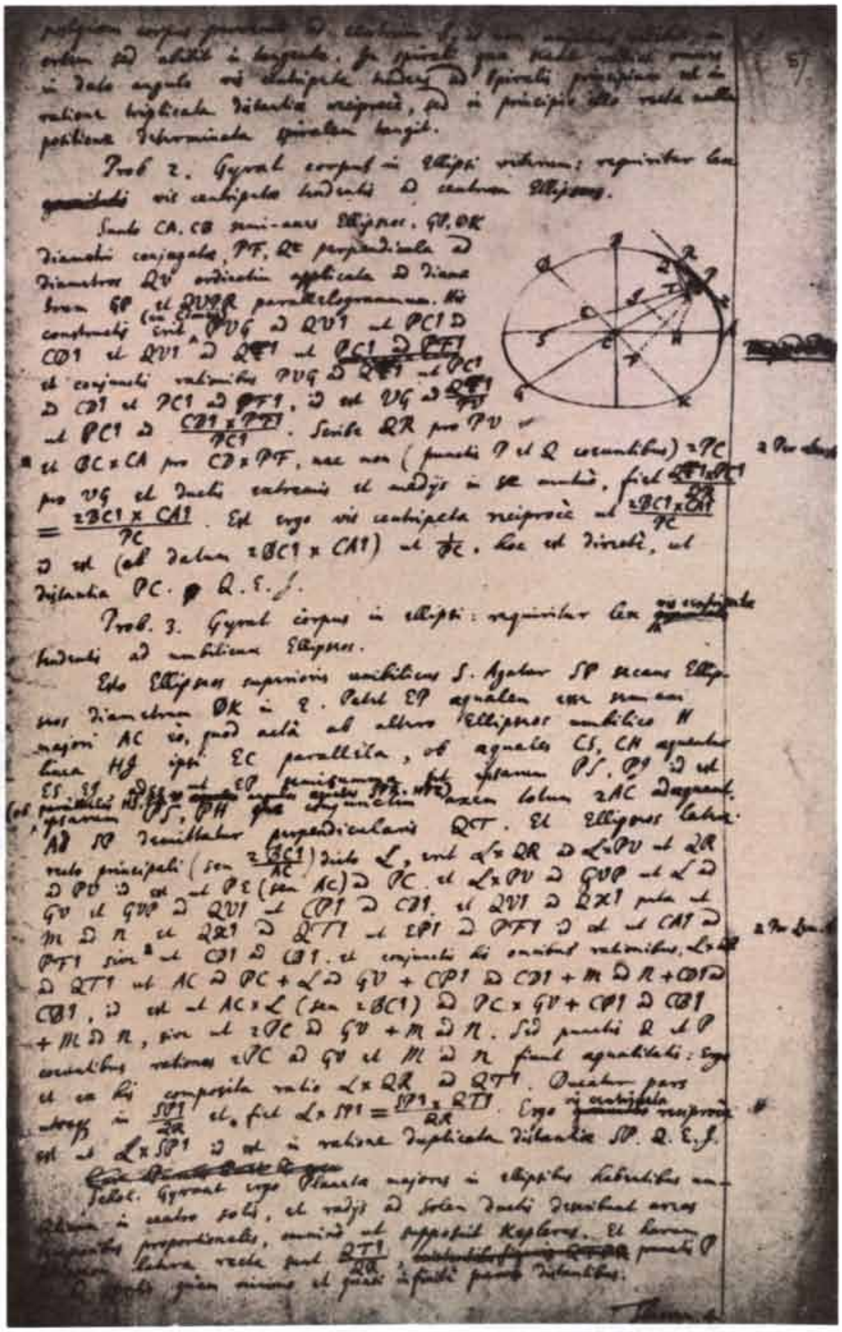
For product innovations, for new applications, for ways to save you money, AMP has better ways to help make your company more productive.

Write for our brochure, "AMP Has A Better Way." AMP Incorporated, Harrisburg, PA 17105.

Barcelona • Brussels • Buenos Aires • Frankfurt • Harrisburg • Helsinki • s-Hertogenbosch • London • Luzern • Mexico City
Montreal • Paris • San Juan • Sao Paulo • Stockholm • Sydney • Turin • Toronto • Tokyo • Valley Forge • Vienna

AMP means productivity™

Jupiter attracts its satellites and similarly the satellites act on one another and on Jupiter, and all the planets on one another. And although the actions of each of a pair of planets on the other can be distinguished from each other and can be considered as two actions by which each attracts the other, yet inasmuch as they are between the same two bodies they are not two but a simple operation between two termini. Two bodies can be drawn to each other by the contraction of one rope between them. The cause of the action is twofold, namely the disposition of each of the two bodies; the action is likewise twofold, insofar as it is upón two bodies; but insofar as it is between two bodies it is



PAGE FROM A DRAFT OF "DE MOTU" that Newton probably wrote in November, 1684, is in his handwriting. In *De Motu* Newton discussed terrestrial and celestial dynamics, including the idea of centripetal force directed toward the focus of an ellipse. The page ends with the scholium, "Therefore the major planets revolve in ellipses having a focus in the center of the sun, and radii drawn [from the planets] to the sun describe areas proportional to the times, entirely as Kepler supposed. . . ." The scholium is false, and the nature of the error indicates that Newton had not yet developed the concept of universal gravitation. As Newton soon realized, the focus of the orbits of the planets is not the sun but the center of mass common to the planets and the sun. Not only does the sun attract each planet but also each planet attracts the sun.

How to reach the people with the longest shopping list in the world.

*科学

****KE XUE**

The People's Republic of China is buying and it's time your company introduced itself.

Our paid circulation of 25,000 includes China's emerging scientific, engineering and industrial leadership; the people who make the future happen. Rarely have advertisers been presented with such a clearly defined and potent audience.

China's shopping list includes most industrial categories: agriculture, energy, aerospace, chemicals, mining, data processing, communications, manufacturing, transportation, resource recovery and heavy machinery. China's ability and inclination to buy will accelerate. Now is the time to present your product, your people and your corporate identity.

SCIENTIFIC AMERICAN is now published in the first languages—English, Italian, Japanese, French, Spanish, German and Chinese—of half the world's population. Worldwide paid circulation exceeds 1,000,000.

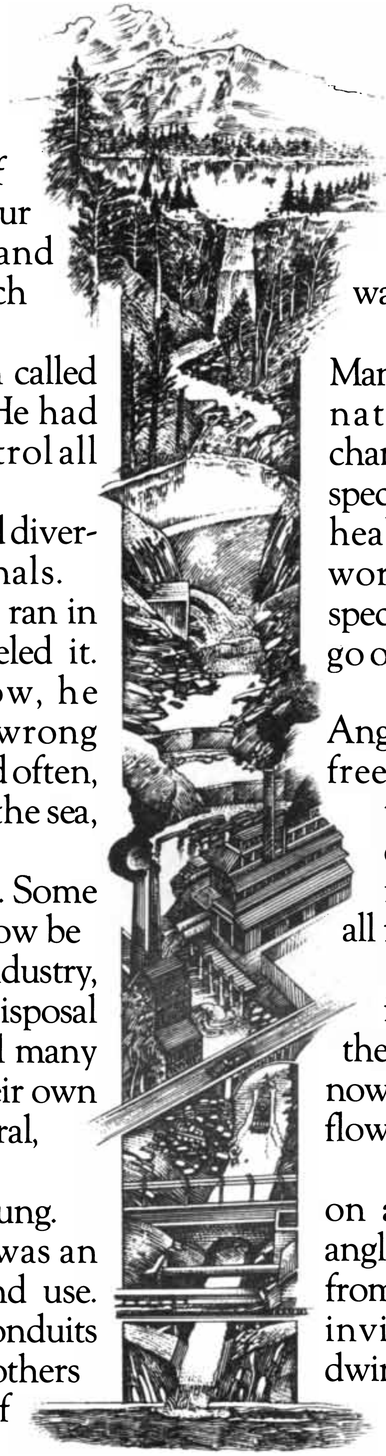
Where will your company be when China makes a purchase? If you want a place on the shopping list, we'll save you a place in the Chinese-language edition of SCIENTIFIC AMERICAN. For more information, contact: Richard Delano, 415 Madison Avenue, New York, NY 10017 (212) 754-0472, or Wallace Burgis, 54 Pall Mall, London SW1Y5JH, England, 011 441 839 1595/6.

***The Chinese Language Edition of Scientific American.**
****English Transliteration; Translation: Exact Knowledge, Science**

SCIENTIFIC AMERICAN

We reach the people who make the future happen.

A FUNNY THING HAPPENED ON THE WAY TO THE OCEAN.



Not so long ago most of the water that fell on our continent ran downhill and back to the ocean, from which it came.

No more. Modern man called that "wasting to the sea." He had other ideas. He would control all that free-running water.

He built dams, dikes, and diversions; levees, locks and canals. Where water inconveniently ran in crooked courses, he channeled it. Where it ran too shallow, he dredged. Where it ran the wrong way, he turned it around. And often, where he did let it return to the sea, he filled it with wastes.

Up to a point it worked. Some of America's waters could now be used for irrigation, cities, industry, power generation, sewage disposal and transportation. And still many of our waters were left to their own devices, providing their natural, less tangible benefits to man.

But then the balance swung.

In many places there was an overkill of manipulation and use. Lakes died. Rivers became conduits of sewage and toxic wastes; others were turned into a series of still ponds. Magnificent streams and awesome canyons were drowned. Fresh waters turned salty.

Fish are among the casualties. Many of them chemically contaminated. Much of their habitat changed—fouled and stilled. Those special conditions necessary for the health and propagation of our worthiest sport and commercial species, gone. As our waters go, so go our fisheries.

The American League of Anglers believes that natural and free-flowing waterways are more than pleasant amenities for fishermen. Their ecological health is essential to man as well as to all forms of wildlife.

We are committed to preserving a proper balance between the need for water manipulation now and the desire for pure, free-flowing waters forever.

We act and speak for this goal on a national level. We urge all anglers, who receive extra benefits from good water, to participate. We invite all those who revere our dwindling stock of natural waters to support our efforts.

THE AMERICAN LEAGUE OF ANGLERS, 810 18TH STREET NORTHWEST, WASHINGTON, D.C. 20006 single and one. There is not, for example, one operation by which the sun attracts Jupiter and another operation by which Jupiter attracts the sun, but one operation by which the sun and Jupiter endeavor to approach each other. By the action by which the sun attracts Jupiter, Jupiter and the sun endeavor to approach each other (by law 3), and by the action by which Jupiter attracts the sun, Jupiter and the sun also endeavor to approach each other. Moreover, the sun is not attracted by a twofold action toward Jupiter, nor is Jupiter attracted by a twofold action toward the sun, but there is one action between them by which both approach each other."

Next Newton concluded that "according to this law all bodies must attract each other." He proudly presented the conclusion and explained why the magnitude of the attractive force is so small that it is unobservable. "It is possible," he wrote, "to observe these forces only in the huge bodies of the planets."

In book three of the *Principia,* which is also concerned with the system of the world but is somewhat more mathematical, Newton treats the topic of gravitation in essentially the same way. First, in what is called the moon test, he extends the weight force, or terrestrial gravity, to the moon and demonstrates that the force varies inversely with the square of the distance. Then he identifies the same terrestrial force with the force of the sun on the planets and the force of a planet on its satellites. All these forces he now calls gravity. With the aid of the third law of motion he transforms the concept of a solar force on the planets into the concept of a mutual force between the sun and the planets. Similarly, he transforms the concept of a planetary force on the satellites into the concept of a mutual force between planets and their satellites and between satellites. The final transformation is the notion that all bodies interact gravitationally.

My analysis of the stages of Newton's thinking should not be taken as diminishing the extraordinary force of his creative genius; rather, it should make that genius plausible. The analysis shows Newton's fecund way of thinking about physics, in which mathematics is applied to the external world as it is revealed by experiment and critical observation. This way of thinking, which I call the Newtonian style, is captured by the English title of Newton's great work: *Mathematical Principles of Natural Philosophy.*

The Newtonian style consists in a repeated give-and-take between a mathematical construct and physical reality. In the development of Newton's ideas on gravity and in his presentation of those ideas in the *Principia,* he started with a mathematical construct that represents nature simplified: a point mass moving around a center of force. Because he did not assume that the construct was an exact representation of the physical universe he was free to explore the properties and effects of a mathematical attractive force even though he found the concept of a grasping force "acting at a distance" to be abhorrent and not admissible in the realm of good physics. Next he compared the consequences of his mathematical construct with the observed principles and laws of the external world such as Kepler's law of areas and law of elliptical orbits. Where the mathematical construct fell short Newton modified it. He made the center of force not a mathematical entity but a point mass. I say a point mass rather than a physical body because he had not yet considered physical properties such as size, shape and mass.

From the modified mathematical construct Newton concluded that a set of point masses circling the central point mass attract one another and perturb one another's orbits. Again he compared the construct with the physical world. Of all the planets, Jupiter and Saturn are the most massive, and so he sought orbital perturbations in their motions. With the help of John Flamsteed, Newton found that the orbital motion of Saturn is perturbed when the two planets are closest together. The process of repeatedly comparing the mathematical construct with reality and then suitably modifying it led eventually to the treatment of the planets as physical bodies with definite shapes and sizes.

After Newton had modified the construct many times he applied it to the system of the world. He asserted that the force of attraction, which he had derived mathematically, is universal gravity. He found that the moon moves as if it were attracted to the earth with a force that is 1/3,600th of the strength of the gravitational force with which the earth pulls on objects at its surface. Since the moon is 60 times farther from the center of the earth than objects on the earth's surface are, the factor of 1/3,600 is consistent with the deduction that the earth's gravity extends to the moon and diminishes with the square of the distance.

The law of universal gravitation explains why the planets follow Kepler's laws approximately and why they depart from the laws in the way they do. It demonstrates why (in the absence of friction) all bodies fall at the same rate at any given place on the earth and why the rate varies with elevation and latitude. The law of gravitation also explains the regular and irregular motions of the moon, provides a physical basis for understanding and predicting tidal phenomena and shows how the earth's rate of precession, which had long been observed but not explained, is the effect of the moon's pulling on the earth's equatorial bulge. Since the mathematical force of attraction works well in explaining and predicting the observed phenomena of the world, Newton decided that the force must "truly exist" even though the received philosophy to which he adhered did not and could not allow such a force to be part of a system of nature. And so he called for an inquiry into how the effects of universal gravity might arise.

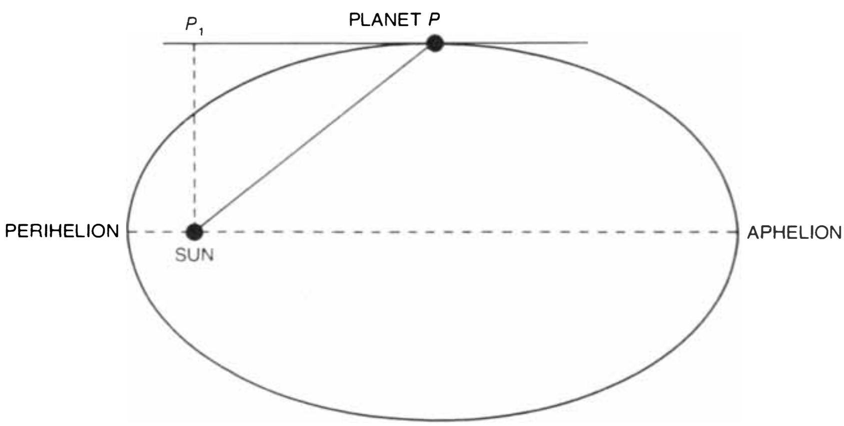
Although Newton at times thought universal gravity might be caused by the impulses of a stream of ether particles bombarding an object or by variations in an all-pervading ether, he did not advance either of these notions in the *Principia* because, as he said, he would "not feign hypotheses" as physical explanations. The Newtonian style had led him to a mathematical concept of universal force, and that style led him to apply his mathematical result to the physical world even though it was not the kind of force in which he could believe.

Some of Newton's contemporaries were so troubled by the idea of an attractive force acting at a distance that they could not begin to explore its properties, and they found it difficult to accept the Newtonian physics. They could not go along with Newton when he said he had not been able to explain how gravity works but that "it is enough that gravity really exists and suffices to explain the phenomena of the heavens and the tides." Those who accepted the Newtonian style fleshed out the law of universal gravity, showed how it explained many other physical phenomena and demanded that an explanation be sought of how such a force could be transmitted over vast distances through apparently empty space. The Newtonian style enabled Newton to study universal gravity without premature inhibitions that would have blocked his great discovery. The 18th-century biologist Georges Louis Leclerc de Buffon once wrote that a man's style cannot be distinguished from the man himself. In the case of Newton his greatest discovery cannot be separated from his style.

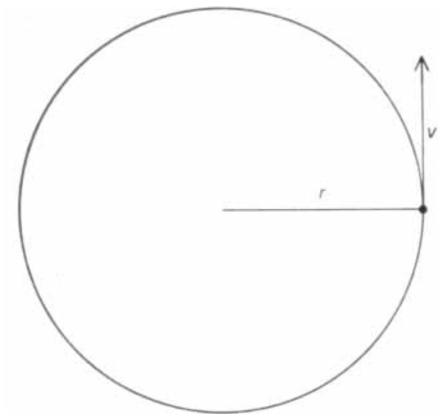
The correspondence between Hooke and Newton clearly shows that Hooke taught Newton how to analyze curved motion. Hooke subsequently made the much stronger claim that he deserved credit for suggesting to Newton the law of universal gravity varying inversely with the square of the distance. Many historians have echoed Hooke's view.

The claim, however, does not hold up. Hooke had merely suggested that the planets are subject to an inverse-square force directed toward the sun. Universal gravitation is much more than a solar-directed force. It also implies an effect of the planets on the sun. What is more, it applies to all objects in the universe. The law of universal gravitation is not merely an inverse-square relation; it is also a mathematical relation between the masses of the attracting bodies. It took tremendous insight to leap from an inverse-square solar-directed force to universal gravitation. And it took the genius of Newton to invent the modern concept of mass.

Newton did not feel he owed Hooke credit even for suggesting that the centripetal force is inversely proportional to the square of the distance. In 1673 Huygens had published a supplement to a book on the pendulum clock in which he states that for circular motion the centrifugal force is measured by v^2/r, where v is the velocity of the orbiting body and r is the radius of rotation. Newton had independently discovered the same relation in the 1660's. Since the mathematical difference between a centrifugal force and a centripetal force is only a matter of direction, the v^2/r relation also holds for a centripetal force. From this relation and Kepler's third law it follows by simple algebra that the centripetal force varies inversely with



ORBITAL SPEED OF A PLANET is inversely proportional not to the direct distance between the sun and the planet but to the perpendicular distance (the distance represented by the broken line between the sun and the tangent to the orbit *PP′*). Only at two points in the orbit (perihelion and aphelion) are the direct distance and the perpendicular distance the same.



$$\frac{v^2}{r} = \left(\frac{2\pi r}{T}\right)^2 \Big/ r = \left(\frac{4\pi^2 r^2}{T^2}\right)\frac{1}{r} =$$

$$4\pi^2\left(\frac{r^2}{T^2}\right)\frac{1}{r} = 4\pi^2\left(\frac{r^2 r}{T^2 r}\right)\frac{1}{r} =$$

$$4\pi^2\left(\frac{r^3}{T^2}\right)\frac{1}{r^2} = 4\pi^2 K\left(\frac{1}{r^2}\right)$$

INVERSE-SQUARE NATURE OF CENTRIPETAL FORCE for circular orbits can be deduced from Kepler's third law of planetary motion and from the law of centripetal force. According to Kepler's third law, r^3/T^2 is a constant K, where r is the radius of the planet's orbit and T is the period of the orbit. The law of centripetal force states that for a circular orbit the centripetal force is v^2/r, where v is the planet's velocity. In time T the planet makes a complete orbit, moving a distance $2\pi r$ (the circumference of a circle), and so the velocity is $2\pi r/T$.

the square of the distance. After Huygens' book was published anyone with a rudimentary knowledge of algebra could have found an inverse-square centripetal force for a circular orbit. Accordingly Newton saw no need to acknowledge Hooke's statement of an inverse-square law.

Both Hooke and Newton were aware that finding an inverse-square law for circular orbits was not the same thing as showing that the law holds for elliptical orbits in which the motion follows Kepler's law of areas. The task, which Newton carried out, was to demonstrate that an inverse-square law of centripetal force corresponds to orbital motion according to Kepler's law of elliptical orbits and his law of areas. In discussing this point in the letter dated January 6, 1680, Hooke made a fundamental error that must have convinced Newton that Hooke did not entirely understand what he was talking about. Hooke said that if the attraction varies inversely as the square of the distance, the orbital speed of a planet will be "as Kepler supposes Reciprocall to the Distance." Yet under the conditions Hooke assumed the orbital speed is not inversely proportional to the direct distance from the sun except at the extreme points of the orbit: perihelion and aphelion. In view of Hooke's error Newton was not about to give him credit for having suggested the inverse-square nature of the centripetal force.

In 1717 Newton wanted to ensure his own priority in discovering the inverse-square law of gravitation, and so he invented a scenario in which he made the famous moon test not while writing the *Principia* but two decades earlier in the 1660's. The documents of the 1660's, however, indicate that he was not then comparing the falling of the moon in its orbit with the falling of objects on the earth but was comparing the "centrifugal endeavor" of the orbiting moon with the "centrifugal endeavor" of a body on the earth's surface rotating along with the daily motion of the earth. He did calculate that for circular planetary orbits the "centrifugal endeavor" would be inversely proportional to the planet's distance from the sun, but he drew no physical conclusions from the calculation.

Newton never published his invented scenario of the early moon test. He included it in the manuscript draft of a letter to the French writer Pierre Des Maizeaux but then crossed it out. Newton also circulated the familiar story that a falling apple set him on a chain of reflections that led to the discovery of universal gravitation. Presumably this invention was also part of his campaign to push back the discovery of gravity, or at least the roots of the discovery, to a time 20 years before the *Principia*.

The real roots of the discovery cannot be put any earlier than December, 1684, when Newton first recognized that if the sun attracts the earth, the earth must attract the sun with a force of equal magnitude. In 1685 he overcame his usual reluctance to write up his discoveries and started to draft the *Principia* for publication by the Royal Society. Perhaps his willingness to present his work for public inspection (and thereby risk possible disapprobation) was motivated first by his momentous discovery of interplanetary perturbations followed by his bold conception of universal gravity. He had within his grasp the foundation of a new system of natural philosophy that could be expounded on mathematical principles. In short, once Newton had something of real consequence to say about celestial dynamics he was willing and even eager to present it to the world.

THE LEADING EDGE

#1 in a series of reports on new technology from Xerox

About a year ago, Xerox introduced the Ethernet network—a pioneering new development that makes it possible to link different office machines into a single network that's reliable, flexible and easily expandable.

The following are some notes explaining the technological underpinnings of this development. They are contributed by Xerox research scientist David Boggs.

The Ethernet system was designed to meet several rather ambitious objectives.

First, it had to allow many users within a given organization to access the same data. Next, it had to allow the organization the economies that come from resource sharing; that is, if several people could share the same information processing equipment, it would cut down on the amount and expense of hardware needed. In addition, the resulting network had to be flexible; users had to be able to change components easily so the network could grow smoothly as new capability was needed. Finally, it had to have maximum reliability—a system based on the notion of shared information would look pretty silly if users couldn't get at the information because the network was broken.

Collision Detection

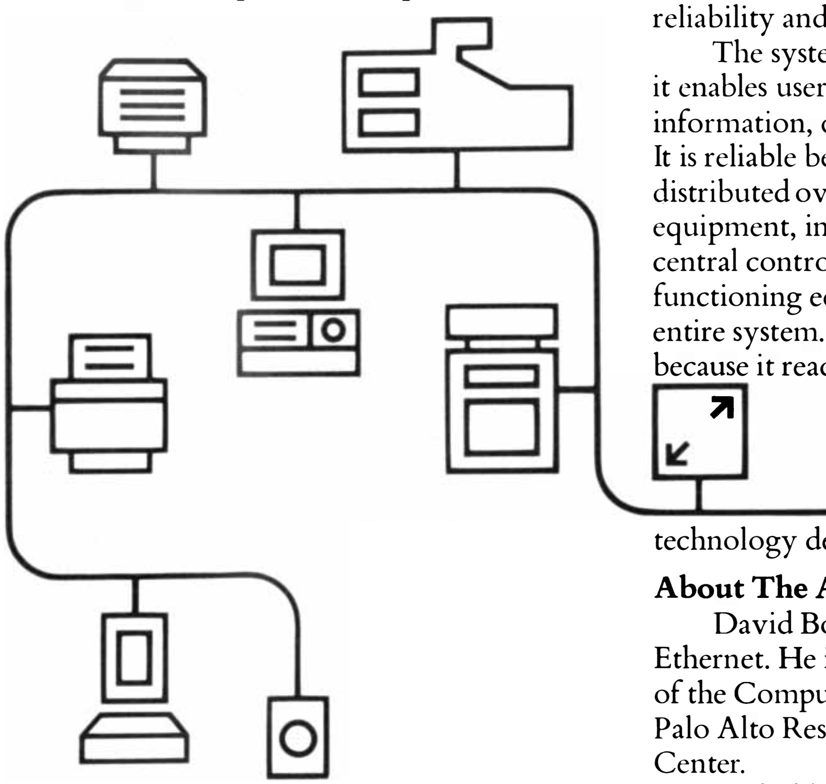
The Ethernet network uses a coaxial cable to connect various pieces of information equipment. Information travels over the cable in packets which are sent from one machine to another.

A key problem in any system of this type is how to control access to the cable: what are the rules determining when a piece of equipment can talk? Ethernet's method resembles the unwritten rules used by people at a party to decide who gets to tell the next story.

While someone is speaking, everyone else waits. When the current speaker stops, those who want to say something pause, and then launch into their speeches. If they *collide* with each other (hear someone else talking, too), they all stop and wait to start up again. Eventually one pauses the shortest time and starts talking so soon that everyone else hears him and waits.

When a piece of equipment wants to use the Ethernet cable, it listens first to hear if any other station is talking. When it hears silence on the cable, the station starts talking, but it also listens. If it hears other stations sending too, it stops, as do the other stations. Then it waits a random amount of time, on the order of microseconds, and tries again. The more times a station collides, the longer, on the average, it waits before trying again.

In the technical literature, this technique is called carrier-sense multiple-access with collision detection. It is a modification of a method developed by researchers at the University of Hawaii and further refined by my colleague Dr. Robert Metcalfe. As long as the interval during which stations elbow each other for control of the cable is short relative to the interval during which the winner uses the cable, it is very efficient. Just as important, it requires no central control—there is no distinguished station to break or become overloaded.



The System

With the foregoing problems solved, Ethernet was ready for introduction. It consists of a few relatively simple components:

Ether. This is the cable referred to earlier. Since it consists of just copper and plastic, its reliability is high and its cost is low.

Transceivers. These are small boxes that insert and extract bits of information as they pass by on the cable.

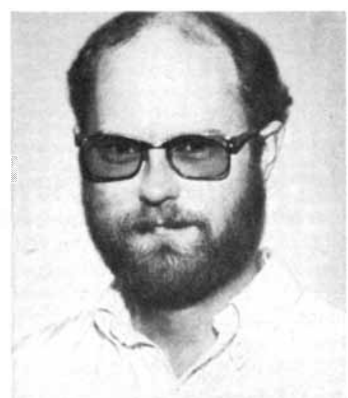
Controllers. These are large scale integrated circuit chips which enable all sorts of equipment, from communicating typewriters to mainframe computers, regardless of the manufacturer, to connect to the Ethernet.

The resulting system is not only fast (transmitting millions of bits of information per second), it's essentially modular in design. It's largely because of this modularity that Ethernet succeeds in meeting its objectives of economy, reliability and expandability.

The system is economical simply because it enables users to share both equipment and information, cutting down on hardware costs. It is reliable because control of the system is distributed over many pieces of communicating equipment, instead of being vested in a single central controller where a single piece of malfunctioning equipment can immobilize an entire system. And Ethernet is expandable because it readily accepts new pieces of information processing equipment. This enables an organization to plug in new machines gradually, as its needs dictate, or as technology develops new and better ones.

About The Author

David Boggs is one of the inventors of Ethernet. He is a member of the research staff of the Computer Science Laboratory at Xerox's Palo Alto Research Center.



He holds a Bachelor's degree in Electrical Engineering from Princeton University and a Master's degree from Stanford University, where he is currently pursuing a Ph.D.

THE AMATEUR SCIENTIST

The physics of spinning tops, including some far-out ones

by Jearl Walker

Spinning tops are ancient toys, but the principles that govern their behavior have come to be understood only in the past century and a half. What keeps an irregularly shaped object spinning on a single point? Why do tops of different shapes behave in such different ways? Here I shall explain some of the mechanics of tops, avoiding the mathematical thickets that obscure the subject in some physics textbooks. I shall also introduce several of the unusual tops made by Donald W. Dubois of the University of New Mexico, who has set spinning such unlikely objects as a golf tee and the stopper from a bottle of India ink.

The behavior of any top is due mainly to the effect of gravity. Every atom of the top's mass is pulled downward by gravity, but the net pull is more readily envisioned as being through the center of mass, which lies somewhere inside the top, usually at its geometric center. The weight of the top can be represented by a vector pointing downward from the center of mass. It seems logical that because of this pull the top would be less likely to remain standing than to topple over (as you would if you were leaning away from the vertical).

The difference is that the top is spinning. As a result the downward pull of gravity gives rise to the surprising rotation of the top about the vertical. This reaction is not easy to visualize because you are more familiar with nonspinning objects. Normally a force on an object causes an acceleration in the direction of the force. When a spin is involved, the force may result in a motion perpendicular to the direction of the force. Such an unfamiliar motion is part of the fascination of tops.

The spinning top has angular momentum. Where linear momentum is the product of the mass and the velocity of an object, angular momentum is the product of the mass distribution (the moment of inertia) and the angular velocity. The top has angular momentum because it is spinning around its long body axis and because it has a mass distributed around that axis. The angular momentum is a vector that lies along the body axis, the axis around which the top is symmetrical.

The only way to change the angular momentum of an object is by means of a torque. A torque is scientifically defined (not too differently from the common usage) as the product of (1) the force on an object and (2) a lever arm that extends from a pivot point to a line drawn through the force perpendicular to the lever arm. The pivot point for a top is obviously at the pointed end touching the floor or a tabletop. The lever arm extends horizontally from that point to a vertical line drawn through the center of mass, which is where gravity is considered to be pulling on the top.

Gravity provides not only a downward force but also a torque to change the angular momentum of a top. The torque does so in a simple way: it redirects the angular momentum, rotating the vector about the vertical axis. Since the vector must continue to be along the body axis, that axis also rotates about the vertical in the motion called precession. It maps out a cone centered on the vertical. (For the present I shall treat the point of the top as being fixed where it touches the floor.)

If the top is spinning counterclockwise as it is viewed from overhead, the precession around the vertical is also counterclockwise. If the top is spinning clockwise, the precession is clockwise. If the top were not spinning, the torque effect of gravity would cause it to fall to the floor, which it actually does during the last stage of its spin because of the effect of friction on the point on which the top is spinning.

Nearly all tops fall somewhat as they start to spin. The reason is an energy requirement. If the top is to precess as a result of the torque due to gravity, it must have kinetic energy. Occasionally it gains energy from the launching. More often the energy must come from the initial fall, during which the decrease in the potential energy of the top is transformed into kinetic energy.

This simple story of the top fails the first time you examine the motion of a real top. The top does not just spin and precess; its body axis nods in what is called nutation. As it nutates the angle between the body axis and the vertical varies between two values determined by the mass distribution and kinetic energy of the top and its initial angle with the vertical. It is possible to illustrate the types of nutation by tracing the position of the upper end of the body axis on a sphere centered on the point where the top touches the floor. The angle between the body axis and the vertical is limited by two circles drawn around the vertical. The circles represent the range through which the top can lean away from the vertical during nutation.

In one type of nutation the body axis weaves between the two limiting circles harmonically, touching each circle tangentially. The precession of the axis around the vertical is always in one direction, either clockwise or counterclockwise as seen from above. In the illustration on page 188 the rotation is counterclockwise.

In the second type of nutation the body axis loops between the two limiting circles but still touches each one tangentially. The direction of travel of the body axis periodically changes between clockwise and counterclockwise. In spite of this reversal the precession has an average value that is in one direction or the other. In the illustration the average precession is counterclockwise.

The third type of nutation traces cusps on the imaginary sphere. The path of the body axis meets one of the limiting circles tangentially as it does in the other types. The path meets the other limiting circle perpendicularly. The precession is consistently in one direction, but the rate of precession varies from maximum at the lower limiting circle to zero at the upper limiting circle.

The type of nutation that occurs (if any) depends on the initial conditions of the top's spinning. Since there are many perturbations in starting a top, you do not always have full control over those conditions. Suppose the launching imparts to the top a precession velocity that is in the same direction as the precession velocity gravity provides. Then the top precesses with harmonic nutation. Regardless of where the top is during its cycle of nutation, either its initial precession velocity or the velocity provided by gravity guarantees its continued precession in one direction around the vertical.

If the initial precession velocity is opposite to the precession resulting from gravity, the nutation is looped. During the lower part of the loop the gravitational precession drives the top around the vertical in the preferred direction. During the higher part of the loop, where the body axis reaches its upper

Renault Presents The Remarkable 18i Sportswagon.

Renault, maker of Europe's best selling cars, presents a remarkable new wagon for America...the Renault 18i Sportswagon.

Remarkable, because it is, at once, a wagon that provides comfortable seating for 5, *and* muscle to move you crisply from 0 to 50...

...a wagon with up to 65.5 cubic feet of load space, *and* the good manners to go where you aim it through precise, quick rack and pinion steering (lock to lock in just 3 turns)...

...a wagon that takes care of the whole family, *and* carefully meters out fuel for outstanding efficiency*...

...a wagon that bristles with innovative technology of the 80's.

38 HWY EST* (25) MPG EST

Bosch L-Jetronic Fuel Injection

This is the Electric Multi-Point Bosch system also used by Porsche 928 and Jaguar XJ-S. It precisely measures out the fuel required for optimum performance and efficiency from the 18i's enthusiastic 1.6 liter (1647 cc, aluminum block and head) 4-cylinder engine.

Slip Stream Aerodynamics

The 18i Sportswagon silhouette is notably slippery. A wide front modesty panel—more discreet than add-on air dams—encourages turbulence to slide beneath the 18i. Slip stream styling effects promote excellent fuel efficiency* as the wagon carves the air.

Road Adhesion

The 18i Sportswagon is masterfully tuned for the road. With 155SR x 13 Michelin steel belted radials, box section "live" rear axle (it flexes in the turns), front and rear sway bars, and beefy helical coil springs (variable-flex in the rear) surrounding long-travel shock struts.

Inner Space Geometry

The 18i's elegantly tailored seats are bio-formed with special support for the lumbar region and upper thighs.

Controls and gauges are strategically angled and positioned for driver access.

A wide rear bench seat folds forward, providing a lushly carpeted flat load bed more than 5½ feet long.

20 Years Experience in Front-Wheel Drive Design.

The 18i front-wheel drive system is a study in compact, lightweight design. Front drive transaxles, engine and 4-speed manual transmission (5-speed, avail.) are deftly mounted to provide a 60/40 weight ratio, front to rear. It also allows quick access for service at any of the more than 1300 Renault and American Motors Dealers.

The Renault 18i Sportswagon. Remarkable? We think so. It provides what wagon users ask for...generous space and great mileage. And something many have been missing...the sheer joy of commanding a responsive, nimble, sensitive, exciting road machine.

*EPA estimated at (25) mpg. 38 mpg highway est. Remember: Compare this estimate with estimated mpg for other cars. Your mileage may differ depending on speed, trip length and weather. Your highway mileage will probably be lower.

Renault 18i

More than just economy

At Renault and American Motors dealers.

limiting circle, the precession due to gravity is exhausted and only the initial precession is left. Until the top can drop and be driven around the vertical again by gravity it precesses in the opposite direction.

The third type of nutation is often seen when a spinning top is initially held at an angle to the vertical and then released. It has no initial precession and so it just falls to the lower limiting circle. The fall is necessary because the resulting decrease in the potential energy of the top provides energy for the precession brought about by the torque exerted by gravity. When continued nutation brings the body axis back to its initial angle with the vertical, the top regains its initial potential energy. Precession then stops momentarily, since no energy is left for it. Afterward the top again falls to the lower limiting circle and precession continues until the top again is raised by nutation.

The limiting circles originate from three severe restrictions on spin, precession and nutation. The total energy (kinetic and potential) of the top must remain constant. (I shall discuss the effects of friction below.) The angular momentum along the body axis must also remain constant in amount, although its direction can change, because there is no torque along the axis to change it. Finally, the angular momentum along the vertical must remain constant for the same reason. (Once precession begins, the calculation of these two angular momentums becomes harder because the angular velocity then includes both the spin of the top and the precession velocity.) Because of these three strict rules the body axis is limited to a certain range of angles with the vertical. If the body axis dropped below the lower limiting circle or rose above the upper one, it would do so only by disobeying the rules. (Tops do rise above the upper circle but only because of friction.)

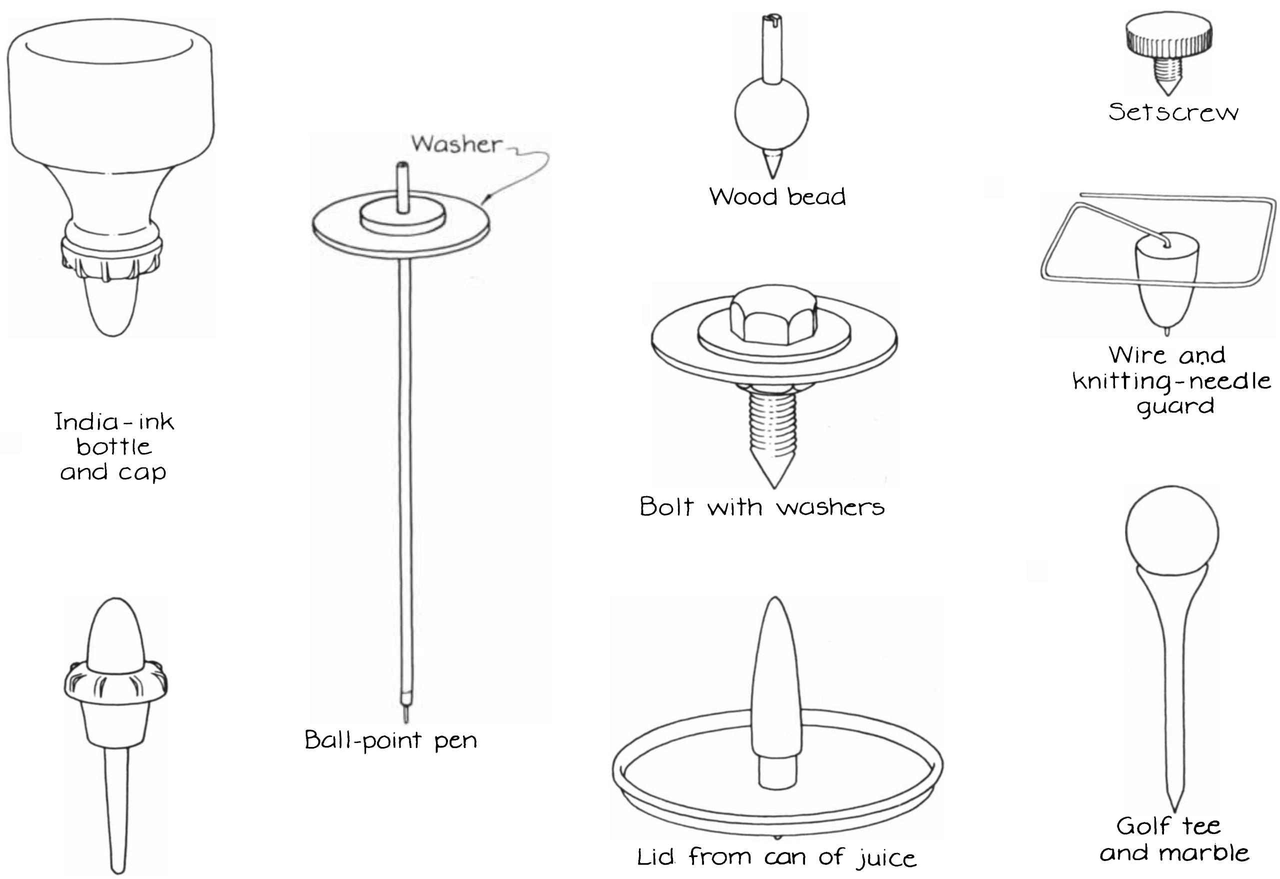
Most mathematical models of tops are devised with the simplifying assumption that the kinetic energy of the spin is much greater than the change in potential energy as the top weaves in nutation. Such a top is said to be "fast." With this assumption several features of the motion can be related to the rate at which the top spins. A lower rate increases the nutation weave but decreases the rate of nutation. The average rate of precession also depends on the spin, being faster with a slower spin. You can easily see these relations when you spin a top on the floor. As friction gradually robs the top of spin the precession rate increases and nutation becomes slower and more pronounced. Finally, just before the end, the top swings sluggishly up and down as it precesses around the vertical faster than ever.

The precession rate and the extent and frequency of nutation also depend on the mass and shape of the top. In general a greater mass increases each of these components of the motion. (Of course, a more massive top is more difficult to spin at a given rate than a less massive one.)

If the top is spinning quite rapidly, the small amount of nutation that should be present might be eliminated by the friction on the point of the top. Then the top would appear to precess uniformly around the vertical without nutation. This uniformity, however, is due only to the intervention of the friction.

Truly uniform precession is possible if the top has coincident limiting circles. If it does, it must be released in such a way that its angle with the vertical corresponds to the angle of the limiting circles. Such a top does not nutate, since it can neither rise above the limiting circles nor fall below them. Since it does not fall, however, it has no way to convert some of its potential energy into energy for precession. Your starting technique or the initial impact of the top on the floor would have to provide the impetus to start precession.

There are actually two possible types of precession. Only the slower one is re-



A few of the objects made into tops by Donald W. Dubois


SIEMENS
Measuring radio waves from outer space depends on a very special electronic component. And on the very special company that makes it.
The component is a backward-wave oscillator, in use at the National Radio Astronomy Observatory in Green Bank, W. Va. And Siemens is the very special company capable of developing this unique device.
Broad Product Line
Siemens makes the broadest line of electronic components in the world. Four million dollars spent each working day on R&D supports this leadership.
Exceptional Services
The Siemens components group in the United States has been growing at twice the industry growth rate and projects similar increases in the future. And because more than half the components sold in the U.S. are also manufactured here, industry is assured of exceptional service.
Siemens Commitment
Siemens, a worldwide electrical manufacturing company with $16 billion in annual sales, has committed its resources to American industry in four major areas: electronics, communications, power engineering, and health care. Almost 8,000 people are employed in five manufacturing locations and field offices serving all 50 states.
Siemens. The technology, the products, the people, the commitment.
Siemens Corporation
Department 0453
186 Wood Avenue South
Iselin, New Jersey 08830
Siemens. Constructive contributions.

Medicine is always changing, now there's a text that changes right along with it.

SCIENTIFIC AMERICAN *Medicine.*

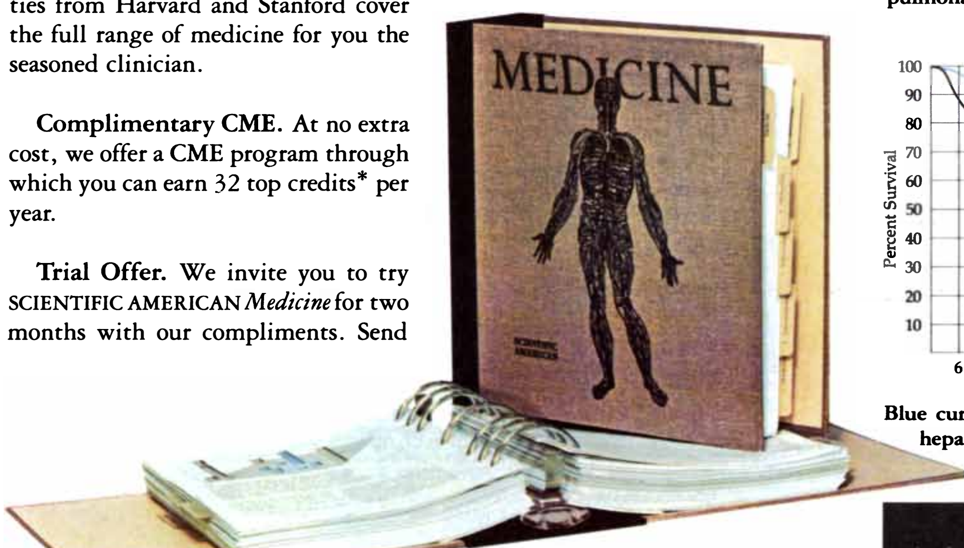
The answers that you get from SCIENTIFIC AMERICAN *Medicine* are current, because each month new chapters replace the old. You will know where to look for the answers, because this advanced text follows the familiar structure of medical knowledge, changing as medicine changes.

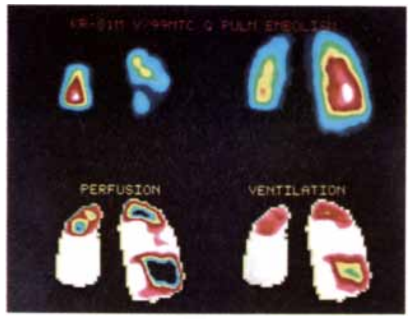
Comprehensive. Leading authorities from Harvard and Stanford cover the full range of medicine for you the seasoned clinician.

Complimentary CME. At no extra cost, we offer a CME program through which you can earn 32 top credits* per year.

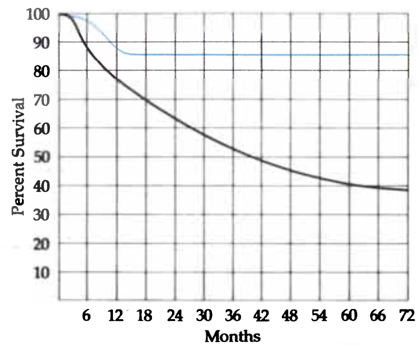
Trial Offer. We invite you to try SCIENTIFIC AMERICAN *Medicine* for two months with our compliments. Send us the coupon and you will receive the two-volume text and two monthly updates. You may also take a CME test for credit. At the end of 60 days, if you decide to continue the subscription, you will owe us $185 for the first 12 months (renewal is $160); otherwise, return the volumes; there will be no charge.

Please mail the coupon today so that we can put medical progress to work for you.





Computerized scintigraphy reveals pulmonary thromboembolism.



Blue curve represents survival in chronic hepatitis patients treated with steroids.

CT scan reveals large meningioma and mild hydrocephalus.

*As an organization accredited for continuing medical education, the Stanford University School of Medicine designates this continuing medical education activity as meeting the criteria for 32 credit hours in Category 1 for Educational Materials for the Physician's Recognition Award of the American Medical Association, provided it has been completed according to instructions.

This program is approved by the American Academy of Family Physicians for 32 elective hours of CME credit.

SCIENTIFIC AMERICAN MEDICINE

415 Madison Avenue, New York, New York 10017

Please enroll me as a subscriber to SCIENTIFIC AMERICAN *Medicine.* On receipt of this coupon you will send me the advanced two-volume text described in your announcement and update it regularly by sending me new monthly subsections. I understand that the annual subscription of $185 for SCIENTIFIC AMERICAN *Medicine* is tax deductible. If I am not entirely satisfied, I may cancel at any time during the first 60 days, returning all materials for a *complete refund.*

- ☐ Please enter my subscription for SCIENTIFIC AMERICAN *Medicine*
- ☐ I shall also enroll in the CME Program
- ☐ I enclose a check made out to SCIENTIFIC AMERICAN *Medicine* for $185*
- ☐ Please bill me

*Please add sales tax for California, Illinois, Massachusetts, Michigan, Ohio, and New York

Name ____________________ MD Specialty ____________
Address ____________________
City ____________ State ____________ Zip ____________
Signature ____________________

Subscribers outside the U.S. and possessions will be charged extra for shipping and handling: updates will be sent by surface routes unless airmail delivery is requested. Please allow 6-8 weeks for delivery. 1V lated to the gravitational pull on the top. The faster one is rarely seen in a top but can turn up under the proper initial conditions. The classical explanation of the faster precession involves a top (usually considered as an ellipsoid in textbook discussions) that is free to roll around on a horizontal plane while it spins. No forces or torques act on this imaginary top, not even gravity. In this situation the top cannot slide into the plane or break contact with it.

Once the top is spinning with its body axis at some angle to the vertical it begins to precess around the vertical in a way that maintains the contact between it and the horizontal plane. The plane is usually called the invariant plane; the path of the top's contact point on the invariant plane is called the herpolhode, and the path of the top's contact point on the top itself is called the polhode. As one writer put it, the motion leads to a jabberwockian statement: The polhode rolls without slipping on the herpolhode lying in the invariant plane. The precession in this idealized case is the fast precession a real top might have.

My colleague James A. Lock has pointed out that fast precession can be seen in a spinning football and a soft-drink bottle thrown into the air. A quarterback throws a football with a spin to keep the ball stable during flight, taking advantage of its streamlined shape. In a long pass the football often has a noticeable wobble, which may be partly due to air drag. Even without drag the ball will wobble if it is thrown to develop fast precession.

The precession is easier to see if a soft-drink bottle is tossed upward with spin. (The bottle should be empty; sloshing fluid would interfere with the spin.) If you toss the bottle with its body axis vertical, you may not be able to see any precession. With a less careful toss the spinning bottle will precess around the vertical in the fast-precession mode. The motion is not caused by gravity because the bottle is in flight. During the precession the bottom of the bottle traces out the herpolhode on an imaginary invariant plane that is perpendicular to the vertical.

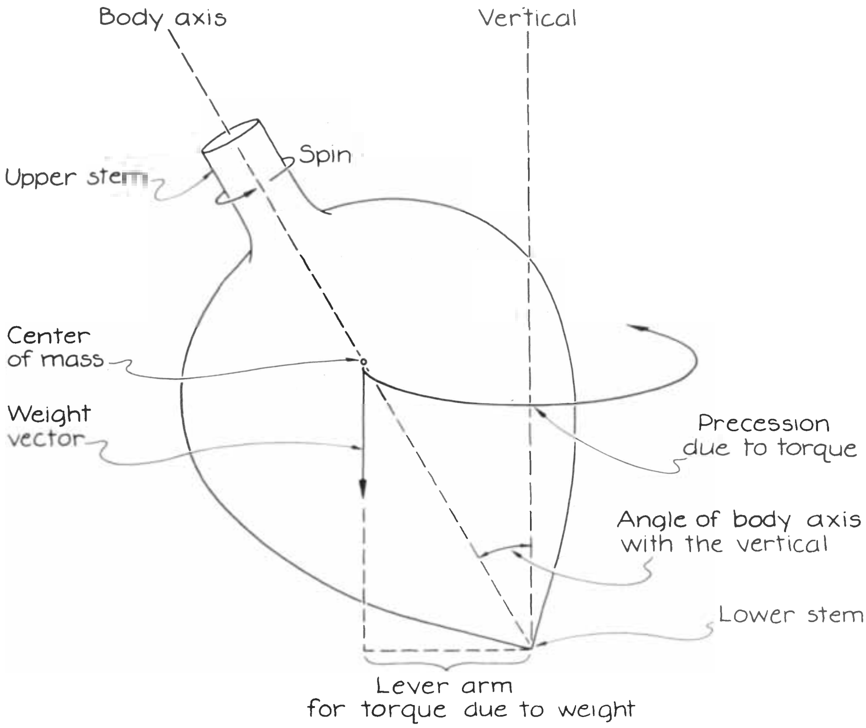
Fast precession is rare with a real top because it is so much faster than the precession that depends on gravity. Much more energy is needed to send the top into fast precession. Unless the top is somehow given that energy during the initial stages (for example, if a fortuitous launching imparts the energy and the required initial motion), it will settle instead into slow precession.

With some care you can set a top spinning so that its body axis is vertical. (The technique of launching a top with two lengths of tape, which I shall describe below, is useful here.) If the spin rate is higher than some critical value, which varies according to the type of top, the top stays upright. At a spin rate lower than the critical value the top nutates between the vertical and some limiting circle corresponding to a larger angle to the vertical.

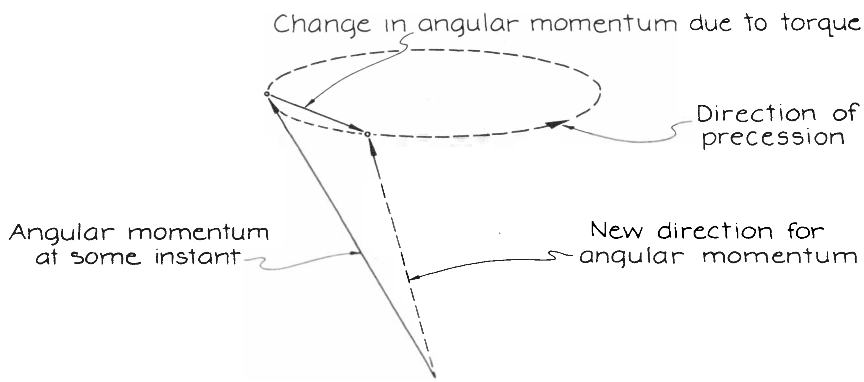
Both kinds of behavior can be seen in a real top with a spin rate that is initially above the critical value and a body axis that is initially vertical. The spin has two stages: first a sedate vertical spinning and then a frantic nutation. The transformation is brought about by the friction at the point where the top touches the floor. Although the initial spin rate is high enough for the vertical spinning, friction eventually pushes the spin rate below the critical value and the top begins to lean and to nutate and finally to topple.

Can a top that is initially tilted ever nutate until it is vertical? In principle it cannot because it will always lack the energy to become precisely vertical. It can come close, however, if the upper limiting circle is quite small. Then the top can nutate upward in such a way that it passes close to the vertical before it nutates downward again. This situation will probably not arise often.

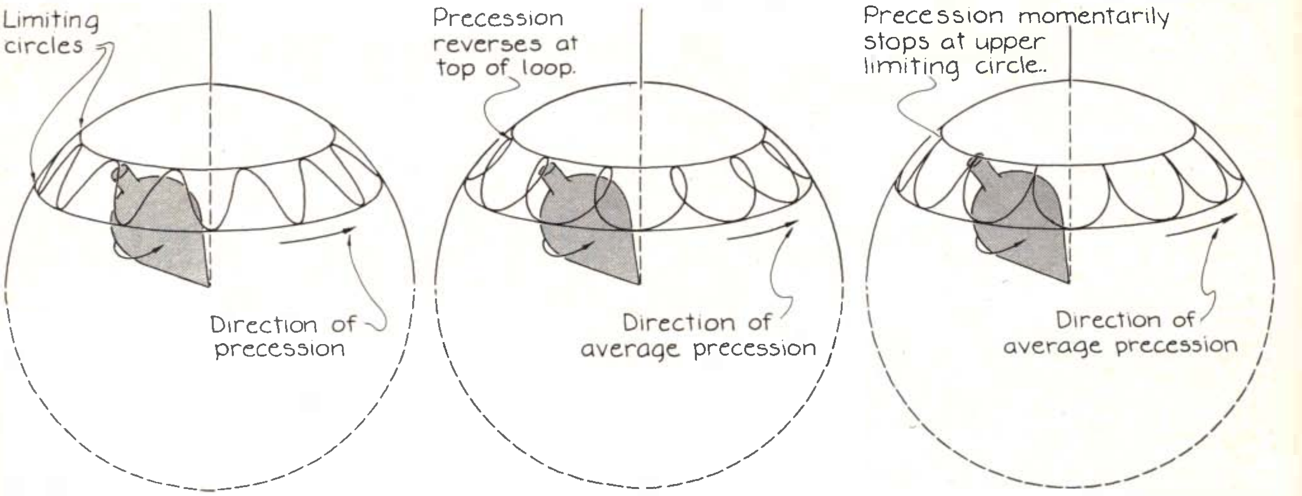
A keen observer of tops would not agree with much of what I have said. Many tops do rise to a vertical spin (such a top is called a sleeper), and so my statements are not entirely correct. A top rises not through a process of nutation but through a subtle interplay of the stem of the top and the friction between the point of the top (the end of the stem on which the top spins) and



Torque on a spinning top



The effect of torque



The nutation patterns of tops

the floor. The stem of any real top is not the infinitesimal point I have been assuming so far. A better model would be a hemispherical stem.

Such a stem would tend to rotate on the floor because of the top's spin. Meanwhile the top is also trying to precess. Since I am now considering a more realistic stem instead of a stationary point of contact with the floor, I must visualize the precession as moving the top around its center of mass, which remains stationary. The contact point on the floor moves, circling below the center of mass, while the upper section of the top circles above the center of mass.

As a result the hemispherical stem is attempting to roll over the floor in two ways, one due to the spin of the top around its body axis and the other due to the precession driving the stem over the floor. The first way is much faster than the second, and so the stem slips. As it slips the friction between it and the floor points in a direction opposite to the direction of the slip, as is shown in the bottom illustration on page 190. It is this friction that rotates the body axis of the top to make the spin vertical. To calculate the torque I treat the pivot point as the center of mass. The lever arm is the distance between that point and the place where the stem touches the floor. The torque raising the top is equal to the multiplication of the lever arm and the friction.

Of course, the friction on the stem is also stealing energy from the top. The trick in designing a good sleeper is to shape the stem so that the top becomes vertical before it loses too much energy. Hence a good sleeper has a stem with a small radius of curvature. Tops with stems that have a more gradual curvature take longer to rise and may not reach the vertical before the spin declines below the critical value needed to sustain a vertical position. A whip top, which is lashed with a whip to add energy to the spin, has a stem with a very small radius of curvature; it rises to be a sleeper within seconds of being lashed.

Most of the commercially available tops are shaped like a pear. Some are spun by the fingertips. Others are wound with string and thrown to the floor. Some handcrafted tops are mounted in a wood holder until the string wrapped around their spindle is unwrapped with a brisk pull; then they fall spinning to the floor.

Although all these types are usually found in toyshops, they by no means exhaust the variety of tops that can be (or once could be) found around the world. A beautiful description of the many kinds of top that have been made can be found in the book by D. W. Gould cited in the bibliography for this issue [*page 194*]. Some tops have multiple faces for the purpose of gambling and telling fortunes. Some are flat disks pierced by a thin rod. Others are flat disks with a central bump on which they spin. (Several of them can be deployed at one time from a single holder.) The most remarkable feature of tops (apart from their rotational mechanics) is how they were developed in such a variety of shapes by people working almost independently around the world.

Dubois recently wrote me about the tops he and his students make from common items such as buttons, needles and machine screws. Often these unlikely objects need only a little fixing before they can serve as tops. A point may have to be ground or a slot may be required in a spindle so that a cord can be inserted. Sometimes several pieces are assembled into a top with glue. (Dubois says epoxy is best.)

Since friction plays a role in the behavior of tops, Dubois experiments on different surfaces. Paper, thin plastic sheets (such as food wrap) and bed sheets work well with some tops but not with others. Dubois has found that a linoleum floor is the only surface on which all his tops work adequately, although most of them work better on some other surface. You can keep track of where the point of a top travels by spinning it on carbon paper, soot-covered paper, ink-covered glass or some other surface on which the point can leave a mark.

Some of Dubois's tops can be spun by a quick snap of the fingers. He holds the upper stem of the top between his thumb and second finger, keeping both the point and his palm downward. Then he snaps those fingers, causing the top to spin outward on the table or the floor. If there is no upper stem, he holds the stem with his palm upward. Then when he releases the top with a snap, the point is properly downward. Another launching procedure is to hold the stem between the forefinger of the left hand and the thumb of the right hand; when the two digits are pulled quickly in opposite directions, the top is spun.

Fast spins can be achieved with some tops by wrapping a cord or a strip of cloth around the stem or some other cylindrical part of the top. A typewriter ribbon works well. (The ink should of course be removed first.) You can slip one end through a notch in the stem and then wrap the ribbon around the stem several times. When you throw the top outward across the floor while holding the end of the ribbon, the top is made to spin rapidly.

This technique does not work well with small or fragile tops. With them Dubois prefers a double wrapping of ribbon around the stem. Both strips are started in the same way in the notch. As the top is turned in the hand the double layer of ribbon is wrapped around it six to 12 turns. Then instead of throwing the top to unwrap the ribbon, pull the two

SCIENCE/SCOPE

A prototype of the system that will serve as radar and radio for NASA's Space Shuttle has met its scheduled completion date and is undergoing tests. As a radar, the system will allow astronauts to rendezvous with orbiting satellites in order to repair or retrieve them. It also can track any payloads released from the Shuttle. As a radio, the system will link with the Tracking and Data Relay Satellite System to let astronauts communicate with stations on earth. Hughes delivered the Ku-band integrated radar and communications system, as it is called, to Rockwell International, builder of the Space Shuttle.

A modified U.S. Air Force F-15 fighter looms as an economical way for NATO to have an aircraft that could attack enemy ground forces in all kinds of bad weather. The proposed two-seat F-15 Strike Eagle would still be able to fly air-to-air missions. But an advanced long-range AN/APG-63 radar with SAR (synthetic aperture radar) modifications would give the F-15 many new capabilities. The crew would have radar maps showing objects as small as 10 feet, and the plane could fly low using terrain masking to avoid detection. A Strike Eagle could operate about 95 percent of the time in Central Europe in winter, compared to about 59 percent for an aircraft equipped with an infrared visual system, and only about 20 percent for daytime visual missions. Hughes supplies the APG-63 radar to McDonnell Douglas Corp., builder of the F-15.

Three communications satellites ordered by AT&T (American Telephone and Telegraph Company) will live longer and handle more long-distance calls than earlier models. The new Telstar 3 satellites will serve 10 years instead of seven and have the capacity for 21,600 simultaneous calls instead of 18,000. These improvements are due chiefly to such technical innovations as solid-state amplifiers, better batteries, and a greater capacity for fuel to keep the satellites on station while in orbit. The first Telstar 3 is set for launch in 1983.

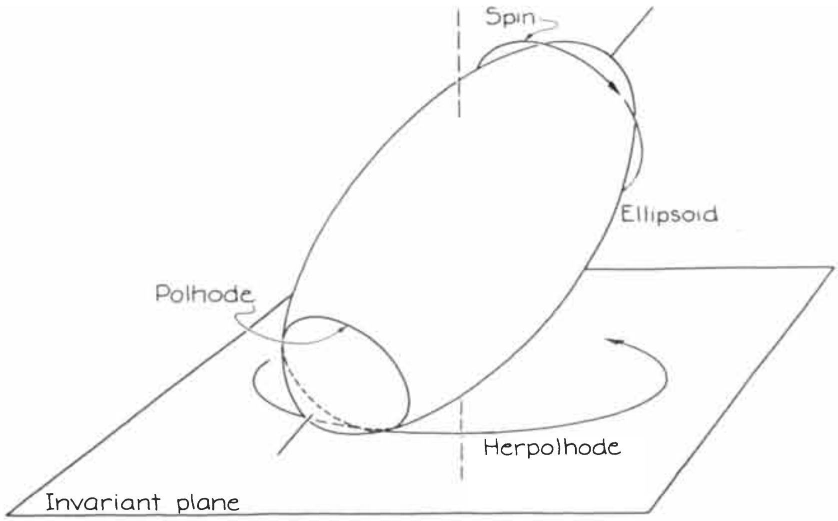
Hughes pushes the frontier of 34 different electronic technologies, from devices to systems, in its Industrial Electronics Group. We offer advantages and opportunities of a small company, but resources of a big one. The diversity includes a choice of facilities in Carlsbad, Irvine, Newport Beach, and Torrance, California. If your interests lie in semiconductors, fiber optics, microprocessors, minicomputers, microwave communications, or lasers, send your resume to John G. Wilhite, Hughes Industrial Electronics Group, Professional Employment, Dept. SSA, P.O. Box 2999, Torrance, CA 90505. Equal opportunity employer.

A device that scans the sky with heat sensors to detect, track, and identify aircraft and missiles is being developed by Hughes for the U.S. Air Force. The device, an electro-optical threat sensor, could be used with ground, ship, or airborne fire control systems. It emits no telltale radiation of its own, it is small, and it can search a wide area rapidly. A signal processor extracts the target signal from the background radiation and feeds this data to a computer, along with the target's relative bearing. An interrogation unit then uses additional sensors to classify targets further. The computer processes the information to classify each target by type and lists them in order of priority.

Creating a new world with electronics

HUGHES

HUGHES AIRCRAFT COMPANY
CULVER CITY, CALIFORNIA 90230



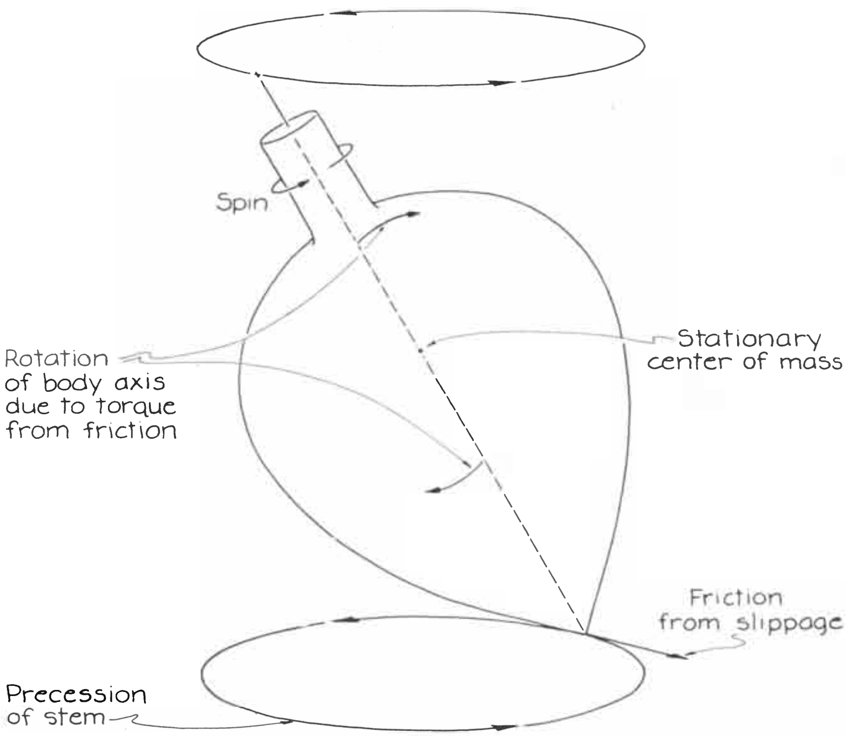
Motions that generate the polhode and the herpolhode

strips horizontally in opposite directions so that the top is twirled. With double-layer wraps of up to 14 revolutions Dubois can make an upholstery tack spin at about 15,000 revolutions per minute. With a finger snap the speed is about 8,000 r.p.m.

If the top has a smooth stem, as an upholstery tack does, the ribbon is difficult to wrap because it slips easily. Dubois coats the surface with a layer of epoxy to provide more friction. The added friction also means that when the end of the ribbon is reached during the launching, the top is deflected by a final tug rather than slipping away freely.

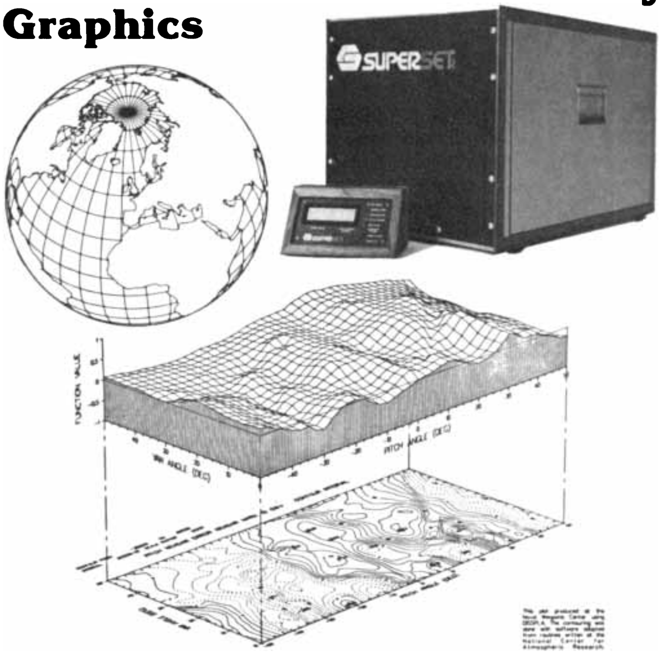
Some of Dubois's tops sing as they spin, either because they scratch the surface on which they are spinning or because they generate turbulence and vibration in the air. The spin of a singing top must exceed 3,300 r.p.m. One of Dubois's faster tops sings two octaves above A (440 hertz). If the top is spinning on a membrane of some kind, the membrane acts as a sounding board to enhance the audibility of the sound. Do not stretch the membrane too tight or it may be punctured by the spinning top.

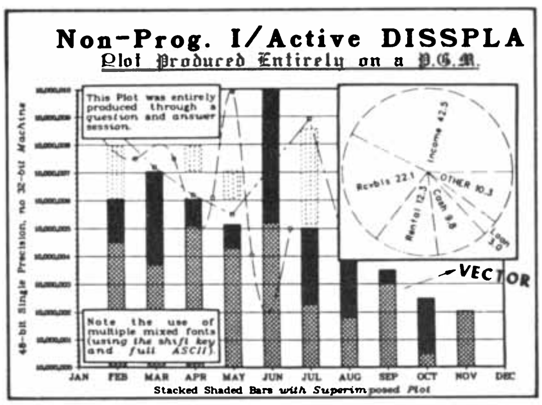
Spin and precession can be monitored



How friction raises a top

48 Bit Virtual Memory Mainframe Power Plus DISSPLA® Publication Quality Graphics





Shaded County Boundaries — The Ultimate Benchmark

The above map was produced with a few simple commands to a custom modified INTERACT version. It represents over 200,000 vectors each of which was projected through Alber Equal Area projection requiring execution of 400 statements per vector for a total of roughly 80 million Fortran statements. The counties comprise over 3,100 polygons, each bounded by about 100 points: **they were accessed in random order.** The map is produced on a 3 x 4 foot plotter for full detail. The program is over 1.5 megabytes. Shaded maps are a standard PGM DISSPLA enhancement. The raw execution speed (no disk swapping) is approximately 100,000 Fortran statements per second. Compilation speed is approximately 1,000 Fortran statements per minute. The map was reproduced courtesy of Zytron Graphic Systems, Cupertino, California.

An Economical Solution To Massive Software Execution. The PGM was initially designed to implement DISSPLA, a large computer graphics package having over 50,000 Fortran statements grouped into approximately 350 subroutines. This package, currently implemented on almost 400 large mainframes, has proved to be beyond 16 bit minis. Not only is the PGM able to handle this massive package with ease, it has proved capable of implementing almost any large scale Fortran system in existence. **It currently compiles, links and executes programs over 2 megabytes** — these are well beyond many mainframes.

48 Bits — Almost Double The Precision Of 32 Bit Machines. The full 48-bit PGM word provides **11+ floating point precision** (39 bit mantissa) compared to 5-6 digits for 32 bit machines. This is a single precision of roughly one millimeter on the earth's circumference.

24 Megabyte Address Space. The virtual memory totally dispenses with program overlay and segmentation. Virtual memory is implemented without cache memory or other expensive hardware crutches: The PGM utilizes the properties of higher level languages to implement a unique virtual memory scheme that is competitive in speed with machines costing orders of magnitude more.

Interactive Textbook Quality Graphics — The Primary Output Medium. Not only is the full DISSPLA package available, but also INTERACT, a **non-programmer** system for generating simple or sophisticated report quality charts and graphs in minutes. INTERACT operates through question/answer menus and any inexpensive digitizer pad. DISSPLA, the leading graphics package, enables the integration of custom sophisticated graphics into user developed programs. INTERACT source code is provided for customization. Also available are DIGIFONT for custom shaded font digitization, POSTGRAPH for storing and plotting identical output on different devices plus other powerful graphics utilities.

Standalone Computer, Distributed Processor or Dedicated Controller. The PGM operates both as a dedicated machine or as a node to a central mainframe. Optional sum checking is provided for data integrity in host communication. RS232, IBM 2780, DCT 2000, and other protocols are supported. **The PGM can provide DISSPLA facility to users of a mainframe without incurring mainframe memory or execution overhead.**

Major Graphics Devices Interfaced. The PGM is **non-partisan to any graphics device manufacturer.** Drivers are available for Tektronix 4010 Series, Chromatics, AED 512, IDT 2000 terminals; H.I. DP 14/15, Zeta 1453, Calcomp 1051, Soltec and Tektonix plotters; Dicomed, COMp80 and FR80 microfilm; Xerox 6500 color hardcopy; H.I. 8400, Versatec and Benson Varian Electrostatics. Color separation is furnished, enabling the use of high resolution monochromatic devices for color printing. Output may be generated on several devices simultaneously. A special parallel interface is available for the Xerox 6500 supporting full 600 X 1000 device resolution. **Vector to raster conversion is conducted in the PGM for raster devices at 500-1000 vectors/second.**

Powerful Virtual Memory Operating System Available. An extensive VMOS, extended Fortran compiler, Multipass Link Editor, Library Management and Trace/Peek/Poke Symbolic debugger provide a complete software development system.

An Elegant Low Entropy Software Engine. The PGM 48-bit TTL bit slice processor is a hardware extension of the compiler. There are **no registers, and no assembly language.** Statements such as A = B + C; A(I)= B(J+K)*C(M,N); IF(A(I)<B(J,K)) GO TO 5 are **single PGM instructions. Polynomial evaluate, block move and search are also provided as single instructions. The PGM achieves its remarkable performance simply by circumventing the superfluous instructions executed by general purpose computers running high level language software. The PGM is the first machine designed totally top down.** An 8-bit microprocessor takes care of all I/O for the 48-bit number cruncher.

Enter The PGM. So if your needs are too big for an Apple™ and your budget is too small for a VAX,™ the PGM is the solution. The PGM provides a graphics based engine with performance rivalling machines costing 100 times more.

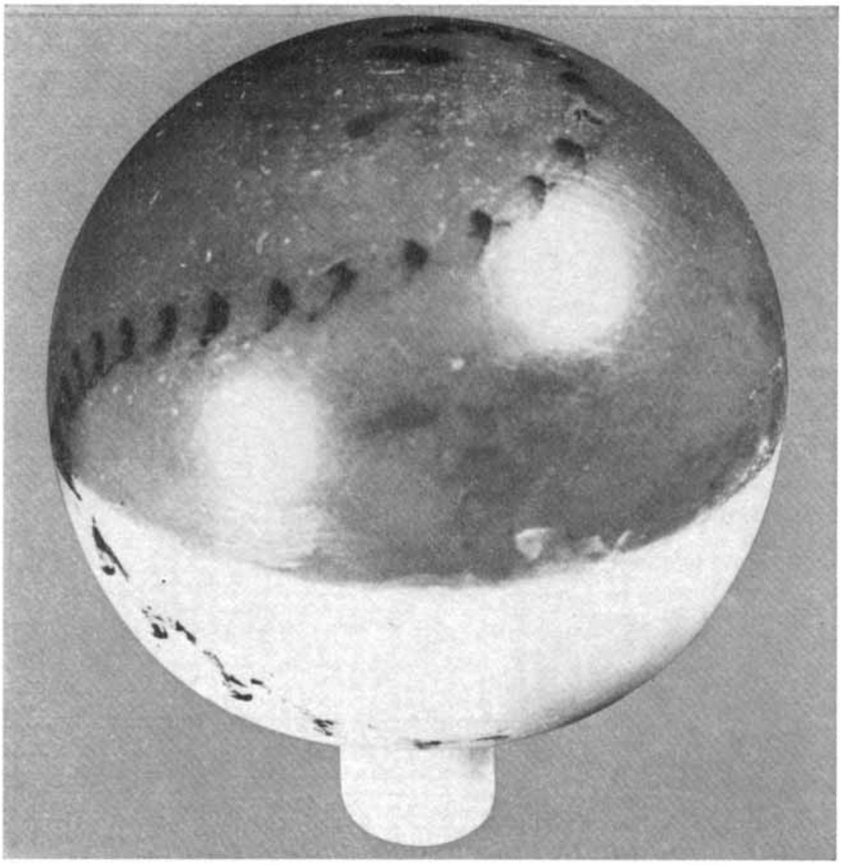
Prices Start At $27,000

Standard PGM includes full 48-bit processor, Intel 8080 I/O frontend, 29 megabyte disk, 300 kilobytes error correcting memory (expandable) and 5 RS232 ports. Options include 9 track tape interface, high speed parallel channel, Xerox 6500 interface and interface for synchronous protocols. Software supplied as extra cost items.

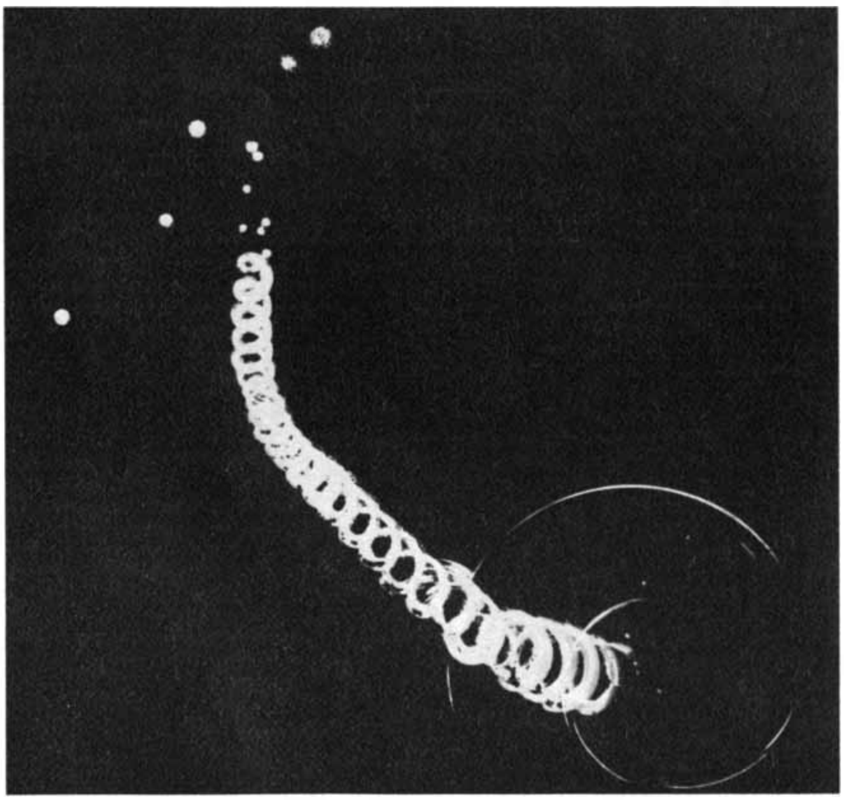
SUPERSET INC

11675 Sorrento Valley Road, Suite L, San Diego, California 92121
(714) 452-8665

DISSPLA is a registered trademark of Integrated Software Systems Corporation.



Frank F. Johnson's photograph of soot marks left on a Tippe Top



The track made by one of Johnson's Tippe tops on soot-covered glass

with a repetitive-flash strobe light. The frequency of the light is varied in order to match either the spin or the precession. Some of the tiny tops Dubois has sent me turn at speeds of more than 100,000 r.p.m., which is possible because the top's moment of inertia around its body axis is small and the energy imparted at launching results in a high rotational speed. The commoner pear-shaped top has a much greater moment of inertia around its body axis, so that the launching energy results in a much lower rotational speed.

The most fascinating top I have ever encountered is the Tippe Top, which I described in this department for October, 1979. The top usually has a hemispherical bottom and a central stem. It is spun with the bottom downward, a logical position because that section is heavier than the stem. Soon after the top is released, however, it inverts so that it spins on the stem. The heavier hemispherical end is lifted, seemingly in defiance of gravity.

The cause of the odd inversion is friction with the surface on which the hemisphere is turning. Spinning the Tippe Top on a soot-covered pane of glass will give a picture of the path of the top. Such a picture, made and analyzed in 1960 by Frank F. Johnson of Hasbrouck Heights, N.J., appears in the bottom illustration at the left. Johnson's analysis is cited in the bibliography for this issue [*page 194*].

What would a top do on a slightly inclined surface? Would its point of contact simply move around the way it does on a horizontal surface? Ledo Stefanini of the University of Bologna recently published a detailed answer to this question. Like other people studying the role of friction in the behavior of a top, he considered the stem to be hemispherical.

If the top is spinning without precession, it will move horizontally across the inclined plane in a straight line. If it is precessing, however, four things are possible, depending on the slope of the inclined plane and how much the top tilts from the vertical. One possible path for the stem is a spiral along a line down the plane. The other patterns resemble the nutation patterns of an ideal top with a sharp stem that does not move across the floor. In some cases the stem traces out loops on the plane as it moves primarily in a horizontal direction. Or it may map out cusps. For another set of angles the stem first climbs the inclined plane and then descends while also moving horizontally across it.

You might like to try various combinations of angles on an inclined plane. Stefanini published a tracing he made from a top weaving across carbon paper he had tacked onto an inclined plane. You can make your own tracings to catalogue the possible motions.

Let onComputing be your guide to personal computing.

Finally, there is a magazine that speaks to the beginner.

onComputing is the new McGraw-Hill quarterly that tells what's ahead — without talking over your head — in the 1980's with personal computers.

onComputing puts you on target with all the applications that go beyond your imagination.

- ☐ Personal computers in the home.
- ☐ Classrooms.
- ☐ They bring computer music.
- ☐ Electronic mail.
- ☐ Robots.
- ☐ They manage your personalized investment.
- ☐ The list goes on and on in **onComputing**.

If you're a beginner with personal computers, get the most understood magazine on computers, **onComputing**. And even if you're an experienced computer user, **onComputing** has the information you need in today's fast-growing personal computer market.

onComputing™

McGraw-Hill

Get onto onComputing. Call 800-258-5485.

onComputing Subscription Dept. P.O. Box 307, Martinsville, NJ 08836

DOMESTIC subscription rate:
☐ U.S. 1 yr. (4 issues) @ $8.50 ☐ Canada & Mexico, 1 yr. (4 issues) @ $10.00

FOREIGN (to expedite service, please remit in U.S. funds drawn on a U.S. bank)
☐ Europe (and all other countries, except above), 1 Yr. @ $12.00 — surface delivery.

☐ Bill Visa ☐ Bill Master Charge ☐ Bill me (North America only)

Card Number ____________ Expiration ____________
Signature ____________ Name (please print) ____________
Street/Apartment Number ____________
City ____________ State/Province/Country Code ____________

Please allow 6-8 weeks for processing

7131

BIBLIOGRAPHY

Readers interested in further explanation of the subjects covered by the articles in this issue may find the following lists of publications helpful.

METAMAGICAL THEMAS

THE THEORY OF GROUPS: AN INTRODUCTION. Joseph J. Rotman. Allyn & Bacon, Inc., 1973.

GROUPS AND THEIR GRAPHS. I. Grossman and W. Magnus. New Mathematical Library, 1975.

THE SERVICE SECTOR OF THE U.S. ECONOMY

GOOD JOBS, BAD JOBS, NO JOBS. Eli Ginzberg. Harvard University Press, 1979.

UNDERSTANDING THE SERVICE ECONOMY. Thomas M. Stanback, Jr. Policy Studies in Employment and Welfare No. 35. Johns Hopkins University Press, 1979.

ECONOMIC GROWTH AND THE RISE OF SERVICE EMPLOYMENT. Victor Fuchs. National Bureau of Economic Research Working Paper No. 486, 1980.

THE GROUND SUBSTANCE OF THE LIVING CELL

CYTOPLASMIC FIBRILS IN LIVING CULTURED CELLS: A LIGHT AND ELECTRON MICROSCOPE STUDY. Ian K. Buckley and Keith R. Porter in *Protoplasma,* Vol. 64, No. 4, pages 349–380; 1967.

THREE DIMENSIONAL FINE STRUCTURE OF CULTURED CELLS: POSSIBLE IMPLICATIONS FOR SUBCELLULAR MOTILITY. Ian K. Buckley in *Tissue and Cell,* Vol. 7, No. 1, pages 51–72; 1975.

MICROTRABECULAR LATTICE OF THE CYTOPLASMIC GROUND SUBSTANCE: ARTIFACT OR REALITY. John J. Wolosewick and Keith R. Porter in *The Journal of Cell Biology,* Vol. 82, No. 1, pages 114–139; July, 1979.

THE CONTROL OF PIGMENT MIGRATION IN ISOLATED ERYTHROPHORES OF *HOLOCENTRUS ASCENSIONIS* (OSBECK), I: ENERGY REQUIREMENTS. Katherine J. Luby and Keith R. Porter in *Cell,* Vol. 21, No. 1, pages 13–23; August, 1980.

THE ERUPTIONS OF MOUNT ST. HELENS

POTENTIAL HAZARDS FROM FUTURE ERUPTIONS OF MOUNT ST. HELENS VOLCANO, WASHINGTON. Dwight R. Crandell and Donal R. Mullineaux. U.S. Geological Survey Bulletin 1383-C, 1978.

VOLCANOES. Robert Decker and Barbara Decker. W. H. Freeman and Company, 1979.

VOLCANOLOGY. Howel Williams and A. R. McBirney. Freeman, Cooper and Co., 1979.

FIRE AND ICE: THE CASCADE VOLCANOES. Stephen L. Harris. Pacific Search Press, 1980.

1980 ERUPTIONS OF MOUNT ST. HELENS. Edited by Peter W. Lipman and Donal R. Mullineaux. U.S. Geological Survey Professional Paper No. 1250, 1981.

THE MILKY WAY GALAXY

STAR FORMATION. Edited by T. de Jong and A. Maeder. D. Reidel Publishing Company, 1977.

THE LARGE-SCALE CHARACTERISTICS OF THE GALAXY. Edited by W. B. Burton. D. Reidel Publishing Company, 1979.

INTERSTELLAR MOLECULES. Edited by B. H. Andrew. D. Reidel Publishing Company, 1980.

STAR CLUSTERS. Edited by James E. Hesser. D. Reidel Publishing Company, 1980.

THE MILKY WAY. Bart J. Bok and Priscilla F. Bok. Harvard University Press, 1981.

SUBMARINE POLLINATION

SUBMARINE POLLINATION IN SEAGRASSES. S. C. Ducker and R. B. Knox in *Nature,* Vol. 263, No. 5579, pages 705–706; October 21, 1976.

BIOLOGY OF AUSTRALIAN SEAGRASSES: POLLEN DEVELOPMENT AND SUBMARINE POLLINATION IN *AMPHIBOLIS ANTARCTICA* AND *THALASSODENDRON CILIATUM* (CYMODOCEACEAE). Sophie Ducker, John Pettitt and Bruce Knox in *Australian Journal of Botany,* Vol. 26, pages 265–268; 1978.

REPRODUCTION IN SEAGRASSES: NATURE OF THE POLLEN AND RECEPTIVE SURFACE OF THE STIGMA IN THE HYDROCHARITACEAE. John Pettitt in *Annals of Botany,* Vol. 45, pages 257–271; 1980.

ANORTHOSCOPIC PERCEPTION

HELMHOLTZ'S TREATISE ON PHYSIOLOGICAL OPTICS. Edited by James P. C. Southall. Dover Publications, Inc., 1962.

POST-RETINAL VISUAL STORAGE. Theodore E. Parks in *The American Journal of Psychology,* Vol. 78, No. 1, pages 145–147; March, 1965.

DISTORTIONS IN MOVING FIGURES VIEWED THROUGH A STATIONARY SLIT. S. M. Anstis and Janette Atkinson in *The American Journal of Psychology,* Vol. 80, No. 4, pages 572–585; December, 1967.

INTELLIGENCE FACTORS IN THE PERCEPTION OF FORM THROUGH A MOVING SLIT. I. Rock and E. Sigman in *Perception,* Vol. 2, No. 3, pages 357–369; 1973.

AN INTRODUCTION TO PERCEPTION. Irvin Rock. Macmillan Publishing Co., 1975.

THE BIOCHEMISTRY OF RESISTANCE TO MALARIA

EVOLUTIONARY RESPONSE TO HUMAN INFECTIOUS DISEASES. George J. Armelagos and John R. Dewey in *Bioscience,* Vol. 20, No. 5, pages 271–275; March 1, 1970.

MALARIA AND HUMAN POLYMORPHISMS. Frank B. Livingstone in *Annual Review of Genetics,* Vol. 5, pages 33–64; 1971.

CULTIVATION OF MALARIAL PARASITES. William Trager and James B. Jensen in *Nature,* Vol. 273, No. 5664, pages 621–622; June 22, 1978.

GENETICS OF RED CELLS AND SUSCEPTIBILITY TO MALARIA. L. Luzzatto in *Blood,* Vol. 54, No. 5, pages 961–976; October, 1979.

NEWTON'S DISCOVERY OF GRAVITY

THE BACKGROUND TO NEWTON'S PRINCIPIA: A STUDY OF NEWTON'S DYNAMICAL RESEARCHES IN THE YEARS 1664–84. John W. Herivel. Oxford University Press, 1965.

FROM KEPLER'S LAWS, SO-CALLED, TO UNIVERSAL GRAVITATION: EMPIRICAL FACTORS. Curtis A. Wilson in *Archive for History of Exact Sciences,* Vol. 6, No. 2, pages 89–170; April 29, 1970.

FORCE IN NEWTON'S PHYSICS: THE SCIENCE OF DYNAMICS IN THE SEVENTEENTH CENTURY. Richard S. Westfall. American Elsevier Publishers Inc., 1971.

THE NEWTONIAN REVOLUTION: WITH ILLUSTRATIONS OF THE TRANSFORMATION OF SCIENTIFIC IDEAS. I. Bernard Cohen. Cambridge University Press, 1980.

THE AMATEUR SCIENTIST

MECHANICS OF THE GYROSCOPE: THE DYNAMICS OF ROTATION. Richard F. Deimel. The Macmillan Company, 1929.

THE TIPPY TOP. Frank F. Johnson in *American Journal of Physics,* Vol. 28, No. 4, pages 406–407; April, 1960.

THE TOP: UNIVERSAL TOY, ENDURING PASTIME. D. W. Gould. Clarkson N. Potter, Inc., 1973.

A REALISTIC SOLUTION OF THE SYMMETRIC TOP PROBLEM. T. R. Kane and D. A. Levinson in *Journal of Applied Mechanics,* Vol. 45, No. 4, pages 903–909; December, 1978.

BEHAVIOR OF A REAL TOP. Ledo Stefanini in *American Journal of Physics,* Vol. 47, No. 4, pages 346–350; April, 1979.

FOUR HEADS ARE BETTER THAN TWO.

JVC's Vidstar has four video heads and the clarity that goes with them.

Admit it. Despite all the fancy features that videocassette recorders have to offer, you're worried about picture quality. Will it be crisp and vibrant? Will you get superb reproduction of recorded material?

Well, stop worrying. When JVC brought you 6-hour mode recording and playback, we equipped Vidstar™ with *four* video heads. Here's why.

Vidstar's 4-Head Construction

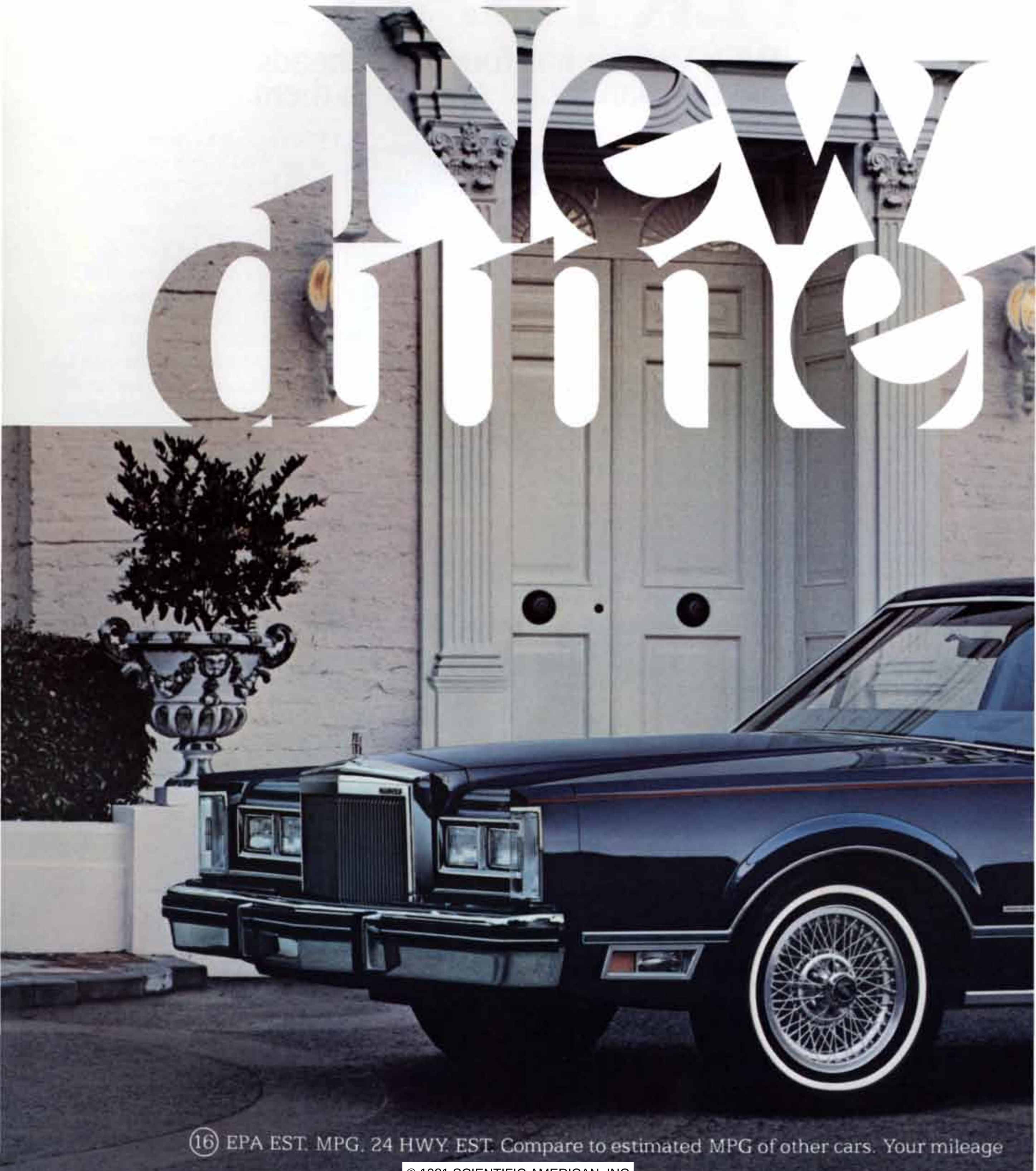
Vidstar uses one pair of recording/playback heads for the 2-hour mode and another pair of heads for the 6-hour mode. Each pair is suited to performing at a specific tape speed. This specialization gives Vidstar unsurpassed picture quality in both the 2-hour and 6-hour modes. That's the kind of performance you'd expect from the people who developed the VHS system. And who developed this four head technology. That's us, JVC.

That's only the beginning. Take a look at our other features: 6-hour recording on a single 2-hour VHS cassette, multi-function remote control unit, stop action and slow motion, double and triple speeds, and a micro-computer-assisted timer that can record up to 42 programs a week. You simply pre-set days, times and channels.

Vidstar is designed to let you add components to your home entertainment system effortlessly. You'll find goof-proof features that prevent recording mistakes and other features that let you add your own audio to recorded material. Vidstar has features you may not even know you want yet.

But you'll want them the minute you get your hands on them. You can do exactly that at your nearest JVC Vidstar dealer.

For JVC dealer names and locations call TOLL-FREE 800-526-5308. In NJ, 201-794-3900.


New
(16) EPA EST. MPG. 24 HWY. EST. Compare to estimated MPG of other cars. Your mileage

Introducing the next logical Lincoln.

New dimensions in size and technology were achieved in 1980. Now, these achievements continue in the new 1981 Lincoln Town Car.

This is a truly remarkable automobile.

Among its technological features:

- Electronic fuel injection and electronic engine controls for fuel efficiency and performance.
- Four-speed Automatic Overdrive Transmission for excellent highway mileage.
- Trim, contemporary size with room for six adults.
- Optional electronic instrument panel with message center that answers twenty different driver inquiries about time/distance/fuel usage.
- Optional keyless entry system that unlocks doors and/or decklid electronically simply by depressing buttons in numerical sequence.

Lincoln Town Car for 1981. A fine sense of tradition...remarkable achievements in technology.

You are cordially invited to experience it at your Lincoln-Mercury Dealer. There, you may exercise your option to lease or buy.


JOHN
JAMESON
EST'D
1780
SINE MET
IRISH WHI
A BLEND
DISTILLED IN IRELA
LIMITED
BOW STREET DISTI
DUBLIN
IMPORTE