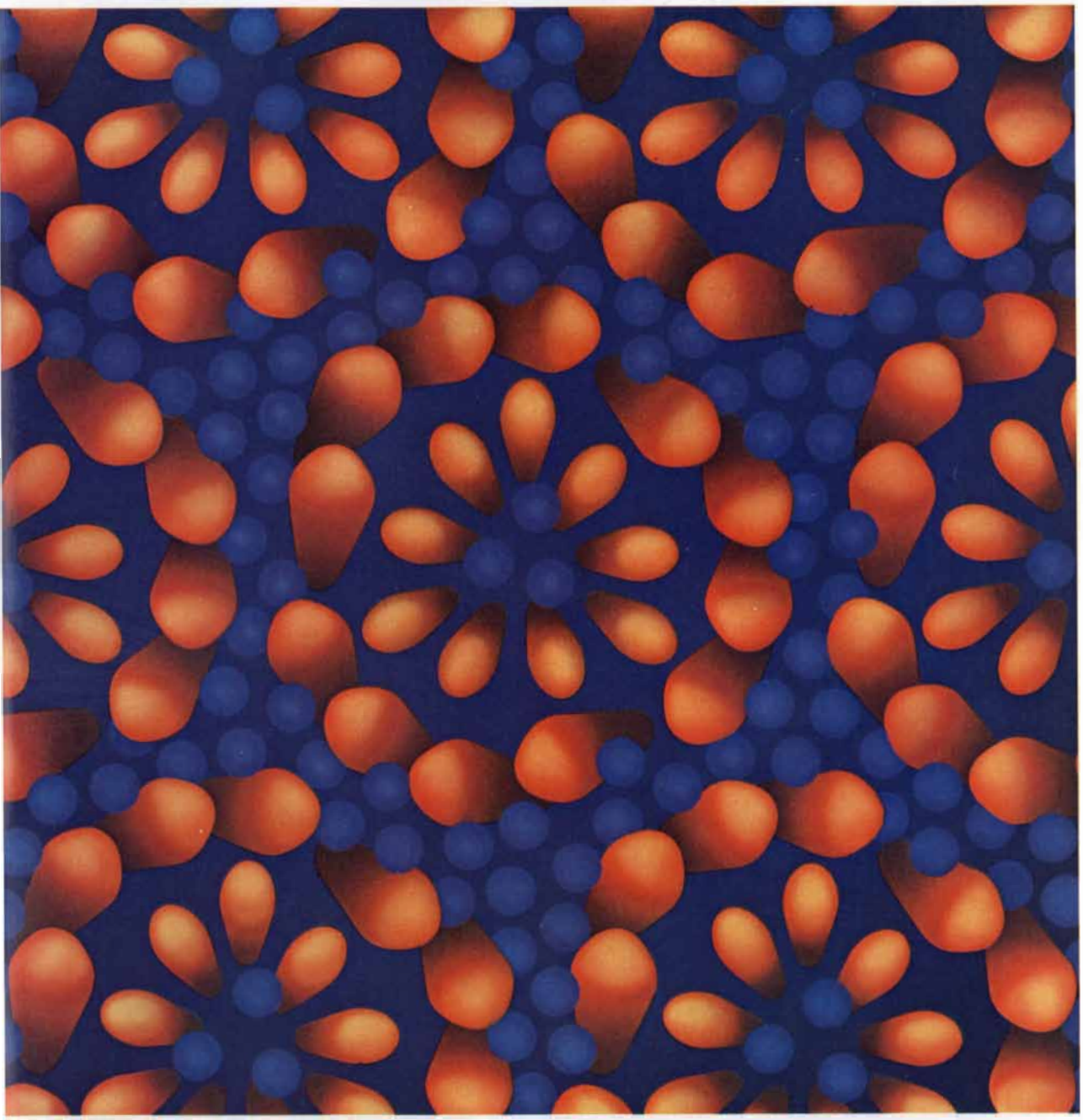
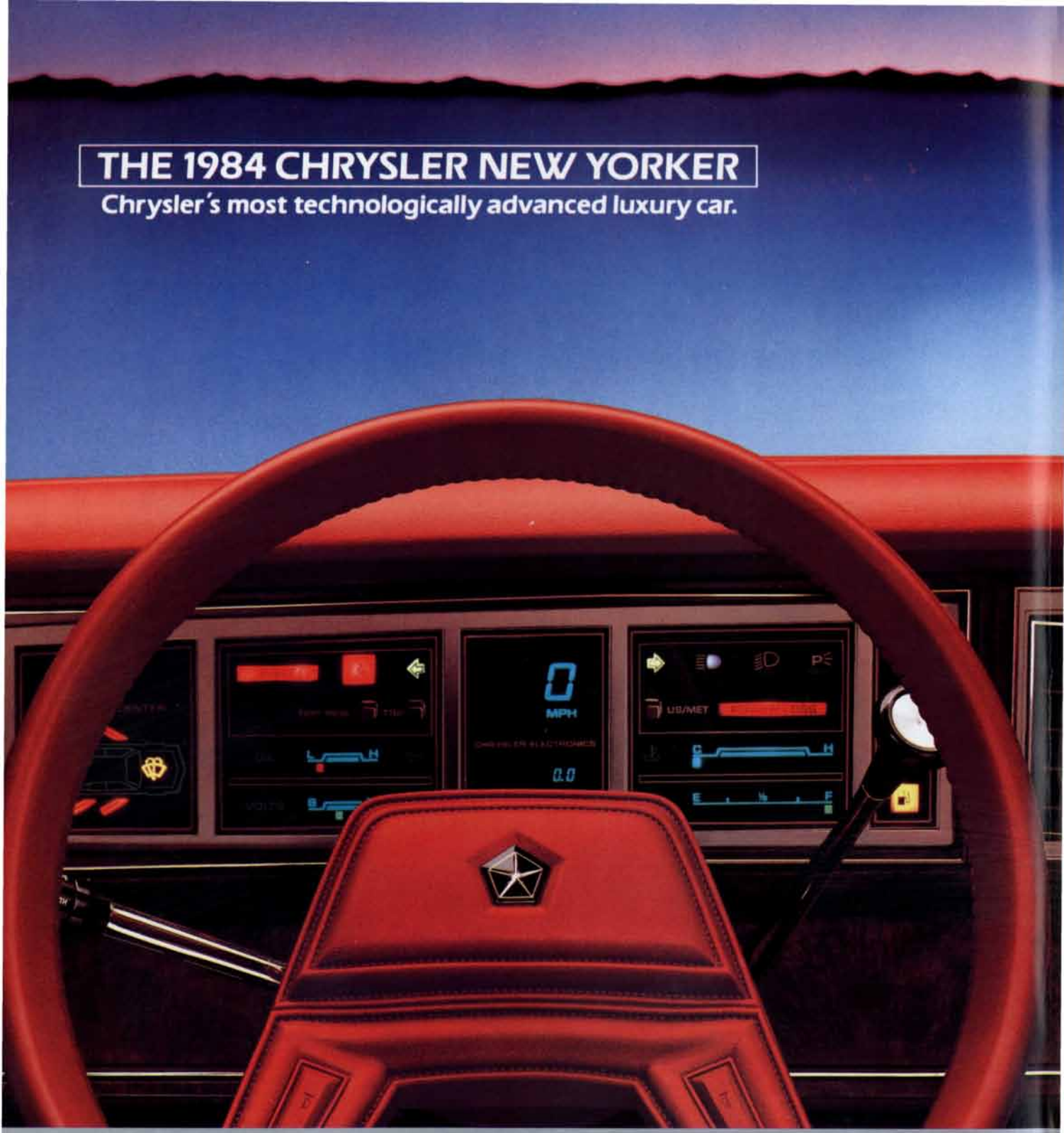
SCIENTIFIC AMERICAN



CELL-MEMBRANE PROTEINS

$2.50

February 1984



You have never been so well informed,

There has never been a luxury car like it. The sleek new Chrysler New Yorker brings you impressive new electronics, a superior Protection Plan and the option of exhilarating new turbopower. Without question, this is Chrysler's most technologically advanced luxury car.

Compare it carefully to other luxury cars. You enjoy front wheel drive for better traction on wet surfaces. You ride quietly and smoothly, surrounded by plush room and a bounty of luxuries. And you are served by ingenious new electronics.

You command new electronics and turbopower.

Enter the New Yorker and take control. The new graphics message center informs you, and the electronic voice alert monitors and reports on 11 vital functions. New electronic fuel injection gives you better performance. And you can add turbocharging, the exciting new way to achieve even more power. Just step on the accelerator and your fuel-injected turbocharged engine delivers true luxury car performance at its finest.

Notice that your car has been carefully fitted, finished and

*Whichever comes first. Limited warranty. Deductible applies. Excludes leases. Ask for details at your Chrysler-Plymouth dealer. **Base 2.2 liter engine. Use EPA National Highway Traffic Safety Administration safety recalls

so comforted and protected in a luxury car.

equipped. Among your standard luxuries are automatic transmission, power windows, power steering, power brakes and dual remote controlled power mirrors. Your front seats are new individual pillow style recliners, and of course, you can choose the luxury of rich leather seating.

5 year/50,000 mile Protection Plan.

The New Chrysler Technology brings you the best built, best backed American cars.† Your car's engine and powertrain are protected for 5 years or 50,000 miles;* and its outer body is protected against rust-through for the same period.

Nobody else offers you 5/50 protection in a luxury car.

And nobody else offers you such a luxury car.

Buckle up for safety.

35 EST. HWY. **25** EPA EST. MPG**

The best built, best backed American cars.

Chrysler

est. MPG for comparison. Your mileage may vary depending on speed, trip length and weather. Highway and California ests. lower.†Based on Corporate percent of for '82 and '83 models designed and built in North America.

Established 1845 **SCIENTIFIC AMERICAN** February 1984 Volume 250 Number 2

ARTICLES

DEPARTMENTS

SCIENTIFIC AMERICAN (ISSN 0036-8733), PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. SECOND-CLASS POSTAGE PAID AT NEW YORK, N.Y., AND AT ADDITIONAL MAILING OFFICES. AUTHORIZED AS SECOND-CLASS MAIL BY THE POST OFFICE DEPARTMENT, OTTAWA, CANADA, AND FOR PAYMENT OF POSTAGE IN CASH. SUBSCRIPTION RATE: $24 PER YEAR IN THE U.S. AND ITS POSSESSIONS; $33 PER YEAR IN ALL OTHER COUNTRIES. POSTMASTER: SEND ADDRESS CHANGES TO SCIENTIFIC AMERICAN, 415 MADISON AVENUE, NEW YORK, N.Y. 10017

SCIENTIFIC AMERICAN

CORRESPONDENCE

Offprints of more than 1,000 selected articles from earlier issues of this magazine, listed in an annual catalogue, are available at $1.25 each. Correspondence, orders and requests for the catalogue should be addressed to W. H. Freeman and Company, 4419 West 1980 South, Salt Lake City, UT 84121. Offprints adopted for classroom use may be ordered direct or through a college bookstore. Sets of 10 or more Offprints are collated by the publisher and are delivered as sets to bookstores.

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

Advertising correspondence should be addressed to C. John Kirby, Advertising Director, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, NY 10017.

Subscription correspondence should be addressed to Subscription Manager, SCIENTIFIC AMERICAN, P.O. Box 5969, New York, NY 10017. The date of the last issue on your subscription is shown in the upper right-hand corner of each month's mailing label. For change of address notify us at least four weeks in advance. Please send your old address (if convenient, on a mailing label of a recent issue) as well as the new one.

Name

New Address

Street

City

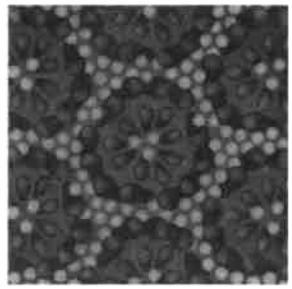
State and ZIP

Old Address

Street

City

State and ZIP



THE COVER

The painting on the cover is an artist's representation of a patch of a cell membrane: that of the salt-loving organisms called halobacteria. The membrane is composed of a bilayer of lipids in which molecules of the protein bacteriorhodopsin (*orange*) are embedded. The configuration of bacteriorhodopsin has been revealed by analysis of electron-diffraction patterns and electron micrographs (see "The Structure of Proteins in Biological Membranes," by Nigel Unwin and Richard Henderson, page 78). The view in the cover illustration is perpendicular to the surface of the membrane. One sees only the heads of the lipid molecules, drawn schematically as small blue spheres; the two long hydrocarbon tails of each molecule extend into the interior of the membrane. The orange objects are alpha helixes of the bacteriorhodopsin. A molecule of the protein consists largely of seven helixes, three of them perpendicular to the plane of the membrane and four of them tilted from the perpendicular. Each roughly circular cluster of alpha helixes seen in the painting is a group of three molecules.

THE ILLUSTRATIONS

Cover painting by George V. Kelvin

Page	Source
14	R. F. Bonifield
16–20	Ilil Arbel
38–39	Andrew Tomko
40–46	Albert E. Miller
47	Nathan Keyfitz, Harvard University
49–52	Gonzalo Vidal, University of Lund
53–54	Patricia J. Wynne
55	Gonzalo Vidal, University of Lund
56	Gonzalo Vidal, University of Lund (*top*); Bonnie Bloeser, Texaco Inc. (*bottom*)
57	Gonzalo Vidal, University of Lund
59–66	Andrew Tomko
79	Richard Henderson, Medical Research Council Laboratory of Molecular Biology
80–81	George V. Kelvin
83	Richard Henderson, Medical Research Council Laboratory of Molecular Biology (*top*); Nigel Unwin, Stanford University School of Medicine, and Richard Henderson, Medical Research Council Laboratory of Molecular Biology (*bottom*)
84	Nigel Unwin, Stanford University School of Medicine
86	George V. Kelvin
91–92	Nigel Unwin, Stanford University School of Medicine, and Richard Henderson, Medical Research Council Laboratory of Molecular Biology
93–94	George V. Kelvin
97	Fritz Goro
98–107	Allen Beechel
111	British Museum
112–116	Alan D. Iselin
117	Vincent Scheil
118	Alan D. Iselin
121–126	Tom Prentiss
128–135	Ilil Arbel
137	Ilil Arbel
138–143	Michael Goodman

LETTERS

Sirs:

The article by Herbert F. York ["Bilateral Negotiations and the Arms Race"; SCIENTIFIC AMERICAN, October, 1983] makes gloomy reading, and it gives the impression that the nuclear stalemate is inevitable and will continue for a long time. I am worried that it may lead to the spread of apathy if people do not realize that the situation *cannot* stay the same: if it does not get better, it will get very much worse. Not only does the number of nuclear weapons continue to increase but also the qualitative arms race has spawned MIRVed missiles and now those with terminal guidance. Such changes (initiated in the West) threaten the ability to retaliate, on which mutual deterrence rests. Nuclear war is therefore becoming likelier.

It is also unfortunate that Dr. York's generally balanced article should end by offering hope only if the Russians change. Since talks have so often broken down over the problem of verification by on-site inspection, it seems to me that the West has a terrible responsibility in at least two ways. First, the U.S. has failed for seven years to ratify the Peaceful Nuclear Explosions Treaty, which allowed some on-site inspection and might have been a prototype for more far-reaching arrangements. Second, the cruise missiles now being introduced are, as a recent article in *Newsweek* puts it, "a verification nightmare" that with other developments may "put an arms-control agreement all but out of reach." I do not know how such developments affect Western credibility in Russian eyes, but they destroy it in mine.

I hope that many of your readers feel, as I do, that if a voluntary moratorium on testing could operate in the 1950's, it cannot be beyond the wit of humankind to adequately verify a total nuclear freeze in the 1980's.

ROBERT WALL

Bedford, England

Sirs:

I am not quite as pessimistic as Mr. Wall's letter implies, and I do not believe that revolutionary change in Russian political practices is a sine qua non for further progress.

Political repression in the U.S.S.R. is not simply an internal problem; it has many ramifications in international affairs, and one of the most important of them is the way it combines with the huge size of that country to make verification of most any kind of arms-control agreement uniquely difficult. Even so, I believe further progress can be made if, on the American side, our leaders adopt a somewhat less paranoid attitude toward the problem, and if, on the Russian side, they fully and frankly recognize that they are a unique case and that they must make special concessions designed to overcome the extraordinary difficulties their political system creates. Indeed, their understanding of the problem and their willingness to at least partially face up to it have slowly but steadily evolved over the years. This is evidenced, as Wall rightly points out, by the provision for on-site inspections (albeit in certain rare and unusual circumstances), by their acceptance of the Peaceful Nuclear Explosions Treaty and, more important, by their acceptance in principle of both on-site inspection and in-country seismic stations ("black boxes") during the now defunct negotiations on a comprehensive test ban. On our side there has also been some evolution in the right direction over the years, although it must be said that the current Administration's attitudes constitute a reversal—I hope temporary—in the longer trend.

HERBERT F. YORK

Science, Technology and Public Affairs
University of California
San Diego

Sirs:

In Edward J. Wasp's article on slurry pipelines [SCIENTIFIC AMERICAN, November, 1983] he discusses the requirement to dewater a coal slurry for use in a power plant. Whereas in the past it was considered essential to separate water from coal before burning slurry in boilers, recent research has shown that this is not necessary. Direct injection of coal-water slurry in the liquid phase through atomizers with satisfactory combustion has been demonstrated in furnaces and boilers. The water in the slurry (30 percent) reduces the flame temperature and modifies the heat transfer and chemical reaction, but ignition and completion of combustion are achieved by appropriate modification to the burner. In addition to the successful burning of slurries in furnaces for industrial and power-plant boilers, tests are being conducted with micronized coal-water slurry for direct injection into gas-turbine and diesel engines. The feasibility of burning coal using liquid handling, transportation and injection systems could lead to a large-scale increase in the utilization of coal as an alternative to petroleum.

NORMAN CHIGIER

Department of Mechanical Engineering
Carnegie-Mellon University
Pittsburgh, Pa.

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

FEBRUARY, 1934: "The huge Van de Graaff electrostatic generator, which recently underwent its first tests at the Round Hill research station of the Massachusetts Institute of Technology, is designed to develop direct current at approximately 10,000,000 volts. Up to the time the Van de Graaff generator was designed, science with the best means at its disposal had not been able to obtain continuous direct current at more than 800,000 volts. The use of this great generator for research is expected to mark the beginning of an era of extraordinary significance in which extremely high voltages will make possible investigation of some of the most fundamental secrets of nature. The generator consists of two separate units, each with a polished aluminum sphere 15 feet in diameter resting on a hollow cylindrical insulating column 25 feet high and six feet in diameter. The generator was built for the specific purpose of atomic research: the bombardment of nuclei, the centers or kernels of atoms."

"The Institute of Geographical Exploration at Harvard University has performed a wonderful feat of making an air map of Massachusetts, the first state to be completely covered. The area of 8,039 square miles was the largest ever surveyed in this fashion. The altitude of the airplane was 15,000 feet, and the scale of the photographs was 1/30,000, or two inches to the mile. Only 800 composite photographs were taken, and the total flying time was limited to 24 hours 40 minutes. This great achievement was made possible by the use of the Fairchild Five-Lens Aerial Camera, which has virtually revolutionized the art of aerial photography."

"The heart of a reflector telescope is the mirror, and for many years the standard mirror coating has been of silver. Even in the best atmosphere a silver coating deteriorates badly and has to be replaced in from six months to a year. A better metallic film is much to be desired, and one with almost ideal properties has now been found. It is of a familiar but most unexpected composition—aluminum! One might fear that this coating too would oxidize and be spoiled, but it covers itself with a uniform, transparent and almost infinitesimal thin film of oxide and undergoes no further change. Aluminum oxide is one of the hardest substances known, appearing in nature as corundum or in its precious forms as ruby or sapphire, and the tenacious coating that automatically forms on the mirror surface has extraordinary protective powers."

"The familiar automobile, when moving at today's normal speeds, exerts more power to overcome air resistance than to conquer ground friction. Yet motor-car streamlining has hitherto been almost neglected by manufacturers. Speed has been achieved by the mere application of brute power—and a tremendous wastage of gasoline. If by 1940 true streamlining of cars has not been adopted, and if average speeds, as projected by present trends, reach 64 miles per hour, the typical energy expenditure will be 64 horsepower. Trusting that gasoline remains at about 20 cents per gallon, the cost of transportation would become $1.62 per hour; of this amount only 31 cents would be devoted to getting over the ground. The rest would be expended on invisible air."

FEBRUARY, 1884: "Professor S. P. Langley, astronomer at the Allegheny Observatory, lately gave the following views on the brilliant sunsets that have been observed in recent months: 'At first I supposed the sunset matter a local phenomenon, but when the reports showed it to have been visible all over the world, it was obvious that we must look for some equally general cause. There is no improbability that I can see in supposing the eruption at Krakatoa may have charged the atmosphere of the whole planet (or at least of a belt encircling it) for months with particles sufficiently large to scatter the rays of red light and partially absorb the others.'"

"Little by little patient workers in European laboratories are tracing the life histories of the minute organisms that are found in the blood and tissues of men and animals affected with various diseases, and are learning to distinguish those that may be said to cause disease from those that only accompany or follow it. It is a shame that while these investigations are being pushed in Germany, France and Great Britain nothing of the sort is going on in this country, the good commencement that was made in this direction by the National Board of Health having been totally abandoned for want of funds."

"With the introduction of gelatine silver paper, which has the property of being extremely sensitive to light, enlarged pictures can now be readily made in a few minutes with artificial light at night. Expensive apparatus and lenses, such as are used in solar printing upon the common albumenized sensitive paper, are dispensed with, and in their place a simple camera or a magic lantern with an ordinary lamp can be employed."

"It is not generally known that pyrites are now extensively used for making sulphuric acid in this country. Like many other chemical industries, the manufacture of sulphuric acid is imported from England, where the acid has been made from pyrites for the past 20 years. As might be expected, the introduction of the use of pyrites in this country at once caused a drop in the price of sulphuric acid, so that now the acid can be obtained almost as cheaply in this country as in the English market. From the fact that sulphuric acid is so largely used in the manufacture of chemicals, and more especially fertilizers, it follows that cheaper acid cannot fail to exert an immense and beneficial influence upon the chemical manufactures of the United States as well as upon the agricultural industries."

"It is reported that on the 10th of September last a well was tapped at Baku in Russia from which petroleum commenced to spout with a jet 300 feet high, at the rate of two million gallons daily. The effect of this sudden outburst is disastrous to the district, pending arrangements for disposing of such a vast quantity of oil. The great local refining firm of Nobel Brothers have 14 wells capped over and idle, it being cheaper for them to buy oil than to use their own. This firm announced that by next spring they will be able to distill 75,000,000 gallons of kerosene and to transport 90,000,000 gallons. As yet the Baku oil has only supplied the Russian, Austrian and East German markets via the Volga, but a new line of railway just opened will convey the product to the southern European markets. The supply is regarded as inexhaustible and is expected to keep down the value of petroleum oils and spirits in Europe."

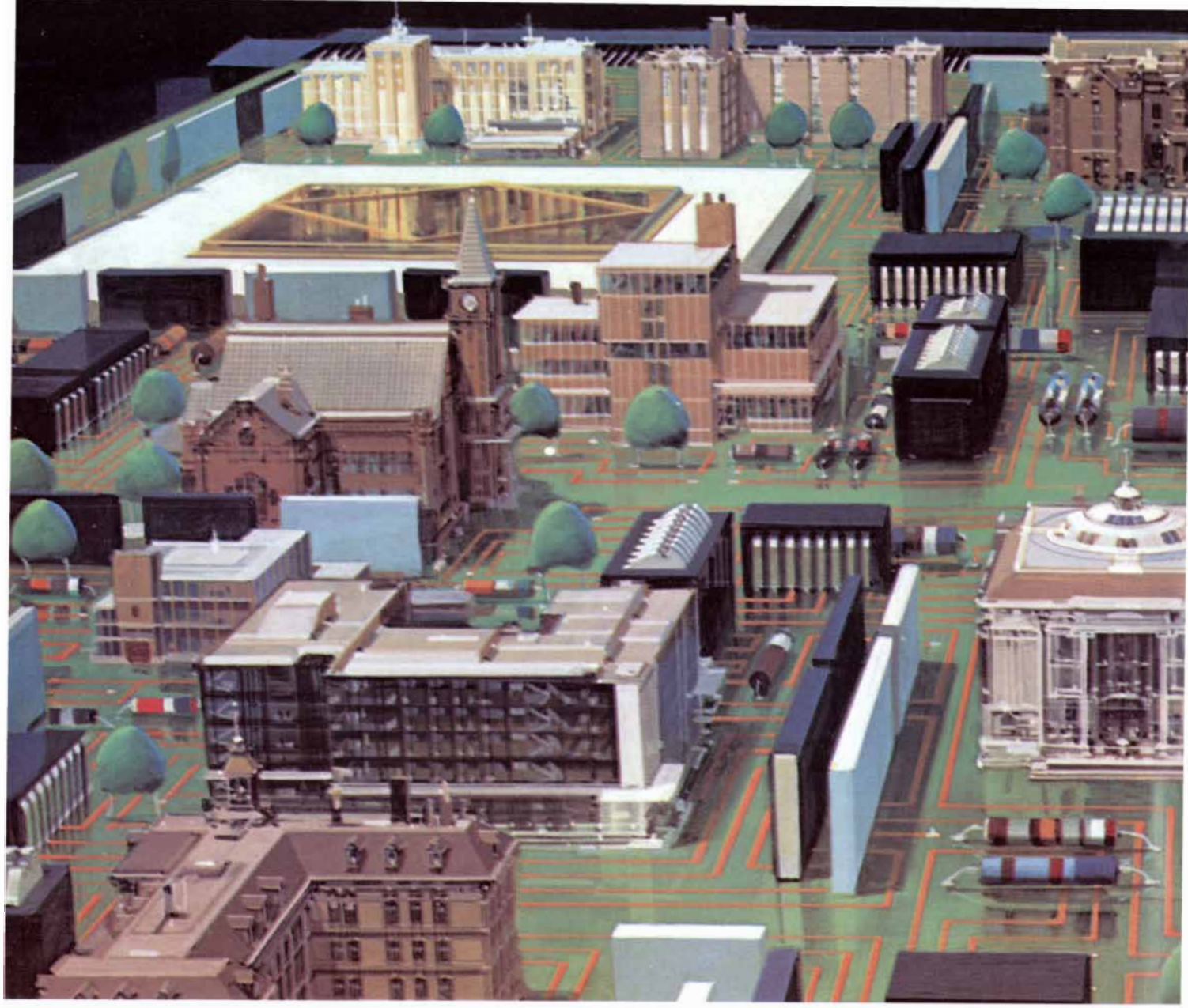
"A body in England calling itself the Society for Psychical Research is addressing a series of interrogatories to the public in relation to hallucinations and dreams. An invitation is issued to all people who think they have seen ghosts or specters to state their experiences. There are some preliminary questions that ought to be asked. Has any group of presumably sane men a moral right to instigate the crazy public to formulate its 'mysterious' experiences? We know that the most disastrous consequences sometimes ensue to weak brains from dwelling too intently on fixed ideas. Such things exist only in the imagination of the persons who are subject to them."

CIMARRON '84
THIS ONE'S GOT THE TOUCH.

The Cadillac touch. It's Electronic Fuel Injection, matched by a tenacious Touring Suspension that comes to grips with the road. It's leather-faced front bucket seats with lumbar support behind a leather-trimmed steering wheel that lets you know you're in control. It's new grille and taillight styling. Laser-accurate quality fits. It's Cimarron '84...with something no other car in its class has. The Cadillac touch.

BEST OF ALL...IT'S A CADILLAC.

Let's Get It Together...Buckle Up.



The IBM 4341 super-mini at universities

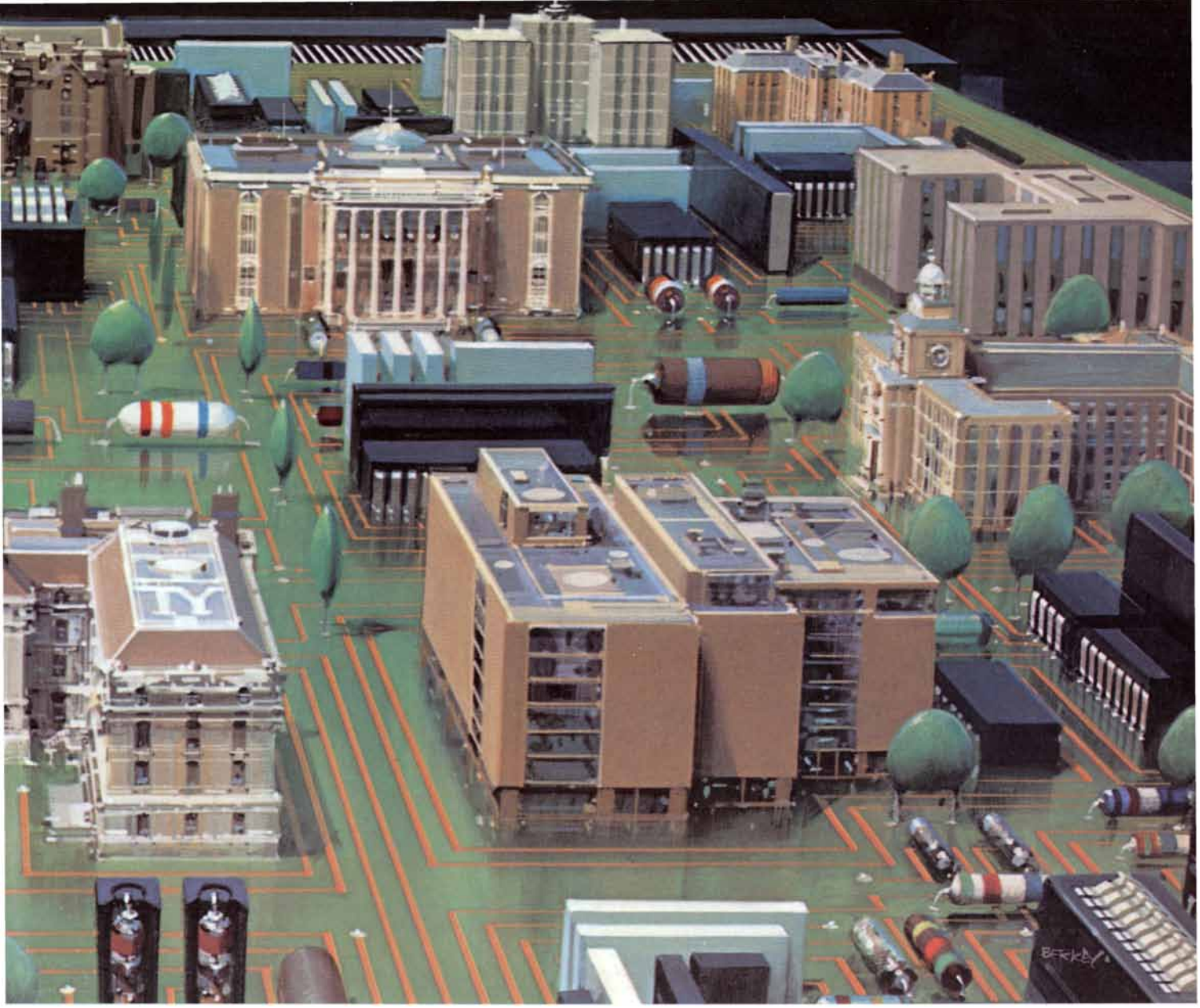
The BMOC, from engineering freshmen to Ph.D. researchers.

That's "Big Machine On Campus" at two big universities.

The University of Maryland

Engineering freshmen start using the IBM 4341 after only a few lectures. "Engineers will utilize the computer all their professional lives. The 4341 is so easy to use our students turn to it instinctively to solve problems," says Professor William Wockenfuss, who developed the introductory course.

Instructors are enthusiastic. "We know students use it. We get a rundown on how often they log in," says John Horan. "Log-in commands help them interact with the 4341 with no technical jargon," adds Aliakbar Elahi.



Dr. Glenn Ricart, Director of the Computer Science Center, remarks, "With the 4341's 'Help' command, students pick themselves up if they stumble. That's critical in how they'll relate to computers in the future."

"The university's 4341s serve as a major resource for interactive computing," says James S. Harvison, Acting Assistant Vice President. "We have five 4341s installed at different campuses serving our university community. They are a definite asset to our instruction and research computing network."

The Johns Hopkins University

The 4341 is used by The Center for Metropolitan Planning and Research. Teams of Ph.D.'s, with Dr. Steven Gottfredson, perform complex analyses of huge data bases. One study takes ten-year census figures and extracts small samples to test the effects of federal programs on inner-city housing. Another breaks down census data to make it useful for planning by fire, sanitation and education departments.

Professor Gottfredson explains that "machine capacity and speed are vital because of the size of our data sets. We also handle many interactive algorithms, and the 4341 gives us fast access to information."

The 4341 is used in other areas at Johns Hopkins, including studies of cancer frequency and astrophysics applications in preparation for the launch of the space telescope. Vice Provost of the university, Richard Zdanis, says, "We decided on a 4341 because it's easy for inexperienced users and can support a variety of peripherals."

IBM systems, using microcode-assisted VM/CMS, help engineers and scientists in industry, research, and universities think about the problem, not worry about the computer. We'd like to help you, too.

For an informative brochure, please write IBM Engineering and Scientific Marketing, 1133 Westchester Avenue, White Plains, N.Y. 10604.

IBM®

Telecommunication architects: The communications revolution is changing everyone's way of doing business. And of doing practically everything else, too. Chances are you'll find one or more Contel companies at the leading edge of these changes. Finding ways to get the "un-do-able" done. And done better. There's a whole group of Contel subsidiaries that plan and design, that construct, install and maintain. Successful innovators in every phase of telecommunication, which has been called "the industry that defines the 80's." At Continental Telecom Inc. you'll find us defining telecommunication. From telephony to satellites.

WRITE TO CONTINENTAL TELECOM INC., DEPT.501, 245 PERIMETER CENTER PARKWAY, ATLANTA GA 30346.

CONTEL

THE AUTHORS

NATHAN KEYFITZ ("The Population of China") is Andelot Professor of Sociology, Emeritus, at Harvard University and Lazarus Professor of Social Demography, Emeritus, at Ohio State University. He is also currently in charge of the population program of the International Institute of Applied Systems Analysis in Laxenburg, Austria. He received a B.Sc. in mathematics from McGill University and a Ph.D. in sociology from the University of Chicago. After a career in Canadian government service, where he was a research statistician, he entered academic life. He has taught not only at Harvard and Ohio State but also at the University of Toronto, the University of Chicago, the University of California at Berkeley and most recently in the department of preventive medicine and biostatistics at Toronto. Although he is primarily a sociologist, he has devoted much of his career to quantitative problems, at first those of statistical sampling and later those of mathematical demography. In 1982 Keyfitz made an extensive tour of China, on which his article in this issue is partly based.

GONZALO VIDAL ("The Oldest Eukaryotic Cells") is associate professor in the department of geology at the University of Lund in Sweden. Born in Spain, he moved to Sweden in 1963. There he received his education at Lund: the Swedish equivalent of the M.S. degree in 1969 and the equivalent of a Ph.D. in 1973. (In 1979 he defended his thesis for the new Swedish doctoral degree and was awarded the title of docent in the geological sciences.) In 1974 he went to New Zealand on a postdoctoral research fellowship in the Department of Scientific and Industrial Research; in 1977 he returned to Lund as a research associate of the Swedish Natural Science Research Council, based in the university's department of geology. He was appointed associate professor in 1980, maintaining a Natural Science Research Council position in pre-Quaternary micropaleontology. Vidal traces his interest in the history of life on the earth to a childhood visit to the natural history museum in Madrid, where he saw his first dinosaur skeleton. He remarks that he has come to enjoy cooler climates and has developed a keen interest in doing field work in arctic regions.

CURT COVEY ("The Earth's Orbit and the Ice Ages") is assistant professor in the department of meteorology and physical oceanography at the University of Miami. Starting out in biology, he received his bachelor's degree at the Massachusetts Institute of Technology in 1973 and his master's at the University of California at Santa Barbara in 1975. Switching to geophysics, he was awarded his Ph.D. at the University of California at Los Angeles in 1982. At the time his main interest was the atmosphere of Venus; he and Gerald Schubert wrote an article on the subject in the SCIENTIFIC AMERICAN issue of July, 1981. In 1982 and 1983 Covey was a postdoctoral fellow at the National Center for Atmospheric Research. There he worked on, among other things, the use of oxygen isotopes in the study of ancient climates and the long-term effects of nuclear war as they are expressed in a model of the general circulation of the atmosphere.

NIGEL UNWIN and RICHARD HENDERSON ("The Structure of Proteins in Biological Membranes") are respectively professor of structural biology at the Stanford University School of Medicine and a member of the scientific staff of the Medical Research Council Laboratory of Molecular Biology in Cambridge, England. Unwin was born in England but grew up in New Zealand; he received his bachelor of engineering degree in metallurgy at the Otago School of Mines. He was awarded his Ph.D. in metallurgy at the University of Cambridge in 1968; thereafter he joined the Laboratory of Molecular Biology, where he began his work with Henderson on applying the technique of low-dose electron microscopy to the structure of proteins in biological membranes. Among Unwin's other interests are the structure of ribosomes and of pores in the membrane of the cell nucleus; he remarks that he would enjoy attempting to develop new physical techniques that would make it possible to describe biological structures in more detail and with greater accuracy. He accepted a professorship at the Stanford School of Medicine in 1980. Henderson was graduated in physics from the University of Edinburgh and did his Ph.D. work on the structure of enzyme molecules with David Blow at the Laboratory of Molecular Biology. After spending three years at Yale University as a Helen Hay Whitney Fellow, Henderson returned to the Laboratory of Molecular Biology. There he has mainly been occupied with the subject of the present article: determining the structure of membrane proteins by means of X-ray diffraction and electron microscopy.

DONALD H. LEVY ("The Spectroscopy of Supercooled Gases") is professor of chemistry at the James Franck Institute of the University of Chicago. His B.A. is from Harvard University

(1961), his Ph.D. from the University of California at Berkeley (1965). From 1965 to 1967 he was a postdoctoral fellow in theoretical chemistry at the University of Cambridge. Levy joined the Chicago faculty in 1967 and was appointed professor of chemistry in 1978.

JÖRAN FRIBERG ("Numbers and Measures in the Earliest Written Records") is professor of mathematics at the Chalmers University of Technology at Göteborg in Sweden. He received his Ph.D. in mathematics at the University of Lund in 1963; he has been at Chalmers since 1967. At first his research was in partial differential equations and approximation theory. Telling a class of liberal-arts students about a Babylonian table of solutions to the Pythagorean equation, he became curious about the purpose of the table and how it was constructed. He is now, as he puts it, "totally engulfed in my research in the history of cuneiform mathematics." Friberg has been a visiting professor at the University of Minnesota and the University of Wisconsin at Milwaukee, and a visiting lecturer and visiting fellow at Yale University.

SUZANNE W. T. BATRA ("Solitary Bees") is a research entomologist with the United States Department of Agriculture. She received a B.A. in zoology at Swarthmore College and a Ph.D. in entomology from the University of Kansas. After working at the University of Kansas, Punjab Agricultural University in India and Utah State University, she joined the Department of Agriculture as a consultant in entomology in 1970; she has been a research entomologist in the department since 1974. One of her main research interests has been the identification of insects with activities beneficial to humankind; among such insects are the solitary bees. She is married to Lekh R. Batra, a mycologist who also works for the Department of Agriculture. Combining their research interests, they wrote "The Fungus Gardens of Insects," which appeared in the November 1967 issue of SCIENTIFIC AMERICAN.

TIMOTHY A. SALTHOUSE ("The Skill of Typing") is associate professor of psychology at the University of Missouri. His B.A. in psychology is from the University of California at Santa Barbara; his M.S. and Ph.D. in experimental psychology are from the University of Michigan. He joined the Missouri faculty in 1976. His primary research interests are in the areas of perceptual and cognitive skills and in the effects of aging on cognitive functioning. Salthouse recently returned from a year spent as a fellow at the Andrew Norman Institute for Advanced Study in Gerontology at the University of Southern California.

COMPUTER RECREATIONS

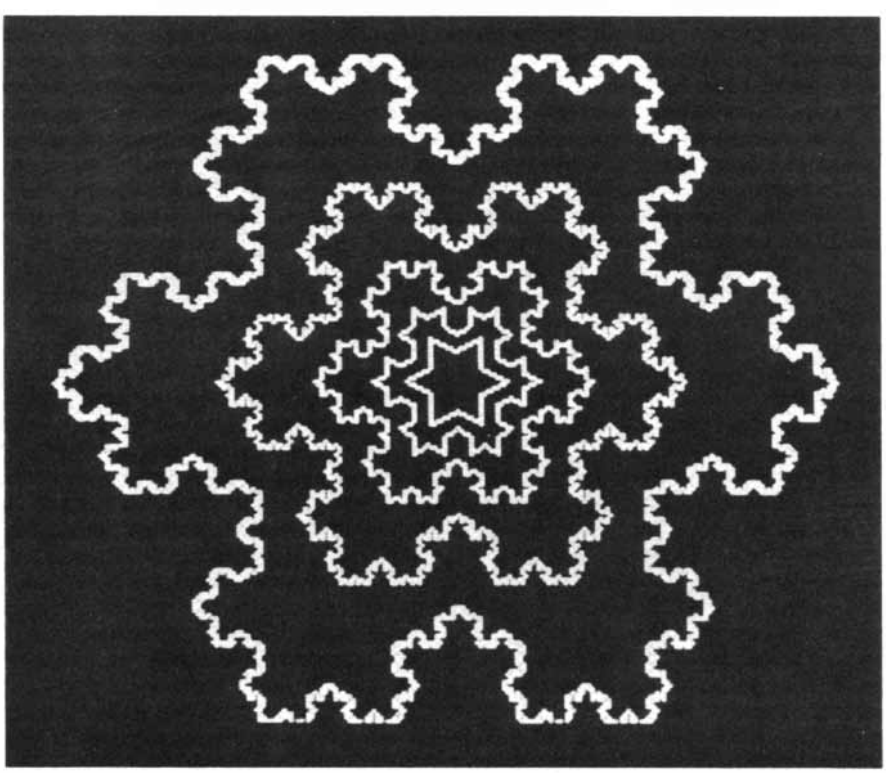
Turning turtle gives one a view of geometry from the inside out

by Brian Hayes

What is a circle? Is it the limiting case of an n-sided polygon as n goes to infinity? Is it the special case of an ellipse whose two foci coincide? Is it the locus of all points in a plane equidistant from a given point?

All these definitions are correct, of course, and more could be given. Consider this one: A circle is the figure formed by walking forward a little, then turning right a little, and repeating this sequence of steps until you have turned exactly 360 degrees. This last definition is distinctively different from the others. Rather than describing a circle or specifying some of its properties, it gives a procedure for constructing a circle. Furthermore, the procedure itself has a special character: it is stated entirely in terms of the "local" properties of the circle. The curve can be created by attending to one's immediate neighborhood only; there is no need for an overview. There is no need to know where the center of the circle lies, or even to determine the radius.

Procedures of this kind form the essence of a new way of thinking about geometry. It has been called experiential geometry, because one is invited to imagine moving around in the world of geometric figures, whereas other approaches to geometry tend to set the figures apart in a space separate from the observer. The difference is somewhat like that between exploring a landscape and reading a map. The experiential approach is particularly well suited to investigating geometric ideas with the aid of a computer. The procedural definitions are readily converted into computer programs, which can be executed to draw the shapes specified.



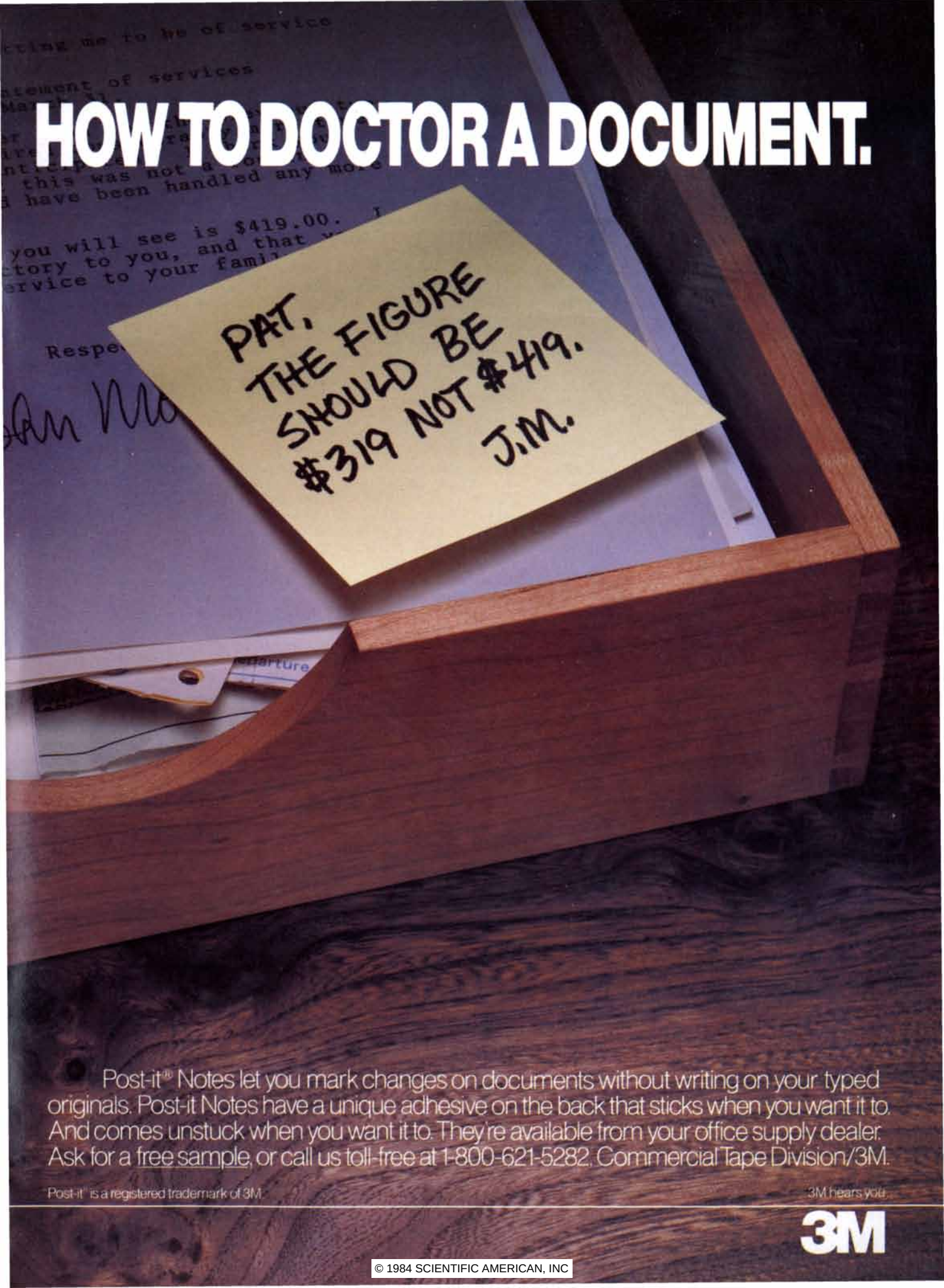
A sequence of nested snowflake curves, drawn by the turtle

The new way of thinking about geometry has come to be known as "turtle geometry." It is closely connected with the programming language Logo, which in turn has its roots in the Massachusetts Institute of Technology. Logo was conceived in the 1960's by Seymour Papert of M.I.T., primarily as a language for introducing children to computers. Many others have since contributed to its development and to its applications both in education and in other fields. Among them are Harold Abelson and Andrea A. diSessa of M.I.T., who have set forth the ideas underlying turtle geometry in a remarkable expository work: *Turtle Geometry: The Computer as a Medium for Exploring Mathematics.*

The original turtle was a mechanical device: a small wheeled vehicle whose movements could be controlled by instructions typed on a computer keyboard. The first such creature was built by the British neurophysiologist W. Grey Walter in the late 1940's. It had a dome-shaped cover somewhat like a turtle's shell. A mechanical turtle can move forward or backward and can change direction by pivoting in place. A pen can be mounted on the undercarriage, so that when the turtle is made to wander over a sheet of paper, it leaves a record of its path. Today such "floor turtles" are less common than "screen turtles," which move and draw on the surface of a cathode-ray tube. The turtle itself is represented on the screen by a simple triangular form, which moves in response to commands or programs entered at the keyboard.

Most methods of drawing with a computer employ a global system of coordinates, generally a Cartesian one, with two perpendicular axes. The position of every point on the screen is defined with respect to some origin, such as the lower left-hand corner, where the imaginary axes cross. Directions in the plane are also absolute. A command for drawing a line might be issued by specifying the two endpoints; for example, giving the points {0,0} and {100,0} might draw a horizontal line 100 units long across the bottom of the screen.

It should be emphasized that in any system of global coordinates the effect of a command does not depend on the sequence of commands that went before. The situation is quite different in a turtle-geometry system. A line 100 units long can be drawn by giving the turtle an instruction such as FORWARD 100. Where the line appears on the screen, however, and what its orientation is depend entirely on the state of the turtle at



Post-it® Notes let you mark changes on documents without writing on your typed originals. Post-it Notes have a unique adhesive on the back that sticks when you want it to. And comes unstuck when you want it to. They're available from your office supply dealer. Ask for a free sample, or call us toll-free at 1-800-621-5282. Commercial Tape Division/3M.

Post-it® is a registered trademark of 3M.

3M hears you.

3M

the time the command is issued. At any moment the turtle has both a position and a heading. The FORWARD 100 command causes it to proceed 100 units from its current position in the direction defined by its current heading (that is, in the direction it is "facing"). Hence the same command has a quite different result if it is given when the turtle is in a different initial state.

A simple turtle-geometry system can be created with just two commands: FORWARD and RIGHT. FORWARD is followed by a number that specifies how many "steps" the turtle is to take. RIGHT is also followed by a value, which gives the number of degrees the turtle is to turn right from its present heading. Any plane figure can be drawn by a sequence of forward moves and right turns. (Note that a 90-degree left turn can be specified in two ways: as RIGHT 270 or as RIGHT −90.) Nevertheless, a practical turtle system generally includes BACK or REVERSE and LEFT commands as well. PENUP and PENDOWN determine whether or not the turtle draws a line along its trail. A few other commands that report on the state of the turtle and offer a means of specifying absolute coordinates are also commonly provided, but they are not essential to turtle geometry.

The kind of geometry done with a turtle is formally called finite differential geometry. It is finite because the turtle can move only in discrete steps. It is differential because all movements are defined in terms of the difference between the present position and heading and the subsequent one. This is another way of saying it is a kind of geometry concerned only with the local properties of lines, curves and surfaces; there is no reference to distant points or to the global properties of a geometric figure. It follows that turtle geometry is most useful for exploring the "intrinsic" properties of geometric figures, those that are defined entirely within the figure.

One idea that can be awkward to formulate in a coordinate system but that comes forth clearly in turtle geometry is that of curvature. In Cartesian coordinates the curvature of a line in a plane might be defined as the rate of change in the slope of the line; the slope in turn can be defined as the rate of change in the y coordinate as a function of the x coordinate. The turtle can keep track of curvature in a much simpler way: it is the total turning per unit of distance traveled. Thus in a circle defined by repeating the instructions FORWARD 1, RIGHT 1 the curvature is everywhere equal to 1.

The concept of total turning leads to other interesting findings. Consider the familiar theorems of Euclidean geometry that give the sum of the interior angles of a convex polygon. For a triangle the sum is 180 degrees, for a quadrilateral 360, for a pentagon 540 and so on. In other words, the interior angle is invariably an integer multiple of 180 degrees, and the integer is equal to 2 less than the number of sides. The total turning of the turtle is measured in terms of a different angle: it is neither the interior nor the exterior angle but the change in heading at each vertex. Theorems about the total turning in a polygon can also be formulated. They are somewhat different from the theorems that pertain to the total interior angle, and they are also more general.

The most fundamental theorem states that the total turning when the turtle traces a polygon must be an integer multiple of 360 degrees. The proof is simple and transparent. If a figure is closed (as every polygon is), the turtle must eventually return to its starting point, and when it gets there, it must have precisely the same heading it had at the outset. (Indeed, this observation serves as an effective definition of geometric closure.) If the heading is the same, however, then the total turning must be zero degrees or 360 or some multiple of 360.

For a convex polygon a stronger result can be proved, namely that the total turning is exactly 360 degrees. (A polygon is convex if a line connecting any two points on its edges never passes outside it.) The theorem itself, however, applies to any polygon, including polygons that are not convex (such as a star-shaped pentagon). It applies even to closed figures that have curved edges and therefore are not polygons.

In coordinate geometry the total interior angle of a convex polygon has few obvious connections with other ideas. In turtle geometry, on the other hand, the concept of total turning is a powerful tool rich in wider applications. For example, Abelson and diSessa show that it can be used to analyze the topology of a closed path. A topologically simple path is one that can have any number of edges, curves and convolutions, provided only that it does not cross itself; like a simple convex polygon, it has a total turning of 360 degrees. Adding a single loop (and a single crossing) to the path makes the total turning either zero or 720 degrees, depending on the handedness, or direction, of the loop. Each additional loop adds or subtracts another increment of 360 degrees. Because of this property, the topology of the path can be determined from a turtle's-eye view of it, with only a microscopic section being visible at any one time.

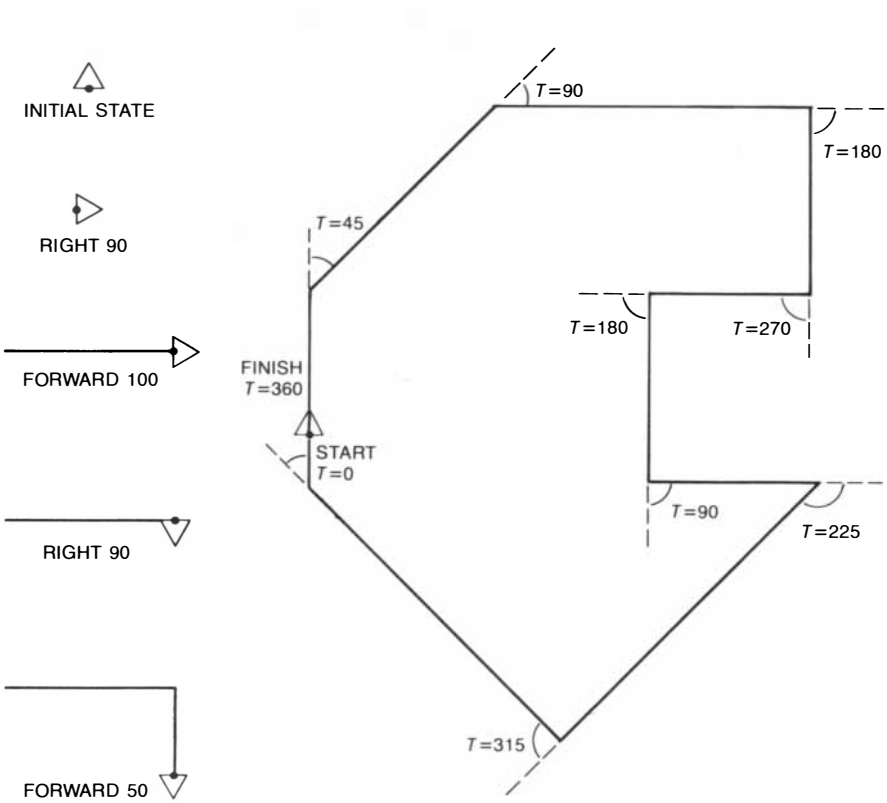
Total turning is also the basis of an algorithm that enables the turtle to find its way out of a maze. Abelson and diSessa state the algorithm as follows.

"1. Select an arbitrary initial direction, call it 'north,' and face that way.

"2. Walk straight 'northward' until you hit an obstacle.

"3. Turn left until that obstacle is on your right.

"4. Follow the obstacle around, keeping it on your right, until the total turn-



Some turtle commands **(left)** *and the concept of total turning* **(right)**

ing (including the initial turn in step 3) is equal to zero.

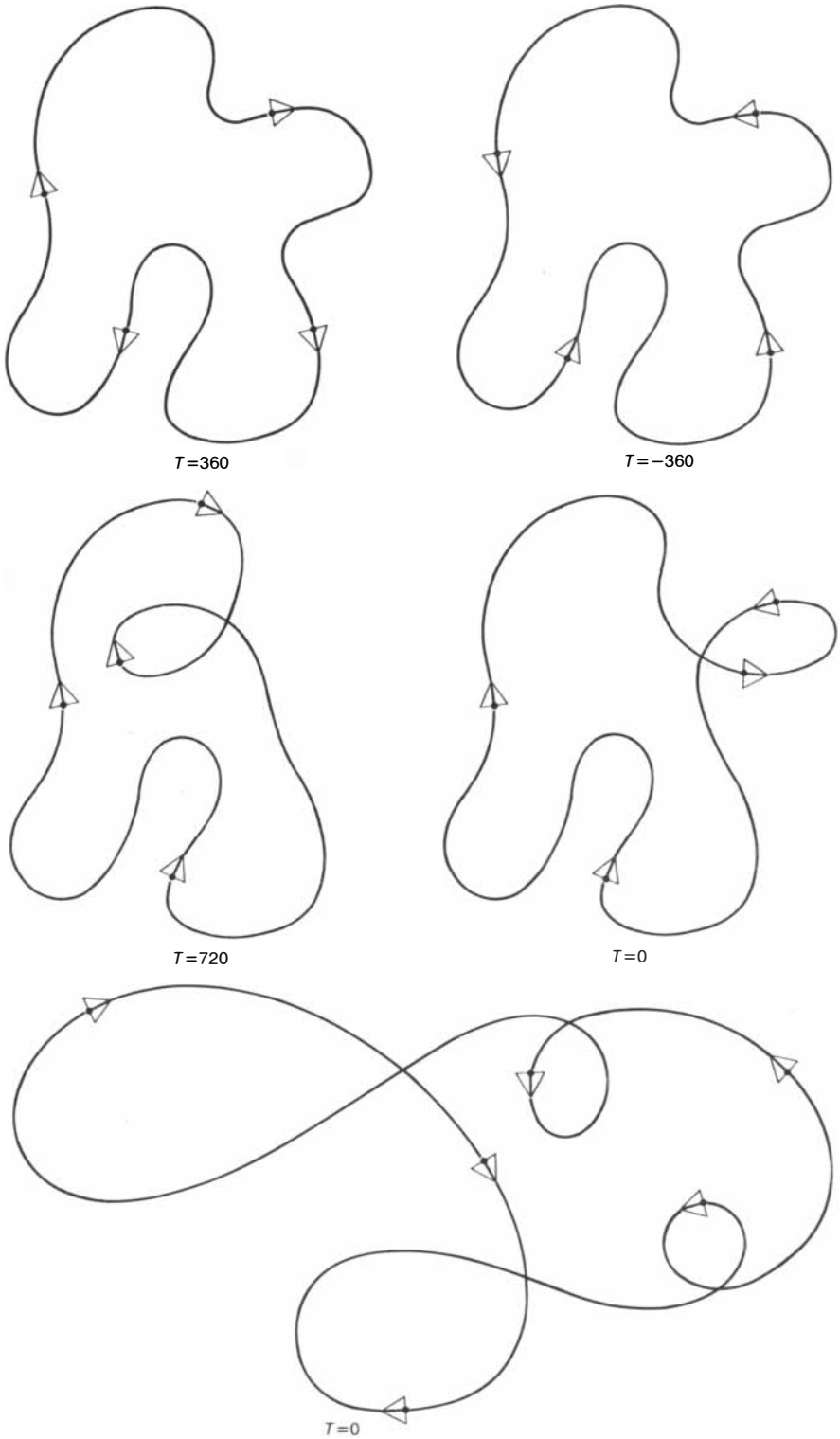
"5. Go back to step 2."

By this method the turtle can solve any fair maze (meaning one that actually has an exit). The procedure works because the only way to trap the turtle is to lead it into an endless loop, and keeping track of the total turning avoids that trap. Note that again a global problem—finding a route out of the maze—has been solved even though the turtle has information about only its local environment; there is no aerial view of the maze. The universal maze-solving procedure is called the Pledge algorithm, after John Pledge of Exeter in England, who developed it at age 12.

Still more applications of the total turning come up when the turtle is allowed to roam on surfaces that are not flat. On a sphere, for example, a closed path is found to have a total turning of less than 360 degrees. Furthermore, the amount of deviation from 360 degrees depends on the length of the path or, to be more precise, on the area it encloses. Readers familiar with the non-Euclidean geometries developed in the 19th century will recognize what is happening here: the deviation, or angular excess, is a measure of the curvature of the surface. In this instance it is important that the measurement can be made with only local and intrinsic information. The curvature one would most like to measure is that of the spacetime of the universe, and it will obviously have to be done by means of observations made within the universe. Abelson and diSessa explore this process in the final chapter of their book, which gives a turtle-geometric formulation of the general theory of relativity.

They also consider the meanderings of the turtle on the surface of a cube, which has the same topology as a sphere but a different geometry. Indeed, the geometry is bizarre. Suppose the cube is 100 units on a side and the turtle is initially in the center of a face and oriented parallel to an edge. Repeat the instructions FORWARD 100, RIGHT 90 three times. The result is a closed figure with three equal sides and three equal angles; it is also an equal-sided polygon whose vertexes are all right angles. Is it an equilateral right triangle or a three-sided square?

Another remarkable object of cubical geometry is the monogon, a one-sided polygon; it is a closed path formed by a turtle walk that has no turning at all. Even without a computer-driven turtle it is possible to see how a simple monogon can be formed: simply draw an "equator" on the cube. A more difficult question is this: Is there any combination of initial position and heading on the surface of the cube that does not close to form a monogon when the line is extended indefinitely? What if the heading must be specified in rational numbers? The answer will be found in *Turtle Geometry* (see chapter 6), but Abelson and diSessa would strongly discourage one from looking there. At one point they post the following warning: "DANGER—The next section contains 'premade' (already discovered) mathematics. It may be harmful to your imagination and should be used only as a last resort."



The topology of closed curves deduced from total turning

The discovery of a principle and the exploration of its extensions or generalizations are characteristic of the style in which turtle geometry is done. The computer greatly facilitates the

process: experiments are easy, and variations on a theme can be tried with little effort. It is a geometry for tinkerers.

Consider the following Logo procedure, discussed by Abelson and diSessa and by other writers on the language:

```
TO SQUIRAL :DISTANCE
FORWARD :DISTANCE
RIGHT 90
SQUIRAL :DISTANCE + 5
END
```

Here SQUIRAL is the name of the procedure and DISTANCE is a variable whose initial value is to be given when the procedure is executed. (The colon is a Logo convention for identifying variables.) The turtle is instructed to move forward this amount and execute a 90-degree right turn; then the SQUIRAL procedure is invoked again, but with a larger DISTANCE value. The result is a "square spiral" that grows outward toward the edge of the screen. (In the procedure as it is given here the spiral goes on growing indefinitely, although only a finite part of it can be displayed.)

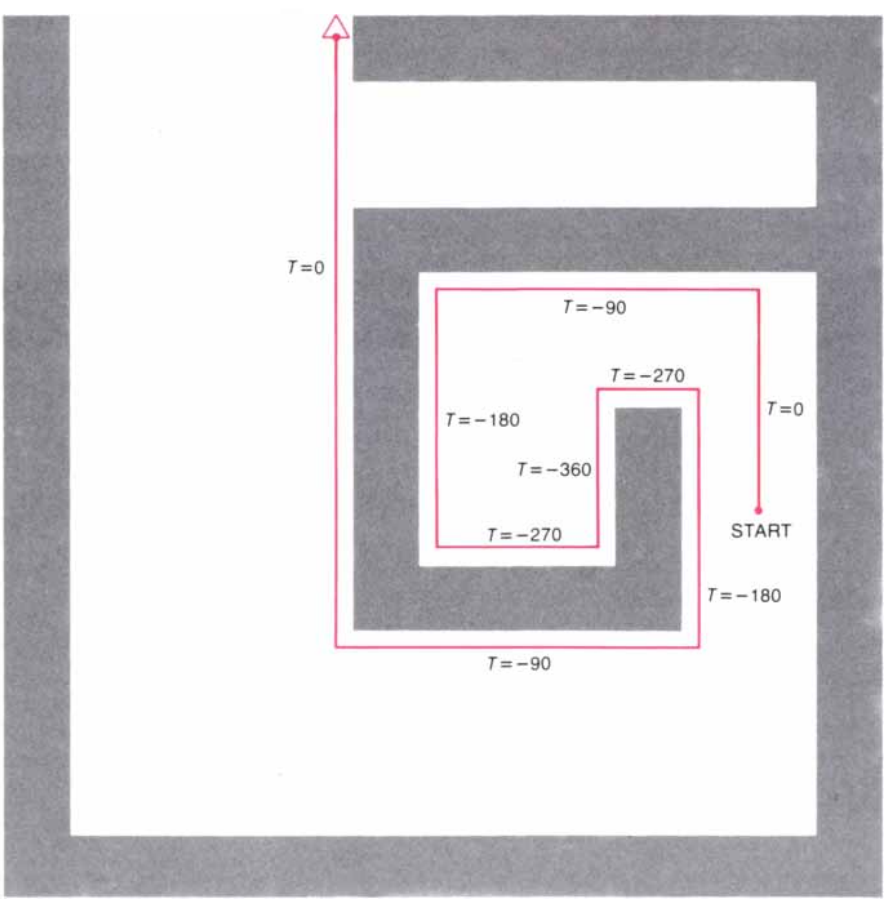
Many variations are possible. Changing the value of the constant that is added to DISTANCE each time the procedure is called merely alters the spacing between successive arms of the spiral. Multiplying by a constant value instead of adding one creates a spiral in which the arms get progressively farther apart in proportion to their distance from the center. Inserting a different constant angle in the RIGHT 90 instruction can convert the square spiral into a triangular or pentagonal or hexagonal one. An angle that differs only slightly from 90 degrees yields a set of nested "squares" that twist around their center, and so the vertexes form secondary spirals. A very small angle yields an approximation to a smooth "circular" spiral.

It is also possible to rewrite the procedure so that what changes with each invocation is the angle rather than the distance. The transformation is remarkable: instead of a single spiral that grows continuously outward the turtle draws an inward spiral and then an outward one of the opposite handedness, then another inward one and so on, creating a symmetrical array of spirals joined by their outermost loops. The underlying reason is that whereas distance can increase monotonically, angular measure is interpreted modulo 360, so that repeatedly adding a constant eventually returns the angle to a small value.

Consider only the subclass of curves formed when the initial angle is zero [*see illustration on page 20*]. All the patterns in this class are similar in basic form. The turtle creates a number of spirals of alternating handedness, then turns, retraces its path and after passing through its starting point creates another identical array rotated by 180 degrees with respect to the first one. Thereafter the turtle continues to retrace the same path indefinitely. The number of spirals in the pattern depends only on the angular increment, but the nature of the relation is not obvious. If the angular increment is not divisible by 8, the number of spirals is equal to one more than the largest factor of the increment that is not also a factor of 360. Can you perceive why? Can you predict the nature of the pattern when the increment is a multiple of 8?

It is the computer that makes possible such an exploratory approach to the family of spiral curves. Similar projects can be undertaken to explore the family of polygons, of tilings of the plane and of recursive figures such as the snowflake curve, which is made up of similar structures repeated at progressively smaller scales [*see illustration on page 14*]. The constructive, procedural nature of turtle geometry makes an important contribution to the process. It would be awkward at best to write an equation to specify the entire structure of the snowflake curve, whereas a procedure to generate the curve is readily broken down into a few elementary steps that are then carried out repeatedly.

Although turtle geometry and the Logo programming language have a strong historical connection, they are by no means inseparable. Abelson and diSessa note that turtles have been installed in at least two other languages (versions of APL and Smalltalk), and the programs in their own book are given not in Logo but in a related language they call Turtle Procedure Notation. They offer advice on the creation of a turtle system in BASIC and Pascal. A new language now being developed by members of the Logo group at M.I.T., called Boxer, also features a turtle.

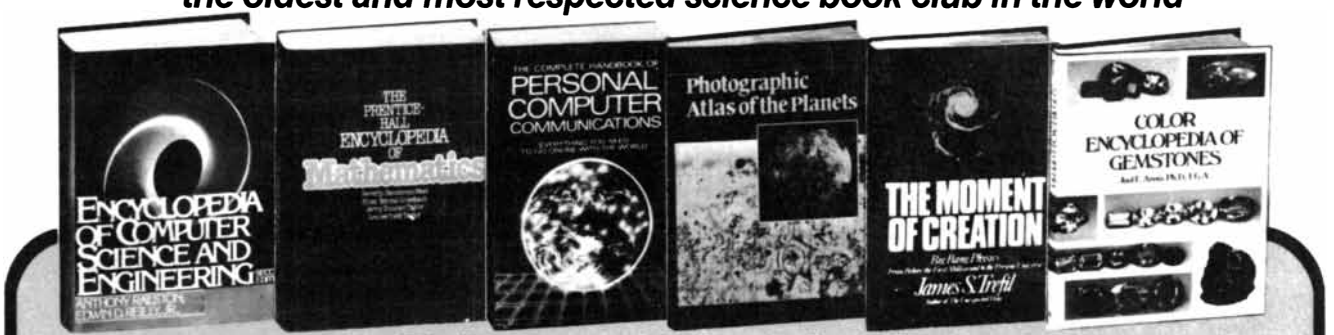


The turtle, by following the Pledge algorithm, escapes from a simple maze

Logo itself is a powerful general-purpose programming language. It has a strong kinship with Lisp, the list-processing language invented more than 20 years ago by John McCarthy. Indeed, the main differences between Lisp and Logo are matters of vocabulary and punctuation. The origin of Logo as a language for children has given it an anthropomorphic aspect that can be unsettling. The programmer writes as if he were speaking directly to the turtle, and occasionally the turtle talks back, issuing an error message such as "I don't know how to squiral." One should not be misled by this coyness into thinking the language is a plaything. The anthropomorphism is deliberate: it is part of a strategy to engage the programmer

The Library of Science

—the oldest and most respected science book club in the world



Take any 3 books for only $1.00 each (values to $74.85)

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

(Publishers' Prices shown.)

39796. THE COMPLETE HANDBOOK OF PERSONAL COMPUTER COMMUNICATIONS. Everything You Need to Go Online with the World. *Alfred Glossbrenner.* Softbound. **$14.95**

44904-3. ENCYCLOPEDIA OF COMPUTER SCIENCE AND ENGINEERING Second Edition. *Ralston and Reilly, Eds.* Thoroughly revised, updated and reorganized. 550 articles. Over 1,670 pages on virtually every aspect of the computer sciences. More than 750 illustrations. *Counts as 3 choices.* **$87.50**

37244. THE CAMBRIDGE PHOTOGRAPHIC ATLAS OF THE PLANETS. *G.A. Briggs and F.W. Taylor.* Authoritative, narrative text with more than 200 magnificent photos from the NASA archives. **$24.95**

70005-2. THE PRENTICE-HALL ENCYCLOPEDIA OF MATHEMATICS. *West et al.* 80 full-length articles. Over 650 pages. More than 700 illustrations. *Counts as 2 choices.* **$35.00**

61637-3. McGRAW-HILL ENCYCLOPEDIA OF PHYSICS. *Sybil P. Parker, Editor-in-Chief.* More than 760 articles. 1,350 pages. 1,035 illustrations. *Counts as 3 choices.* **$54.50**

52310-2. HANDBOOK OF PHYSICAL CALCULATIONS. Second Edition. *Jan J. Tuma.* A concise reference for the major tools of applied physics and engineering. *Counts as 2 choices.* **$39.95**

62913. THE MOMENT OF CREATION: An Introduction to Big Bang Physics: From Before the First Millisecond to the Present Universe. *James S. Trefil.* **$15.95**

39570-2. COLOR ENCYCLOPEDIA OF GEMSTONES. *Joel E. Arem.* The first full reference source covering *every* known species and variety of gemstone with color photos. 64 pages of color plates. *Counts as 2 choices.* **$39.50**

35529. THE BASIC HANDBOOK. Second Edition. *David A. Lien. The* reference-encyclopedia of the BASIC language. Softbound. **$19.95**

36391-2. THE BIRTH OF PARTICLE PHYSICS. *Brown and Hoddeson, Eds.* Based on a Fermilab Symposium. New glimpses into the intellectual development of modern physics. *Counts as 2 choices.* **$44.95**

85899. THE VOICES OF TIME: A Cooperative Survey of Man's View of Time as Expressed by the Sciences and Humanities. Second Edition. *Edited with a new Introduction by J.T. Fraser.* Softbound. **$15.00**

65664. 100 BILLION SUNS: The Birth, Life, and Death of the Stars. *Rudolf Kippenhahn.* A composite biography of the stars of the Milky Way. **$24.95**

50552. THE GRAND DESIGN: Form and Color in Animals. *Sally Foy and Oxford Scientific Films.* With 362 color photos. **$24.95**

37219-2. THE CAMBRIDGE ENCYCLOPEDIA OF EARTH SCIENCES. *Edited by David G. Smith.* A massive reference by 32 eminent scientists. Over 500 illustrations, maps, and diagrams, with half in full color. *Counts as 2 choices.* **$35.00**

61230-2. THE MATHEMATICAL EXPERIENCE. *Davis and Hersh.* "Truly an exceptional book, written by masters of exposition."—*Library Journal.* Illustrated with charts, diagrams, and photos. *Counts as 2 choices.* **$27.95**

62616. MICROCOMPUTER GRAPHICS. *Roy E. Myers.* The essential mathematics and programming techniques for applications in BASIC. Softbound. **$12.45**

68751-2. THE PHYSICAL UNIVERSE: An Introduction to Astronomy. *Frank H. Shu.* With over 500 illustrations including a special color section with NASA photos of Saturn, Jupiter and Mars. *Counts as 2 choices.* **$28.00**

46570-2. EVOLUTION OF THE VERTEBRATES: A History of the Backboned Animals Through Time. Third Edition. *Edwin H. Colbert.* Illustrated. *Counts as 2 choices.* **$28.50**

82233-2. TALKING COMPUTERS AND TELECOMMUNICATIONS. *John A. Kuecken.* The latest methods and devices for automating in situ telephone networks. *Counts as 2 choices.* **$26.50**

81404. THE SUN, OUR STAR. *Robert W. Noyes.* An up-to-date history including new studies on solar phenomena and their effect on Earth. 153 illustrations. **$20.00**

74102. REVOLUTION IN MINIATURE. *Braun and MacDonald.* Revised and updated second edition of the classic account of the history and impact of semiconductor electronics. **$24.95**

32465-2. ALBERT EINSTEIN'S SPECIAL THEORY OF RELATIVITY: Emergence (1905) and Early Interpretation (1905–1911). *Arthur I. Miller.* The background, an English language version, and an analysis of the theory. Softbound. *Counts as 2 choices.* $27.50

37385. CATASTROPHE THEORY AND ITS APPLICATIONS. *Poston and Stewart.* The first complete treatment of this revolutionary theory. Softbound. **$24.95**

55322. INFINITY AND THE MIND. *Rudy Rucker.* "Informal, amusing, witty and profound."—Martin Gardner. **$15.95**

46583-2. THE EVOLUTION OF THE EARTH. Third Edition. *Dott and Batten.* Illustrated throughout with photos, drawings, diagrams, and maps. *Counts as 2 choices.* **$28.95**

50204-2. GENERAL RELATIVITY: An Introduction to the Theory of the Gravitational Field. *Hans Stephani.* From the fundamentals of Einsteinian spacetime to evolving theories of gravitation. *Counts as 2 choices.* **$49.50**

57839. LAST AID: The Medical Dimensions of Nuclear War. *Edited by Chivian et al.* Illustrated. Softbound. **$9.95**

60775. MATHEMATICAL DISCOVERY: On Understanding, Learning, and Teaching Problem Solving—Combined Edition. *George Polya.* Softbound. **$20.95**

64910. NUMBERS AND INFINITY. A Historical Account of Mathematical Concepts. *Sondheimer and Rogerson.* **$15.95**

66518. OTHER WORLDS: Space, Superspace and the Quantum Universe. *Paul Davies.* A non-mathematical introduction to the world of subatomic physics and the quantum theory revolution. **$11.95**

69749-2. PRACTICAL EXPERIMENT DESIGNS FOR ENGINEERS AND SCIENTISTS. *William J. Diamond.* Everything from fundamentals of experiment design to designs with up to 127 variables. Illustrated. *Counts as 2 choices.* **$31.50**

61649. THE MEASURE OF THE UNIVERSE. *Isaac Asimov.* Our foremost science writer looks at the world large and small. Illustrated throughout by Roger Jones. **$15.95**

49070. FROM ATOMS TO QUARKS. *James S. Trefil.* How muons, quarks, and other strange particles are related to the physical world we touch. **$12.95**

59843. MAN-MADE LIFE. *Jeremy Cherfas.* The awesome potentials of genetic engineering. With 50 photos and drawings. **$15.95**

32061. THE ACCIDENTAL UNIVERSE. *Paul Davies.* Probes the "apparently miraculous" accidents of nature. **$19.95**

53622. HOW LIFE LEARNED TO LIVE. *Helmut Tributsch.* Filled with superlative photos and illustrations. **$19.95**

50218. THE GENESIS AND EVOLUTION OF TIME: A Critique of Interpretation in Physics. *J.T. Fraser.* "A penetrating, questioning analysis." —P.C.W. Davies. Illustrated. **$20.00**

68786. THE PHYSICS OF SOUND. *Berg and Stork.* A comprehensive introduction to the science of acoustics. Extensively illustrated. **$22.95**

64881. NUCLEAR POWER: Both Sides. *Kaku and Trainer, Eds.* The best arguments for and against the most controversial technology. **$14.95**

63929. THE NATURE OF MATTER. *Edited by J.H. Mulvey.* Using high energy particles to uncover the secrets of the nature of matter. **$15.95**

48966. FRAMES OF MIND: The Theory of Multiple Intelligences. *Howard Gardner.* A kaleidoscopic view of human competences. **$23.50**

86847. WHEELS, LIFE, AND OTHER MATHEMATICAL AMUSEMENTS. *Martin Gardner.* Brain-teasers that illustrate and test problem-solving methods. **$15.95**

73181. QUARKS: The Stuff of Matter. *Harald Fritzsch.* An exploration of the subnuclear world. Illustrated. **$19.00**

4 Good Reasons To Join

1. The Finest Books. Of the hundreds and hundreds of books submitted to us each year, only the very finest are selected and offered. Moreover, our books are always of equal quality to publishers' editions, *never* economy editions.

2. Big Savings. In addition to getting 3 books for $1.00 each when you join, you keep saving substantially—up to 30% and occasionally even more. (For example, your total savings as a trial member—including this introductory offer—can easily be over 50%. That's like getting every other book free!)

3. Bonus Books. Also, you will immediately become eligible to participate in our Bonus Book Plan, with savings of up to 70% off the publishers' prices.

4. Convenient Service. At 3–4 week intervals (16 times per year) you will receive the Book Club NEWS, describing the Main Selection and Alternate Selections, together with a dated reply card. If you want the Main Selection, do nothing and it will be sent to you automatically. If you prefer another selection, or no book at all, simply indicate your choice on the card and return it by the date specified. You will have at least 10 days to decide. If, because of late mail delivery of the NEWS, you should receive a book you do not want, we guarantee return postage.

If reply card has been removed, please write:
The Library of Science–Dept. 2-AY7, Riverside, N.J. 08075 to obtain membership information and application.

Scientific American 2/84

(whether child or adult) in the turtle's experience of geometry. When one is baffled by a program, one is told that the answer is to "play turtle."

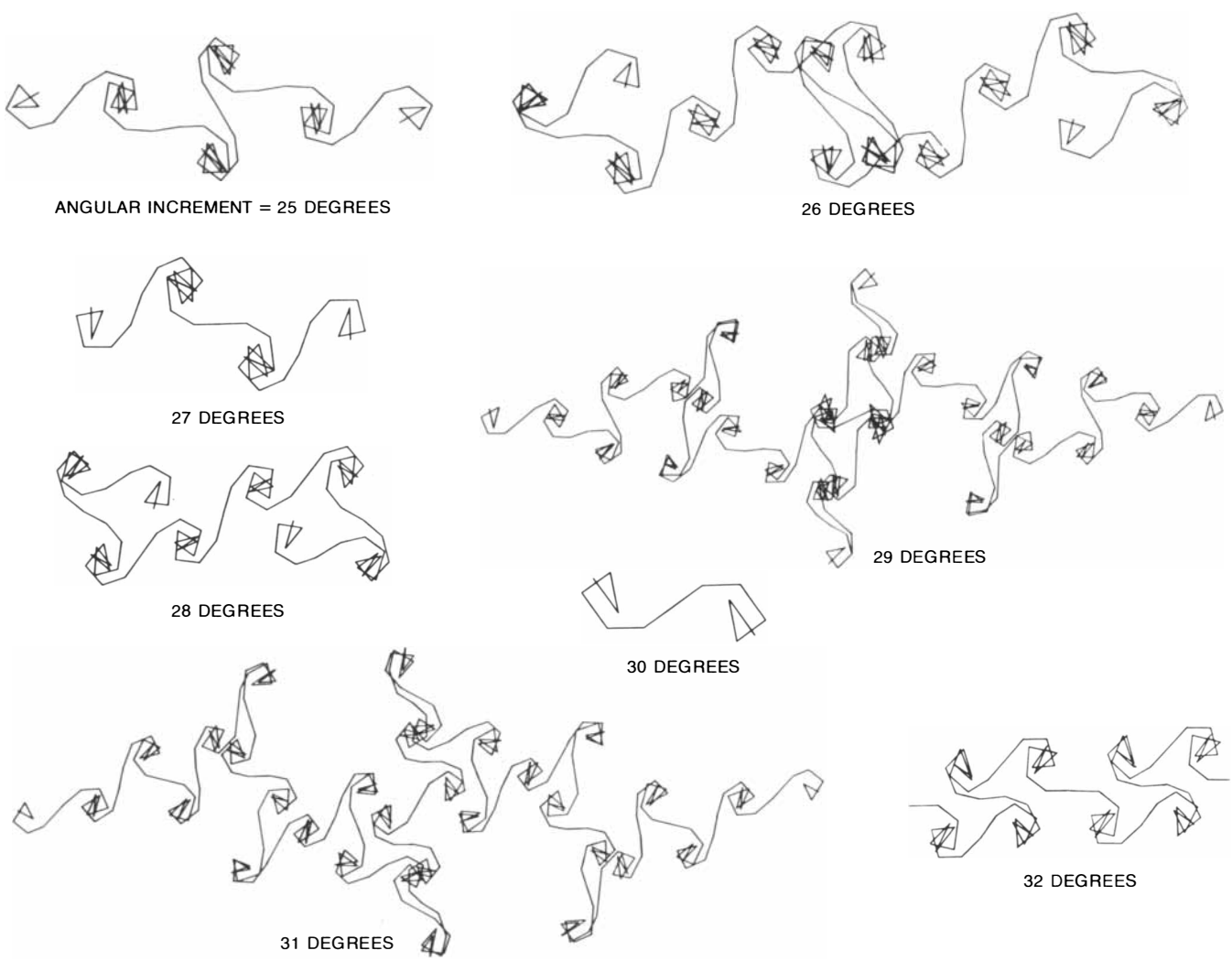
The deepest connection between turtle geometry and Logo is that they spring from a common philosophy of education. It is a philosophy based in large part on the work of Jean Piaget, which places the highest value on the student's own discoveries. Papert, who worked with Piaget for five years, declares his ambitions in *Mindstorms: Children, Computers, and Powerful Ideas.* "Programming the Turtle starts by making one reflect on how one does oneself what one would like the Turtle to do. Thus teaching the Turtle to act or to 'think' can lead one to reflect on one's own actions and thinking.... The experience can be heady: Thinking about thinking turns the child into an epistemologist, an experience not even shared by most adults."

The discussion of randomized prose in November elicited a number of comments I should like to pass on for the benefit of anyone considering a similar project. One algorithm I described called for searching through a text for each instance of a given sequence of characters, then building a frequency table for the letters that follow the sequence. The entire procedure was to be repeated for each letter of random text generated. Several readers proposed more efficient methods.

One approach, suggested by Judith E. Dayhoff and Stephen C. Locke, is to employ a data structure called a hash table. Each sequence of letters in the text can be taken to encode a numerical value, which can serve as an index pointing to an entry in the table. The entry gives the frequencies of the characters that follow the index sequence. Only sequences that actually appear need to be included. The procedure should be quite fast because the hash table is constructed once and subsequent references to it require only a calculation of the index, not a search of all the entries.

Bobby Bryant, James W. Butler, Ronald E. Diel, William P. Dunlap and Jim Schirmer pointed out still another algorithm that is not only faster than the one I gave but also appreciably simpler. It eliminates frequency tables entirely. When a letter is to be selected to follow a given sequence of characters, a random position in the text is chosen as the starting point for a serial search. Instead of tabulating all instances of the sequence, however, the search stops when the first instance is found, and the next character is the selected one. If the distribution of letter sequences throughout the text is reasonably uniform, the results should closely approximate those given by a frequency table.

An important historical precedent for work of this kind was brought to my attention by Sergei P. Kapitza, editor of *V Mire Nauki,* the Russian-language edition of SCIENTIFIC AMERICAN. The procedure for selecting a letter in the random-text program is known in probability theory as a Markov process, after the Russian mathematician Andrei A. Markov. Kapitza points out that Markov presented his first discussion of the process in terms of randomizing text. Markov's paper "On the Sequence of Letters in Eugene Onegin" asks to what extent Pushkin's poem remains Pushkin's when the letters are scrambled.



Eight arrays of spirals created by adding a fixed increment to an initial angle of zero

PHILADELPHIA

A GROWING TRADITION OF TECHNOLOGY

The rapid progress true science now makes, occasions my regretting that I was born so soon.

Benj. Franklin 1780

There's a sense of acceleration here. Some call it momentum. Others say it's growth, nurtured by a tradition of success.

It doesn't really matter which label you use as long as this one fact rings out loud and clear.

Philadelphia offers a spectrum of financing and resources that few areas in the country can come close to matching.

And when it comes to the business of advanced technology, that fact takes on even more significance.

Our history here is rich in the pure process of planning and development. Things have been managed well; there's order: defined, deliberate and decisive.

That structure is solidly anchored. And it's spawned a broad economic base. The financial dimension is complemented by human resources. An ambitious, educated and willing work force. An abundance of universities and the benefits that an academic atmosphere brings to any environment.

The result of all this is tangible. It's created the momentum and growth, and kept it moving.

Come see for yourself. Feel it. Hear what's going on. Talk to the people. Make your own evaluation. Decide.

That process can begin with one phone call. And we urge you to make it now. To the number listed below.

There's new success and achievement waiting here for you.

Philadelphia: Built by success, positioned for growth.

PIDC Philadelphia

PHILADELPHIA INDUSTRIAL
DEVELOPMENT CORPORATION
One East Penn Square, Suite 1800
Philadelphia, PA 19107 (215) 568-4448
G. CRAIG SCHELTER,
Executive Vice President

Call: 215-568-4448

There's a surge of new technology across southeastern Pennsylvania. Advances in computer applications, electronic control and biotechnology are altering the economic, industrial and educational framework of Philadelphia and its four surrounding counties.

In 1946, Dr. J. Presper Eckert and the late Dr. John W. Mauchly designed, built and successfully operated the original Electronic Numerical Integrator and Computer, known as ENIAC, at the University of Pennsylvania. UNIVAC, the first computer designed for commercial use, soon followed.

Philadelphia was the birthplace of the first computer company. The Eckert-Mauchly Computer Corporation became the Computer Systems operations of the Sperry Corporation in Blue Bell, an attractive suburb in Montgomery County.

In August 1982, Robert E. Mittelstaedt Jr. and Thomas A. Penn of the Innovation Center in the Wharton School of the University of Pennsylvania published a study on venture capital.

Mittelstaedt and Penn discovered an unexpected and exciting development in the Philadelphia area economy. Some 25 advanced technology firms had clustered along U.S. Route 202 and the Pennsylvania Turnpike north of Philadelphia, and a similar number were found in the City.

As awareness of the growth of advanced technology in the region has increased, relationships between industry, educational institutions, government and the investment community have been transformed and a common view of the technological and economic future of Philadelphia and southeastern Pennsylvania has arisen.

An Urban Research Center

The Philadelphia area has the second-largest concentration of colleges and universities in the nation, including six medical schools and several research-oriented institutions. This academic-technical complex is integrated by a research and development facility that also connects it with business, government and the investment community.

Situated along West Market Street just north of the University of Pennsylvania and near Drexel University, the University City Science Center is a 19-acre research park owned by 28 universities, colleges and medical institutions. More than 70 technically oriented companies, government agencies and research groups, employing 5,500 people, operate out of its nine buildings.

The Science Center has served as the home for more than 40 new businesses. The Science Center's 100 full-time research professionals do contract research for government and industry. It also provides assistance from the faculties of member institutions.

According to UCSC President Dr. Randall M. Whaley, the purpose of the Science Center is "to use resources of the region and its institutions to improve quality of life and to address the problems of the community at large." This can be accomplished, Whaley says, through technology transfer: shifting the products of individual and university research to the commercial sector, where they can be applied to manufacturing and service industries and produce jobs.

The executives of a number of mature companies are directors of the Science Center.

John R. Selby, president and chief executive officer of SPS Technologies in Newtown, is a director of the Science Center. He believes that southeastern Pennsylvania has "an excellent balance between corporate and institutional

research and development interests in the area of advanced technologies. And each shows a willingness to invest in the disciplines necessary to support these interests."

The Science Center has served as the temporary home for a number of advanced technology companies formed as the result of industry-supported university research. The genesis of Biological Energy Corporation is an illustration.

GEVENCO, the wholly owned venture capital subsidiary of General Electric Company, established BEC to develop and commercialize industrial biotechnology.

GEVENCO invested $2 million, and BEC began operating in the University City Science Center.

A pilot plant is operating successfully, and a $14 million pre-commercial plant has been designed, according to BEC President Dr. E. Kendall Pye. The Company, having "incubated" at UCSC, is now located in Valley Forge.

"Centers of Excellence"

The Science Center also is part of Pennsylvania's efforts to accelerate development of advanced technology industries statewide. It is the home of the Advanced Technology Center of Southeastern Pennsylvania, one of four centers set up by the state under the Ben Franklin Partnership, a program that promotes the growth of advanced technology business through cooperation between the educational and private business sectors.

Each center is located in a different region of the state and is operated by a consortium of leading universities, business, and labor. The other centers are based at Lehigh University in Bethlehem, Carnegie-Mellon, the University of Pittsburgh, and Pennsylvania State University. Each center conducts joint research and development in concert with area industries.

Governor Dick Thornburgh has made a strong, long-term commitment to the Ben Franklin Partnership.

"Our Southeastern Pennsylvania Advanced Technology Center, created through the state's Ben Franklin Partnership, draws upon research and educational communities, as well as on private sector resources, to provide job-creating new enterprises."

The Ben Franklin Partnership is headed by the State Department of Commerce's deputy secretary for technology, Dr. Walter Plosila. According to Plosila, its mission is to "mobilize the resources of the state so that regional features benefit the entire state. Our goal is to establish 'centers of excellence,'" he said.

"The program is different," Plosila says, "in that the private sector companies make the decisions first. This provides a better basis for creating jobs. Job creation doesn't come with simply throwing money at universities to do research, but through building linkages, commitments and incentives for the private sector to work from the beginning with the university in a partnership." The centers estimate current projects could create 10,000 jobs in four years.

The Advanced Technology Center of Southeastern Pennsylvania, based at UCSC, is supported by the University of Pennsylvania, Drexel University, Temple, Thomas Jefferson University, Pennsylvania Hospital and 12 other public and private colleges and universities and 136 private sector firms. The Greater Philadelphia First Corporation has provided business community leadership. The Center received a total of $11.75 million available in 1983-84 for 75 projects in advanced sensor technologies, human adaptability in space, and advanced biomedical technologies.

"Our mission is economic development," said Phillip A. Singerman, Technology Center director.

"Our goal is not to create a new institution," said Singerman, "but to work with existing organizations to achieve the objectives of the Ben Franklin Program."

Singerman sees these objectives as "job creation, along with training to take advantage of new opportunities, and assisting businesses in bringing new products and new technologies to the marketplace in a relatively short period of time."

A research complex about 30 miles north of Philadelphia represents an effort to transfer academic research to industry which began in 1953. The Merck Sharp & Dohme Research Laboratories (MSDRL) in West Point, Pennsylvania, are comprised of 13 facilities in which more than 700 scientists do biomedical research. Clinical research is done in Blue Bell, Montgomery County.

"Over the 50-year span, we have developed one of the world's leading biomedical research laboratories," said Merck chairman and chief executive officer John J. Horan. "From the beginning, talented scientists have come here from leading universities and research institutions."

Dr. Maurice Hilleman, senior vice president of MSDRL and adjunct professor of virology at the University of Pennsylvania Medical School, outlined some developments at the West Point facility: "The laboratory is proud of its basic breakthroughs, exemplified by the discovery of SV 40 virus and its role in cancer induction, the discovery and classification of the rhinoviruses that cause common colds, and discovery of the mechanism for interferon induction by double stranded ribonucleic acid."

Thomas Jefferson University, the largest private medical school in America, works closely with the Science Center. University faculty participate in applied research and technology transfer, as University vice president John D'Aprix explains: "Commercialization of intellectual properties, the practical applications which emerge from our research enterprise, is given very active management support.

"TJU is currently managing a portfolio of 50 properties. Included are inventions like an artificial larynx, a medium for storing blood platelets, and a stroke recovery system. Our philosophy is that intellectual capital should be managed to generate a return. Companies here and around the country have asked us to bring them our ideas. Their doors—and ours—are open."

A Direct Descendant

The origin of advanced technology in the area is symbolized by the computer systems operations facility of Sperry Corporation in Blue Bell, a suburb 20 miles northeast of Philadelphia. The operation is the descendant of the company founded by Eckert and Mauchly.

The first home of the Eckert-Mauchly Computer Corporation was in downtown Philadelphia. The operation was acquired by Remington Rand, Inc., in 1950 during the development of UNIVAC I. In 1955, Remington Rand and the Sperry Corporation were consolidated to form the Sperry Rand Corporation. A Univac Engineering Center was established in 1961 in Blue Bell, and, in 1966, the site became world headquarters for Sperry Rand's Computer Systems Operations. The division name was changed to Sperry Univac in 1973.

Eckert, who continues to consult with Univac management, recently discussed progress in computer technology. Holding one of the original four-foot ENIAC chassis, he pointed to the rows of resistors, capacitors and tube sockets. "About 10 times the amount of information contained in this piece can now be fit into the space of a period on a typewriter key," he said.

Sperry pioneered development of real-time computers, which enable users to find information virtually instantaneously. Sperry computers installed at the God-

dard Space Flight Center in Greenbelt, MD, link NASA's worldwide communications network (NASCOM).

Space-Age Spin-Offs

In the late 1960s, the GE Space Center in King of Prussia (off Route 202) was conducting research on environmental chambers for astronauts. In 1970, several engineers involved in that research decided to start Retnord Safety Products, located in the Great Valley Corporate Center. According to Gerald Corkery, Retnord operations manager, space research technology was transferred to the underwater environment. The firm designs and manufactures sophisticated underwater breathing apparatus using electronic devices to allow divers to work at greater depths.

In Upper Merion Township in Montgomery County, a structure is taking shape that will unite city and suburbs in a large-scale technical, commercial and humanitarian enterprise (see cover photograph).

Dr. Stanley Crooke is president of research and development for SmithKline & French Laboratories, the pharmaceutical division of SmithKline Beckman Corporation. According to Crooke, the $200 million Upper Merion complex for research and development in pharmacology and biotechnology is "the largest building project ever undertaken in southeastern Pennsylvania."

By 1985, Crooke said, all of SmithKline's R&D programs in the United States will be consolidated within this one-million-square-foot complex. SK&F currently has more than 1,100 scientists and technicians in Philadelphia. A $28 million expansion of a pathology/toxicology facility in the Upper Merion complex was dedicated a year ago. A $25 million molecular biology building, dedicated last October, is occupied by 128 scientists and technicians. Construction of nine additional buildings is underway.

Upon completion of the project, SmithKline Beckman will have linked its corporate headquarters in center city Philadelphia to its huge R&D facility. "What we're doing in Philadelphia," Crooke said, "is a reflection of the strategic approach of SK&F, which is driven by knowledge and foresight." The company's commitment to the city and southeastern Pennsylvania, he said, is based on the fact that "Philadelphia has the largest concentration of academic medicine in the country. Its excellent medical schools have developed enclaves of outstanding scientists which, in turn, attract additional scientists."

In the past few years, original research by SK&F has resulted in the introduction of several innovative specialty drugs. For example, after almost a decade of devel-

"ADVANCED TECHNOLOGY HAS A FRIEND IN PENNSYLVANIA"

Advanced technology is crucial to Pennsylvania's economic development strategy.

The strategy was derived from "Choices for Pennsylvanians," Pennsylvania's economic blueprint developed by the State Planning Board with the input of 185,000 Pennsylvanians and the strong support of Governor Dick Thornburgh.

The strategy encourages new business start-ups, particularly those on the leading edge of technology; the incorporation of the latest technological processes into traditional manufacturing; and a hospitable business climate to attract new firms to the state.

Resources Pennsylvania offers technology-driven firms include:

Financial Assistance—The Pennsylvania Industrial Development Authority provides low-interest loans for land and buildings with 25 percent of its annual resources targeted to advanced technology firms. The Pennsylvania Capital Loan Fund provides loans, particularly to smaller firms. Through Pennsylvania's Revenue Bond and Mortgage Program, one of the nation's most aggressive industrial development bond programs, tax-exempt bonds may be approved through private sector financing sources for capital needs.

Job Training—Pennsylvania offers customized job training. Technology firms locating or expanding in the state work with local educational institutions to develop tailor-made training programs. Grants cover the cost of equipment, classroom, and/or on-site training.

Small Business Assistance Services—The Pennsylvania Small Business Action Center is a one-stop center offering free assistance for any state-related business question. The Center provides information on state forms and applications required to start-up or expand a business. And with its network of 35 expediters in state agencies, the Center assures prompt answers. Also, Small Business Development Centers (SBDC) located at colleges and universities throughout the state offer managerial and start-up assistance. Many SBDCs are involved with the state's Ben Franklin Partnership Advanced Technology Centers.

Research and Development—Pennsylvania is one of four states with a program of research "seed" grants, given directly to small firms for R&D, leading to the introduction of new products. Joint research and development efforts with the private sector are under way through each of four Advanced Technology Centers to commercialize R&D into new products or processes, creating new firms or helping existing companies, and creating and retaining jobs.

The Ben Franklin Partnership brings all these services together. Beyond joint programs in research and development, the four Centers help modify courses throughout the educational system to assure the technological literacy of Pennsylvania's work force. Small business incubator facilities associated with each Center provide low-overhead rental space during start-up for fledgling R&D firms. Access to venture capital and other financial institutions is provided through each of the Center's entrepreneurial development services.

Pennsylvania's work force of 5.4 million people represents the fifth largest labor pool in the nation and is characterized by a high work ethic, low turnover, and a wide range of skills. More than 150,000 scientists and engineers live and work in the Commonwealth. Pennsylvania has four of the nation's top 50 graduate research universities, ranks third in corporate contributions to university R&D, and ranks fifth in corporate R&D expenditures.

Pennsylvania's proximity to more than 60 percent of the nation's personal and corporate purchasing power, along with a sophisticated transportation system, is an important plus for business.

Pennsylvania is the third largest producer of electricity in the nation, with enough resources to supply 100 percent of its own needs, as well as some of the needs of neighboring states.

The state's two largest metropolitan areas, Pittsburgh and Philadelphia, rank, respectively, fourth and sixth in the nation as desirable places to live and raise a family, according to Rand McNally's *Places Rated Almanac*. Every Pennsylvania resident lives within 25 miles of a state park. Pennsylvania also offers a variety of excellent recreational and cultural facilities.

Because of the state's commitment to the growth of advanced technology industries, a special position has been established within the Pennsylvania Department of Commerce: the Office for Technology Development. For additional information contact: Dr. Walt Plosila, Deputy Secretary for Technology, Pennsylvania Department of Commerce, 433 Forum Building, Harrisburg, Pennsylvania 17120 or call (717) 787-3003.

opment and marketing, Tagamet™ has become standard therapy for ulcer disease. Ridaura™, an oral medication containing gold, was developed in 1974 for treatment and control of rheumatoid arthritis. Recent research has concentrated on peptides, which, according to Crooke, are protein chains that occur naturally in the human body and carry a large amount of molecular information. For example, Vasopressin™, a peptide hormone that affects renal and cardiovascular functions, is used in the treatment of hypertension.

Many breakthroughs have been made possible by SK&F's investment in advanced technology equipment. "Most of the work we do here could not be done without the computer," Crooke said. "We deal with an extraordinary amount of information." The research effort relies on approximately $60 million worth of high technology hardware including nuclear magnetic resonance scanners, ultracentrifuges and Fourier spectrophotometers.

To Crooke, however, the real estate, buildings and equipment have value only in terms of SK&F's mission: "This organization is committed to the aggressive pursuit and discovery of new drugs and is based on a recognition of the necessity to do outstanding science efficiently."

Such a commitment, Crooke acknowledges, demands applying high standards in recruitment. "The key investment is people. The technology is useless without the associated investment in people."

He places great value on SK&F's relationship with Philadelphia's universities and medical schools. "The relationship serves to enhance the capability of academicians to do research, and it contributes to the education of future scientists," which Crooke describes as a "sacred responsibility."

Referring to the new molecular biology building in Upper Merion, he says, "The building is a sculpture that represents the opportunity to construct the future, and to make sure that the present and the future are well cared for."

From Transistors to Superchips

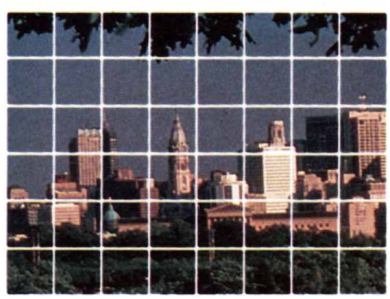
The first commercial transistor was produced at Western Electric's Allentown Works north of Philadelphia.

In 1981, Bell Laboratories developed the first 256K random access memory (RAM) chip; production began in Allentown in 1982. The silicon chip has over half a million microscopic components yet is small enough to fit on a fingertip. According to Bell scientists, it can hold 262,144 bits of information.

The effects of these new technologies can be seen in the business, medical and environmental fields.

Southeastern Pennsylvania, for instance, has the industrial automation and control technology to cause significant change in methods of production and marketing in the area. United States Robots in King of Prussia manufactures and markets factory automation systems "with the ability to interact intelligently," according to President George S. Crosby. USR was founded in 1980 by Peter Chance, a native Philadelphian, and Mitchell Weiss; in 1982, it became a subsidiary of Square D Company of Illinois and now employs 60 people.

Development of biotechnology in the area promises innovations in the detection and treatment of disease. Centocor, Inc., develops and markets biological substances for use in diagnostic and therapeutic products. Chairman Michael A. Wall said that the company recently made three important breakthroughs: a hepatitis test, a blood test for pancreatic cancer



and a test for ovarian cancer.

The hepatitis test, Wall said, is capable of detecting more positive cases than any other test. The test for ovarian cancer was favorably reviewed recently in the New England Journal of Medicine. "None of these breakthroughs would have been possible," Wall said, "without the monoclonal antibody technology developed by Centocor." Originally established at the Science Center, Centocor has expanded to larger quarters at Rouse Associates' Great Valley Corporate Center in Malvern; according to Rouse Marketing Director Jill Felix they're "starting to outgrow the present quarters here as well."

An Outstanding Location

"Philadelphia offers inexpensive incubator-space, in close proximity to major universities, an intensive labor pool, and substantive financial programs to enhance the development of high-technology companies," says G. Craig Schelter, executive vice president of Philadelphia Industrial Development Corporation.

Hewlett-Packard has its headquarters in Palo Alto in Silicon Valley. In the mid-60s, the growing corporation was attracted to the Philadelphia suburb of Avondale.

According to Avondale Division general manager Mason Byles, the division employs 800 persons in the manufacture of gas chromatographs and laboratory automation.

Byles adds, "There are a number of factors which make southeastern Pennsylvania an outstanding location for firms, considering the density of industries in the Northeast that require the use of analytical instrumentation."

AMP Products Corporation, a subsidiary of AMP Inc. of Harrisburg, located in Berwyn long before development of the Route 202 advanced technology corridor.

"Berwyn offered AMP Products Corporation easy access to major transportation routes, and a supply of skilled and dedicated workers," said vice president Oscar B. Rudolph. "Philadelphia, with its impressive roster of universities and institutions, is close by."

"We originally were attracted by the area's potential," said Ronald A. Wallace, president of Sorbus Service in Frazer, a nationwide computer-maintenance firm.

Wallace sees the area's usefulness in "the activity of a talented workforce, educational and cultural facilities, local talent, good communication and transportation networks, and pleasant living conditions. Philadelphia's western suburbs are by their very nature a 'natural resource' for the development of a high-tech service company like ours."

Technology Park

One of the largest concentrations of new advanced technology business is found along the Route 202 corridor in the Great Valley Corporate Center in Malvern in Chester County. The 210-acre commercial park was built in 1974 by Rouse & Associates, a development, construction and property management corporation.

"Great Valley and Route 202 are not accidents," says principal partner Bill Rouse. "Certainly the attractive natural environment has always been here. So have the universities and the service infrastructure."

Shared resources are a hallmark of any Rouse & Associates operation. "We work to learn what our client corporations need in common," says marketing director Jill Felix, "and we provide it.

"Clients told us they wanted convenient access to off-hours continuing education. We found educational institutions in the region who had excess capacity. Result: a range of courses to help high-tech operatives retain and improve their skills."

The Great Valley Corporate Center has 90 tenants employing around 3,500 in 23 buildings. The firm recently broke ground for a $200 million, 200-acre, high technology park adjacent to the existing center.

RIGHT NOW,

Philadelphia has in place the most sophisticated telecommunications network in the world. With this network Philadelphia leads the country in telecommunications technology.

Right now, Philadelphia has 100% electronic switching capabilities. And has a 100% digital network that can handle from 56,000 bits to 1½ megabits of information per second.

Right now, Philadelphia's lightguide facility has one of the leading optical fiber systems in the country capable of transmitting at speeds up to 90 megabits.

Right now, our lightguide and software enhancements are opening new markets such as videotex, teleconferencing, electronic message systems, online polling and packet switching, to name a few. And right now, Philadelphia is connected to the world via fiber optics, digital microwave and satellite.

This telecommunications network is the vital resource linking our city and its business community with all the benefits of today's technology.

Philadelphia businesses can communicate more information, with more speed and accuracy, and in more ways than anyplace else in the world. Right now.

And this is just the beginning.

Bell of Pennsylvania

A Bell Atlantic Company

An Advanced Technology Network

According to Joseph J. Houldin, Debra E. Buck and W.M. Cristman III of the Philadelphia Industrial Development Corporation (PIDC), the Wharton study was the key force in the recent organization of venture capital availability in Philadelphia. PIDC is a 24-year-old quasi-public, nonprofit economic development agency. It sponsors a Technology Assistance Program that is funded through the Federal Community Development Block Grant Program.

Buck administers the Technology Assistance Program, designed to "strengthen and maintain the city's capacity for economic growth and stability. PIDC's Technology Assistance Program provides an incentive for commercialization of new technologies through financial assistance to small firms." Funds may be used to finance start-up costs after research and development has been completed, Buck said.

PIDC's G. Craig Schelter points, with some pride, to the relocation of one growing technology firm away from a suburb to a downtown location—to facilitate planned growth. Sonex/Philmont Electronics' own equity contribution was joined to a PIDC mortgage loan and a tax-exempt loan from a commercial bank.

Houldin, assistant director for client relations, believes that the technology network that has arisen spontaneously in Philadelphia has advantages over enterprises in other regions. "An advanced technology center can arise via an official master plan, in which each sector has an assigned role, or by allowing the universities, venture capital institutions and public agencies to approach it in their own ways.

"Here in Philadelphia everybody is conscious of advanced technology and is addressing it in their own context. This approach has much more grass-roots effectiveness."

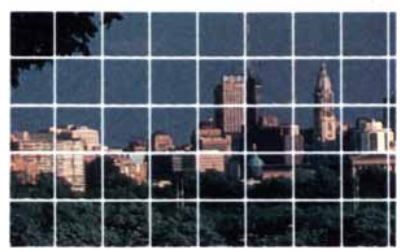
To illustrate the coordinating process currently underway, Houldin cited recent meetings between PIDC, the University City Science Center, the Franklin Institute and the Wharton School.

An important junction in the communication network is the Technology Council of the Greater Philadelphia Chamber of Commerce. Cristman, PIDC vice president for research and corporate strategy, describes the council as a "think-tank effort that has resulted in a mini-explosion of joint venture activity."

A Technology Exchange

Since the Technology Council was founded by Dr. Bowen Dees, former president of the Franklin Institute, as the Committee on Technological Innovation in 1980, it has sought to heighten the region's awareness of the importance of technology.The Council acts as a "high-tech communications exchange" for all communities. Its monthly newsletter, Technology Council News, is the first communications network for technology in southeastern Pennsylvania.

"The whole attitude of the universities has changed," said Fred Lipman, Technology Council chairman. "They now permit more liberal consulting by professors. They're extremely anxious to participate and are a fertile source of spin-off business."



Investing In Technology

As venture capital is vital to the development of advanced technology, technical knowledge is an aid in making prudent investment decisions, according to Geoffrey S.M. Hedrick, who is president and chief executive officer of Smiths Industries Aerospace & Defense Systems, Malvern. As an electrical engineer, a holder of 38 patents and a lecturer at Wharton, Hedrick has an excellent perspective on the technical and financial elements of advanced technology development.

Hedrick believes advanced technology ventures call for technical counseling of investors. "The rapid development of technology has placed new demands on venture capitalists," he said. "Substantial technical depth is often required to adequately assess the products of entrepreneurs."

Atty. Raymond R. Rafferty Jr. is chairman of the Chamber's Delaware Valley Venture Group. According to Rafferty, the Delaware Valley Venture Group is "an association of professionals—including accountants, bank representatives, corporate finance specialists and other financial intermediaries—involved in the capital investment process." It holds monthly seminars to discuss issues associated with commercial development of research-derived products.

Rafferty also is cochairman of Century IV Partners. The partnership was organized in 1982 by the Chamber of Commerce and the Greater Philadelphia Part-

nership, in response to the Mittelstaedt-Penn study.

As of September 1983, said Rafferty, the partnership had received in excess of $20 million, with more than $17 million coming from Philadelphia institutions such as SmithKline Beckman Corporation, Rohm & Haas Corporation, Fidelity Mutual Life Insurance Company, Provident Mutual Life Insurance Company, and Thomas Jefferson University.

The Seed Company, formed in 1982, grew out of the Wharton study. Its president is Thomas A. Penn, co-author of the study. The Wharton Innovation Center (WIC) assists Seed in the formation of venture partnerships.

Seed's first partnership, Schuylkill Venture Partners, is an association seeking selective investments in growth-industry firms.

"At present," Penn noted recently, "the WIC at the University of Pennsylvania is approached with 400-500 new deals per year. Evidently, of the four factors necessary for a successful entrepreneurial environment—universities, labor force, entrepreneurs with ideas, and dollars—the only one that the Delaware Valley lacks is dollars. The deal-stream is in place, and we are tapped into it."

Greater Philadelphia First

According to Executive Director Ralph Widner, Greater Philadelphia First is an association of area corporations that was formed "to unite city and suburban business, labor, civic and government leadership in Greater Philadelphia in a common effort to advance the economic development of the region and to formulate and maintain an overall economic development strategy." It recognized, he said, the special relationship between the Pennsylvania Turnpike, the Route 202 corridor and Center City Philadelphia's universities and businesses and institutions. The association is affiliated with the Chamber, the Greater Philadelphia Economic Development Coalition, and PIDC.

"As a major center of education, research and medicine in a highly strategic location and with an exceptional quality of life, Greater Philadelphia can realistically continue to develop as an 'international city,'" Widner said.

Counties' Initiatives

Delaware, Chester, Montgomery and Bucks counties all have industrial development corporations, which assist new advanced technology firms.

Gary Smith, director of the Chester County Industrial Development Corporation, says, "Years ago—when the in-migration of advanced technology firms had not begun—we knew no better. But now

"PHILADELPHIA: A GROWING TRADITION OF TECHNOLOGY"
IS SPONSORED BY ITS ADVERTISERS AND BY:

Great Valley Corporate Center
A Rouse Associates Development
10 Great Valley Parkway
Malvern, PA 19355
Jill Felix,
Marketing Director
(215) 647-7300

RDC, Incorporated
115 E. 15 St.
Chester, PA 19013
Judd L. Howard,
Marketing Director
(215) 874-5555
Creative financial strategies for new enterprise.

One Reading Center
Opening June 1984
Spectacular 32-Story Office Tower
Northwest Corner/11th & Market Sts.
Leasing Inquiries Invited:
(215) 625-2500

Pennwalt Corporation
KYNAR Piezo Film Group
King of Prussia, PA 19406
J. Victor Chatigny,
Manager
(215) 337-6710

Chester County Development Council
750 Pottstown Pike, Exton, PA 19341
Gary W. Smith, Executive Director
(215) 363-6110
Complete assistance for high-tech companies on the move.

"Philadelphia: A Growing Tradition of Technology" was written by Gene Gorsky of the Public Relations Division of Lewis, Gilman & Kynett, Philadelphia, with graphics by Sherin & Matejka, Inc., New York. PHOTO CREDITS: SmithKline Beckman Corp.; PIDC; One Reading Center Associates; City of Philadelphia. ILLUSTRATION CREDIT: Glen Wolf.

we're all part of the Advanced Technology Center.

"This new cooperation is better for us and for the industries we serve."

Riverfront Development Corporation (RDC) is a private-sector, nonprofit industrial development organization in Chester, Delaware County. According to President and Chief Executive Officer Carl E. Russell, it provides not only seed capital, venture capital and other financing, but also technical assistance through the corporate sector and the universities.

"We're most useful to ventures that are either too small, or too underorganized, to approach conventional financial sources. We package new ideas and new firms, and help them reach a level where they're of interest to the money-markets."

Uniting Philadelphia and the five counties in southeastern Pennsylvania in an economic and social, as well as a technological network, is the Philadelphia Electric Company. According to Jim O'Brien, manager of area development, PE began development work shortly after World War II, attempting to stimulate the area's economy and create manufacturing jobs. "We act as a catalyst to put things together and make things happen," O'Brien says.

The Choice for Hope

"Technology is inexorably advancing," says Dr. Randall Whaley. "It will remarkably affect the way we do business with one another; and we are fortunate to be in a region that can adapt to change. Educational institutions can respond to technical and practical needs to maintain a skilled labor pool. Understanding the needs of emerging and expanding businesses by our financial community will lend credence to the hospitable environment of the region. State and federal research and development assistance programs which couple universities and industry are in place.

"There is already satisfactory evidence that business and industry and educational institutions are willing participants in the advanced technology center concept. We feel that southeastern Pennsylvania is building in such ways as to be supportive to all our citizens — whatever their level in the workplace."

University-industry linkages are top priority for Temple University. UCSC is establishing the Temple Science Center Campus as a resource for local entrepreneurial business. Temple is securing patents and licenses on its own, and its faculty's, behalf.

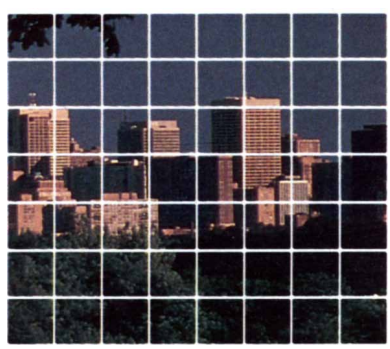
Temple is involved with the work of Dr. Jack Kolff on the total artificial heart (see "An Artificial Heart Inside The Body," SCIENTIFIC AMERICAN, November 1965). Dr. Kolff is chief of the cardio-thoracic surgery section at Temple University Hospital and a professor of surgery there as well. With Temple's assistance Dr. Kolff has assembled a new effort to refine a smaller, more useful artificial heart. The effort is supported by the Ben Franklin Partnership, the American Heart Association, and private companies. "We've pulled together an excellent team," says Dr. Kolff. "SmithKline Beckman, for example, has given us the modern equipment we need to properly monitor our clinical cases.

"We've married the skills and facilities we have here at Temple with the skills and facilities of industry."

Says Dr. Tim Weckesser, special assistant to Temple's provost, "Temple is proud of its contributions to science and to our environment."

Transition, Not Revolution

Robert G. Wilder is president and chairman of Lewis, Gilman & Kynett, Inc., Philadelphia, the largest full-service marketing/communications agency in the region. Mr. Wilder is a past chairman of Greater Philadelphia Chamber of Commerce and is a board member of the Opportunities Industrialization Center.



"The rise of the new technology," Wilder states, "has meant replacement of some traditional manufacturing companies by new firms providing sophisticated hardware and services. By planning this change—through the Ben Franklin Partnership and organizations such as UCSC and the Technology Council—we are experiencing a transition, rather than a revolution."

According to Wilder, "The key is recognizing that an area like ours is a fabric of economic, cultural and personal threads to be rewoven over time and not abruptly torn and pieced together. We're developing advanced-technology business while encouraging growth and maintenance in other areas. And we're reinvigorating older industries by incorporating into them technological innovations, thereby forming a growing, integrated economy. We here—as professional communicators—are an important thread in this broad new fabric."

LGK's client list includes many that are technology-based. One major new agency client, AT&T Western Electric, is a world leader in computer, microelectronic, and lightwave technologies. Western Electric will sponsor the 1984 Summer Olympics and is designing a sophisticated electronic messaging system for the Games.

Philadelphia is headquarters for Bell Atlantic Corporation, new parent company for telephone operating companies in a seven-state region. One of those companies is Bell of Pennsylvania, whose division manager for network planning, Ken Weaver, is enthusiastic about Philadelphia's technological growth.

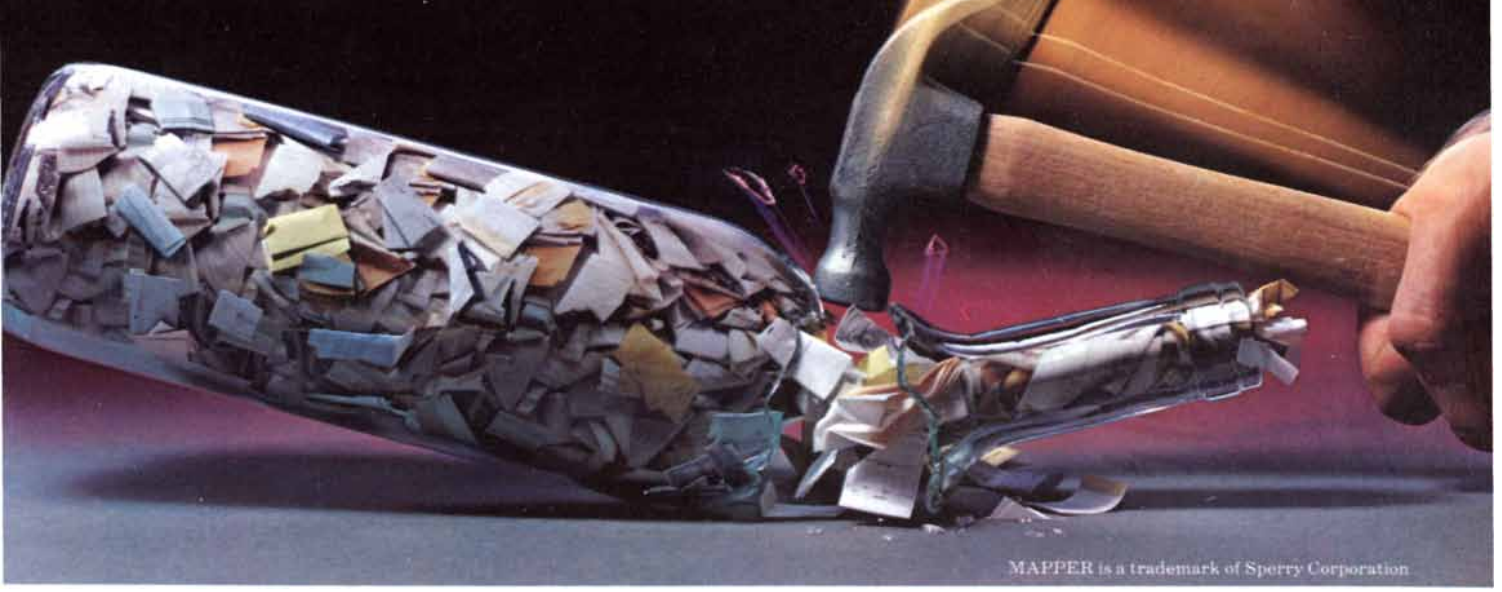
"In the telecommunications industry, deployment involves as much advanced technology as does invention. For example, we're implementing a new 'interface' based on fiber-optics here. It allows our customers to obtain access to the new inter-city carriers like MCI and Sprint in the most modern ways.

"Like any corporation working at the newest limits of technology, we're concerned with learning what our large business customers need and want, and then letting them know what Bell of Pennsylvania can do for them. That process works more smoothly in Philadelphia, where things are happening in technology in a wide range of disciplines."

One example of promising new technology comes from Pennwalt Corporation, founded in 1850 as the Pennsylvania Salt Manufacturing Company. West of Pennwalt's Center City world headquarters, and near Route 202, Pennwalt manufactures Kynar™ Piezo film, a polymer transducer with many applications. One engineer described the product, available for application for just 18 months, as "the most important transducer technology ever."

"Bell Laboratories is located just north of our King of Prussia technology center," said Pennwalt manager J.V. Chatigny. "Bell Labs has tested Kynar Piezo film through more than 15 million test actions, and a fair number of other engineers consider it highly reliable." The film is even being studied as a replacement for damaged human skin.

Philadelphia's new mayor, W. Wilson Goode, has already announced a five-part City plan to create jobs in high-tech industries. "These industries have the potential to play a major role in a 'Second Renaissance' in Philadelphia and to put more of our people to work," the mayor said. The mayor's program includes an active Scientific Advisory Committee "to help the City government locate and lure commercially viable technological research to Philadelphia."



A sure cure for the problem of information bottleneck.

Programming backlog is the source of the problem. Eliminate it, and you eliminate the bottleneck.

Which is precisely what Sperry has done with the MAPPER™ System.

With MAPPER, you work with the computer directly. You ask questions in plain English. And you get immediate answers.

And if the information you get raises other questions, you can ask them right away. With no delay for programming. Or reprogramming.

MAPPER is that powerful. It allows you to manipulate information in almost any way you want. And, interesting to note, it can even help your programmers become more productive.

AN AFFORDABLE SYSTEM.

You don't have to be a large company to have a MAPPER System. MAPPER can be scaled to the real and present needs of just about any size company. The cost of a MAPPER System makes it practical for even a department within a company to own its own system. Or you can time-share through a Sperry service bureau.

What you get for your money is a whole new order of efficiency in your day-to-day management tasks. Because you'll have the information you need right at your fingertips. Literally.

SEEING IS BELIEVING.

We've made some promises here that may sound extravagant. But if anything, our claims are on the conservative side. And to prove it, we offer you the opportunity to see a demonstration of MAPPER at work.

800-547-8362

But first, you might want to look over our MAPPER brochure. A copy is yours for the asking. Call toll-Free: 800-547-8362 (9 a.m. to 5 p.m. E.S.T.). Or send us the coupon.

MAPPER™ SYSTEM

Sperry Corporation
Computer Systems
Department 100
P.O. Box 500
Blue Bell, PA 19424

Please send me a brochure on the MAPPER System.

SA2/84-13

Name_____

Title_____

Company_____

Address_____

City_____ State_____ ZIP_____

Telephone_____

We understand how important it is to listen.

SPERRY

Philadelphia and SmithKline Beckman: the synergy of place and people

SmithKline Beckman's Pathology/Toxicology Research Laboratory, part of our new Research and Development Center.

Throughout our 150 year history, SmithKline Beckman has been associated with the Greater Philadelphia area. That association goes deeper than our roots or civic pride. It is based also on the synergy between us.

SmithKline Beckman has contributed to the rich technological environment surging across southeastern Pennsylvania. In turn, we have drawn on the resources, knowledge and people that make up this scientifically fertile area. Our work on molecular-based biotechnology alone has required the addition of several hundred scientists to our new $200 million research complex.

We salute the growing tradition of technology so much a part of the Greater Philadelphia area. It is exciting to be part of it.

SmithKline Beckman

CORPORATION

World Headquarters: One Franklin Plaza, Philadelphia, PA 19101

BOOKS

Crops and their wild relatives, a timetable from the big bang to A.D. 2000, vortex flows

by Philip Morrison

THE DIVERSITY OF CROP PLANTS, by J. G. Hawkes. Harvard University Press ($20). Fifty varieties of chili pepper, bead-size to hand-size, lie in three long ranks to illustrate the morphological diversity of crops. The next page marshals an even more unexpected lineup: bottle gourds from northeast Brazil, golf balls to trombones. These minor crops are not, to be sure, anyone's daily bread, but they are no trivialities. By their antiquity alone they command respect; the two earliest domesticates found by archaeologists in one long dated series of three dozen species are these two, cultivated for 11,000 years there along the valley road from Mexico City south to Oaxaca.

Reflective, quiet and clear, this small book by a distinguished English botanist and geneticist presents an example of applied archaeology. The use of the distant past to form the powerful myths of nationalism is familiar enough, but here prehistory is placed in direct economic service. In our populous time we need the genetic diversity of crops, their grand pool of genes for resistance to adversities and for yield, for new range and for desirable product, perhaps more than ever. Where are those half-hidden legacies? How can we best find and preserve them? These are the bread-and-oil questions of the day whose answers must in part be given by ethnobotanists and archaeologists.

Crops are weedy plants, weeds not by the self-centered complaint that they grow in the wrong place for human purposes but by a deeper ecological definition, plants able to thrive in disturbed soil around the busy hearths and camps of people, less successful in open competition with the established vegetation of their region. Crop plants have grown under conscious and accidental selection by their human growers. In many crops selection is marked for gigantism, against protective mechanisms (the wild African eggplant is a thorny vine) and against dispersal (the common bean had a wild ancestor able to disperse its seeds by an explosive twist of the pod, and the wild poppy had fruit with pores to cast out the fine poppy seeds).

The delightful visible forms and colors of chili and gourd may reflect an ancient aesthetic principle of selection. Crop plants are largely statistical artifacts of the knowing women who tended gardens long ago. More than half of the book amounts to an expert's survey of the varied explanations and conjectures—they are far from proved—that surround the question of where and how crop plants arose. Professor Hawkes does not neglect the root and tuber crops either; they are mostly of tropical origin, so that their preservation on site is relatively poor. He suggests, partly from research of his own with the potato, that there is greater similarity between seed and root crops both in the epoch and in the process of domestication than others have held.

Diversity is not always visible. There are hidden rich biochemical differences. For instance, electrophoresis disclosed that every single European potato variety possessed a set of proteins different from those of all other varieties. Plant breeders have long argued the merits of single-gene immunity as against nonspecific resistance to disease; crop varieties resistant on the basis of a single gene often act mainly as a screen for the disclosure of new genetic combinations among pathogenic fungi in the "permanent cold war between the plant breeders and the fungus pests." Longer-lasting results follow from a polygenic approach, giving complete immunity against no single pathogen but more durably effective overall. The issue is complex; there is no one path, and the diverse genetic statements are essential raw material.

Diversity is of course dwindling as wild habitats are lost, and local farmers everywhere are tempted to adopt the selected strains with superior properties offered them. The sophisticated farmer has gone far in optimized uniformity; even the law requires genetic purity in seeds offered for sale, where a couple of generations ago even U.S. farmers grew land races, mixtures of genotypes that "could not even be called varieties." This is much more of a problem for the major crops, wheat, maize, soybeans and rice, say, than it is for the roots, shrubs, climbers and herbs that flourish in the dooryard gardens of the tropical villager from Puerto Rico to central Java; those diverse plots are almost as varied as the forests outside.

The danger of genetic erosion has long been seen. In 1936 H. V. Harlan, an American plant breeder, wrote: "In the hinterlands of Asia there were probably barley fields when man was young. The progenies of these fields with all their surviving variations constitute the world's priceless reservoir of germ plasm.... When new barleys replace those grown by the farmers of Ethiopia or Tibet, the world will have lost something irreplaceable." The news, however, is good. There is a worldwide strategy at work. Collection, evaluation, standardized documentation and careful storage are under way. There are lists of priorities by crop and region, and each of the 14 centers of the international crop-research program has a few crops under its wing and its funds from the World Bank; the effort is coordinated by the United Nations Food and Agriculture Organization. The Peruvian center, for example, has stored 12,000 primitive cultivated forms of potato and is now reaching out to the wild species. A flood of booklets has appeared telling crop by crop what is on hand and what and how to find more. There is a growing training scheme for needed specialist workers at every level. There are national and regional centers too, including strong seed banks in eastern Europe, such as the one in Leningrad, with extensive collections since the time of the pioneer expeditions of N. I. Vavilov.

Everywhere seeds are held refrigerated, dried to the right moisture content and sampled for germination to decide the time for replanting. Such seeds might last in careful storage for 50 to 100 years. The limit is set by internal oxidation of the lipids in the cell embryo. The most important species, particularly the "recalcitrant" seeds of many tropical tree crops, from citrus to oil palm, are conserved much more expensively by orchard plantings or in natural reserves. New possibilities are offered by tissue cultures held in liquid nitrogen, and perhaps by seedlings dormant under controlled low light for a long time, as when they wait in the forest for some sunny opening to appear in the canopy above. The economic microflora are not treated in this engaging volume, nor is the power of recent recombinant techniques. Surely DNA splicing and copying will become valuable; genetic information retrieval is a DNA-based game.

TIMESCALE: AN ATLAS OF THE FOURTH DIMENSION, by Nigel Calder. The Viking Press ($19.95). We unwind the spiral from near its center, one power of 10 to each turn. The turn is taken in five equal segments (their ratios always as

10 : 16 : 25 : 40 : 63 : 100) from the big bang to 1984, adjusted to offer nine full turns and a bit more toward an end point in the convenient bimillennial year 2000. The thread we unwrap is more the thread of the Fates than a single bland dimension; it is the array of all that has been. Two text pages, each flanked by a relevant photograph or graph or two, narrate coolly enough what events each segment held, whether it was the entry of the "ruthless sting of oxygen" into the atmosphere of an earth where life was solely microbial, dated 1,800 million years past, or whether it was the reusable spaceship and the eradication of smallpox, dated nine turns later, just 18 years before the conventional end point.

More than 300 events are signaled by arrowheads along a logarithmic time scale that runs across the bottom of the 90 pages. There is a prologue, particle-physics events in the first microsecond or so of time, in the bold views of today's theories. Then the true logarithms begin, with a dozen markers to the origins of the sun, by way of various element-forming events. Earth history, geology in the large, is spelled out in a couple of dozen more events, past the gold boom (geological) and the iron boom (bacteriological) to the coming of animals. Another 60 events, plenty of continental encounters and asteroid splashdowns, before the grass comes, now identifiable as "the master of the planet Earth."

Another 30 arrowheads and we come to the first scavenging ancestors of humankind. It takes somewhat more events, climate oscillations and certain human accomplishments before the breakout of modern humanity, ascribed to the genetic facilitation of language only 40 millenniums ago. After that genetics is outstripped; it is culture that changes species and moves mountains. After that the events are cultural, too swift to enumerate, with many inventions in our populous world (although no more per head per year than in the deep human past). The logarithms now pay off in a somewhat human-centered way; the last 30 events string back from 1982 only as far as *Das Kapital* and the telephone.

A section of world maps plots the drifting continents of the dim past and the human movements of history to the present pattern, the Pacific now rival to the Atlantic for economic focus, Japan and California salient instead of London, the Ruhr and the U.S. Northeast. The time-scale section follows, and the book ends with a remarkably dense and diverse reference index, crammed with explication, cross-references and citations, and accented by no fewer than three dozen closely set timetables.

These schedules display the assigned dates of the main junctures in geology, paleontology and human history. Some are quite expected lists, their entries standard, although here given an unusual consistency. Others are fresh: information technology from Paleolithic lunar calendars to optical-fiber links; epidemics from Roman times onward; cosmic impacts from the oldest shield scars to the Tunguska fall of 1908; transport from floating logs to space flight. Breakouts are a key part of the author's view of prehistory and history: human groups suddenly spread both geographically and through cultural influence. That timetable is the longest, and it compiles assigned dates for the coming of the modern hunter-gatherers 400 centuries ago and the spread of blacksmiths and charioteers, on to the Americans and the Japanese, with a history of national expansion in a few wry lines for each.

The entire product is a tour de force by the experienced author. Like the best fireworks, it is a brilliant display of what reasoned expertise can do, with now and again a little too much noise and smoke. The material is fully up-to-date; more than one item is based on a single recent paper in *Nature* or *Science*. Always stimulating, those conclusions are not always enduring.

The form of the book all but demands decisive conclusions. Little is left open to doubt; too little, if precedents are any guide. Calder thinks like the physicist he is; he openly admires the easy detachment of physicists and their tone of strict causality. Indeed, if there is a fault here, it is the absence of chance. In history epidemic disease plays a major part; in prehistory the genetic mutation that allows the digestion of lactose by most people of European extraction does the same. It is open to doubt that milk-based nutrition was more important than cannon and nationalism in the breakout of the palefaces. Milk may be more efficient than meat in the conversion of energy from grass, but one would like to see some evidence that the naval gunners of the fleet of the Duke of Albuquerque, say, showed signs of milk-fed vigor. Microorganisms can predigest lactose very well; quite a few people innocent of both Indo-European verbs and enzymes enjoy the proteins of fermented milk. It is witty to cite the Opium War of 1839–42 as marking the Chinese cowed; perhaps they were instead hoist by their very own gunpowder. In much the same vein Calder bets heavily on molecular dating against the fossil record, on human genetics against other causes—or chance—in the broad sweep of history and on asteroid collisions for much of large-scale evolution.

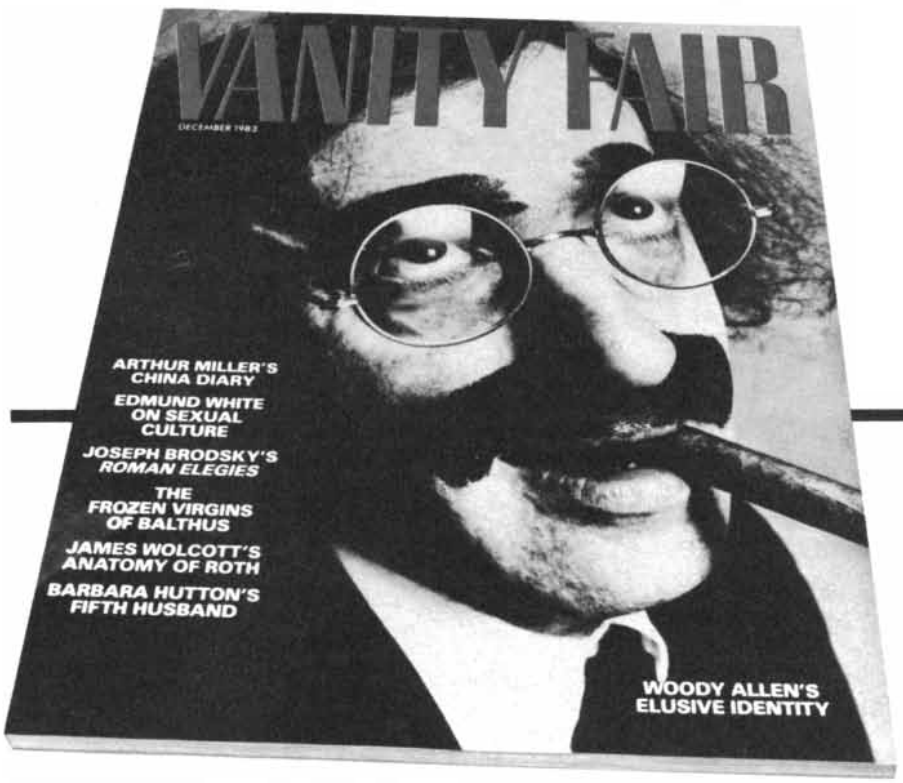
But this is carping in the presence of an enriching and ample book. Recognize the author's penchant for firm dates even as the evidence remains conflicting and you have the needed caution for enjoying these tightly argued and marvelously catholic pages, never provincial or petty. The writing, occasionally strained, is tart. An observant Martian announces the first law of human biology: Nutrition goes inversely with skin pigmentation. "Exceptions in the form of fat blacks and scrawny whites show that this is not a genetic but a social phenomenon. Every two seconds a small child dies of poverty, and from Mars the planet Earth is just South Africa written large." This is no book for the gentlest reader, but for those who can temper its sharpness with a few grains of deferred conclusions it is an exciting book to read and a lasting resource on the shelf.

There are few unclarities in the swath cut across science and history. The complex volume is well designed for use; its high level of authority flows alike from the author's experience and perspicacity and his knowing contacts with more than 100 scientists and other scholars as expert consultants. It is a shrewd example of the current state of the art of unraveling the entire past, except that almost every debate in progress has been shifted backstage.

WORLD OCEAN ATLAS, VOLUME 3: THE ARCTIC OCEAN. Edited by Sergei G. Gorshkov, Commander-in-Chief of the U.S.S.R. Navy. Department of Navigation and Oceanography, Ministry of Defense, Union of Soviet Socialist Republics. Distributed by Pergamon Press ($400). THE TIMES ATLAS OF THE OCEANS, edited by Alistair Couper. Van Nostrand Reinhold Company ($90.50). The hefty volume from Moscow, standing a foot and a half tall, is the third of its set to appear; the Pacific volume came out in 1976 and the volume treating the Atlantic and the Indian Ocean together in 1979, each priced above $400. This is a major production of the mapping agency of the Russian navy, with the scientific collaboration of the appropriate research institutes and an editorial board including a clutch of admirals under Admiral of the Fleet Gorshkov. The map pages are handsomely colored and well printed, with shaded relief for the land masses and all the other apparatus of a fine atlas. It is no book for the casual English-reading user, since it is entirely in Russian, in Cyrillic characters, apart from a title page in English.

If the Arctic Ocean is not Mare Nostrum to the U.S.S.R., it comes close to it. The heavy maritime use made of the long Arctic coastline in summer remains a persistently adventurous note in Russian economic life. The historical section of this atlas, perhaps the part of most interest for the general reader, makes the tradition plain. The first airplane ever to fly in the Arctic was that of Ya. I. Nagursky, who flew over the barren island of Novaya Zemlya in 1914. The drifting ice stations, the merchant

The return of a legendary magazine



Please accept a complimentary issue.

From 1914 to 1936, a most memorable magazine called *Vanity Fair* epitomized the elegance and style, artistry and wit of a now-vanished age.

In March of 1983, Condé Nast Publications presented the premiere issue of the new *Vanity Fair*. A magazine that captures the sparkle and excitement of the '80s as no other magazine today.

Literary and liberated. Adventurous, skeptical, irreverent. As complex and contradictory as the times in which we live. A magazine of literature and the arts. Politics and popular culture. Private lives and public events. Films and theater. Music and dance. Art and photography. TV and travel. Books and economics. Poetry. Food. Sports. Just about everything worth doing, seeing, talking about, knowing about.

Our contributors include some of today's most original novelists, journalists, poets, artists, photographers, critics. We offer them a unique freedom and breadth of expression. They, in turn, offer us their wit, new perceptions, provocative points of view. From its essays, reviews and profiles to its photography, paintings and cartoons, *Vanity Fair* is a magazine of excellence and innovation. A magazine whose time has come... again.

To receive a complimentary issue of *Vanity Fair* and enter your one-year subscription (a total of 12 issues for only $19.97), simply mail the accompanying card. (If card is missing, write to: *Vanity Fair*, P.O. Box 5228, Boulder, Colorado 80322.) This limited offer may be withdrawn at any time.

VANITY FAIR

convoys and the icebreakers came later; in 1977 the nuclear-powered icebreaker *Arktika* reached the North Pole, the first visit there by any surface vessel, and in 1979 seven young men on skis went across the early spring ice from the New Siberian Islands straight to the pole. This reader, as perhaps other Americans will, found that our media had offered surprisingly little familiarity with such agreeable if transient leading stories in the Russian press in recent years. The atlas by no means emphasizes Russian exploration; it is a full and fair historical summary in maps, from old Eric the Red onward to Robert Peary and the U.S. drifting stations Bravo and Charlie.

History takes up less than a tenth of the book. The ocean floor, its relief and its structure, the climate month by month and altitude by altitude, the regime of rain, snow and cloud, the circulation of the winds and the air masses, the physical, chemical and dynamical properties of the ocean waters, the fauna and flora, and the ice, the ice, the ice, both land-fast and floating—all appear here in detailed quantitative mappings, summaries of the state of present knowledge. The aurora and the earth's magnetic field are not omitted, and even the times of sunshine are carefully plotted. One delightful inset shows the position of the magnetic pole every five or 10 years since 1900; it is moving out through the Canadian islands straight toward the pole, although not quite as neatly as those skiers did.

This is a research atlas, and one hopes that the naval and merchant ships' officers for whom the preface says the work is intended can use it well. The pages that map month by month the probability contours for encountering floating ice are likely to draw anxious eyes on shipboard.

Arctic reference libraries will have to have it; oceanographic ones will surely want the entire set (although probably the Arctic volume is the freshest, since the Russian physical data are so important for the area). It is fortunate that Pergamon is able to distribute this valuable work outside the U.S.S.R., but the experienced multinational publisher has somehow piloted the Russian navy into shoal commercial waters. Doubtless the work is heavily subsidized; it never appeals to the civil servants of any treasury to ship such a subsidy abroad, even as fine heavy books. Perhaps that somehow justifies the 1,200 percent dollar markup from the 1980 domestic price of 25 rubles printed right on the book. For the $400 Pergamon should have found some way to bind into the volume the useful English translation of the section introductions and the full index that it now supplies only as a separate offset pamphlet, of no great beauty or permanence. That title page in English is no more than a buoy adrift.

The Times Atlas of the Oceans is a less imposing reference volume of smaller size and wider scope, much more suited to use by the general reader. It treats all the oceans, the polar ones as well, at a less detailed level. It is not so much a specialized atlas of the seas as it is a richly illustrated and comprehensive introductory treatise on the economic geography of the oceans, with a sturdy supporting base in physical and biological oceanography. The big pages display plenty of images, most of them in bright color, from up-to-date drawings of the main types of merchant ships to photographs of krill and squid and the plankton they consume. The only systematic mapping in full-page size is in the opening 10 pages, which display the sea bottom worldwide, with its named trenches, plains, seamounts and fracture zones.

There are hundreds of other maps, all insets to textual spreads describing a wide variety of interesting salty topics: environment, resources, trade, strategy, exploration, law. Here is the track of the Trolltunga iceberg, a fragment of the southern ice cap, first seen by satellite in 1967. It was then a giant floe about half the area of Long Island; it floated slowly northward, to dwindle into "little more than a collection of small bergs" south of the Cape of Good Hope a decade later. There is a world map studded with close to 200 straits and canals, and good-size detailed maps of the half-dozen most important of them, such as Hormuz, a candidate for the headlines in any troubled year.

There are maps of ocean cables, of the drill stations of the *Glomar Challenger*, of the moving front of a typical tsunami, of worldwide shipwrecks for two years and even of the two dozen major surfboarding sites (three million participants, the legend says, on all but the polar oceans). The Falklands-Malvinas war is mapped as one of nearly 200 naval incidents around the world since 1946. U.S. involvement leads, whether in defense or in attack: we keep more naval ships at sea in harm's way.

This volume is as full of general interest as it will be a first resort for reference; its Cardiff editor and contributors (that university in Wales claims most of the authors) have reason for satisfaction, to be shared by the production groups in England, Denmark and Italy.

VORTEX FLOW IN NATURE AND TECHNOLOGY, by Hans J. Lugt. John Wiley & Sons, Inc. ($49.95). "Vortices are ordered structures of fluid motion, which nature prefers over chaos in many situations." From the tiniest turbulent eddy to the entire galaxy the rotation of a multitude of material particles about a common center is a pattern of motion that has fascinated the eye and the mind's eye for a very long time. Lugt is a distinguished hydrodynamicist at the David W. Taylor Naval Ship R & D Center outside Washington. He is at home alike with big ship models towed down mile-long tanks, with the historical development of understanding and with the powerful computer codes of the day. In this book he has compiled a comprehensive story of the vortex at a serious descriptive level, without equations but copiously documented with tables, graphs and diagrams. The coverage is so wide (he has a first chapter on ancient views of the vortex, including myth and the turbulent musical metaphors in the Bach Passions, and a final one on the vortices of the extraterrestrial cosmos) and the essential geometry is so far from intuitive that the book demands and rewards close study. He has added a brief mathematical supplement, carefully sequestered, to document the remarks of the expository section for students familiar with vector fields and the equations of fluid dynamics.

The chief text has two parts. In the first part the motions are those that gyre from inner causes: the effects of walls, edges and wings, and of the instability and turbulence of most laboratory-size flows. The treatment amounts to an overview of the theory of flow of viscous fluids. In the second part Lugt examines the larger flows of wind and water, at a scale adequate to exhibit the related consequences of the rotation of the earth and the stratification induced in air and salt sea by gravity. The two sections, somewhat more than 100 pages each, present a broad survey of the phenomena in the laboratory and outside it, causally described and quantitatively assessed through a couple of dozen examples.

Try a small sample. Three or four large pages, with one striking photograph of a turbine intake and half a dozen diagrams done in line, are given to the intake and discharge vortices, kin to the bathtub whirls. Surely these ordered and stable motions that evince such compact power without any easily seen origin must be a widespread source of wonder. The conservation law is invoked by a simple numerical estimate, to show how a tiny initial motion can explain the swift central whirl. The bathtub vortex is carefully included in the account, and three or four hardly perceptible sources of the input vorticity are named. Then the experiments are cited that under exquisitely controlled conditions (although not at all in everyday experience) have been able to reveal the majestic spin of the earth with a humble piece of plumbing. In Sydney the symmetrically prepared tub drains clockwise; in Boston it drains the other way.

The most interesting and novel sections of the first part treat of the vortex as it arises from flow collisions and joinings, the instabilities and turbulence of large or fast flows. This work is current;

that ordered vortex patterns appear even as the flow is steadily evolving toward the confused eddies of turbulence is one more example of the modern realization that chaos is not to be entered chaotically. A marvelous photograph from Göttingen only a couple of years old shows a turbulent vortex, looking like a rather ragged smoke ring, shedding infant vortices from its core into a grainy, spin-filled wake as it moves along.

The second part treats of the vortex-laden atmosphere and ocean of our spinning earth. Of course, it has plenty to say about tornadoes and dust devils, about the Gulf Stream and jet streams. It offers a good general account of the modern views of the circulation in sea and sky; although the jet streams are very convincingly mapped and diagrammed (there are two to a hemisphere, the familiar subpolar one and a subtropical one), their origin does not come clear to this reader. On the other hand, the vortex progeny of the Gulf Stream are neatly made rational, and with some surprise we read that the domain central to the North Atlantic gyre, which we know as the quiescent pool called the Sargasso Sea, has a counterpart in the North Pacific circulation, around the islands of Hawaii.

The regimes of flow governed by force ratios expressed as the numbers of Reynolds, Rossby, Ekman and Rayleigh are made physically quite reasonable. Unstable rotation in odd-shaped vessels gives rise to potholes of surprising variety and order, with thanks to Sir Geoffrey Taylor. The chapter that closes the work is rather too ambitious, seeking in a dozen pages to give an account of everything from the Great Red Spot of Jupiter to the spin of the galaxies. Lugt hunts the origin of the angular momentum manifest in the realm of the galaxies, but he does not much entertain the most popular view, which is that galaxies whirl simply because they have taken in one another's spin. Overall there is no sign of net angular momentum in the cosmos.

The book is certainly a first-rate resource at a nonmathematical level for anyone who has a genuine interest in fluid flow. Students and teachers of the topic will want it handy, not least for its 500 references. One hopes that general readers too will dip into its compendious chapters; it was so greeted by a serious audience of readers of German when it first appeared in that language a couple of years back. Its readers will never forget that every darting fish and slow-turning submarine, every falling leaf and summer breeze, leaves behind it a strange wake, a subtle if invisible trail of vortices of varied shape and size, a pattern of motion more organized and perhaps even more ubiquitous than waves.

Established 1845 SCIENTIFIC AMERICAN February 1984 Volume 250 Number 2

The Population of China

The recent census confirms that it stands at a little more than one billion. Through strong measures to reduce the birth rate the government is seeking to have it level off at 1.2 billion

by Nathan Keyfitz

It is frequently observed that one-fourth of the world's population is Chinese. What is noted less often is that 80 percent of the population of China live in the rural areas of the country and hence one-fifth of the world's population is made up of the Chinese peasantry. The demographic trends in this very large social group therefore have a considerable significance not only for China but also for the world as a whole. In many economically developing countries the rate of population growth has fallen faster among the educated urban population than it has among the peasantry. The peasant economy encourages demographic growth: many children provide many hands to work the small plot of land the typical peasant family owns or rents. Children also ensure that the land will be worked when the parents are old. Such conditions prevailed for centuries in China before the revolution of 1949 and the rate of population growth was high.

The collectivization of agriculture that followed the Communist revolution removed much of the economic incentive for large families on the farm. The rate of population growth, however, remained high because profamily cultural attitudes persisted, because the death rate had declined sharply and because little attention was given to birth control. Not until after the death of Mao Zedong in 1976 was the control of demographic growth pursued vigorously as an official policy. The political structure of Chinese society, however, is such that policy decisions can be quickly implemented. As a result the rate of population increase has declined faster in China than it has in any other large economically developing country. Moreover, contrary to the experience of many developing nations the reductions in the birth rate have taken place in rural areas as well as in the cities.

Although a decrease in the rate at which the population is growing is on the whole beneficial to a developing country, a rapid decline in the birth rate has significant social costs. The official position of the Chinese government is that the largest population China can support is 1.2 billion, or 200 million more than the population counted in the most recent census (taken in 1982). I have made projections showing the effect of an immediate decrease in the birth rate to a level commensurate with

ADMINISTRATIVE UNIT	POPULATION (THOUSANDS)
BEIJING	9,231
TIANJIN	7,764
HEBEI PROVINCE	53,006
SHANXI PROVINCE	25,291
INNER MONGOLIA AUTONOMOUS REGION	19,274
LIAONING PROVINCE	35,722
JILIN PROVINCE	22,560
HEILONGJIANG PROVINCE	32,666
SHANGHAI	11,860
JIANGSU PROVINCE	60,521
ZHEJIANG PROVINCE	38,855
ANHUI PROVINCE	49,666
FUJIAN PROVINCE	25,931
JIANGXI PROVINCE	33,185
SHANDONG PROVINCE	74,419
HENAN PROVINCE	74,423
HUBEI PROVINCE	47,804
HUNAN PROVINCE	54,009
GUANGDONG PROVINCE	59,299
GUANGXI ZHUANG AUTONOMOUS REGION	36,421
SICHUAN PROVINCE	99,713
GUIZHOU PROVINCE	28,553
YUNNAN PROVINCE	32,554
TIBET AUTONOMOUS REGION	1,892
SHAANXI PROVINCE	28,904
GANSU PROVINCE	19,569
QINGHAI PROVINCE	3,896
NINGXIA HUI AUTONOMOUS REGION	3,896
XINJIANG UYGUR AUTONOMOUS REGION	13,082

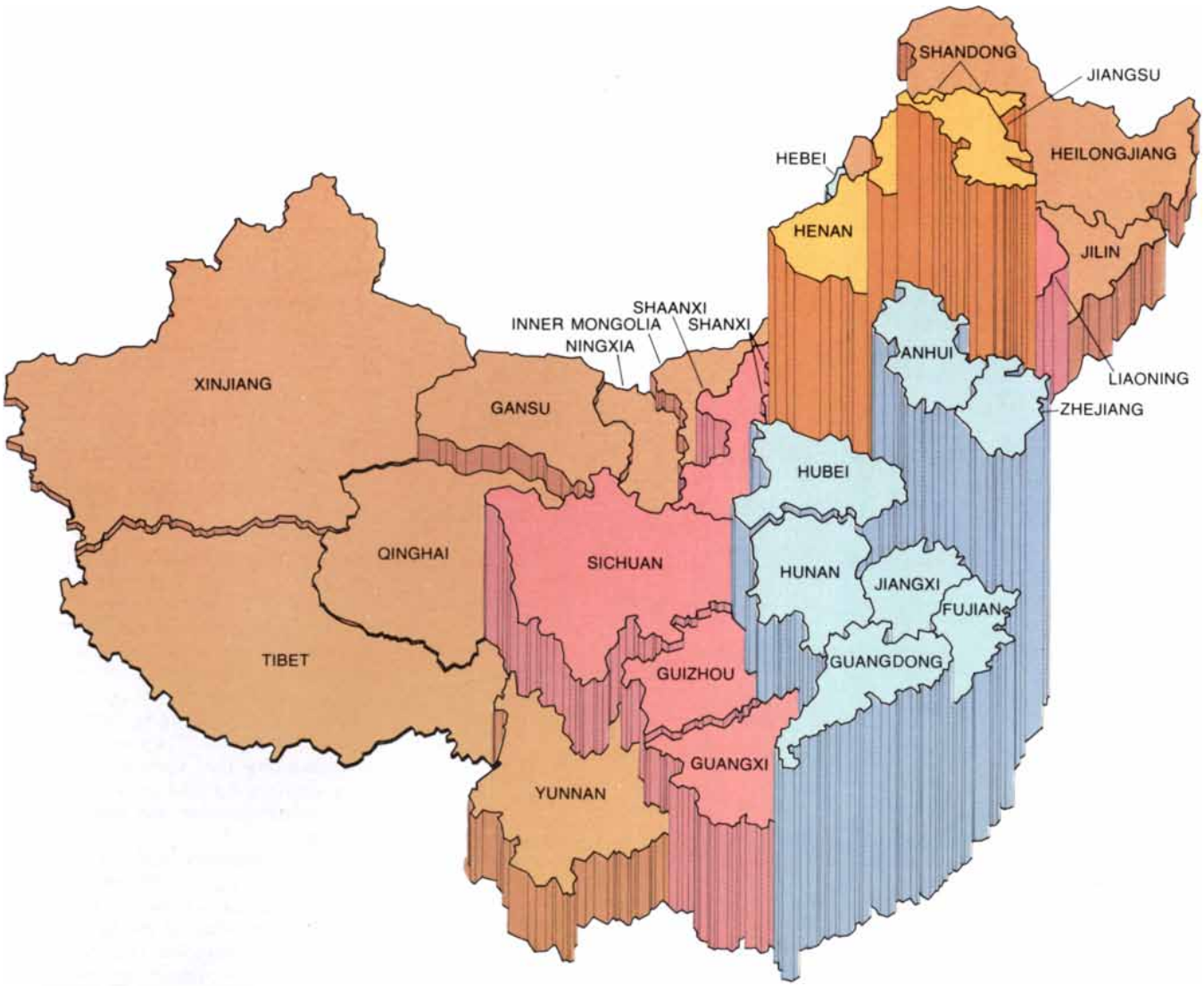
POPULATION DENSITY OF CHINESE PROVINCES is shown on a map where the height of each area is proportional to its population density. Yellow stands for an average density of more than 400 people per square kilometer, blue for 200 to 399 people, pink for 100 to 199

an ultimate constant population of 1.2 billion. The projections show that such a rate could have at least two substantial implications. The first implication is that in the middle of the 21st century there will be a disadvantageous distribution of the Chinese population according to age: a relatively small group of workers will be supporting a large group of retired people. The second is that if China is to achieve a constant population of 1.2 billion, Chinese married couples must average fewer than two children for much of the 1990's, a figure lower than that in most economically developed countries.

The authorities of the People's Republic of China are well aware of such implications. The projections present Chinese officials with some difficult choices. If the birth rate were to drop faster, the ultimate pressure on natural resources would be reduced but the resulting age distribution would be even more unfavorable and the social discipline necessary to impose family planning on the population would be even stronger. If the decline in the birth rate were slower, a less stringent social discipline would be needed and the age distribution would be more favorable but the pressure on natural resources and on the capital available for economic development would be substantial. Whatever choices the Chinese regime makes, it is evident that one of the most intriguing demographic experiments in human history is being conducted in the largest society on earth. The results will have implications for the parts of the developing world where the birth rate is still high and for the human population as a whole.

There have been many changes in the People's Republic of China since the death of Mao but none of them has been more dramatic than the reversal in population policy. Mao believed that the Chinese people, no matter how numerous, could construct the material basis for prosperity; it was said that "revolution plus production" would solve all problems. In the regime of Deng Xiaopeng, however, population growth is perceived as the main obstacle to improving the standard of living. The emphasis given to birth control by the current regime is reflected in many official statements and actions. For example, in December, 1982, the National People's Congress adopted a new national constitution that includes a provision making family planning the duty of all married couples.

My interest in demographic policy in China was sharpened by a monthlong visit there in January, 1982, under the auspices of an academic exchange pro-

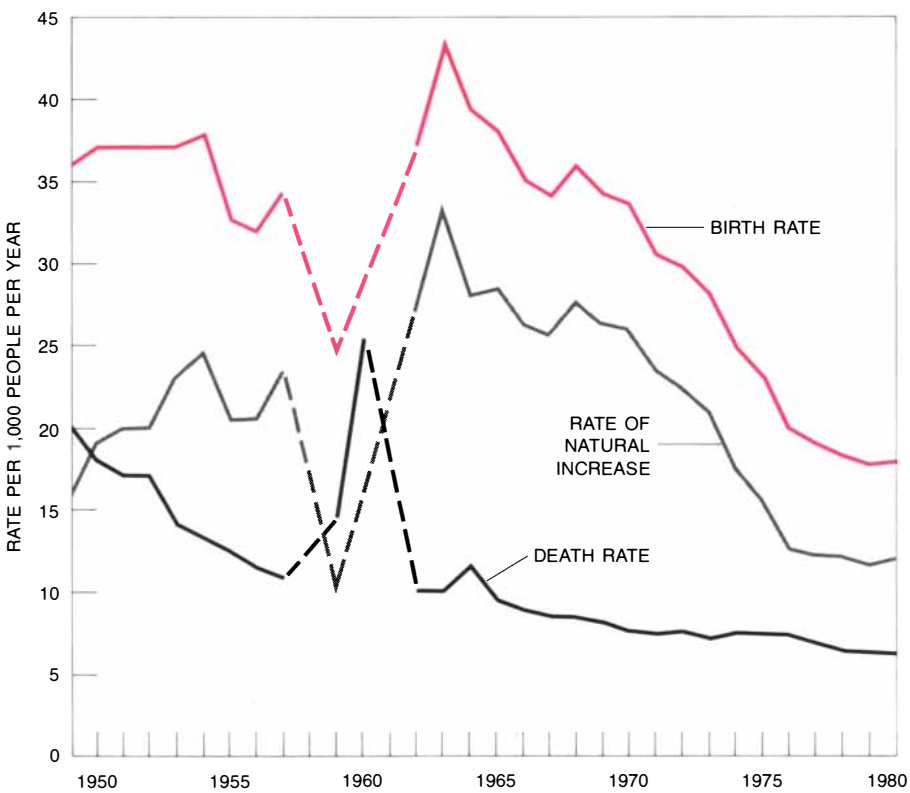


people and brown for fewer than 100 people. The east coast with its large ports and concentration of industry is the most densely populated part of the country. The table at the left shows the population of the provinces, the autonomous regions and the three independent municipalities (Shanghai, Beijing and Tianjin) as counted in the census of 1982. Shanghai, with 11.9 million people, is the largest city.

gram sponsored by the National Research Council of the National Academy of Sciences. The visit took me to Beijing, Chengdu and Canton and to several rural areas as well. I lectured and interviewed people at many levels of Chinese society, including factory workers, peasants, managers and government officials. From such experiences I gathered a sense of how the problem of demographic growth is perceived in the People's Republic. These impressions augmented the objective information that is given by censuses and numerical projections.

The most frequently expressed official reason for controlling the growth of the Chinese population is the scarcity of agricultural resources. Although per capita food consumption has increased since the revolution, it is said by officials that 10 percent of the population are underfed. Moreover, most of the country consists of mountains, deserts or lands otherwise unsuited for agriculture. The 11 percent of the total land area that is now under cultivation includes virtually every square meter of arable land. A recent article on the front page of *People's Daily*, the newspaper of the Chinese Communist Party, asserted that if the population continues to grow at the rate of the past 20 years, demographic pressure will threaten the improvements in the standard of living that have been achieved since 1949.

Whereas food consumption per capita has increased only slowly, there has been a remarkable change in the distribution of agricultural products. Unlike many other developing nations, China has chosen an egalitarian policy of food distribution. An attempt has been made to provide a healthy, if austere, diet for all. Furthermore, the rationing system keeps people from spending increases in pay on food, and as a result differences in income are not immediately converted into differences in diet. In many developing countries as much as 80 percent of additional income is spent on food; in China the figure is said to be about 20 percent. In particular, meat consumption has been held down. As long as people are hungry the available grain is employed to feed people rather than livestock. The people are urged to spend their income on manufactured goods, which stimulates industrial development. Nevertheless, there is a substantial pent-up demand in China for food of higher quality. Rapid population growth would increase the demand and also perhaps generate political resistance.



DEMOGRAPHIC RATES have fluctuated sharply in China since the revolution of 1949 because of economic, political and public-health factors. The rate of natural increase is the excess of the birth rate over the death rate. Broken lines correspond to years in which the demographic rates had to be estimated. As a result of the more egalitarian distribution of food and universal, albeit rudimentary, medical care the death rate has declined continuously except for a brief period of economic disruption caused by the "Great Leap Forward" of 1960. The birth rate, which was high in prerevolutionary times, declined in the 1950's, then rose during the Cultural Revolution of the mid-1960's. In the 1970's population control began to be pursued vigorously as official policy and the birth rate declined rapidly. The One-Child Campaign adopted in 1979 has led to further decreases. The rate of natural increase has fallen, but the population is still growing. Recent reports from China suggest the birth rate has risen since 1980.

Although the well-defined limits of agriculture are the strongest motivation for family planning in China, other considerations have strengthened the collective will to limit population growth. China has a considerable potential for expanding its industry and its output of energy. Economic development, however, requires substantial capital. Chinese officials have thoroughly assimilated an argument advanced more than 20 years ago by Ansley J. Coale and Edgar M. Hoover of Princeton University. Coale and Hoover maintained that if in a developing country the birth rate is high, the capital needed to develop industry will be expended in supporting and educating the younger generation. A lower rate of population growth would make it easier to accumulate the capital for development.

Another disadvantage resulting from a high rate of population growth is the presence of many young people who need jobs. The problem of employment for the young is particularly acute in two groups: the educated urban populace and the residents of rural areas where mechanization has reduced the demand for labor. Of the two groups the more serious political threat comes from the educated. Intellectuals have traditionally enjoyed respect in China and have provided leadership for many political opposition movements, including the revolution of 1949. The government needs the knowledge, expertise and prestige of well-educated city dwellers yet is skeptical of their political reliability.

The problem is exacerbated by the shortage of jobs suitable for educated people. The regime is attempting to limit the size of the bureaucracy, which is already too large for maximum efficiency. Denied entry into official positions, many urban young people now have what is called provisional employment: low-level jobs that make little use of their qualifications. They work with the understanding that they will be given more demanding and prestigious jobs when openings occur, but there are few openings.

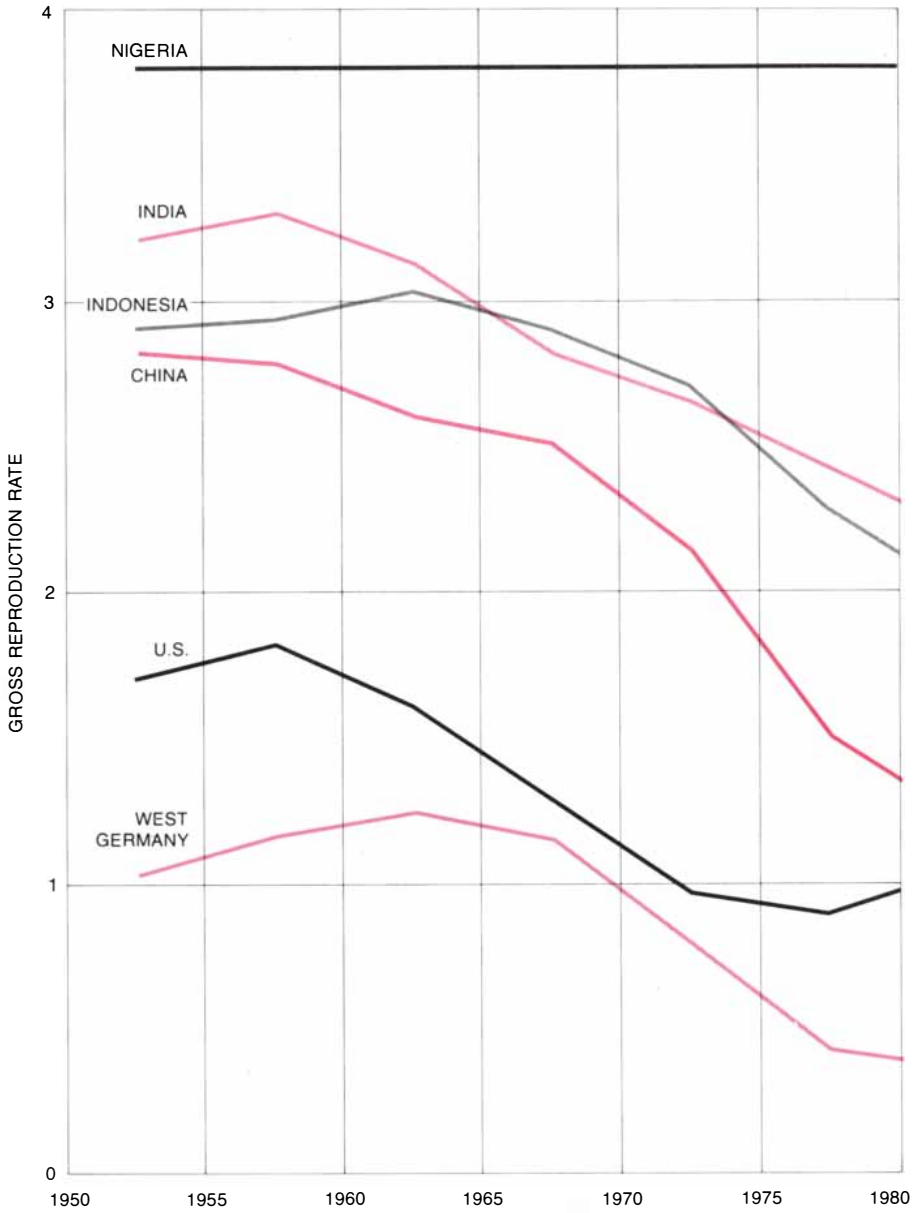
One of the reasons there are few openings is that the elderly are entrenched at the top of the career ladder. The political organizers who led the fight against the Kuomintang in the 1930's are currently in their seventies, and they still occupy positions of authority. At the end of their careers they hold the most conspicuous and responsible jobs and

stand in the way of the advancement of the next two generations. The regime has required the resignation of some senior officials as a symbolic concession to the young ones whose chances for promotion are blocked. Such gestures, however, do little to open up careers for millions of new graduates. Slower population growth would result in fewer underemployed urban youths in the future.

Surplus population is not as serious a threat in the rural areas as it is in the cities, but the countryside also has its problems. The mechanization of agriculture is not currently being promoted as vigorously as it once was. The introduction of rather crude farm machinery, however, has increased the output per person in the fields. Since the amount of arable land is quite limited, increases in productivity reduce the demand for labor. The share of the gross domestic product contributed by agriculture declined from 47 percent in 1952 to 26 percent in 1970. In order to accommodate those thrown out of farm work the government has tried to stimulate the development of light industry in the countryside, but finding the capital for such projects exerts considerable pressure on the fragile budgets of agencies in the lower levels of government.

In addition to the dilemmas of resources and employment, uncontrolled population growth could aggravate tensions between the city and the country. Developing nations have difficulty limiting the size of their cities, in particular their largest city. In China the flow of people to the urban centers has been restrained by the apparatus of government control. Only people with jobs can legally settle in cities and obtain ration cards, housing and coupons authorizing the purchase of medical care and consumer goods, including the all-important bicycle. As a result of such control, squatter colonies, which exist in the cities of many poor countries, are not seen in China. Even the stringent official policy, however, does not prevent some Chinese from living illegally in urban areas with their relatives. The cities currently hold 21 percent of the population, and limiting the exodus from the farms has been a constant struggle, with the authorities sometimes winning and sometimes losing.

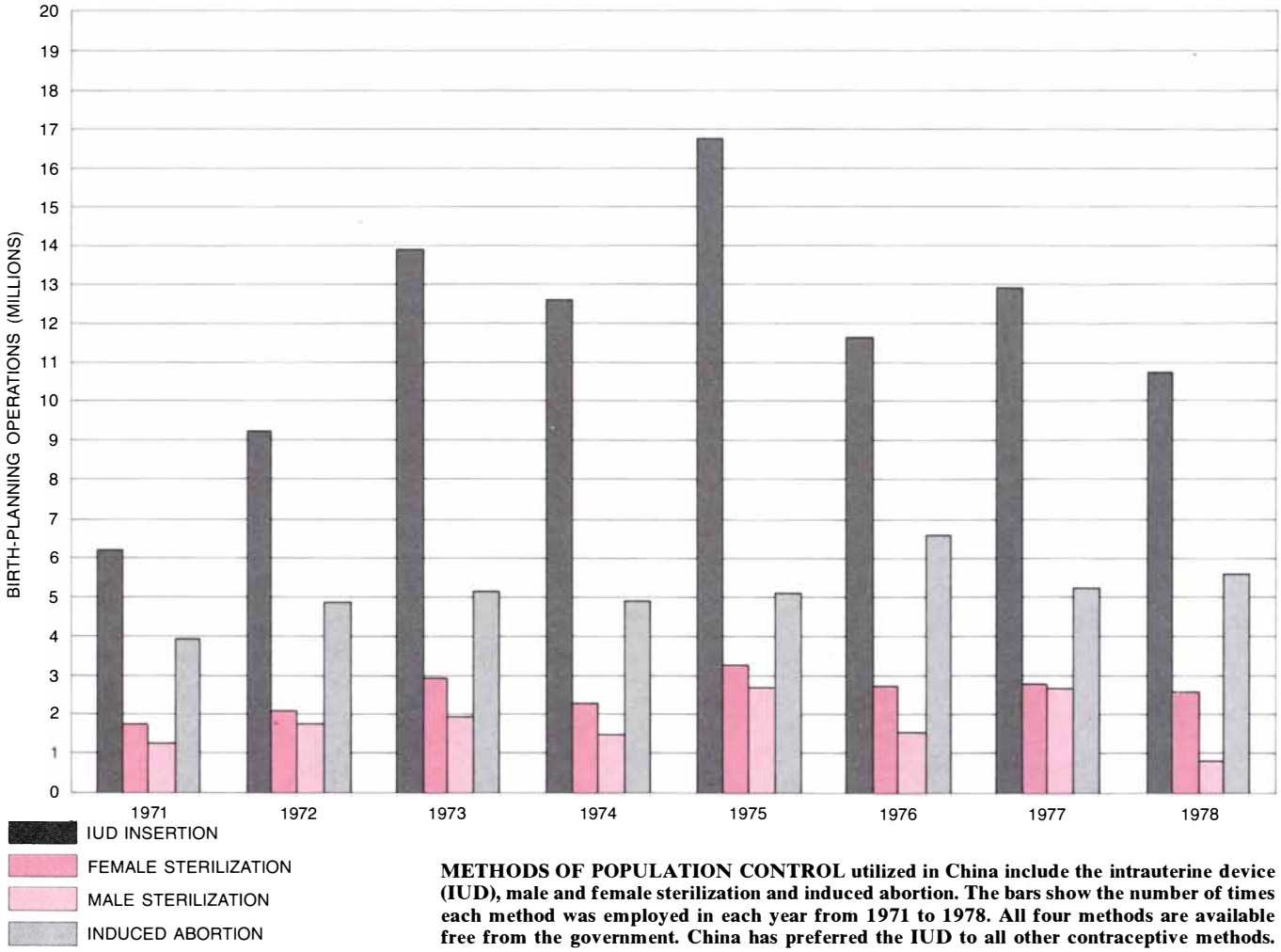
If the urban population increases faster than industrial production in the urban areas, the burden on the farm population will become heavier. The residents of the cities are fed largely with about 50 million tons of grain per year called commodity grain. The commodity grain is sold at a controlled price under the rationing system. The grain is collected in part by a tax in kind levied on the agricultural collectives. To increase the tax burden in the rural areas would incur the risk of renewing the conflict between the urbanites and the farmers.



PACE OF CHILDBEARING has decelerated more rapidly in China than in any other large developing country. The curves show the gross reproduction rate for several countries. The gross reproduction rate is the number of daughters each woman would have on the average if all women survived to the end of their childbearing years. In China the rate has declined from about 3.0 in 1950 to less than 1.5 in 1980. In Indonesia, a developing country with a highly organized family-planning program, the rate is still higher than 2.0. In the U.S. and other economically developed countries it can be 1.0 or less. In Nigeria, where there has been no decline, the rate is about 4.0, implying that healthy women are having an average of eight children.

Thus the desire to limit population growth, which is motivated by scarce resources, is reinforced by several social and political factors. The national census taken in July, 1982, did nothing to allay official concern. Indeed, the census results that have been released hold few surprises. The total population count of 1.008 billion was close to the estimates made beforehand. The high priority given to population matters was made clear in the time and energy devoted to the census. About 5.7 million census takers were appointed, and a member of each household was required to come to a census official and answer 19 questions. Some 100,000 clerks are now preparing the 300 million census forms and entering them in computer files.

Only summary figures from the count of population have been published; they include figures for the population of the 26 provinces and autonomous regions. The final census results, with such detailed information as the distribution of the population by age, have been promised before the end of this year. The



METHODS OF POPULATION CONTROL utilized in China include the intrauterine device (IUD), male and female sterilization and induced abortion. The bars show the number of times each method was employed in each year from 1971 to 1978. All four methods are available free from the government. China has preferred the IUD to all other contraceptive methods.

published figures on individual provinces confirm the sharp differences in population density in China. Eastern China, with its large ports, remains the most densely populated part of the country. Shanghai, which has a population of 11.9 million, is still the largest city. The strip that includes Shanghai and Beijing continues to be the most heavily industrialized area in spite of official efforts to promote the development of industry in the western region of the country.

The centralization of industrial production carries with it large economies of scale, and in attempting to disperse industry the government is opposing underlying economic and social factors. The poor success of the decentralization presents a striking contrast to the resounding success of the family-planning program.

Between 1949 and 1980 the birth rate fell by half, a drop unprecedented in a large developing country. The decline in fertility, however, has not been smooth or steady. The birth rate has fluctuated sharply with changes in economic organization and social policy. In 1949 the rate was 35 births per 1,000 people per year, a level that had prevailed for many decades. As a result of the elimination of private ownership of agricultural land and early birth-control campaigns the rate fell to 25 per 1,000 in the late 1950's. During the Cultural Revolution and other political upheavals of the 1960's birth planning was abandoned as an official goal, whereupon the birth rate rose rapidly to a maximum of 44 per 1,000 in one year.

It was in the 1970's that the full effect of organized population control began to be felt. The Third Birth Planning Campaign, with the slogan of "Later, Longer, Fewer," advocated later marriage, a longer interval between births and fewer children. It was succeeded in 1979 by an even more energetic program: the One-Child Campaign. By 1980 the birth rate had decreased to 18 per 1,000. Official reports may have underestimated the number of births; there has also been a recent increase in fertility. Hence the current birth rate could be about 20 per 1,000.

The magnitude of the decrease that has been achieved in fertility is impressive. The One-Child Campaign appears to be taking hold, at least in some parts of the country. In Shanghai, Beijing and five of the 26 provinces and autonomous regions between 80 and 90 percent of all births are births of first children.

The effectiveness of the recent birth-control campaigns results chiefly from the political structure of Chinese society. The structure is the legacy of the 20-year struggle for power waged by Mao and his associates. In the struggle cadres were carefully selected and trained in secret party schools; each cadre then became a link in the unified command structure. Many cadres were recruited from the peasantry, and their job was to transform the rural populace from the unprotesting, taxed people that Asian peasants have tended to be into conscious agents of change.

When the Communist Party took power, the increased political consciousness and organization that had been developed in war were turned to civilian purposes, including family planning. In part because the political structure originated in an armed conflict, it has a hierarchical form much like the form of a military organization. The hierarchy has six main levels. At the top is the nation as a whole and at the bottom are the production teams, each with from 250 to 800 members. Every ad-

ministrative body, regardless of its level, includes an agency of the executive branch, a Communist Party unit and a unit responsible for the delivery of health services. Birth-control services are integrated into the health-care system, and in each administrative body there are personnel whose job is family-planning education. In small discussion groups guided by party cadres the concept of population control is passed down from the highest executive agency to married couples.

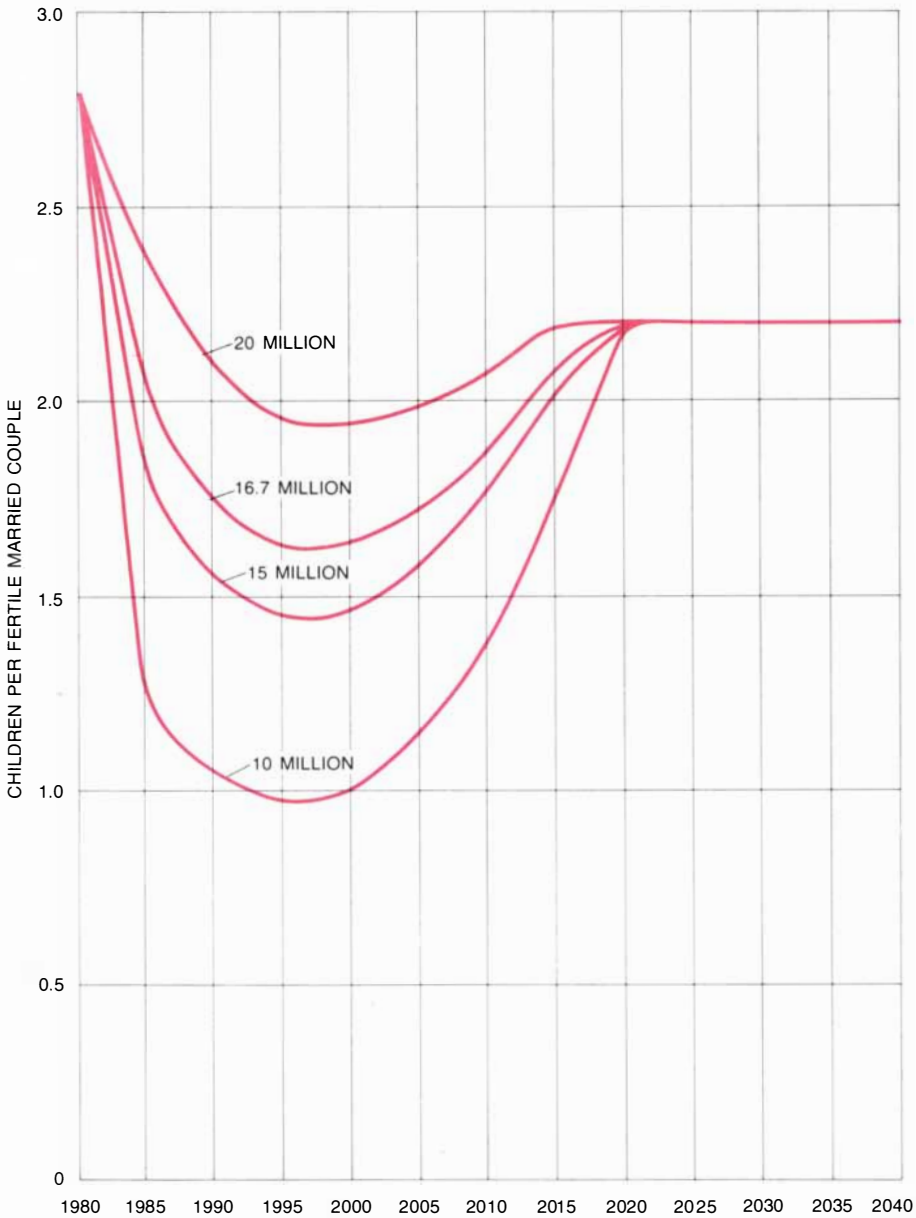
In addition to arguments for population control, birth-control devices and rewards and punishments are delivered through the administrative apparatus. Contraceptive pills, intrauterine devices (IUD's) and sterilization operations are available free to married couples, as are premarital physical examinations that include explanations of contraception. Unlike most other nations, the Chinese have favored the IUD over all other contraceptive devices.

The means of persuading married couples to limit their fertility is left to the local authorities, but in all localities there are rewards for family planning and punishments for having too many children. Newly married couples are asked to sign a statement committing them to having only one child. By the middle of 1981 more than 11 million couples had signed the agreement, which entitles them to free hospital delivery of their child, free medical care and education for the child, special consideration for better housing and an extra month's salary each year. Single children are given preferential treatment in school and in obtaining jobs, particularly if the child is a girl.

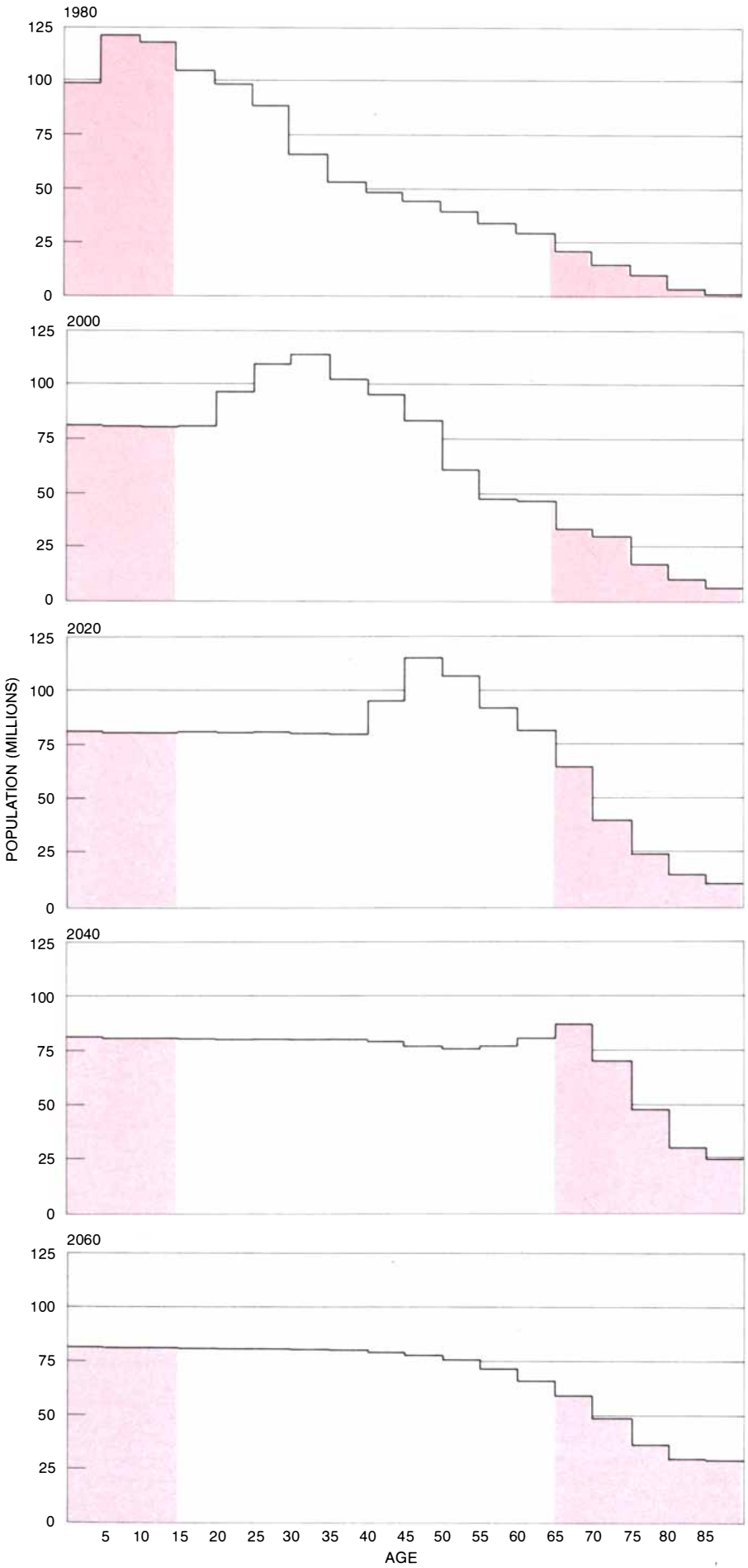
The penalties for unsanctioned pregnancies can be substantial. The birth of a second child to a couple that has signed the one-child document results in an official reprimand. The birth of a third child results in a pay cut. Many women who are pregnant without permission are persuaded to have an abortion. Having an abortion entitles a woman to a paid vacation, as does a contraceptive sterilization operation. The decentralized execution of political decisions, with zealous local party cadres competing to make a good showing, can lead to coercion in both abortion and sterilization.

The combination of free contraceptives, the system of rewards and punishments and the well-organized apparatus of social persuasion has led to one result that is quite unusual. In most countries where the demographic transition (from high birth rates and death rates to low birth rates and death rates) has begun the reduction in fertility has been greatest in the areas with the highest level of urbanization and income and the lowest death rate. Work done by H. Yuan Tien of Ohio State University shows that until recently there was a correlation in China between urbanization, income and low mortality on the one hand and reductions in fertility on the other. By 1980, however, the relation was much weaker and indeed was no longer significant in the statistical sense. The family-planning campaign has apparently now reached the poor rural areas without the health, wealth and urbanization that have been associated with the demographic transition in Europe and many developing countries.

Notwithstanding the effectiveness of the birth-control program there are two factors that tend to sustain the increase of the population. The first is the rapid decrease in the death rate. Like other poor countries, China must choose between, on the one hand, a meat diet and sophisticated medical care for its middle class and, on the other, a universal diet



AVERAGE NUMBER OF CHILDREN each Chinese couple could have if there were a constant number of births each year is shown in these curves. The top curve corresponds to 20 million births per year, the curve below that to 16.7 million births, the next curve to 15 million births and the bottom curve to 10 million births. The number of births in China has fluctuated greatly in the past. Therefore the number of couples who will mature and begin to have children at various times in the future will also fluctuate. If a constant number of births is to be maintained, the average number of children per married couple must vary during a transitional period, reaching a low point in the 1990's, when the large cohorts born from 1960 to 1980 are of childbearing age. With 15 million births per year the average Chinese couple could have slightly less than 1.5 children during much of the 1990's. A level of 10 million births per year corresponds to an ultimate population of 700 million. If there were 10 million births per year Chinese couples would have to average less than one child during much of the 1990's.



of cereals and rudimentary universal medical care. Unlike most developing countries, China has chosen the egalitarian alternative.

As a result mortality declined from 20 deaths per 1,000 people per year in 1949 to six per 1,000 in 1980. The death rate in the U.S. is about eight per 1,000. The comparison is a misleading one, however, because the population of China is on the average much younger than the population of the U.S. A more accurate comparison is that the average life expectancy in China is estimated by the United Nations to be approaching 70 years, whereas in the U.S. it is 73 years. Nevertheless, the improvement in the past three decades has been remarkable: in 1950 the life expectancy in China was 47 years.

Because the death rate has declined quickly, the population has increased even though fertility has begun to fall. The rate of natural increase, or the excess of births over deaths, was five per 1,000 people in 1949. After rising to 20 per 1,000 in the 1960's the rate has recently fallen, but only to 12 per 1,000. If that rate were maintained, the population would double every 60 years.

The second factor that tends to keep the population of China growing is the one known as demographic momentum, which is the delayed effect of a recent history of rapid population growth. In a population that has recently grown rapidly there is a distribution of people by age that is heavily skewed toward the young, who will in a short period be ready to marry and have children. Consider what would happen if all married couples in such a population immediately reduced their fertility to the level necessary for replacement.

For a human population to replace its members each fertile couple must have on the average 2.2 children. The figure is slightly more than 2.0 because of infant and child mortality, infertility and people who do not marry. Even if all married couples in the population reduced

CHART OF AGES shows what would happen to the Chinese population if the number of births declined immediately to 16.7 million per year and remained there. The curves show the distribution of the population according to age at 20-year intervals beginning in 1980. The working population (*white*) must support the dependent population (*color*), consisting of children and retired people. The immediate effect of a reduction in births would be beneficial: the large group born from 1960 to 1980 would enter the work force and there would be few children to support. Beginning in about 2020, however, the number of retired people would increase. The retirees would have to be supported by the small cohorts born after the number of births became constant.

their fertility to the replacement level, the population would not stop growing immediately because the large cohorts of children born recently would mature and have children. In a transitional period of about 50 years the population would continue to increase. The ultimate stationary, or constant, population could be as much as 60 percent larger than the population was when the decline to replacement-level fertility took place.

If it is assumed that the decline to replacement-level fertility in China came about in 1980 and that the death rate will continue to decline at the pace of the past few years, the population in the year 2000 will be 1.2 billion. Demographic momentum would continue to operate in the 21st century, and by 2075 the population would be more than 1.5 billion. There is much debate in China about how many people the country's resources can support. Chinese estimates of the capacity range from 700 million to 1.4 billion. In 1980 Hua Guofeng, chairman of the Communist Party, said the goal of the party was to limit the population to 1.2 billion by the end of the 20th century. Furthermore, the government's goal is that the population should then be stationary.

Since the immediate decline to replacement-level fertility would yield an ultimate population of more than 1.5 billion, it can readily be seen that if a stationary population of 1.2 billion is to be achieved, fertility must be below the replacement level in the period before such a population is achieved. One way to arrive at a stationary population is for there to be a constant number of births each year. If the right annual number is chosen, any particular stationary population can be achieved. The size of the population will fluctuate as the cohorts born before the number of births was stabilized age and die, but when the first of the constant cohorts reach the end of their lives, the population will be stationary at the target level.

In most countries asking what the effects on society of having a constant number of births each year would be is idle speculation because the number of births cannot be controlled. In China, however, the regime appears to be able to control the annual number of births with considerable precision. Therefore the question is of practical interest.

I have prepared a set of projections of the population of China over the next several decades starting from the assumption that the annual number of births will be constant from now on. Several figures for births have been employed, with each figure corresponding to a different ultimate population.

A somewhat more modest goal than the one announced by Hua Guofeng is that the population in the year 2000 should be 1.2 billion and that the constant population should be 1.2 billion but that in the period between 2000 and the achievement of a stationary level the population should be allowed to rise somewhat before returning to the stationary level. Such a goal could well be practical with the current means of birth control in China.

I have assumed that the death rate will continue to drop and that the average life expectancy over the next two generations will be 72 years. If the death rate follows this course and a stationary population is achieved by means of the pattern described in the preceding paragraph, an ultimate constant population of 1.2 billion corresponds to 16.7 million births per year. For the sake of simplicity I have assumed that the decline to the constant annual number of births took place in 1980.

From the point of view of allocating resources one of the most significant demographic features of a society is the distribution of the population according to age. The age distribution determines the ratio of people who are working to those who must be supported by the population of workers. The dependent group includes both children and older people who have retired from the labor force. An increase in the fraction of very old or very young people implies that more resources must be devoted to taking care of dependents and that hence less capital is available for investment in industry. Such considerations are particularly significant in a country where the pace of economic development is limited by the available capital.

What would the ratio of dependents to workers be if there were 16.7 million births every year in China? In 1985 there would be 81 million surviving children less than five years old. In the 15 years preceding 1980, however, there were considerably more than 16.7 million births per year.

The large cohort made up of those born between 1965 and 1980 will for many years constitute a peak in the curve representing the distribution of population according to age. As the large cohort ages and moves through the phases of the working life of its members it will be followed by the small cohorts born after the annual number of births declines to 16.7 million. The immediate effect on the ratio of workers to the dependent population would be favorable. The decrease to an annual level of 16.7 million births per year represents a reduction in the number of children that more than compensates for the initial increase in the number of older people. The fraction of the population in the working ages from 15 to 60 would continue to rise until 2005: the large recent cohorts would be entering the labor force as the small cohorts of the 1920's retire and the small cohorts of the 1980's and 1990's are in school.

After 2005, however, the fraction of retirees would begin to increase sharply. By 2040, when the last of the cohorts born between 1960 and 1979 reach 60 and retire, the ratio of the retired population to the working population would reach its highest level: there would be 32 retirees for every 100 workers. Such a ratio would entail a substantial cost. For the purpose of comparison, in 1980 there were nine retired people for every 100 workers.

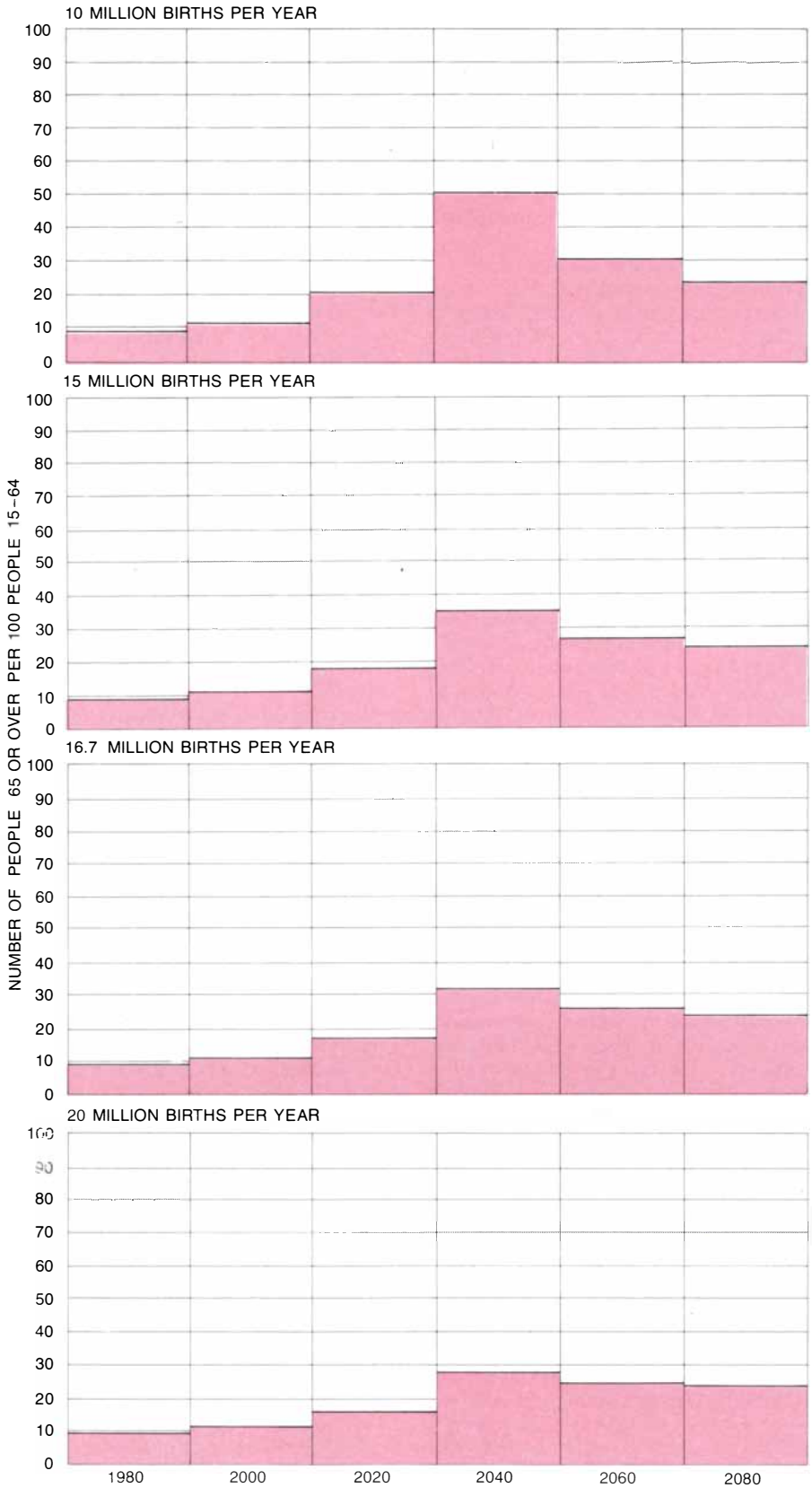
If the annual number of births were to fall below 16.7 million, the shifts in the age distribution would be even more disadvantageous. Song Jian, the president of the Chinese Society for Automation, has estimated that if every person is to have 85 grams of protein per day, of which 75 to 80 percent comes from animal sources, then China can support only 700 million people.

A stationary population of 700 million corresponds to 10 million births per year. If the number of births had fallen to 10 million per year in 1980 and remained at that level, by 2040 there would be 50 retired people for every 100 workers. Every couple in the working population would be supporting one older person, and the financial burden would be heavy. With 20 million births per year, corresponding to a stationary population of 1.4 billion, the shifts in the age distribution would be much less disadvantageous.

There is another kind of cost that must be taken into account when the consequences of a population goal are considered: the social tensions that arise from the need to impose a stringent limit on how many children people can have. As I have mentioned, any stationary population of less than 1.4 billion requires that for a period married couples must on the average have fewer than 2.2 children. If a very low stationary population is chosen, then during a transitional period the average fertility would have to decline far below the replacement level.

For example, an ultimate constant population of slightly more than one billion corresponds to 15 million births per year. For there to be a constant level of 15 million births per year the fertility of the average married couple in the early 1990's would have to be less than 1.5 children. No married couple could have more than two children; half of the couples could have two children and half would have to settle for one child.

If the extreme goal of 700 million is selected, the limits on fertility are even more draconian: in the early 1990's married couples could not average even one child each. Together with the unfavor-



RATIO OF WORKERS TO RETIRED PEOPLE who must be supported depends on the birth rate. The bars show the ratio of those 65 or older to those from 15 to 64 if there are 10 million births per year, 15 million births, 16.7 million births or 20 million births. The current ratio of older people to workers is about nine to 100. The ratio will reach a maximum in 2040, when the large groups born between 1960 and 1980 have retired. How high the ratio rises depends on how many people are of working age in 2040, that is, on how many births there are each year for the next few decades. With 10 million births per year the ratio would reach 50 to 100 in 2040: every two workers would be supporting one retiree. With 16.7 million births the maximum would be 33 to 100. With 20 million births the maximum would be about 27 to 100.

able age distribution such a harsh limitation on childbearing suggests why a stationary population of 700 million is currently an impractical goal. With 16.7 million births per year the lowest level of fertility would be about 1.7 children per couple.

Officials of the People's Republic have already begun to evaluate methods of resolving the problems caused by a rapid decline in the birth rate. Difficult choices will have to be made, particularly with respect to employment. One way to solve the problem of too many dependents in an aging population is to defer retirement. If the age at which people can retire is raised, the ratio of retired people to workers will obviously decline. Indeed, if retirement age is raised enough, the current ratio could be maintained in the future even though the average age of the population will have risen considerably.

The current official retirement age in China is 60. If there are 16.7 million births per year, the ratio of retirees to workers that prevailed in 1980, which as we have seen was nine to 100, could be maintained by raising the age of retirement to 62 in 2000, to 67 in 2020 and to a maximum of 73 in 2040. Thereafter the retirement age could be slowly lowered.

There would undoubtedly be objections to such a plan. Some would come from the old, who would think it unfair to have to work past 70 when their parents had retired at 60. The chief objection, however, would come from the young, who would perceive the elderly as holding on to the good jobs. Whatever solution is chosen, the old must be provided for, and if there are many elderly people in relation to the size of the labor force, a choice must be made between allowing them to retain their jobs and retiring them and supporting them with tax money.

The main current problem, however, is finding employment for young people born in the 1960's who are entering the labor force in large numbers. Part of the solution that has been chosen is to decentralize authority and make local units, including communes and production teams, responsible for employment. The local administrative units are expected to maintain a financial balance and also a balance of human resources utilizing all the labor that is offered. In the rural areas the employment balance has been threatened by the introduction of trucks and mechanized plows that increase the productivity of agricultural labor.

To absorb surplus labor and increase the production of food the government has recently begun to allow some peasants to work private plots and sell the produce they raise in state-sponsored

CHINESE FARM MACHINERY appears primitive by the standards of the economically developed countries but has reduced the demand for labor on the collective farms. The photograph was made by the author in the province of Sichuan during a visit to China in 1982. The front of the vehicle shown is a gasoline-powered unit that can pull a wagon or pull a plow with a man walking behind it. There is little arable land in China that is not currently being cultivated. Therefore an increase in the productivity of agricultural labor such as is yielded even by crude machines challenges the government to create industry to employ the laborers no longer needed on the farm.

markets. Such privatization, however, tends to reinstate the social arrangements of prerevolutionary peasant life and hence to increase the economic value of children. The policy of private cultivation could undermine the efforts of the government to control population growth.

Therefore some recent social tendencies suggest parallels to prerevolutionary society. The revolution of 1949, which was made largely by peasants, was the last of many peasant revolts. Each successful revolt eliminated many large landlords and brought to power a new dynasty with an austere and uncorrupted court. As the new dynasty became entrenched taxes increased and freeholders had to borrow to meet their obligations to the government. Sooner or later the farmers were unable to pay and they sold their lands, becoming tenants of large landlords who could resist the government tax collectors.

As the court became more corrupt the pay of government employees declined. The poorly paid army was ineffective and eventually a new revolt of hungry tenant farmers overthrew the existing regime. Since the system of land tenure and taxation was not changed, the cycle began again.

Demographic growth was a central feature of the cycle. In a statement that is still quoted in China, Han Fei-tzu wrote 2,500 years ago: "People think five sons are not too many and each son has five sons also. Therefore people are more numerous and wealth is less; the people work hard and receive little." Jack Goldstone of Northwestern University has shown that in the Ming dynasty there was a correlation among rising food prices, demographic growth and revolution. From the late 14th century to the late 16th century the population of China more than doubled although the area of land under cultivation increased less than 50 percent. In the same period the price of rice more than tripled, and there were many popular uprisings. The current regime is well aware of such precedents and is working to avoid the fate of earlier rulers.

The change of government in 1949 abolished the form of land tenancy that drove the prerevolutionary cycle. Even more significant, it transformed the Chinese peasantry into a watchful, politically active group; the transformation was continued by the expansion of the educational system after the revolution. The profound social change is in large part responsible for the success of organized family planning in China. The reduction of demographic growth, however, brings with it complex social problems, including an unfavorable age distribution and the necessity for strong social discipline.

The regime has considerable support from the Chinese people and its guidelines elicit a degree of spontaneous compliance. Nevertheless, newspapers and other periodicals in China disclose and criticize many elements of compulsion. A large fraction of the national product, at times as much as 30 percent, is skimmed off as capital, not all of which is well invested. Part of the sacrifice that the 30 percent represents goes to support an overlarge bureaucracy. Urban incomes average three times as high as rural ones. With regard to the success of family planning, John S. Aird of the United States Bureau of the Census observes that "the single most important factor is direct administrative pressure on individual Chinese families from family-planning cadres and other officials who are themselves under pressure to achieve quick results." Accounts are published of forced abortions, and there are cases of infanticide when the firstborn turns out to be a girl. Protests are heard when a female only child and her parents are given exceptional privileges. One can expect this form of resistance and others to increase as the economy becomes less regulated under the policy of self-reliance.

The Oldest Eukaryotic Cells

All higher organisms are composed of eukaryotic, or nucleated, cells. A review of the fossil record indicates that the eukaryotes originally evolved in the form of unicellular plankton some 1.4 billion years ago

by Gonzalo Vidal

According to the modern interpretation of the fossil record, the first traces of life appeared on the earth about 3.5 billion years ago, within a billion years or so of the planet's formation. The earliest identifiable life forms were prokaryotes: primitive unicellular microorganisms that, like present-day bacteria, lacked a distinct nucleus. Multicellular organisms, which are composed entirely of eukaryotic, or nucleated, cells, did not make their first appearance until almost three billion years later. When in the immense span of geologic time that intervened did the first eukaryotic cells evolve from their prokaryotic ancestors?

Until quite recently it was widely thought that direct evidence of this momentous evolutionary event would most likely be found among the fossilized remains of benthic, or bottom-dwelling, aquatic microorganisms. It now appears, however, that the original eukaryotes were plankton, buoyant creatures that floated freely in the sea. The microscopic fossils identified as the remnants of the earliest eukaryotic cells have been classified for some time as acritarchs, a compound term from the Greek meaning "of uncertain origin." Acritarchs are found in various sedimentary rocks from many parts of the world, some of which have been estimated on the basis of radioactive-dating techniques to be at least 1.4 billion years old.

Before 1930 the oldest fossils known were those recovered from sedimentary rocks dating from the Cambrian period, the first interval of the Paleozoic era, which according to the currently accepted chronology began 570 million years ago. For years the division between Cambrian and Precambrian time was considered an enigmatically sharp discontinuity. Rocks of Cambrian age yielded comparatively abundant fossils of extinct organisms representing virtually every major biotic group extant today. Precambrian rocks of the late Proterozoic era, although quite similar in most cases to their immediately younger Cambrian counterparts, were found to be conspicuously barren of fossils. The only notable exceptions were the laminated limestones called stromatolites, which are now interpreted as sedimentary structures produced by superposed communities of bacteria; although stromatolites had long been known to extend far back in time into the Precambrian, they were not regarded by most paleontologists as definitive evidence of early life. Published claims of the discovery of various other forms of Precambrian life were made as early as the late 19th century, but they too were generally viewed with skepticism.

The first convincing evidence of Precambrian animal life came in 1930 with the finding by the German paleontologist G. Gürich of traces of fossilized multicellular organisms in late Proterozoic rocks from Namibia in Africa. In 1947 the Australian geologist R. C. Sprigg discovered numerous imprints of primitive aquatic animals—including jellyfish, various kinds of worms and possibly sponges—in the sandstone beds of the Ediacara Hills in South Australia; at the time the rocks were believed to date from the early Cambrian period, but subsequent study showed that they too belonged to the late Proterozoic. Since then similar animal fossils have been found in sedimentary rocks of comparable age in Britain, Canada, Sweden and the U.S.S.R. The discovery of these "Ediacarian" fossils pushed the oldest record of multicellular life back to about 670 million years ago. Except for the equivocal stromatolites, however, no trace of earlier Precambrian life forms was found until the 1950's.

The existence of Precambrian microbial life was first established in 1954 by Elso S. Barghoorn of Harvard University and Stanley A. Tyler of the University of Wisconsin, on the basis of their study of microscopic bodies in stromatolitic rocks from the Gunflint Iron Formation in southern Ontario. Radioactive dating of minerals in the rocks showed they were formed about two billion years ago. The most important step in the uncovering of early Precambrian life had been taken. From this point on the investigation of fossil microorganisms was to lead to a veritable flood of published reports on findings in Precambrian deposits throughout the world. Comparable studies of Precambrian and early Paleozoic microfossils were undertaken at the same time by Boris V. Timofeev of the Academy of Sciences of the U.S.S.R., and a similar deluge of publications followed his pioneering work.

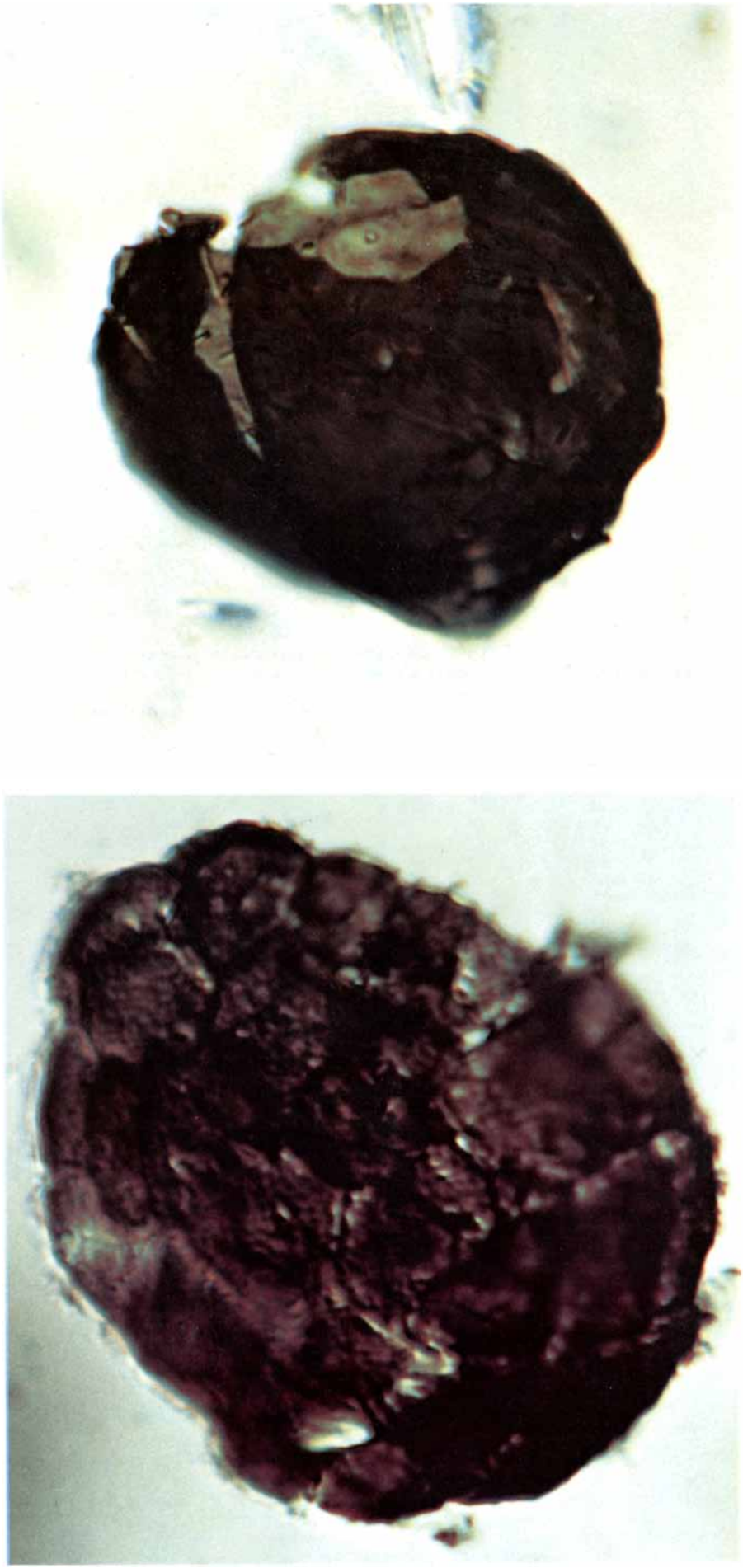
The early studies of Precambrian microfossils in the U.S. and the U.S.S.R. soon diverged into two quite separate lines of inquiry, differing in both the kinds of questions asked and the methodology employed. The American line followed the original hypothesis of Charles Doolittle Walcott, who first ascribed a biological origin to stromatolites at the beginning of this century. In the U.S. and to some extent in Europe the investigation of stromatolites and of the communities of microorganisms that form them became the focus of a multidisciplinary research effort. The Russian line of inquiry, on the other hand, grew out of the long-established European school of palynology: the study of both living and fossil spores, pollen grains and other microscopic structures. Palynological studies of microfossils in acid-resistant residues of early Paleozoic rocks were initiated in 1931 by Alfred Eisenack of the University of Tübingen. The application of Eisenack's technique to Precambrian microfossils constituted a logical extension of his work on the early Paleozoic.

The Russian line of inquiry led naturally to the investigation of the well-preserved Riphean, or late Proterozoic, sedimentary deposits that are widespread in the U.S.S.R. The Russian investigators gave priority to geologic questions; they sought to utilize the microfossils for the purpose of relative dating, in the tradition of biostratigraphy. Questions of a paleobiological nature—those that asked about the rela-

tions of microfossils to one another, to their habitat and depositional environment and to such global factors as paleoclimatology and the evolution of the atmosphere—tended to be ignored by the Russian school. Indeed, some of the initial paleobiological conclusions reached by the Russian palynologists were later shown to be erroneous.

On the other hand, the American line of inquiry, because it relied primarily on the study of petrographic thin sections, dealt with only a limited part of the available fossil record. Nevertheless, the American school came to take the lead in paleobiological studies. Skepticism outside the U.S.S.R. about the paleobiological interpretation of the Russian findings tended to obscure their importance for many years. Although the two lines of inquiry are still quite distinct, there is now general agreement about the basic outlines of the origin and evolution of early life.

A significant part of the Precambrian fossil record is derived from the study of stromatolites in which the original calcium carbonate has been replaced by silica. Stromatolites are formed by the accumulation of fine sediment, usually calcium carbonate, trapped by photosynthetic microbial communities. The carbonate, which is porous, does not preserve the structure of the organisms that make up the communities, whereas the silica sometimes does. The microfossils discovered by the American investigators and their European colleagues from silicified stromatolitic sediments represent for the most part shallow benthic communities consisting almost exclusively of nonmotile prokaryotic microorganisms, chiefly cyanobacteria (formerly known as blue-green algae). Living stromatolites, that is, microbial mats dominated by cyanobacteria that are in the process of lithifying, have been recognized in both saltwater and freshwater environments. These structures are invariably shallow benthic microbial communities.



FOSSILIZED REMAINS of microorganisms identified by the author as the earliest eukaryotes are magnified about 2,500 diameters in these photomicrographs. Both specimens, known as acritarchs, were recovered in the form of acid-resistant organic residues from their enclosing matrix of sedimentary rock. The roughly spherical microfossil at the top is thought to be a cyst, or comparatively sturdy outer membrane, formed by a unicellular plankton during a resting stage in its life cycle. Collected from the southern Ural Mountains by Tadas V. Jankauskas of the Lithuanian Institute of Geology, it is estimated to be between 1.6 and 1.4 billion years old. The object at bottom is a cluster of similar forms collected from the Chamberlain Shale Formation in Montana by Robert J. Horodyski of Tulane University. It is about 1.4 billion years old.

MOUND-SHAPED STRUCTURES called stromatolites consist of superposed layers of limestone accumulated by successive mats of photosynthetic microorganisms, chiefly cyanobacteria (formerly known as blue-green algae). In this case the calcium carbonate of the original sediments has been replaced in part by silica, which tends to preserve the microfossils better. The rock, which is from a dolomite formation in northern Norway, is 800 million years old.

THIN SECTION OF A SILICIFIED STROMATOLITE is viewed in transmitted light in this photomicrograph, revealing the fossilized sheaths of both filamentous and spherical cyanobacteria. Microorganisms of this type, which lack a distinct nucleus, are classified as prokaryotes. They are usually preserved as finely disseminated organic matter and can rarely be extracted from the siliceous rock matrix. The specimens, which are from the Gunflint Iron Formation in southern Ontario, are some two billion years old. The magnification is about 4,000 diameters.

Many fossil microorganisms from Precambrian stromatolites exhibit a remarkable morphological resemblance to living cyanobacteria. Stromatolitic microfossils have been found in Precambrian rocks of various ages, the oldest of which may be more than three billion years old. Specimens are known from every continent except Antarctica. Studies of fossil stromatolites and their living counterparts have greatly expanded paleobiologists' understanding of ancient benthic communities of microorganisms. Because these predominantly cyanobacterial communities have been remarkably stable for billions of years, however, they have provided little information on the paleobiology of organisms living in open waters. Nevertheless, as I shall undertake to show here, it was probably in open-water planktonic communities that most major evolutionary changes in the Precambrian took place. Most important, it was apparently there that the first eukaryotic cells appeared and rapidly diversified. The techniques of the Russian school and the interdisciplinary American approach have both been instrumental in elucidating the fossil history of the earliest plankton.

Whereas the study of benthic microfossils has been limited to thin sections of fine-textured siliceous rocks such as cherts, a wide variety of sedimentary rocks—sandstones, siltstones, shales and limestones, to name a few—are known to yield the fossilized remains of microorganisms with a comparatively sturdy organic wall, resembling the cyst, or capsulelike outer membrane, formed by certain extant microorganisms during a resting stage in their life cycle. By comparison with sediments deposited under present conditions, rocks that harbor these microfossils can be shown to have been formed in shallow to moderately deep marine basins overlying continental platforms. The microfossils recovered from these rocks exhibit a variety of shapes and structures. Classified as acritarchs, they are uniformly distributed in the matrix of the hardened sediments. Although acritarchs have occasionally been reported in thin sections of rocks, they are usually too sparsely distributed in the rock matrix for them to be studied under those conditions.

The organic walls of the acritarchs are resistant to most inorganic acids and bases. Hence they can be readily concentrated and studied in acid-resistant residues of the enclosing rock matrix. The planktonic origin of many acritarchous microfossils can be inferred from the types of sediment that enclose them. The fact that acritarchs are often found uniformly distributed in shale, sandstone, siltstone and carbonate suggests they had a free-floating dispersive way of life in open water. Their planktonic character is also indicated by the cosmopolitan distribution of

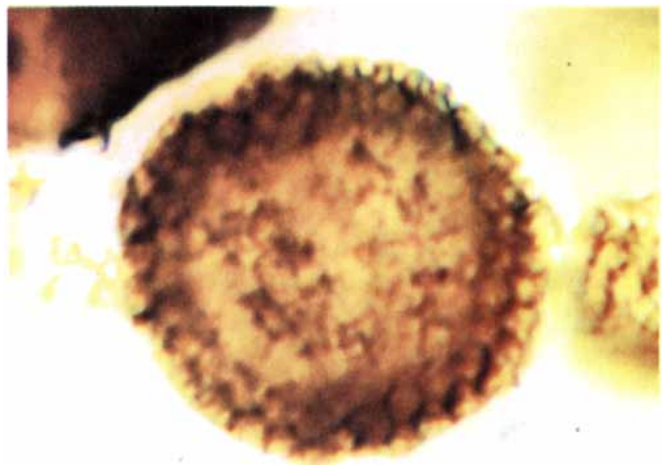
many forms in contemporaneous layers of sediment. In both their external appearance and their uniform distribution in the enclosing sediments the Proterozoic acritarchs resemble much younger microfossils that are unquestionably planktonic.

Most acritarchs were probably cyst-forming, unicellular, photosynthetic eukaryotes—algae of some kind—although in a few cases they appear to be characteristically prokaryotic in organization. For example, some acritarchs resemble the existing colonies of spherical cyanobacteria called chroococcaleans. These forms are typically found in large numbers in what are otherwise quite barren marine sediments laid down under glacial conditions during the late Proterozoic episode known as the Varangerian ice age, which took place about 650 million years ago. By analogy with the "blooms" of morphologically similar microorganisms observed today, one can conclude that these abundant microfossils represent the abrupt proliferation of certain cyanobacteria in periods of generally low planktonic productivity and diversity. Such conditions are known to be caused by raised levels of water-soluble organic and inorganic nutrients, chiefly nitrogen and phosphorus. The Varangerian assemblage is evidently related to a comparatively sudden influx of such nutrients, released as a result of widespread glaciation.

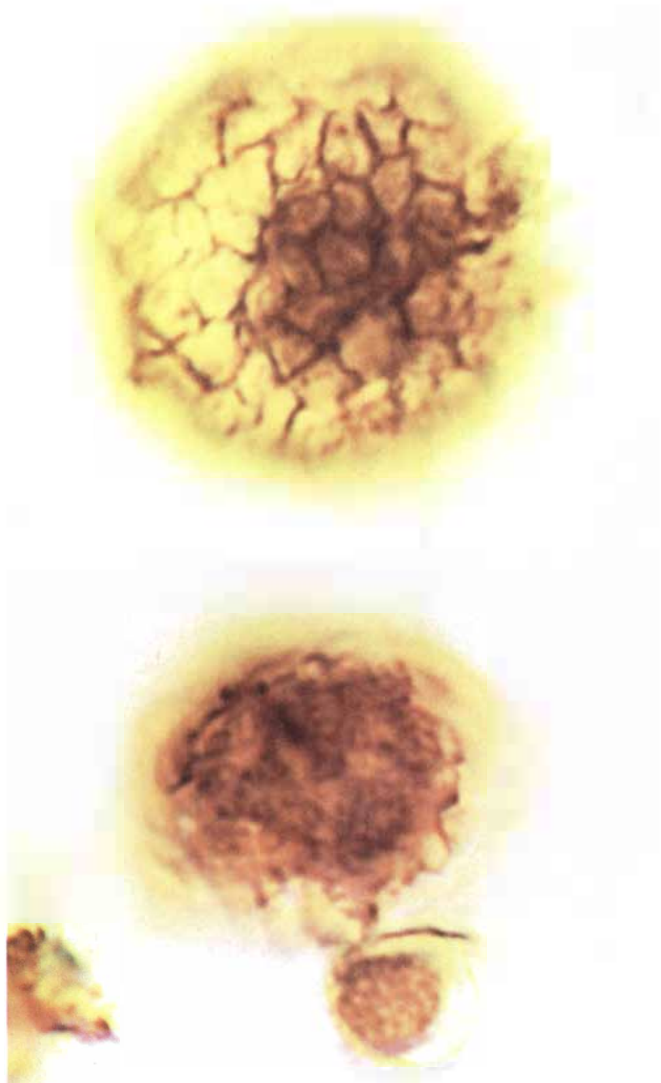
J. William Schopf of the University of California at Los Angeles has shown that Precambrian microfossils in general exhibit a striking tendency to increase in size as they decrease in geologic age. The evolution of eukaryotes entailed not only an increase in the amount of DNA per cell and the segregation of the bulk of the genetic material into a distinct nucleus but also the development of intracellular organelles such as mitochondria and chloroplasts. Under the circumstances one would expect the process to be accompanied by an increase in cell size. A significant increase in the size of fossil acritarchs some 1.4 billion years ago has been interpreted to mean that eukaryotic organization first emerged at about that time.

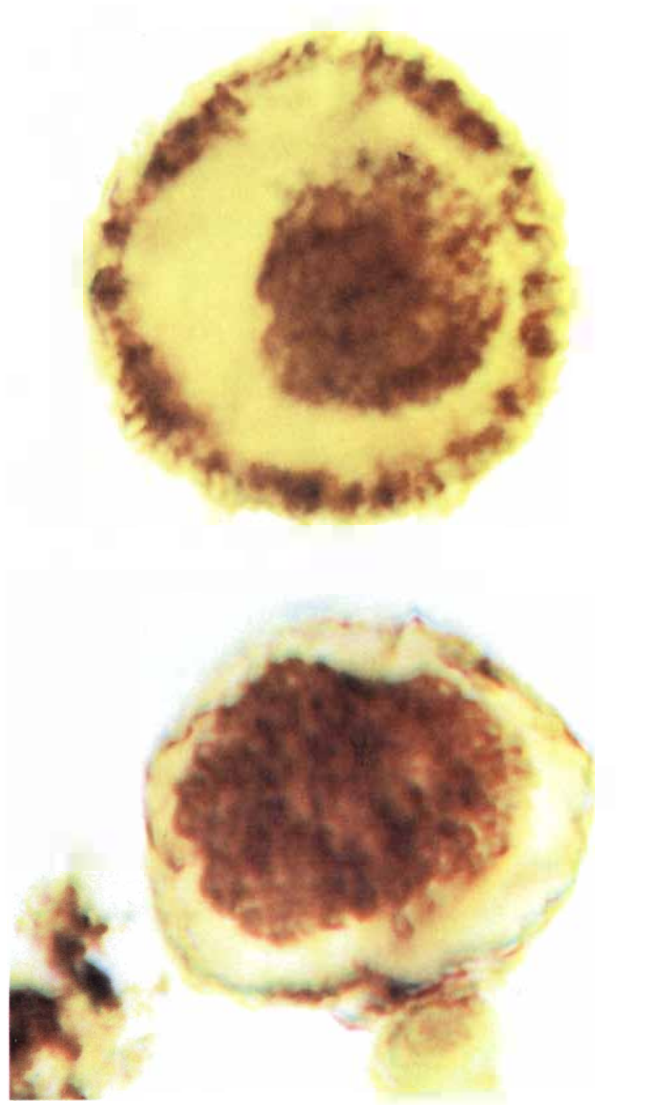
Much of the debate on the origin of eukaryotic cells has centered on the interpretation of certain spherical stromatolitic microfossils that have internal bodies reminiscent of nuclei or other cellular organelles. In addition some

EXPOSED SEDIMENTARY BEDS on the Varanger peninsula in northeastern Norway represent a more or less continuous stratigraphic record of the late Proterozoic era and early Cambrian period. The rocks contain the fossilized remains of both benthic (bottom-dwelling) and planktonic (free-floating) life forms, ranging from unicellular prokaryotes and eukaryotes to primitive multicellular animals.



FOSSILIZED CELL COLONY dating from the late Proterozoic era (about 700 million years ago) is viewed at two different depths of focus in this pair of photomicrographs. The microfossils, which were found in a shale formation in southern Norway, resemble the characteristic cell colonies formed by extant species of prokaryotic cyanobacteria. The cells are magnified approximately 1,500 diameters.



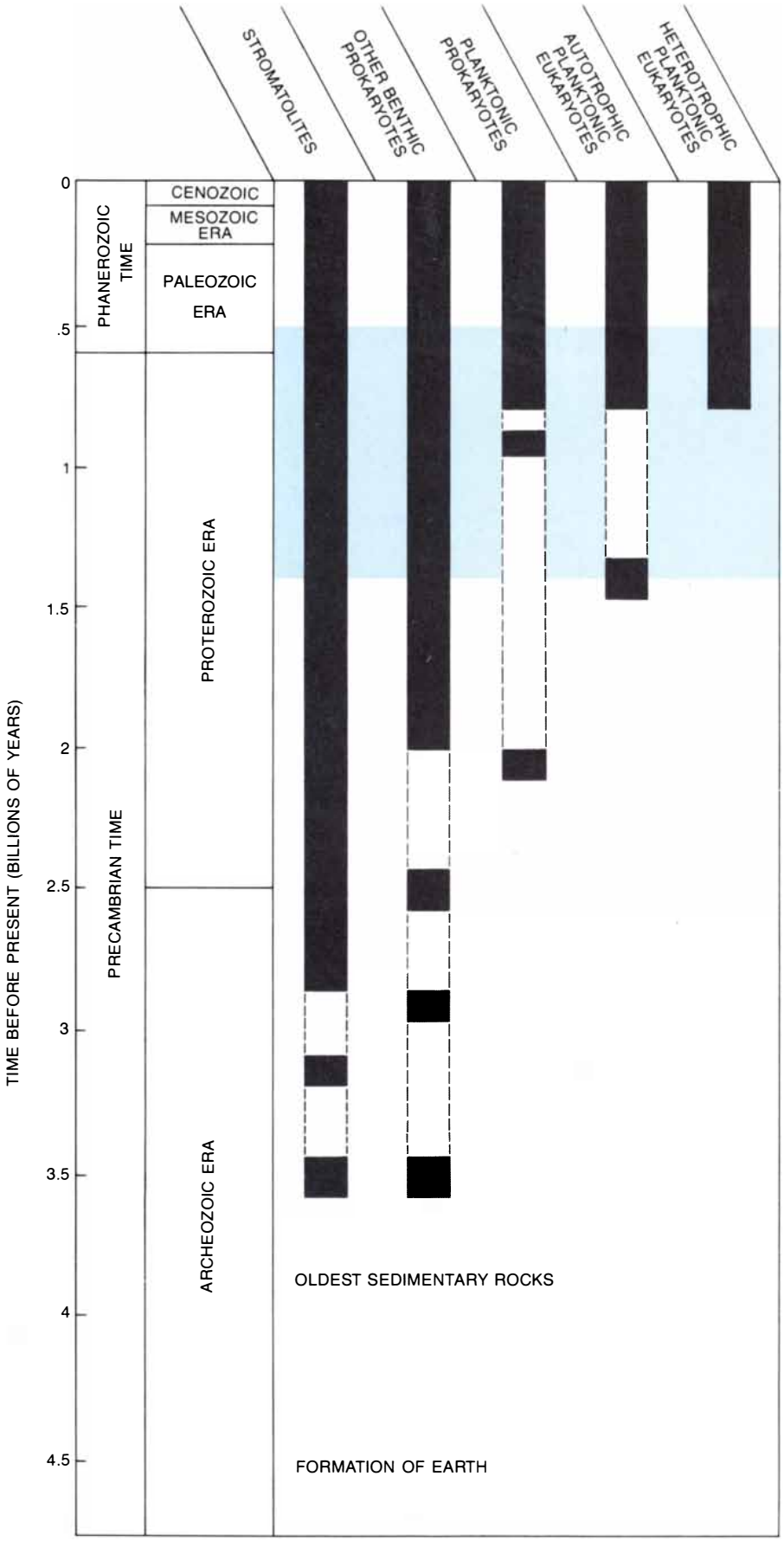


PRESENCE OF INTERNAL BODIES is not always a reliable indicator of an acritarch's cellular organization. Both microfossils shown at different depths of focus in these paired views, for example, show the coagulated remains of intracellular material. The acritarch at the top has been identified as the remnant of a eukaryotic microorganism, whereas the one at the bottom is probably a prokaryotic cyanobacterium. Both specimens are from the same late Proterozoic rock layer in southern Norway. Magnification is about 1,500 diameters.

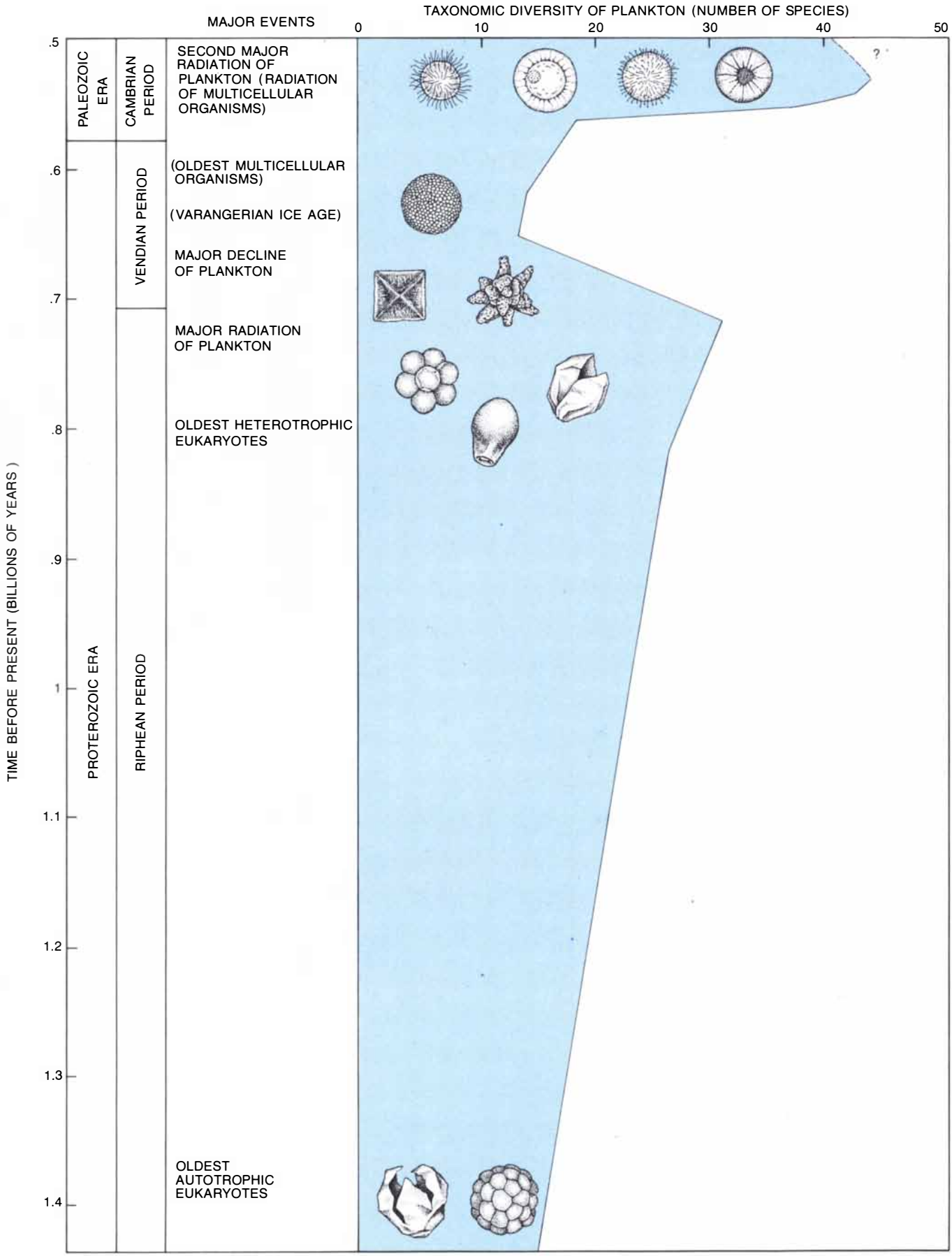
spheroids found in close tetrahedral groups have been viewed as fossil eukaryotes caught in the act of cell division. The interpretation of such spheroids as eukaryotes is now generally considered unconvincing. Most if not all of them appear to be the products of degradational processes or the fortuitous arrangement of the remains of spherical prokaryotes. In contrast, the record of planktonic acritarchous microfossils, even the oldest ones, is unambiguous: it clearly represents the initial evolutionary radiation of eukaryotic microorganisms. The paleobiological record of Precambrian and early Paleozoic plankton is particularly significant in that it is the most representative sample available of the global evolutionary status of life during those eras. This role is enhanced by the cosmopolitan distribution of most acritarchs over quite sharply delimited periods of time.

The earliest acritarchous microfossils, identified by Tadas V. Jankauskas of the Lithuanian Institute of Geology, come from rocks estimated to be early Riphean in age, that is, between 1.6 and 1.4 billion years old. These remains, collected from rocks in the southern Ural Mountains, resemble two kinds of younger fossil eukaryotes probably related to existing kinds of green algae. An assemblage of microfossils from the Chamberlain Shale Formation in Montana, estimated to be about 1.4 billion years old, includes not only filamentous cyanobacteria and thin-walled spheroidal sheaths of prokaryotes but also characteristically thick-walled spheroidal acritarchs. The chief investigator of this assemblage, Robert J. Horodyski of Tulane University, regards these acritarchs too as eukaryotic remains, again by analogy with younger forms.

Most acritarchs are fairly large (between 10 and 500 micrometers in diameter), thick-walled, distinctively shaped and ornamented with various external structures. The largest recognized acritarchs, measuring about three millimeters (or 3,000 micrometers) across, are clearly visible to the unaided eye. A number of them have fissures in their walls, which are thought to represent openings through which the planktonic organisms were released during the excystment stage, in which they shed their outer walls. Alternatively, certain circular holes have been interpreted as openings in the cell wall through which elongated cell processes such as undulipodia (eukaryotic flagella) might have protruded. In any event most of the late Proterozoic and early Paleozoic acritarchs appear to be cystlike structures associated with nonmotile stages of green algae. A few polygonal forms have been interpreted as colonies of green algae, as have other colonies consisting of thick-walled spherical



COMPARATIVE AGES of the earliest life forms are displayed in this version of the geologic time scale. The early part of the fossil record tends to be incomplete, as is indicated by the broken segments of the bars. Autotrophs are organisms such as photosynthetic plankton that thrive solely on inorganic nutrients. Heterotrophs are organisms that consume organic substances. The section of the chart marked with color is reproduced in greater detail on the next page.



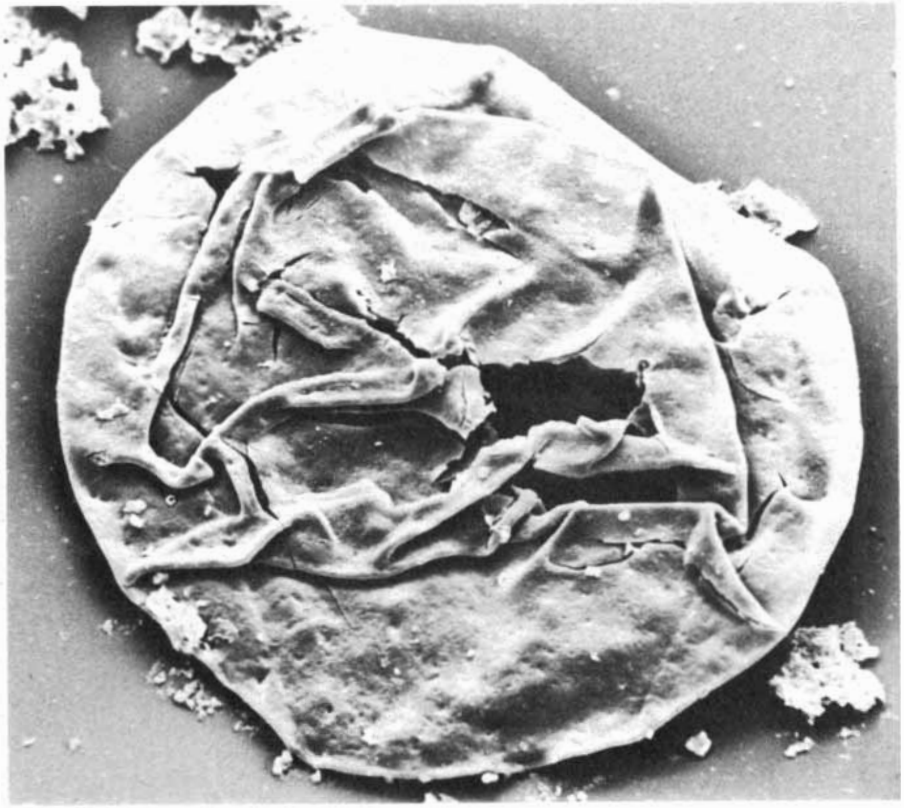
EXPANDED SECTION of the geologic time scale shows some major events in the evolution of life during the late Proterozoic era and the early Cambrian period. The colored area at the right reflects the fluctuation in taxonomic diversity observed among planktonic microfossils during this interval. The abrupt decrease in diversity observed about 650 million years ago coincides with the beginning of the Varangerian ice age. Some representative microfossils of the time are indicated by the drawings, which are not done to the same scale.

vesicles. In short, the evidence shows that a variety of planktonic eukaryotes have existed for at least the past 1.4 billion years.

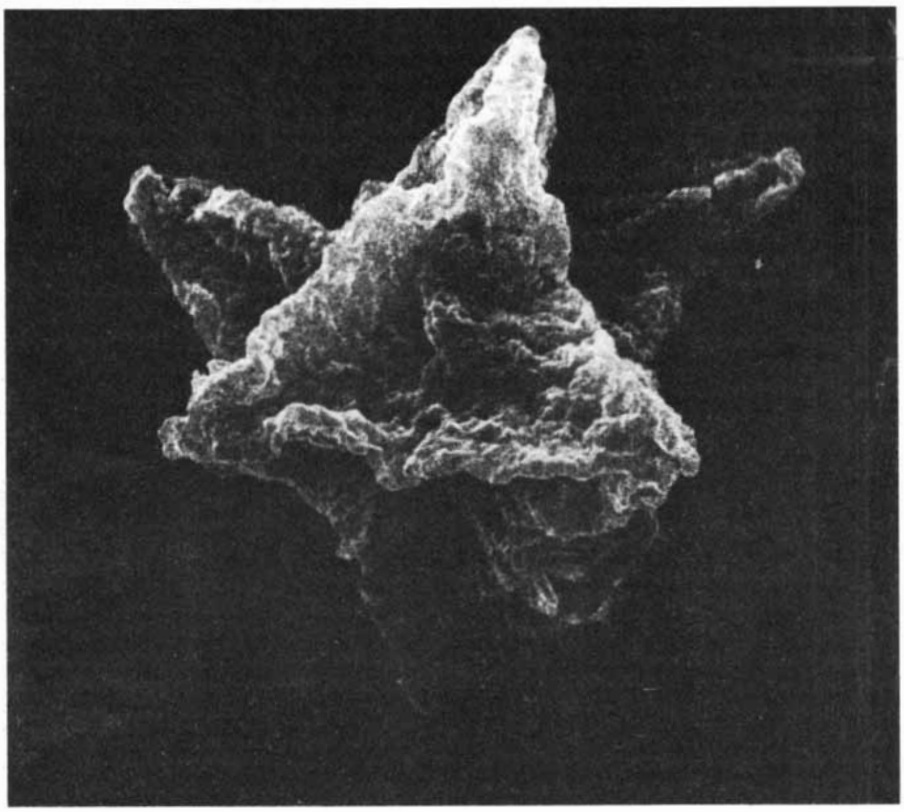
The planktonic mode of life may go back much further in geologic time. The oldest reasonably certain planktonic forms are rather large spherical unicellular microfossils from the cherts of the two-billion-year-old Gunflint Iron Formation. It seems likely that these microfossils are the remains of prokaryotes. The planktonic realm was probably first populated by photosynthetic prokaryotes. Gas vacuoles—internal cell structures found in many cyanobacteria, photosynthetic bacteria and other free-floating aquatic microorganisms—apparently served as primitive flotation devices. They represent an early evolutionary solution to the problem of how to move vertically in open water. In today's world prokaryotic microorganisms play a rather minor role in marine planktonic associations. They have presumably been outcompeted since the Proterozoic by eukaryotic plankton.

The preceding model of the evolution of early plankton is in good agreement with current ideas about the rise of the oxygen level in the earth's atmosphere over geologic time. Cells with a eukaryotic organization may have emerged by one of two evolutionary pathways: either as self-generated organisms or as symbiotic ones. (According to the latter view, put forward most prominently by Lynn Margulis of Boston University, the first eukaryotic cells were assembled from preexisting prokaryotic cells.) In either case the survival of the eukaryotes required the prior establishment of aerobic metabolism in an oxygenated environment. The photosynthetic, oxygen-generating cyanobacteria were responsible for the initial rise in the oxygen level of the atmosphere, but they were poorly adapted for survival at high oxygen levels. Some apparently increased their tolerance for oxygen by surrounding themselves with a thick protective layer of mucilage.

The first significant amounts of oxygen generated photosynthetically by anaerobic cyanobacteria must have accumulated in restricted shallow-water environments. Increasing levels of this bacterially produced oxygen probably forced the development of new oxygen-protective strategies, such as the colonization of anoxic environments, the production of oxygen-protective extracellular layers and the development of specialized cells such as nitrogen-fixing heterocysts. The irreversible oxygenation of the near-shore waters probably exerted considerable adaptive pressure toward the colonization of the first oxygenated offshore waters. Nitrogen fixation by means of thick-walled heterocysts is still accomplished by a number



CHARACTERISTIC FOLDS evident in this scanning electron micrograph of an acritarch recovered from an 800-million-year-old rock layer in eastern Greenland are the result of compaction in the original sedimentary bed. The microfossil, identified by the author as the abandoned cyst of a free-floating eukaryotic algae, measures roughly 300 micrometers in diameter.

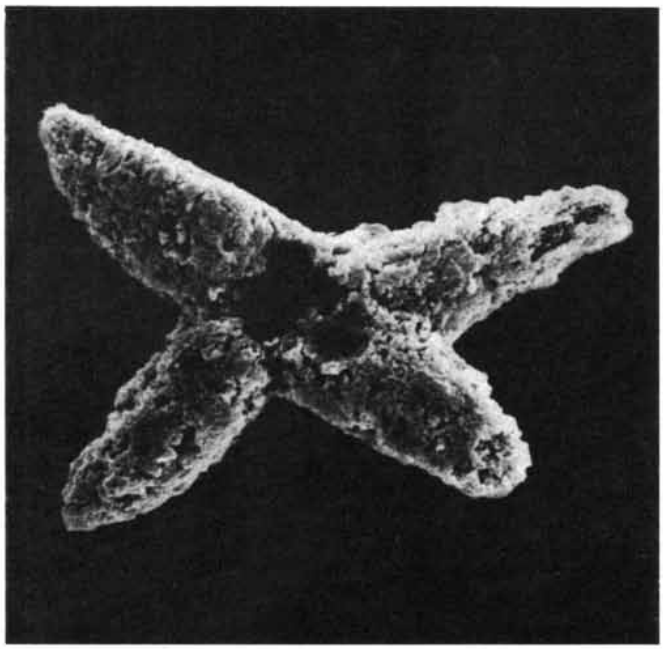


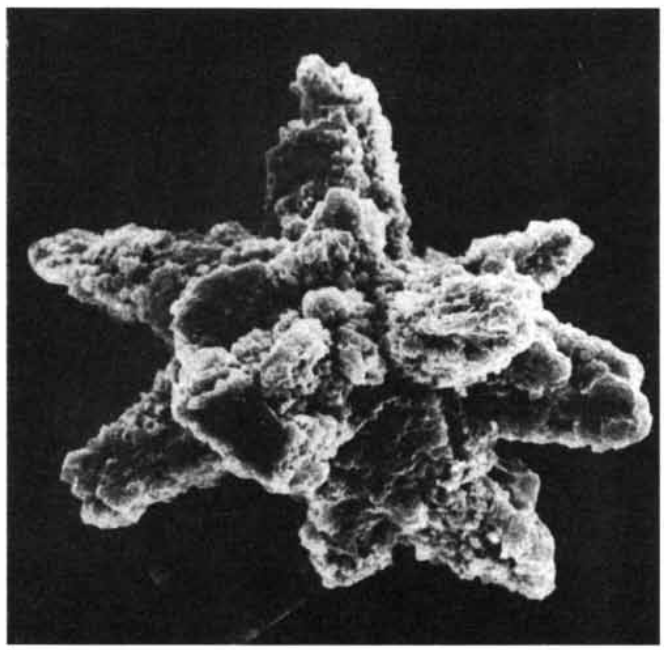
COMPLEX OCTAHEDRAL MORPHOLOGY distinguishes this late Proterozoic acritarch, recovered from a 700-million-year-old shale in southern Sweden. Comparable specimens have been found in rocks of the same age in Greenland, Norway and the U.S.S.R. Such microfossils are also common in much younger Paleozoic rocks. Magnification is about 1,900 diameters.

of planktonic cyanobacteria. Hence the planktonic mode of life is probably an exceedingly old but biologically expensive adaptation among such prokaryotic organisms.

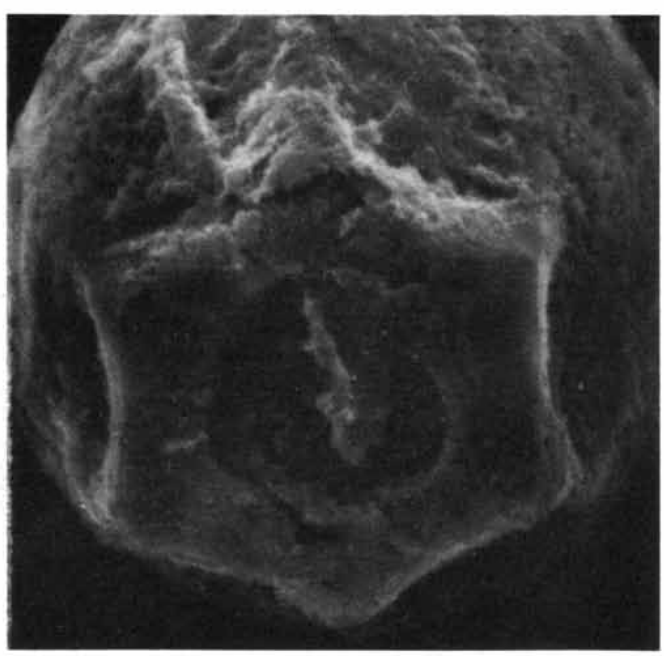
At present cyanobacteria have limited success in the planktonic realm. The evidence provided by the fossil record indicates that even in the early Proterozoic, before competition with eukaryotes, cyanobacteria formed a rather insignificant part of the record of planktonic life. Indeed, prokaryotes in general have probably never been well adapted to life in open water. Their tiny, dense, mucilaginous cells tend to make them sink. The cyanobacteria originated in an anoxic environment that was accumulating increasing levels of life-threatening oxygen. The eukaryotes, which not only tolerated a high oxygen level but also required it, were increasingly competitive in the newly oxygen-rich planktonic realm. By about 800 million years ago both autotrophic and heterotrophic forms (consumers respectively of inorganic and organic substances) had arrived in force.

Plankton are constantly being starved by inadequate supplies of mineral nutrients. They must have availed themselves first of erosionally derived nitrogen and phosphorus by colonizing





STAR-SHAPED ACRITARCHS were recovered from sedimentary beds in the Varanger peninsula estimated to be about 700 million years old. Similar microfossils are also known from contemporaneous beds elsewhere in Scandinavia and in the U.S.S.R. Measuring about 50 micrometers across, the elaborate structures are probably the remains of radially branching colonies of spherical green algae.



VASE-SHAPED MICROFOSSIL (*left*) is representative of an abundant type of plankton that appeared in shallow coastal waters roughly 800 million years ago. This particular specimen is from the Grand Canyon in Arizona. Similar microfossils have been found in Greenland, Brazil, Saudi Arabia, Sweden and Norway. The triangular opening at the base (*right*) is apparently plugged, suggesting that the cells, which were presumably heterotrophic eukaryotes, moved by means of undulipodia (the eukaryotic equivalent of bacterial flagella).

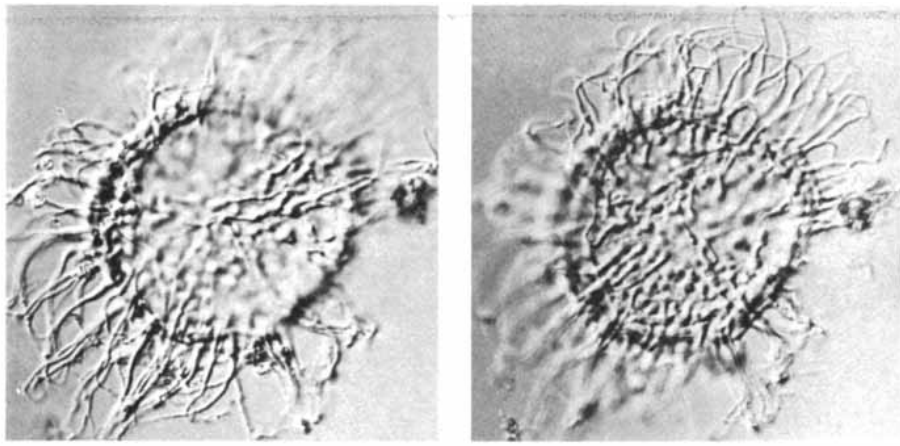
shallow marine environments. At a later stage evolutionary pressure may have acted in favor of the recovery of nutrients from deeper waters and from those more remote from continental shelves. A conspicuous rise in the diversity of large, spheroidal, cyst-forming acritarchous eukaryotes in late Riphean times—between 900 and 700 million years ago—can be seen in the fossil record. This trend coincides with a period in which the major preexisting continents broke up, forming shallow intracontinental basins. The formation of such basins through widespread geologic faulting is amply demonstrated in both North America and Europe.

The large marine basins resulting from this process accumulated thick deposits of sediments. The late Riphean geologic regime undoubtedly provided the plankton with new sources of mineral nutrients along the continental margins. Acritarchous microfossils from the period were buried in enormous numbers in sediments formed in shallow marine environments. Most of them are diversely ornamented, spheroidal forms of cyst-forming eukaryotes. Proterozoic turbidites—rocks that formed in deeper marine settings—have very few fossils of any type. Nevertheless, they do contain some acritarchs that seem to have adapted to a buoyant way of life.

The external morphology of late Proterozoic acritarchs (that is, those between 800 and 700 million years old) resembles that of some early Paleozoic forms. The similarity is observed among acritarchs in comparable environmental settings even if they belong to different taxonomic groups. Although they may be only remotely related, their similar means of dispersal and their simultaneous appearance in widely scattered sedimentary beds are the result of recurrent adaptive solutions to similar environmental problems.

By the late Proterozoic the basic adaptation in favor of the recovery of "easy" sources of water-soluble nutrients had been tried and the available innovations were probably exhausted. Additional adaptive pressure apparently led to the development of buoyancy-promoting features. The earliest solutions to the buoyancy problem are evident in microfossils from sediments estimated to be more than 800 million years old. Several kinds of plankton apparently developed blisterlike structures, probably filled with lighter liquids or gases, surrounding a thick spherical wall. The clustering of cells into motile colonies, with the addition of undulipodia, may have provided another solution to the buoyancy problem.

Abrupt departures from simple spherical symmetry, leading to polygonal walled structures, appeared in the late Proterozoic about 700 million years ago. The same adaptation resulted in microorganisms with a similar morphology but very different biology during later periods. Such adaptive similarities, called convergences, are comparable to later examples of parallel evolution, for example the formation of hydrodynamically shaped limbs in such diverse marine animals as dinosaurs, whales, seals and penguins.



NUMEROUS APPENDAGES covering the surface of an acritarch discovered in a 550-million-year-old rock in Sweden are indicative of the increasing variation in form observed in specimens dating from the early Cambrian period. The magnification is about 1,500 diameters.

About 800 million years ago a new class of planktonic microorganisms appeared in great abundance all over the globe. These are represented in the fossil record in the form of vase-shaped bodies with massive tests, or external coverings, which in some cases may have been mineralized, most likely with calcium carbonate or silica. Some had flangelike collars. These plankton are interpreted to be heterotrophic eukaryotes that inhabited coastal waters and fed on abundant algal plankton. Some of the vase-shaped cells have plugged openings at one end, suggesting that they moved by means of undulipodia. The evolution of even larger eukaryotes is reflected in the fossil record in immediately younger deposits, estimated to be slightly more than 600 million years old. The same rocks also yield the earliest scattered traces of soft-bodied animals.

Highly ornamented, greatly varied and presumably more efficiently buoyant microfossils appeared early in the Cambrian period, some 550 million years ago. This major event of planktonic diversification, giving rise to the first true spiny acritarchs, coincides with the first appearance of trilobites: the characteristic shelled arthropods of the Paleozoic era.

Estimates of the diversity of planktonic populations in Proterozoic and early Paleozoic times can only be approximate. Nevertheless, like the diversity estimates made for the Paleozoic era as a whole, they do represent fluctuations in morphology that can be interpreted to reflect some taxonomic variation. A major decline in planktonic diversity apparently occurred during the late Proterozoic Varangerian episode. The widespread eutrophication of waters in major basins resulting from the Varangerian glaciation probably led to an abundance of mineral nutrients and concurrent bacterial blooms. At the end of the Varangerian glaciation normal nutrient conditions were apparently restored, a development that is correlated with the rapid radiation of new acritarchous microfossils in the early Cambrian. This new radiation may be related to worldwide fluctuations in sea level that are known to have occurred at the time.

One of the main goals of the study of Precambrian microfossils is a biostratigraphic system of relative dating that allows worldwide correlations. Any progress tending toward the subdivision of Precambrian rock sequences into a number of biologically distinctive strata is of great significance. Impressive results in this direction, based on microfossils such as the ones described in this article, have already been achieved. Such micropaleontological research has led to the subdivision and correlation of late Proterozoic and early Cambrian rocks in widely separated areas. Because the late Proterozoic and early Cambrian acritarchs had a planktonic mode of life, they achieved the widespread geographic distribution required of such diagnostic fossils. Thus the Russian line of inquiry, emphasizing the use of these microfossils as biostratigraphic indicators, and the American line, concentrating on the paleobiology of the organisms themselves in the reconstruction of ancient ecosystems and their evolution, have finally come together. The late Proterozoic can be considered the golden age of unicellular life.

The Earth's Orbit and the Ice Ages

Periodic variations in the geometry of the earth's orbit have long been considered a possible cause of ice ages. The idea is now supported by a more reliable chronology of glaciations

by Curt Covey

Eighteen thousand years ago one-third of the land surface of the earth was covered with ice. In North America the ice was several kilometers thick in places and extended as far south as Oregon and New York. Antarctica and part of Eurasia and Greenland were also covered with ice. Glaciers in mountain ranges from the Alps to the Andes had expanded. So much water was locked up in the ice that the sea level was almost 100 meters lower than it is today; altogether about 5 percent of the earth's water was in the form of ice on land. As might be expected, the earth was significantly colder. The temperature at the earth's surface, averaged over the globe and over the seasons, was roughly five degrees Celsius lower than the average today. (For purposes of comparison, annual fluctuations in mean global temperature are only a few tenths of a degree C.)

The glaciation of 18,000 years ago was the most recent of about 10 ice ages in the past million years. Although many theories have been proposed to explain the ice ages, a growing body of evidence strongly supports one idea: that they are caused by small changes in the tilt of the earth's axis and in the geometry of the earth's orbit around the sun. The theory is called the Milankovitch theory, after Milutin Milankovitch, a Yugoslav astronomer who championed it in the first half of the 20th century. It asserts that orbital variations change the climate by altering the amount of solar energy the earth receives at different latitudes and seasons.

For some time the Milankovitch theory was considered untestable, largely because there was no sufficiently precise chronology of the ice ages with which the orbital variations could be matched. Recently, however, new types of geologic evidence have established a better chronology. A correspondence has been demonstrated between the history of the ice ages and the orbital variations. Furthermore, reasonable mechanisms have been proposed to explain how small variations in solar radiation could cause large changes in climate. These developments have renewed interest in the theory and led to its widespread acceptance.

The first persuasive evidence that large parts of the mid-latitude continents were once covered with ice was compiled by the Swiss-American geologist Louis Agassiz early in the 19th century. Agassiz found rocks that had been scratched and polished by the passage of ice, and he identified deposits of sand and rocks, including boulders too large to have been carried by water, as detritus left behind after the retreat of glaciers. With the realization that land forms and sediments in ice-free regions could have been produced by the action of ice rather than by the working of wind and water, the traces of glaciers were recognized throughout North America and Europe.

In mapping glacial deposits geologists soon found that there are several distinct layers; the earth has clearly undergone many glaciations in the past million years. Eventually the study of even older glacial deposits, particularly tillites (mixtures of pebbles, boulders and clay that were consolidated into rock), led geologists to the conclusion that there were at least two or three glacial episodes much earlier in the history of the earth. The first extensive glaciation may have been in the Precambrian, more than 500 million years ago.

It should be noted in passing that the term "ice age" reflects in its lack of precision the history of glacial geology. It can refer either to a single advance of the ice sheets during the period of relatively cold climate that began about a million years ago (the Pleistocene epoch) or to the entire Pleistocene and the earlier cold periods. Because the Milankovitch theory explains fluctuations in ice volume once the climate is relatively cool but not the onset of a colder climate and because the evidence of earlier glacial cycles is comparatively vague, "ice age" is generally used here to mean a time of maximum extent of the ice sheets during the Pleistocene.

Although it has long been clear from the geologic record that there were many Pleistocene ice ages, the exact number, date and duration of the glaciations have proved difficult to pin down. It is a recent discovery that has finally made it possible to establish a better chronology. The global volume of ice at any given time can be deduced from the ratio of two isotopes of oxygen in ocean sediments deposited at the time.

The most abundant isotope of oxygen is the one with an atomic mass of 16 (in which each atomic nucleus is made up of eight protons and eight neutrons). Almost all the oxygen in water is oxygen 16, but a few molecules out of every thousand incorporate the heavier isotope oxygen 18 (in which each nucleus has two extra neutrons). The heavier molecules tend to be left behind when water evaporates from the ocean surface. As a result rain and snow and the ice formed from precipitation over land have less oxygen 18 than ocean water does. When an ice age begins, the continental ice sheets grow, removing water from the ocean. The ocean water left behind thereby becomes enriched in the heavier oxygen 18.

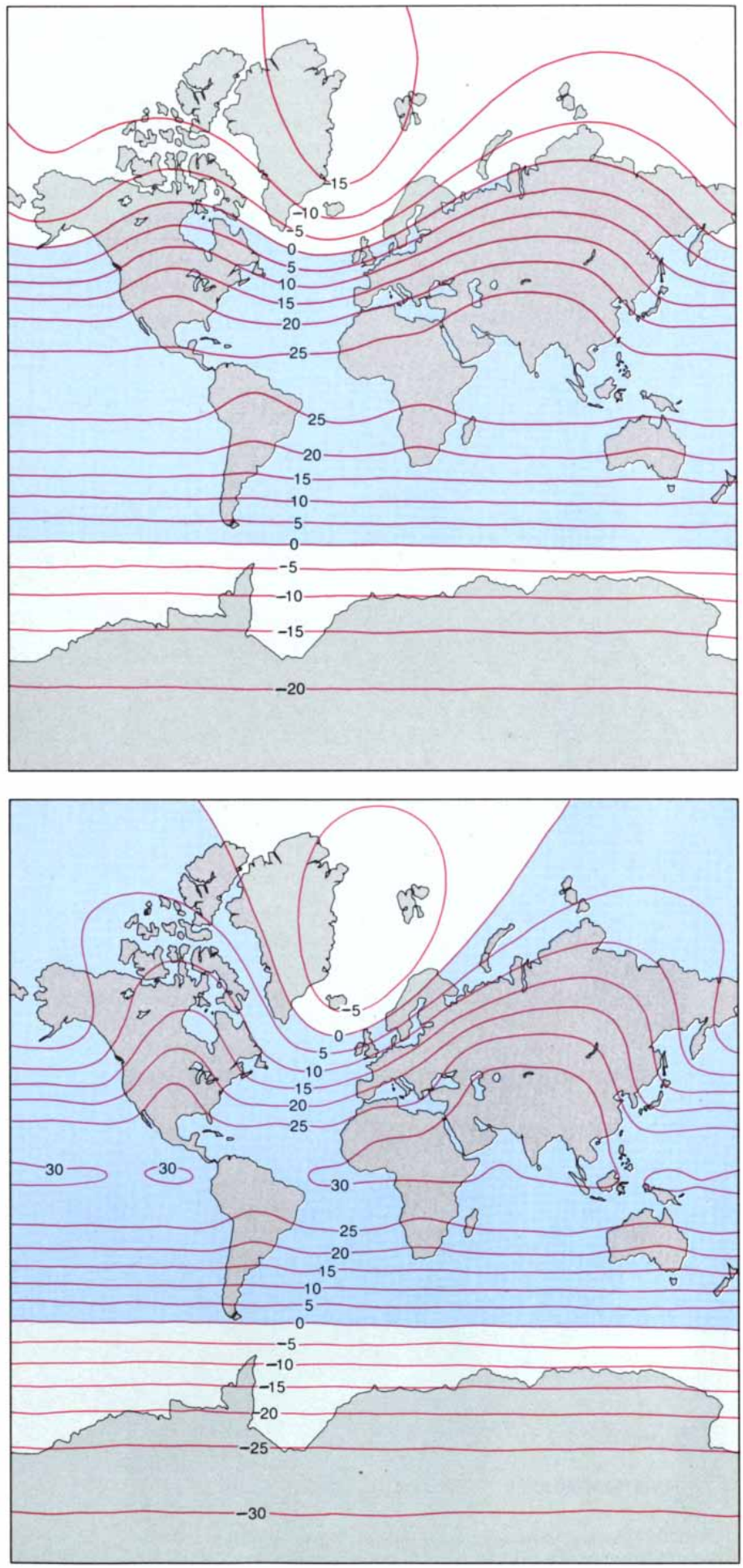
Of course, specimens of seawater from a remote era cannot now be recovered or identified. The isotope ratio can be determined, however, from the calcium carbonate ($CaCO_3$) shells of marine organisms, which are constructed with the oxygen atoms in seawater and reflect the water's isotopic composition. The shells of the organisms fall to the ocean floor and accumulate to form sediments. The higher the ratio of oxygen 18 to oxygen 16 in a sedimentary specimen is, the more land ice there was when the sediment was laid down.

The complexity of the sequence of events the isotopes record limits the conclusions that can be drawn from the sedimentary evidence. The degree to which the evaporation of water and the deposition of snow sorts the isotopes is dependent on many poorly understood factors. For example, Stephen H.

Schneider of the National Center for Atmospheric Research and I have found that the degree of isotopic separation depends in part on the vertical distance the water vapor travels before it condenses as rain or snow. How the circulation of the atmosphere during an ice age differed from that of today is not known, however, and so the altitude factor adds some uncertainty to the interpretation of the isotopic data. In addition the isotopic composition of carbonate shells is known to depend on the temperature of the water in which the shells formed. More oxygen 18 is incorporated into the shells when the oceans are cooler. The temperature effect is thought to account for about a third of the variation in oxygen-isotope ratios.

Because of effects such as these, the isotope record cannot reveal precisely how much land ice there was at a given time in the past. Nevertheless, it is agreed that a shift to a higher oxygen-18 level indicates a shift to a colder and more glacial climate, and the record does provide the best way of determining the sequence of ice-sheet advances and retreats over the past several hundred thousand years. For this purpose the isotope record has two advantages over the more traditional geologic data. First, it is a global record: the ratios in a given core of sediments extracted from the sea floor are thought to reflect the total amount of ice on land the world over. There is remarkably little variation among cores taken from different locations. Second, it is a fairly continuous record: the sequence of ocean sediments has far fewer gaps than the record of rocks on land.

The shifts in climate recorded in ocean sediments can be dated with sufficient accuracy to test the Milankovitch theory by radiometric methods. The methods also call for a measurement of isotopic ratios, but the isotopes—unlike oxygen 16 and oxygen 18—are radioactive ones whose steady decay marks the passage of time since the sediments were formed. By dating a few points in each core one can construct a time line on



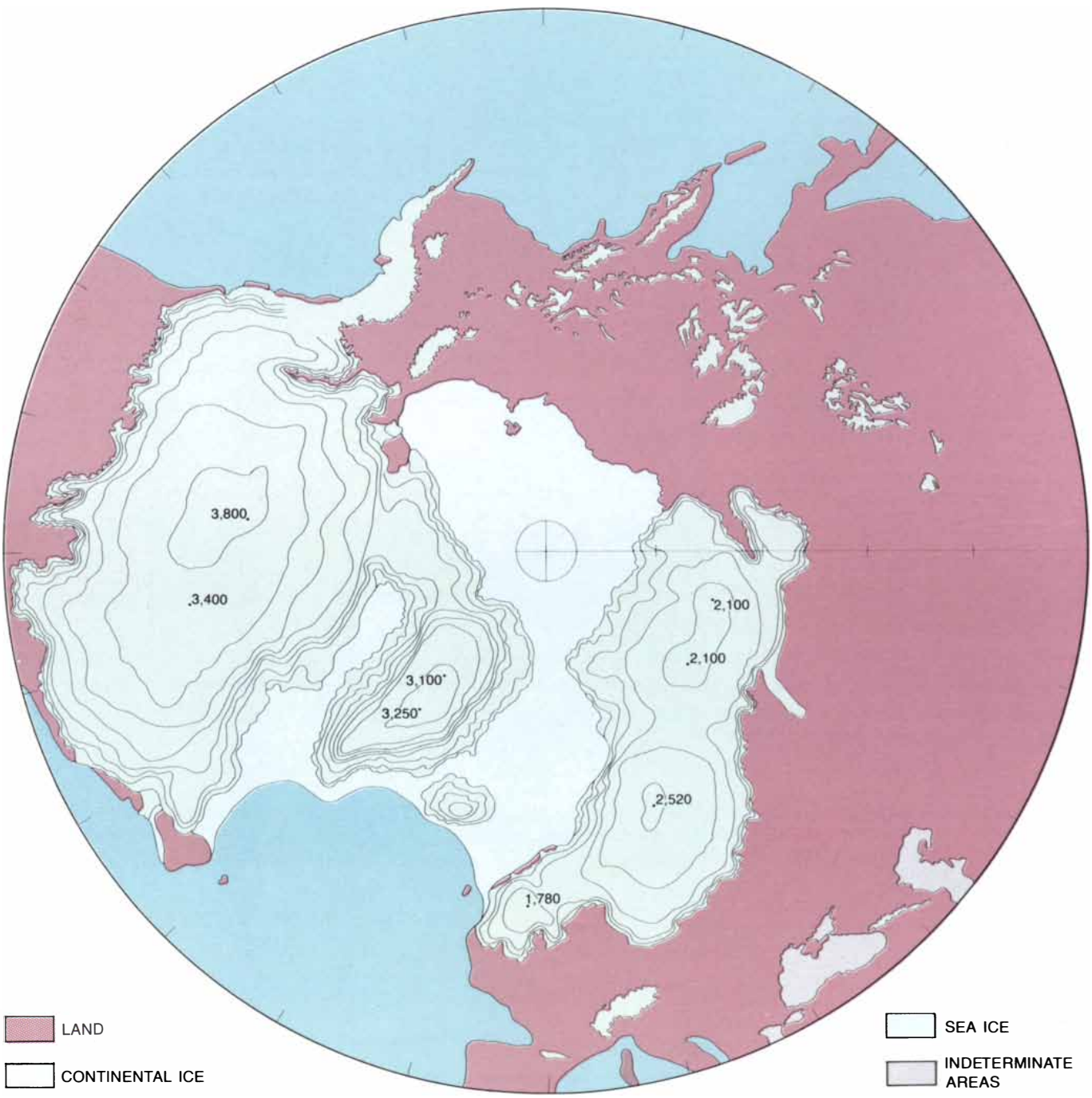
SIMULATION OF CLIMATE suggests that small, cyclic variations in the earth's orbit may be sufficient to bring on ice ages. The changes in the orbit alter the amount of sunlight reaching various parts of the earth's surface. The most important factor is the insolation in summer at high northern latitudes; if the summer temperature remains below zero degrees Celsius, ice can accumulate from year to year. The maps show the temperatures in July predicted by a climate model developed by Gerald R. North, David A. Short and John G. Mengel. In the configuration of the orbit 115,000 years ago (*upper map*) the model predicts ice-age conditions for much of North America and for part of Eurasia. With the current parameters of the earth's orbit (*lower map*) the July temperature is below the freezing point only in Greenland and Antarctica.

which the oxygen-isotope variations can be plotted.

The dated isotope record constructed in this way shows that the global ice volume has increased and decreased repeatedly in the past several hundred thousand years. The fluctuations have a pattern: roughly speaking, an ice age comes once every 100,000 years. The pattern is not simple, however. The peaks in the curve have an asymmetric, sawtooth shape, indicating that the ice takes much longer to build up than it does to disappear. In addition many smaller fluctuations are superimposed on the dominant 100,000-year cycle.

The pattern suggests that the climate has responded to some continual "forcing" oscillation. The most obvious candidate for a forcing factor would be a variation in the amount of sunlight received by the earth, and the most obvious cause of such a variation would be oscillations in the energy output of the sun. Satellite-mounted instruments have only just begun to establish an accurate record of the energy flux the earth receives from the sun, however, and as yet there is no firm evidence of suitable long-term variations in solar output. The Milankovitch theory proposes a mechanism that is much more accessible to observation and calculation. Even if the output of the sun is constant, global insolation varies with the distance between the earth and the sun, and the insolation at any given latitude varies with the orientation of the earth with respect to the sun.

A familiar example of changing insolation and its effect on climate is the cycle of the seasons. The earth undergoes seasonal changes mainly because the axis of its rotation is tilted with respect to a line perpendicular to the plane of the earth's orbit around the sun. During June, July and August the Northern



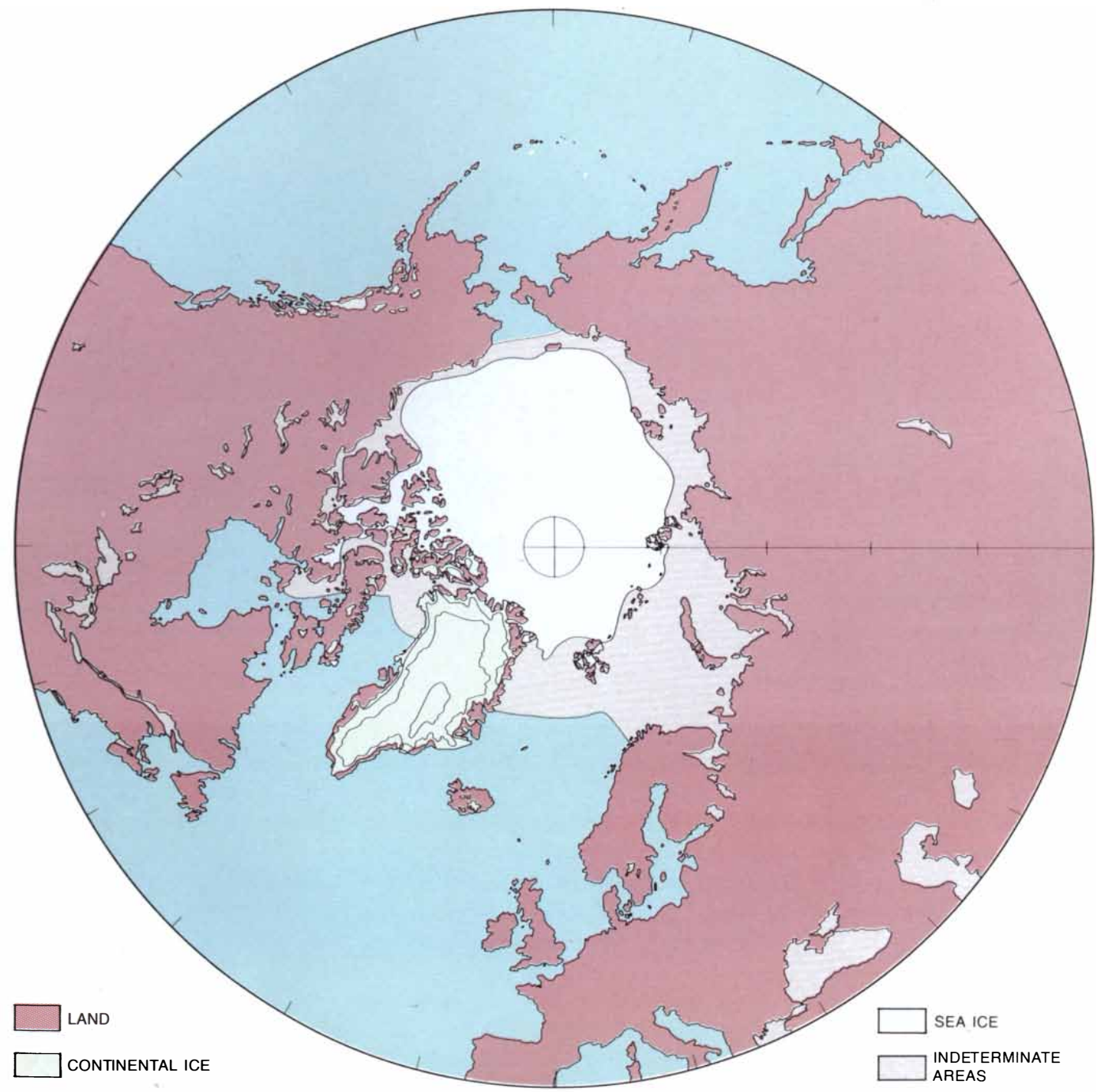
ICE SHEETS several kilometers thick blanketed the Northern Hemisphere 18,000 years ago, at the height of the most recent ice age (*left*). So much water had been transferred from the oceans to the ice that Siberia and North America were connected by a land bridge. Today

Hemisphere points toward the sun and receives more sunlight than it would if the axis were perpendicular to the plane of the orbit; the Southern Hemisphere receives less sunlight. Six months later, when the earth has moved 180 degrees along its orbit, the situation is reversed.

An additional factor modifies the seasonal cycle slightly. The earth's orbit is an ellipse rather than a circle, and so the distance from the earth to the sun varies with the time of year. In January the earth is closest to the sun (perihelion) and in July it is farthest from it (aphelion). Accordingly the global insolation is slightly greater in January than it is in July. This effect gives the Northern Hemisphere warmer winters and cooler summers than the hemisphere would have if the earth's orbit were circular.

Three quantities are needed to specify the earth's orbit for the purpose of discussing insolation. First, there is the obliquity, or angle of tilt, of the earth's axis with respect to the plane of its orbit; the obliquity today is 23.5 degrees. A second quantity, the orbital eccentricity, measures the departure of the orbit from a perfect circle. A circle has zero eccentricity, and an ellipse twice as long as it is wide has an eccentricity of .866. At present the eccentricity of the earth's orbit is .017. Third, as the earth spins, the direction in which its axis points slowly precesses, describing a circle with respect to the fixed background of distant stars. The axis currently points toward Polaris, the polestar.

The earth's orbital parameters change with time as a result of the small gravitational pull of the moon and of the other planets. The nearly periodic changes during the past several hundred thousand years can be calculated with great accuracy. The tilt of the earth's axis varies between 22.1 and 24.5 degrees with a period of about 40,000 years. The eccentricity of the orbit varies be-

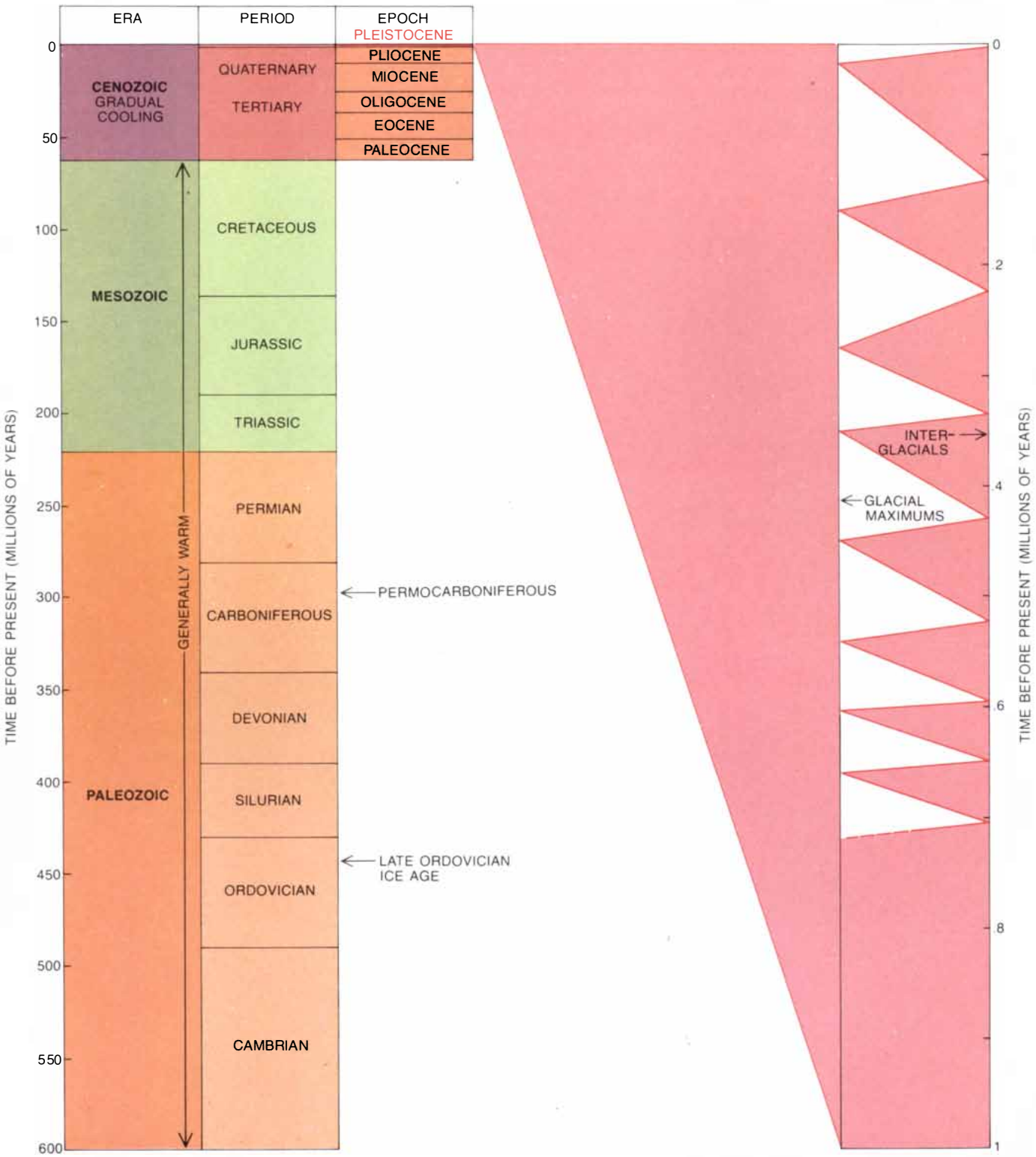


(*right*) only Greenland and Antarctica are covered with ice. The maps are based on ones prepared by members of the CLIMAP project. The distribution of the ice during the last ice age is based on the distribution of glacial land forms and sediments and on fossil evidence.

tween about .005 and .06 with a period of about 100,000 years. It takes about 26,000 years for the earth's axis to describe a complete circle as it precesses.

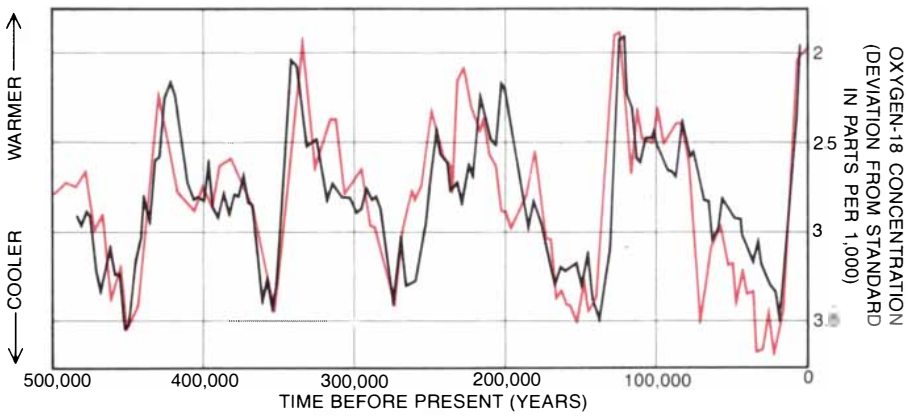
The variation in the earth's orbital parameters leads to a corresponding variation in the amount of sunlight received at a given latitude and time of year. The idea that these slow changes in the incoming solar energy are responsible for the advance and retreat of glacial ice was proposed almost as soon as the ice ages themselves were discovered in the 19th century. Most of the credit for the orbital hypothesis, however, belongs to Milankovitch. In the 1920's and 1930's he laboriously calculated the change in insolation that results from orbital variation and suggested ways in which it could affect climate.

Changes in the earth's orbital parame-



TIME LINE summarizes the earth's climate during the past 570 million years of its 4.6-billion-year history. In the distant past the climate was generally warmer than it is today, although there is some evidence that there were ice sheets during parts of the Precambrian era and the Ordovician and Permocarboniferous periods. The climate began to cool at the end of the Cretaceous, some 65 million years ago; the first permanent ice sheets probably appeared in the Oligocene epoch of the Tertiary. The most recent cycle of ice ages, the Pleistocene ice ages, began a million years ago. There have been about 10 Pleistocene ice ages; the earth is currently in an interglacial period.

ters lead primarily to changes in the intensity of the seasonal cycle. Increasing the tilt of the earth's axis increases the amplitude of the cycle: winters are colder and summers are hotter in both hemispheres. The effect of precession on insolation is less straightforward but together with the rotation of the ellipse of the earth's orbit around the sun it determines the time of perihelion. The variation in the time of perihelion has a period of approximately 20,000 years, somewhat shorter than the period of precession itself. In general a change in the time of perihelion increases the intensity of the seasons in one hemisphere and decreases the intensity in the other. Perihelion is now reached on January 3, and so the Northern Hemisphere winter is slightly warmer than it might be and the Southern Hemisphere summer is cooler. In about 12,000 years, when the earth's axis is pointed toward Vega rather than Polaris, the situation will be reversed.



RATIO OF TWO ISOTOPES OF OXYGEN in ocean sediments is a measure of past global ice volume. Fewer of the water molecules in glacial ice than in ocean water include the heavy oxygen isotope oxygen 18. Thus when continental ice sheets grow, ocean water and the shells of the marine organisms in ocean sediments are enriched in oxygen 18. The higher the oxygen-18 concentration in a sedimentary layer is, the more land ice there was when the layer was deposited and the colder the climate was. Two sets of isotope data are shown; each set is a composite of measurements from several cores of sedimentary material extracted from the ocean floor. One set was compiled by James D. Hays of the Lamont-Doherty Geological Observatory of Columbia University (*black*) and the other by Cesare Emiliani of the University of Miami (*color*). The curves are remarkably similar even though the cores were from widely separated places; the sedimentary isotope ratios are therefore thought to reflect the global volume of ice. Note that there is less ice today than there has been at any time in the past 120,000 years.

The eccentricity of the earth's orbit has two distinct effects on insolation. The greater the eccentricity of the orbit, the greater the difference between the maximum and the minimum distance of the earth from the sun and hence the greater the amplitude of the precession effect. The total annual insolation also varies with the eccentricity of the orbit. This effect, however, is a small one; the global insolation has changed by at most .3 percent over the past million years because of the change in eccentricity. According to most calculations, a change of .3 percent in global insolation would change the average temperature by only a few tenths of a degree.

In sum, eccentricity gives rise to only a small change in annual global insolation, and tilt and precession alter only the distribution of insolation with latitude and season and not the total insolation. How, then, could orbital variations be responsible for the major changes in climate that accompany the advance and retreat of ice sheets? Milankovitch argued that the force driving the ice ages is not the total global insolation but rather the insolation at high northern latitudes in the summer, which varies by as much as 20 percent, much more than the total insolation does. If this argument is accepted, it becomes much more reasonable to suppose orbital variations cause ice ages.

Summer insolation is more important than winter insolation because snow accumulates at high latitudes even during a mild winter. What matters is not how much snow accumulates in the winter but how much is preserved through the summer. If less sunlight than normal is received at high latitudes in the summer, the snow will tend to be preserved. If the situation persists over many years, ice is formed as the snow is compressed under its own weight, continental ice sheets begin to form and the earth enters an ice age. If the amount of sunlight received during the summer increases, on the other hand, more ice melts during the summer than is replaced by the snow in the winter and the earth returns to a climate similar to the present one.

Northern Hemisphere insolation is more important than Southern Hemisphere insolation because glaciers can form only on land and there are currently large land masses at high northern latitudes. In the Southern Hemisphere the only continent at high latitudes is Antarctica, which is covered with ice even today. A decrease in Southern Hemisphere insolation would accordingly be less effective in pulling the climate toward an ice age.

The orbital theory of the ice ages has long been subject to skepticism for two reasons. First, it was not clear whether the theory is consistent with the geologic record. Second, it seemed implausible that the small changes in insolation could be responsible for the large changes in climate. Only within the past 10 years has enough evidence accumulated to convince the majority of earth scientists that there is indeed a connection between the ice ages and the orbit.

The evidence includes periodicities in the lake-bed sediments called varves, whose deposition is controlled by the seasons. Periodicities have also been observed in ocean sediments made up of species of marine organisms that prefer warmer or cooler water. The decisive evidence, however, has come from the examination of isotope ratios in ocean sediments. In 1976 James D. Hays of the Lamont-Doherty Geological Observatory of Columbia University, John Imbrie of Brown University and Nicholas J. Shackleton of the University of Cambridge published a study of the isotopic record. Hays, Imbrie and Shackleton wanted to establish how much of the fluctuation in the global ice volume could be attributed to the 100,000-year cycle of orbital eccentricity, how much to the 40,000-year cycle of axial tilt and how much to the 20,000-year cycle of precession. They approached the question by subjecting the ice-volume curve to Fourier analysis, the technique commonly employed to detect cyclic variations and to define the period or periods of oscillation.

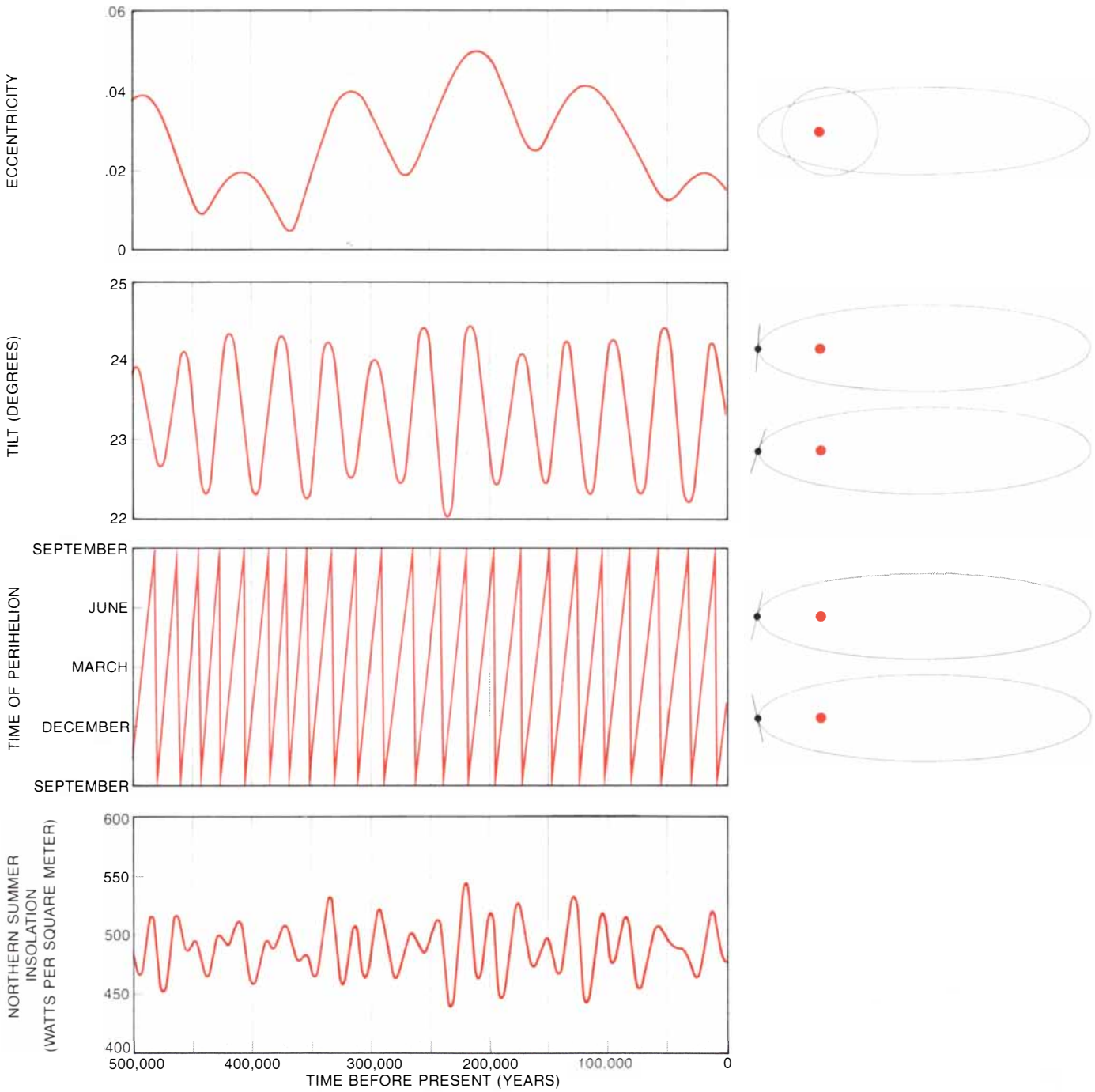
The French mathematician Jean Baptiste Joseph Fourier proved in 1807 that the graph of almost any function can be approximated to any degree of accuracy by superposing sine and cosine curves of various phases, amplitudes and frequencies; this is the process of Fourier synthesis. In the converse process, which is generally done with the aid of a computer, a complex curve is decomposed into elementary sine and cosine waves; the amplitude and phase of each single-frequency component are chosen so that the sum of the waves best approximates the measured curve. The result of a Fourier analysis is usually displayed as a spectral plot, which gives the total amplitude as a function of frequency or period. If there is a sharp peak in the spectrum at, say, a period of 100,000 years, much of the information can be explained by a 100,000-year cycle.

Spectral analysis brings into relief a number of distinct frequencies in the ice-volume curve. The most prominent

frequency is indeed a 100,000-year cycle, but there are lesser peaks at periods of about 40,000 and 20,000 years. Altogether, about 60 percent of the fluctuation in the curve can be expressed as the sum of functions with the same periods as the orbital parameters. This is powerful evidence in favor of the Milankovitch theory.

Although the geologic evidence is compelling, it is still only circumstantial; it does not identify the mechanism by which a change in insolation causes a change in ice volume. Plausible mechanisms are being tested with mathematical models of the climate.

There are a great variety of climate models, ranging in complexity from a single equation for globally and annually averaged temperatures that can be solved with a calculator to general cir-



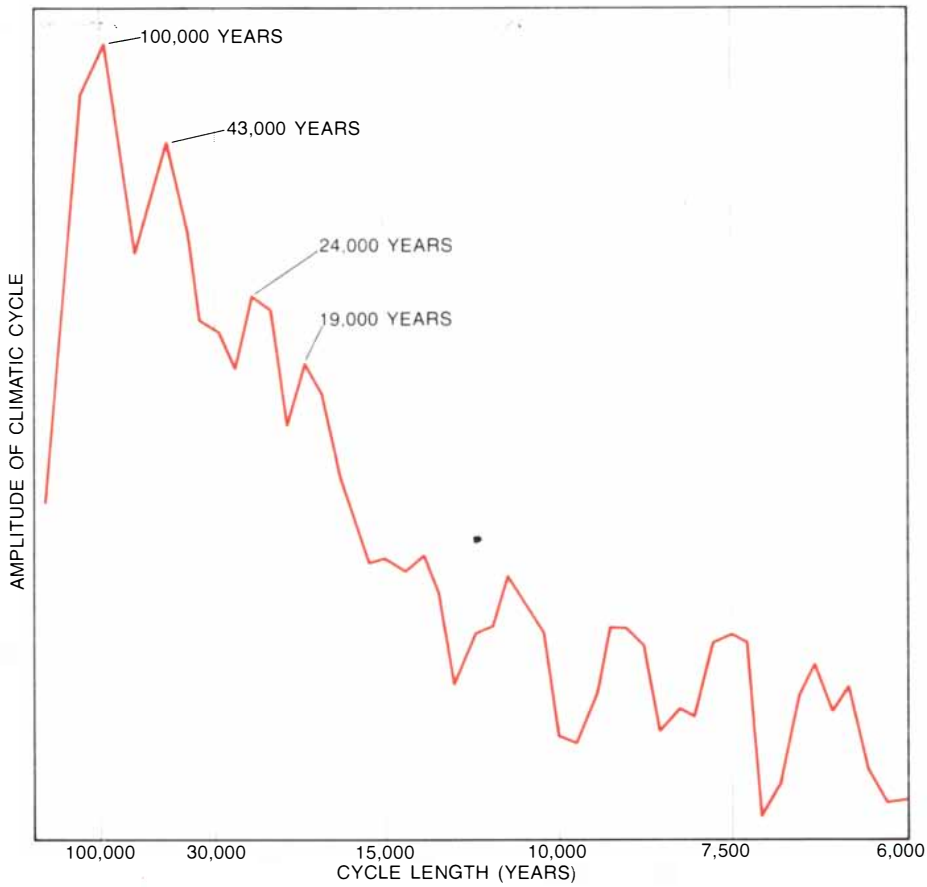
MILANKOVITCH THEORY, formulated by the Yugoslav astronomer Milutin Milankovitch, attributes the onset of ice ages to variations in three parameters of the earth's orbit. The eccentricity is the degree to which the orbit departs from a perfect circle. The tilt angle is the angle between the earth's axis and a direction perpendicular to the plane of its orbit. The time of perihelion determines the direction in which the axis points when the earth makes its closest approach to the sun. Each of the parameters changes slowly under the influence of the gravitational attraction of the moon and of the other planets. Changes in the orbital parameters lead to changes in the amount of sunlight received by the earth; there is little change in the annual global insolation but there are relatively large changes in high-latitude summertime insolation. The variations in eccentricity, tilt and the time of perihelion were calculated by a method developed by André Berger. Times of maximum eccentricity are separated by periods of roughly 100,000 years, although the eccentricity has a second period of roughly 400,000 years. The cycles of tilt and time of perihelion have periods of approximately 40,000 and 20,000 years. The time of perihelion is also not a simple sine curve; it has two cycles of 19,000 and 23,000 years. Variation in the amount of sunlight received is shown for the zone between 60 and 70 degrees north latitude in July. Insolation was calculated with a computer program written by Tamara S. Ledley of the Massachusetts Institute of Technology and Stanley L. Thompson of the National Center for Atmospheric Research.

culation models that offer a detailed representation of temperatures, winds and other meteorological variables and their dependence on time, latitude, longitude and altitude. Most climate models require extended numerical calculations; with general circulation models it takes the fastest computers hours to simulate the development of climate over the course of a year. Clearly the evolution of climate over the past hundreds of thousands of years can be studied only with simplified models.

The kind of climate model that is most suitable for testing theories of the ice ages is an energy-balance model, one based ultimately on the law of conservation of energy. In the model physical conditions are defined only at discrete points on the earth's surface, generally in a gridlike pattern. The rate of change of the temperature at each point is determined by repeatedly subtracting the outgoing infrared energy from the incoming solar energy and adding to the difference any energy transferred from adjacent points. The simplification that makes it possible to solve energy-balance models quickly enough to simulate glacial cycles lies in the treatment of energy transport between points on the grid. The transport is assumed to follow a simple law such as "Heat flows from warmer points to cooler points at a rate proportional to the temperature difference." Actually the winds and currents that transport heat follow complex and often unpredictable patterns. Nevertheless, energy-balance models successfully simulate variations in temperature with latitude and season; they may therefore be capable of simulating changes in climate that have the same cause as the seasonal cycle.

Recently experiments with energy-balance models have shown both that orbital forcing could cause large changes in climate and that it could give rise to a 100,000-year climate cycle. A model developed by Gerald R. North and David A. Short of the Goddard Space Flight Center of the National Aeronautics and Space Administration and John G. Mengel of the Applied Research Corporation demonstrates that small decreases in insolation can indeed lead to extensive glaciation. The key, as Milankovitch suggested, is the insolation at high northern latitudes.

North and his co-workers wondered if climate models were failing to predict correct changes in ice volume because the continents were typically represented as one large land mass. They found that a model with a more realistic distribution of land was more sensitive to changes in insolation. Moreover, it predicted distributions of ice like those seen in the geologic record. For example, when the insolation of 115,000 years ago, at the beginning of an ice age, was put into the model, it predicted that ice sheets would grow at northern latitudes, particularly in North America. Less ice was predicted to accumulate in Eurasia because it is larger. The interior of Eurasia, far from the ocean's moderating influence, undergoes greater seasonal temperature extremes, and the relatively warm summers melt the ice that is built up in the winter.



SPECTRAL PLOT shows the amplitude and frequency of the sine waves whose sum best approximates the curve derived from the ratios of oxygen isotopes. The components with the largest amplitudes have frequencies close to those of the orbital parameters; the four labeled components can account for about 60 percent of the fluctuation in the oxygen-isotope record.

This finding still leaves unexplained the prominence of the 100,000-year oscillation in the cycle of the ice ages. In many ways this is the most puzzling part of the problem. The variation in the eccentricity of the earth's orbit does have a 100,000-year period, but the distribution of insolation, which seems to be the driver of climatic change, is affected by tilt and precession much more than it is by eccentricity. Why does the record show that the 100,000-year cycle is stronger than the 40,000-year tilt cycle and the 20,000-year precession cycle?

The answer may lie in the phenomenon of resonance, a property of many physical systems. A pendulum gives a simple example. If the pendulum is set in motion and then left alone, it swings with a frequency that depends only on its length and the strength of the force of gravity. The pendulum can also be forced to oscillate by a series of timed pushes; in this case the frequency depends on the frequency of the pushes and does not have to be the same as the natural frequency. Even so, the amplitude of the swing, or the response of the system, increases as the frequency of the forcing approaches the natural frequency of the pendulum. It is this increase in the amplitude of the response that constitutes a resonance.

It is possible that the climate system, composed of the atmosphere, the oceans, the continents and the ice sheets, may have a natural or preferred frequency of oscillation with a period of roughly 100,000 years. Eccentricity may change insolation less than tilt or precession does, but the changes may be more effective in driving the climate because they have a frequency near the natural frequency of the system.

A number of investigators, notably Imbrie and Michael Ghil of New York University, have advocated resonance as the explanation of the 100,000-year climate cycle and have examined computer models incorporating factors that might give the climate this natural frequency. In my opinion the most reasonable of the models is one that takes into

account the response of the underlying bedrock to the massive ice sheets. The bedrock sinks under a growing ice sheet, but it takes some time to adjust to the load, at least thousands of years. Meanwhile, as the ice sheet thickens, its surface rises to higher altitudes, where snow is likelier. As a result ice accumulates even faster. When the bedrock finally sinks, the entire ice sheet is lowered, which favors melting. Thus the response of the bedrock can amplify a tendency toward ice-sheet growth and then later encourage the opposite process of melting. After melting begins the bedrock slowly rises in response to the lighter load, raising the remaining ice and causing the cycle to begin again.

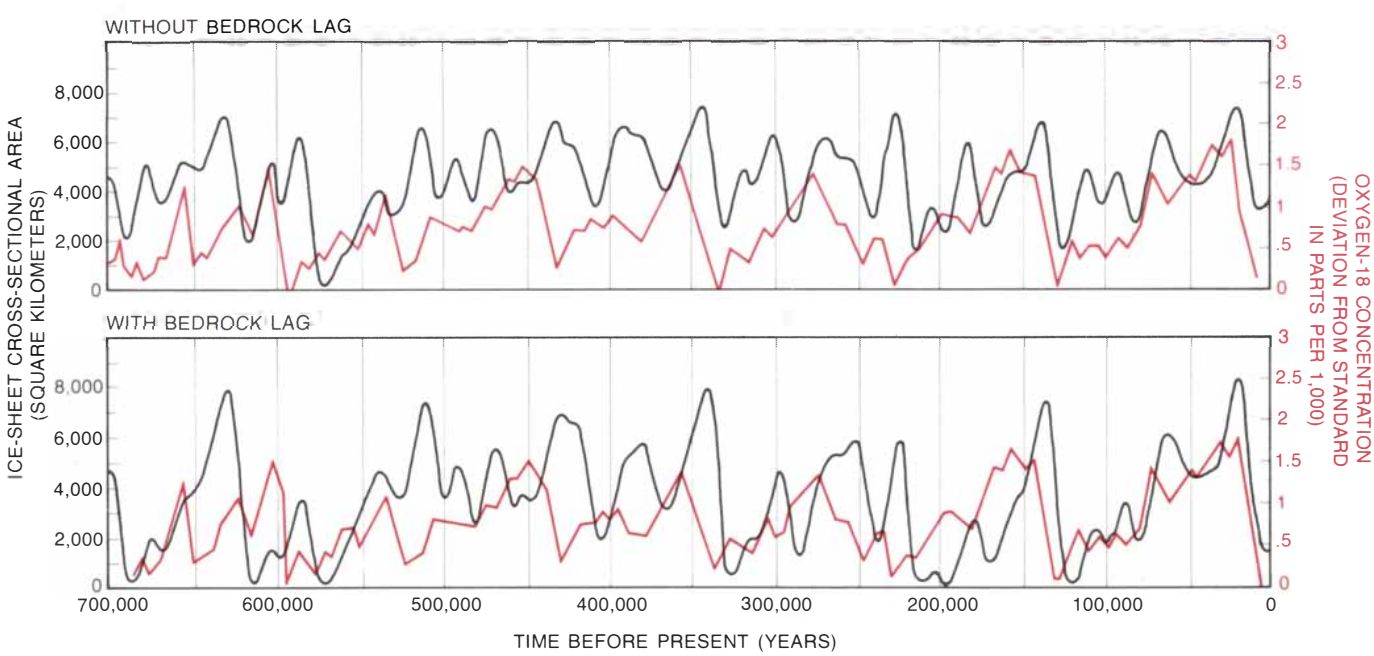
The effect bedrock loading could have on global ice volume was demonstrated by an ice-sheet model developed by David Pollard of Oregon State University. In the model it was assumed initially that the growth of the ice would be inversely proportional to the insolation; without taking into account the response of the bedrock the model predicted ice volumes dominated by the 40,000-year tilt cycle and the 20,000-year precession cycle. When the response of the bedrock was included in the model and was assumed to lag behind the loading and unloading of the rock by about 10,000 years, the model predicted changes in ice volume dominated by the 100,000-year cycle. The response of bedrock to ice sheets is actually far more complex than Pollard's simulation of it, and so the issue cannot be considered settled. Nevertheless, bedrock loading appears to be a reasonable explanation of the natural frequency of the climate.

Although it seems clear that there is a connection between the earth's orbit and the ice ages, a word of caution is in order. Orbital forcing is only one of many factors that could potentially have changed the climate. For example, it is possible that particulates injected into the upper atmosphere by large numbers of volcanic eruptions could have absorbed or scattered enough sunlight to cause a substantial global cooling. The advantage of the Milankovitch theory is that it is testable; the changes in the earth's orbit and hence in insolation can readily be calculated by applying Newton's law of gravity to progressively earlier configurations of the bodies in the solar system. Climatologists naturally prefer to experiment with well-defined forcing factors such as orbital variations, but the lack of information about other possible forcing factors does not mean they are any less important.

In any case the Milankovitch theory does not explain why ice ages began to occur at the beginning of the Pleistocene epoch. Throughout most of the earth's history there was little or no land ice and the climate was significantly warmer than it has been in the past million years. What caused the cooling that led to the ice ages? One answer may be the movement of the continents across the earth's surface. The climate may have begun to cool when the North American and the Eurasian continents reached their present position a few tens of millions of years ago and new atmospheric and oceanic heat-transport patterns began to establish themselves. The process appears to have been gradual. Glacial ice formed on Antarctica about 30 million years ago, after it had separated from South America and was cut off from the circulation of warm equatorial waters. There is evidence that high-latitude glaciers began forming in the Miocene.

Will there be more ice ages in the future? The fluctuations in the isotope record show no sign of abating, but the planet may nonetheless have seen its last ice age because of the effect of human activities on climate. As industrialization continues and more fossil fuel is burned the concentration of carbon dioxide and other trace atmospheric gases increases. By the end of the century these gases may cause an appreciable warming of the earth, an increase in average temperatures large enough, if it were to continue, to overwhelm any cooling due to orbital variations.

The earth is not the only planet whose climate may be affected by variations in its orbit and axial tilt. Mars has much greater orbital fluctuations because it is closer to the massive planet Jupiter. The tilt of Mars's rotational axis varies between 12 and 38 degrees with a period of about 120,000 years. The resulting changes in insolation must have a significant effect on the Martian climate. Some of the geologic features of Mars recently detected by spacecraft, such as layered deposits near the poles, may record these changes.



DELAYED RESPONSE OF THE BEDROCK may explain why the dominant period in the ice-volume curve is 100,000 years although high-latitude insolation is affected less by eccentricity (which has a 100,000-year period) than by the orbital parameters that have shorter periods. The graphs show the ice volumes predicted from the insolation by an ice-sheet model developed by David Pollard. In one experiment with the model (*upper graph*) the bedrock was assumed to sink immediately under the weight of accumulating ice. In another experiment (*lower graph*) it was assumed that the bedrock requires about 10,000 years to adjust to the load, maintaining the surface of the ice sheet at higher and colder altitudes. The second experiment came closer to duplicating the peaks in the isotope record (*color*).

What's data storage coming to?

We've found a way to stem the tide of the "Paper Sea."

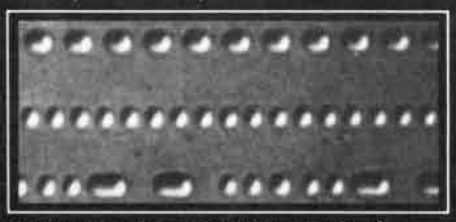
Kodak scientists recently unveiled a 14-inch optical disk with the capacity to store *over 100 times more data* than is now possible with the most sophisticated magnetic disk of equal size. It can integrate microfilm and computer technology, provide a powerful mass storage medium, and add a new dimension of flexibility to rapid transmission and receipt of images.

To accomplish this feat, we called upon our century of expertise in creating visible records and our understanding of the electromagnetic spectrum. We used our knowledge of polymers and dyes to develop coatings of amorphous materials that react differently at different wavelengths and allow for extremely high resolution. (See photo inset.)

This technology is incorporated into an optical disk which affords designers the option of using single or dual wavelengths and has the capacity to store more than 100,000 pages of information.

Its potential is tremendous. Our new optical disk, using single-wavelength laser record/read ability, delivers good-contrast readout and provides enhancement to current computer magnetic peripherals. Its dual-wavelength capability offers exciting opportunities for the future generation of information-handling products.

For more information on optical disk technology, the latest application of our interdisciplinary capabilities, send for the paper, "Digital Optical Recording in Infrared-Sensitive Organic Polymers." Write: Eastman Kodak Company, Dept GB SA-9, 343 State Street, Rochester, NY 14650.



Scanning electron micrograph of disk surface: coded information, representing a document, is precisely recorded along a track less than one-micrometre wide.

Kodak. Where technology anticipates need.

Kodak

SCIENCE AND THE CITIZEN

The Poor Get Poorer

The continuing global recession, caused in large part by anti-inflationary measures adopted by the governments of the world's richest countries, is exacting a heavy toll among those least able to afford it: the citizens of the poorest countries. According to the latest annual report of the World Bank, a combination of factors beyond the control of the developing countries has placed severe constraints on their prospects for economic growth. The overall drop in demand in the industrialized countries since 1980 has reduced both the amounts and the prices of primary exports from the developing countries. At the same time the restrictive monetary policies of some of the major industrialized countries, notably the U.S., have kept real interest rates at high levels, thereby increasing the burden of debt service for the developing countries. Some of them have been unable to meet their payments on debts, and many have been forced to reduce their borrowing from private sources and to curtail their imports in an effort to control the deficits in their accounts.

The ensuing lower rate of growth in the developing world as a whole, the World Bank reports, "has undoubtedly resulted in increases in unemployment and underemployment, and hence additional poverty. In some countries, particularly in the more developed parts of Latin America, the recession has caused the progress made during the 1970's to be abruptly interrupted. In others, particularly in sub-Saharan Africa, the recession has accelerated the decline in living standards that began in the 1970's." For hundreds of millions of people in the poorest countries, the report states, "the current 'cyclical downturn' is, in reality, a tragedy of no small proportion. Emergence from this condition for these countries will be a slow process."

In view of the "urgent need" for more development assistance in such cases, the International Development Association (IDA), the affiliate of the World Bank set up to aid the poorest developing countries, had planned to raise a total of at least $12 billion over the next three years from its more developed members, thereby matching the total raised for the preceding three-year period. IDA funds are disbursed to its less developed members in the form of interest-free credits with repayment periods of up to 50 years. More than 50 countries are eligible for such credits under the present criterion: an annual per capita gross domestic product of less than $796 (in 1981 dollars).

The scheduled replenishment of the IDA funds suffered a serious setback in December, when the Reagan Administration announced its intention to cut the projected U.S. contribution to $750 million annually for the next three years. The Administration had previously decided to spread the $3.2 billion pledged by President Carter in 1979 over four years rather than three, in effect reducing the U.S. share of the last IDA refinancing to $800 million annually. A number of other countries that contribute to the IDA thereupon set up a special fund to replace some of the money lost as a result of the U.S. action.

Since the U.S. is committed to providing 25 percent of the new IDA replenishment, the Administration's latest decision means that the total amount available for IDA credits over the next three years would be reduced to $9 billion, or $3 billion less than the amount that had been planned. Representatives of the 32 other donor countries in the IDA immediately appealed to President Reagan to restore the U.S. commitment to the previously projected level of $1 billion annually.

Bypassing the Bypass

The creation of artificial channels around obstructions in the arteries that supply blood to the heart is now the major surgical operation most frequently performed in the U.S. In 1981, the most recent year for which reliable figures are available, 159,000 coronary-artery bypass graft procedures were done at a cost of about $20,000 each. Thus in that year some $3 billion was spent on the operation, not including the substantial cost of the diagnostic tests preceding it. In spite of the frequency of bypass surgery, the evidence that the operation prolongs the life of those who have had it has until recently been sparse and ambiguous. A thoroughgoing study of the effects of bypass surgery by the National Heart, Lung, and Blood Institute now shows that for people with mild or moderate coronary-artery disease surgery does not yield an increase in survival rate over current nonsurgical treatments. The study recommends that people with nonsevere coronary-artery disease be treated without surgery unless their condition deteriorates. If this advice is followed, the number of bypass operations done in the U.S. each year will be reduced by at least 25,000.

Coronary-artery bypass surgery in its current form was devised in 1968. The operation was accepted rapidly by surgeons and their patients. In 1971, 24,000 operations were done, and the number rose steadily through the 1970's. Part of the reason for the rapid increase in the number of bypass procedures is the high incidence of diseases of the coronary arteries. Heart disease causes more deaths in the U.S. than any other illness, and about three-fourths of the deaths caused by heart disease are due to coronary-artery disease. In 1980, 566,000 people in the U.S. died from conditions that affect the coronary arteries.

The symptoms and life-threatening consequences of coronary-artery disease originate with the reduction in the flow of blood to the muscular tissues of the heart. The reduction is caused by a progressive decrease in the lumen, or inner diameter, of the coronary vessels. Four main vessels supply blood to the ventricles, which are the pumping chambers. The left main coronary artery, the largest of the four vessels, branches into the left circumflex coronary artery and the left anterior descending coronary artery, which are the chief sources of blood for the left ventricle. The right coronary artery is the chief source of blood for the right ventricle.

Any or all of the four vessels can be obstructed. The severity of the disease depends largely on which artery is narrowed and on the degree of the obstruction. The most serious conditions result from the narrowing of the left main coronary artery, because the vessel carries most of the blood for the left ventricle. Hence a narrowing of the left main coronary artery can lead to impairment of the function of the left ventricle. The left ventricle, which is the most muscular part of the heart, propels the blood through the body. (The right ventricle propels it through the lungs.) If the left main coronary artery is not affected, the severity of the disease depends on how many of the other three vessels are narrowed and the degree of the narrowing.

The most dangerous consequence of the occlusion of the coronary arteries is myocardial infarction: the death of a region of heart tissue. In many instances, however, long before infarction develops the patient experiences angina pectoris, a recurrent and occasionally incapacitating pain. The pain is generally felt in the chest but may be "referred" to other parts of the body.

In increasing the flow of blood to the heart the bypass operation is quite effective in relieving angina. The majority of patients who have the operation experience an immediate reduction in the frequency and severity of their attacks. This is probably the main reason bypass surgery was widely accepted.

The capacity of the surgery to prolong patients' lives, however, has not been as clearly established. Whereas the reduc-

Only one disk guarantees safe passage through the torrid zone of drive heat. Maxell.

A lifetime warranty. And manufacturing standards that make it almost unnecessary.

Consider this: Every time you take your disk for a little spin, you expose it to drive heat that can sidetrack data. Worse, take it to the point of no return. Maxell's unique jacket construction defies heat of 140°F. And keeps your information on track.

And Maxell runs clean. A unique process impregnates lubricants throughout the oxide layer. Extending media and head life. How good is Gold?

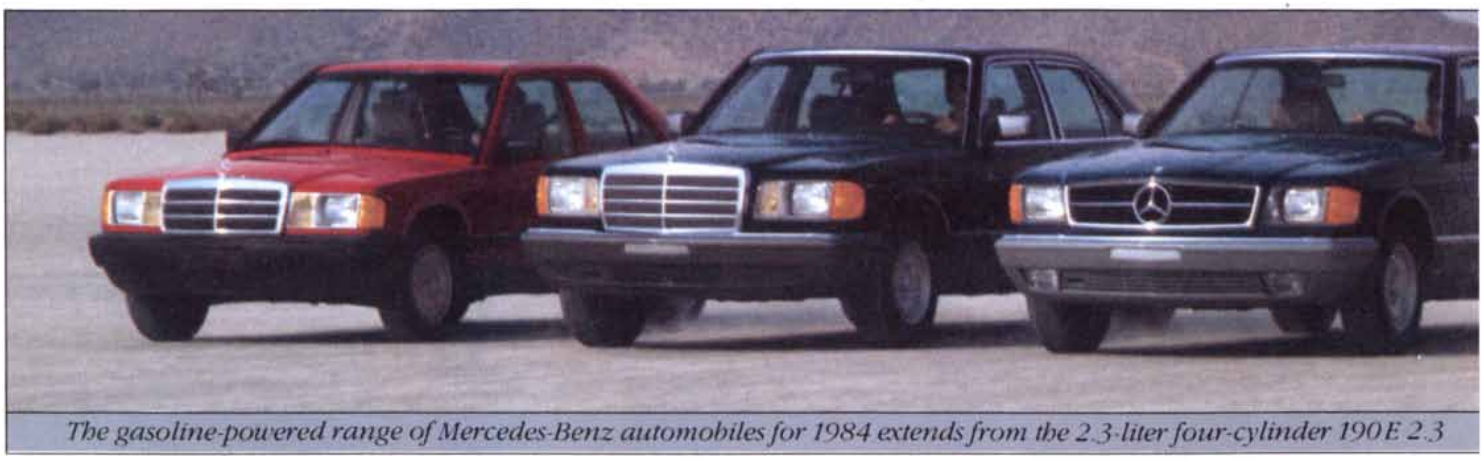
Maxell's the disk that many drive manufacturers trust to put new equipment through its paces. It's that bug-free.

So you can drive a bargain. But in accelerated tests, Maxell floppys lead the industry in error-free performance and durability. Proving that if you can't stand the heat you don't stand a chance.

maxell®

IT'S WORTH IT.

Maxell Corporation of America, 60 Oxford Drive, Moonachie, N.J. 07074 201-440-8020



The gasoline-powered range of Mercedes-Benz automobiles for 1984 extends from the 2.3-liter four-cylinder 190E 2.3

For 1984, five of the more potent gasoline-powered automobiles sold in America are sold by Mercedes-Benz.

CHANGING CONDITIONS never seem to catch the engineers of Mercedes-Benz napping.

Diesel cars, years in advance of the efficiency stampede.

Turbodiesels that anticipated a demand for efficiency plus performance. And it should now come as no shock that Mercedes-Benz greets the current performance renaissance with a range of *five* muscular gasoline performance cars.

In fact, gasoline-powered models now represent half the Mercedes-Benz line. Four of them–the 190E 2.3 Sedan, the 380SE Sedan, the 500SEL Sedan and the 500SEC Coupe–are new to America for 1984.

They help form the most vivid group of performance automobiles Mercedes-Benz has ever built for sale.

The most spirited Mercedes-Benz is unquestionably the stunning new 190E 2.3 Sedan "–perhaps the best in the world in its size class," in *Road & Track's* opinion. And perhaps the equal, in pure driving exhilaration, of any sports sedan now sold.

BOLD PERFORMANCE, MUTED NOISE

The 190E combines its trim 2,655 lbs. and a four-cylinder, 2.3-liter fuel-injected engine to reach test track maximums of 115-mph-plus. Yet it is so finely engineered that vibration and noise are as muted as in larger Mercedes-Benz sedans.

Its performance character is underscored by the five-speed manual gearbox that can be ordered as an alternative to the four-speed automatic.

"Transmission designers from around the world," *Road & Track* suggests, "should take a sabbatical to spend time at Stuttgart learning the Mercedes way of building gearboxes."

A V-8 LIKE NO OTHER

The new 380SE Sedan exploits the technology of the eighties to shine as one of the most confidently responsive passenger cars of the eighties.

Turbine-smooth and jackrabbit-quick, its 3.8-liter V-8 engine evolves from a new design generation meant to deftly balance power and efficiency.

Note that its block is even made of a different substance than most V-8's–cast not in iron but in an exotic alloy of aluminum, magnesium and silicon that radically cuts weight. The pistons in this engine move not against metal cylinder walls but against a surface of slippery, smooth, seemingly indestructible silicon crystals.

The 380SE's torque-converter four-speed automatic gearbox is so well integrated with this power plant that it seems to function almost as part of the engine. To the benefit of satisfying performance motoring.

What that 3.8-liter engine does for the 380SE it does for the 380SL Roadster, the two-seater descendant of Mercedes-Benz' immortal SL sports car series, carried over intact for 1984.

*Approximate suggested advertised delivered price at port of entry. ©1984 Mercedes Benz of N.A., Inc., Montvale, NJ.

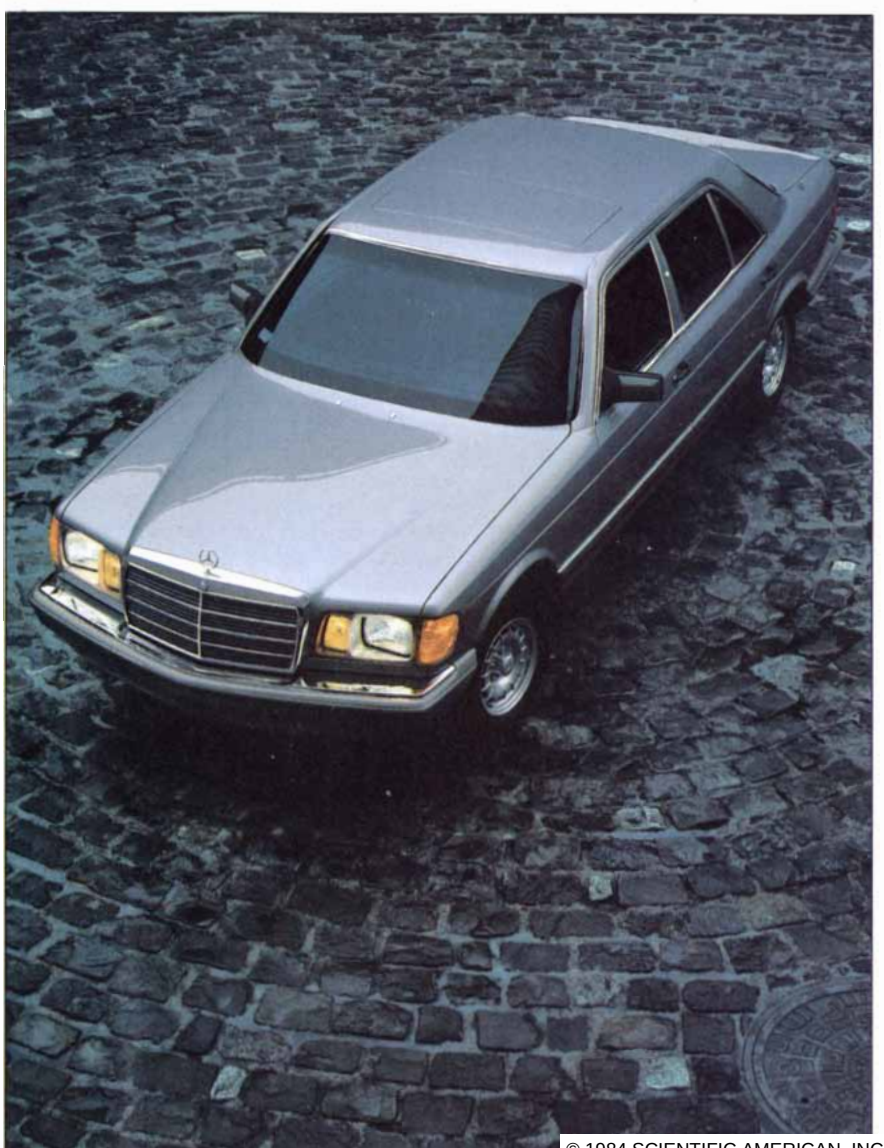
Sedan, at $24,000. to the 5-liter V-8 500 SEC Coupe, at $57,000.* Left to right: 190 E 2.3, 380 SE, 500 SEC, 500 SEL, 380 SL.*

Mercedes-Benz launches two new flagship models for 1984–the 500 SEL Sedan and the 500 SEC Coupe. They top off the gasoline performance line in formidable fashion.

The 500 SEL, featured below, is a five-passenger sedan of near limousine status. The 500 SEC is its sybaritic four-passenger touring coupe counterpart.

Each is propelled by an Olympian performer of an engine: the five-liter, light-alloy Mercedes-Benz V-8, new to America for 1984 but already somewhat of a legend on the highways of Europe.

So massive are its power reserves that, though not a practice recommended for normal driving, it could quietly cruise at 55 mph in *second* gear. At 55 mph in top gear it is turning at a lazy and barely audible 1,886 rpm. Acceleration from zero to 55 mph is jet-smooth. Acceleration from almost any point in the speed range, in any gear, is vivid.



HIGH-PERFORMANCE ENGINEERING

The automobiles powered by these engines are painstakingly engineered to high performance levels, from suspensions to steering to brakes. And to the stern safety standards imposed by Mercedes-Benz on every car it builds.

These five gasoline-powered Mercedes-Benz models for 1984 are meanwhile complemented by five of the most advanced *diesel*-powered automobiles you can buy–each an outstanding performer in its own right. Making for the most varied, the most exciting, the most timely lineup of models Mercedes-Benz has ever sold in America.

Engineered like no other car in the world

tion of angina can be observed by an individual physician, measuring an increase in survival rate calls for studies that include many patients and last for a substantial period. Moreover, if surgical treatment is to be compared with nonsurgical treatment, a group of patients whose pathology is about the same must be assigned at random to the treatments and observed for a considerable period. Such a randomized prospective study is complex to administer, is costly and raises ethical issues.

Because of such problems, only two large randomized studies of bypass surgery were done before the current study. One of them was a collaboration among several European medical institutions and the other was done by the Veterans Administration in the U.S. Both studies found that surgery prolonged the life of patients who showed a narrowing of the left main coronary artery.

Such patients, however, are only a fraction of all those who undergo bypass surgery, and the results with patients showing a less severe pathology were not clear-cut. The current study, called the Coronary Artery Surgery Study (CASS), was directed at patients with mild or moderate forms of disease: those who had some narrowing of coronary arteries other than the left main coronary artery and those who had suffered a myocardial infarction that had subsequently healed but who did not suffer from angina.

A group of 780 patients from 15 hospitals who met these criteria were assigned at random to surgery or to nonsurgical treatment. Nonsurgical treatment consists chiefly of the administration of drugs, including nitrates and beta-blocking agents. The 780 patients were observed for as long as five years.

The results of the CASS study, which were published in the journal *Circulation,* show that in general surgery does not extend the lives of patients with less than severe coronary-artery disease beyond what might be expected from drug treatment. After five years the average annual mortality rate among patients who had had surgery was 1.1 percent and the rate among those who had had drug treatment was 1.6 percent. The difference was not statistically significant. When the patients were divided according to whether one, two or three coronary vessels were affected, the differences in survival were also not significant.

The study confirmed the earlier finding that surgery is effective in relieving angina. The fraction of surgically treated patients who experienced complete relief from pain was significantly larger than that of the nonsurgically treated patients. The difference was confirmed by the patients' reports on their condition and by the capacity to exercise. There was little difference between the groups, however, in the fraction who continued to hold jobs during the study, or in the fraction who exercised regularly. The study finds that "the decision to proceed with bypass surgery in these patients [those with mild or moderate symptoms] can be reduced to a question of whether the current level of symptoms is acceptable to the patient."

Thus for many people suffering from coronary-artery disease surgery appears to lead at considerable expense to a reduction in symptoms but not to an increase in longevity or to a substantial change in the style of life. Furthermore, although mortality from bypass surgery has decreased as surgeons have mastered the technique, 1.4 percent of CASS patients who had been operated on died in the course of the operation or within 30 days of it.

In assessing the implications of the new findings Eugene Braunwald of the Harvard Medical School predicts a decrease in the frequency of coronary-artery bypass surgery. According to Braunwald, several factors could contribute to such a decrease. In addition to the results of the CASS study such factors include "the greater effectiveness of nonsurgical therapy [and] the increasing efforts to contain the costs of medical care, especially of procedures whose indications have not been clearly established." Braunwald concludes: "I suggest that after 15 years of increasing enthusiasm for coronary-artery bypass surgery, if appropriate education of the health professionals and the public concerning this procedure is carried out, the pendulum will swing back to a more appropriate equilibrium."

Sharpened Bits

About one light-second, or 300,000 kilometers, of optical fiber has now been shipped for installation in the U.S. telephone network. Of this quantity about 60 percent has already been installed, most of it in regional centers but also in two long trunk lines: one between Washington, D.C., and New York and the other between Sacramento and San Jose, Calif. Each pair of fibers necessary to complete the circuit of a conversation can transmit 90 million bits of information per second, the equivalent of about 1,400 telephone calls. In spite of this remarkable transmission rate there are two major bottlenecks in the system, which until recently had seemed unavoidable. Workers at Bell Laboratories now report that by taking advantage of two nonlinear interactions of light with the molecules that make up the fiber both bottlenecks can be eliminated.

The first bottleneck in the fiber-optics system is caused by the physical characteristics of the digital pulse of light, whose presence or absence within a short interval during the reception of the signal constitutes one bit of information. Under ordinary circumstances such a pulse tends to broaden as it propagates along the fiber. The broadening comes about because the pulse is made up of numerous simple sine- or cosine-wave components at various frequencies and amplitudes, which add by superposition to generate the pulse. The components travel along the fiber at different velocities; this velocity dispersion causes the phase, or distance between wave crests, of each component to change in such a way that the shape of the pulse is smeared. The calculated broadening of the pulse at the receiving end of the fiber determines the minimum length of the time interval that can be assigned to carry one bit of information: if the interval were too short, the broadening could cause two pulses to overlap and become indistinct.

The second bottleneck is caused by the system of repeaters that is designed to overcome the small but finite loss of light from the fiber. The repeater currently in service converts the light pulse into an electric pulse, amplifies the electric pulse by accelerating the charged particles that make up the pulse and finally reconverts the amplified electric pulse into a pulse of light. The device is relatively expensive; indeed, repeaters make up about half of the cost of hardware needed for a fiber-optics network. Moreover, the response time of the repeater is relatively slow, and so the device imposes a limit on the rate at which information can be transmitted.

The amplitude, or strength, of the light signal in fibers now being manufactured leaks out of the fiber at a rate of only about 5 percent for each kilometer of transit. Accordingly it has often been assumed that the disadvantages of repeaters can be minimized by placing the repeaters at intervals of up to 150 kilometers. Unfortunately this scheme itself limits the data-transmission rate. For pulses shorter than 100 picoseconds (trillionths of a second) the spreading of the pulse is so severe that over a distance of 150 kilometers the pulse, without amplification, would become indistinguishable from random optical noise.

It now appears that nonlinear processes can circumvent all these problems. When light passes through a material, the oscillating electric field of the light photons causes vibrations of the electric charges in the material. For the light intensities usually encountered in optical systems such as mirrors and lenses the displacement of the electric charges is nearly proportional to the strength of the electric field of the light, and so the interaction is called a linear one. In the optical fiber, however, the transmitted light is confined to a core whose diameter is only about five micrometers; the intensity of the light in the core can be quite high. Moreover, the interaction of light with the fiber can build up over the

entire length of the fiber. Hence in a fiber the strict linear proportionality between the electric field and the displacement of the charged particles does not hold. The nonlinear relation gives rise to diverse and exotic phenomena that until recently were not accounted for in the design of fiber-optical systems.

Perhaps the best-known nonlinear effect is Raman scattering (named for C. V. Raman, who first observed it in 1927). Akira Hasegawa of Bell Laboratories has proposed that the Raman effect can be exploited to make a repeater that does not depend on the interconversion of light and electrical energy. Hasegawa likens the Raman effect in an optical fiber to the effect on an oscillator such as an ordinary rope swing that is not pumped, or driven, at its natural resonant frequency. For a boy standing on the seat of a swing to pump the swing efficiently he must shift his weight up and down in both the forward direction and the reverse one; in other words, the most efficient pumping frequency of the swing is twice the natural frequency at which the swing would oscillate without any pumping. Two such oscillators, namely the swing and the pumping of the boy, are said to be coupled nonlinearly because their oscillation frequencies do not match.

The swing responds to the pumping at two frequencies: one response is the sum of the pumping frequency and the natural frequency of the swing, and the other response is the difference between the pumping frequency and the natural frequency. When the boy pumps at just twice the natural frequency of the swing, the difference between the pumping frequency and the natural frequency is equal to the natural frequency. Hence although the coupling is nonlinear, one response frequency of the swing resonates with its natural frequency, and the energy added by the pumping tends to be stored in the natural vibrational motion of the swing. The amplitude of the swing becomes increasingly high.

According to Hasegawa, if light with a frequency slightly higher than the frequency of the signal pulse is injected at regular intervals into a fiber, the same kind of resonant, nonlinear coupling will increase the amplitude of the signal. The frequency of the pumping light corresponds to the frequency of the pumping by the boy in the swing. The pumping frequency must be adjusted so that it is equal to the sum of the signal frequency and the natural frequency of the vibrational motions of the molecules in the fiber. The fiber will then respond nonlinearly at a frequency equal to the difference between the pumping frequency and the vibrational frequency of the molecules; in other words, the response frequency to the pumping light is equal to the frequency of the signal, and so the response of the system resonates with the signal light. Raman amplification of light could be relatively inexpensive, and it would impose no limitations on the data-transmission rate.

Hasegawa, Yuji Kodama, Linn F. Mollenauer and Roger H. Stolen, also of Bell Laboratories, have investigated another nonlinear effect that could eliminate the limitation on transmission rate imposed by pulse spreading. When light passes through a solid medium, its speed varies with the intensity of the light; this variation is a nonlinear phenomenon, but high intensities are required before it can be observed.

In an optical fiber the effects of the phenomenon build up as a light pulse travels down the fiber. The frequencies of the components that make up the leading half of the pulse are lowered, whereas the frequencies of the components that make up the trailing half are raised. Although the redistribution of the components would not affect the shape of the pulse in the absence of the dispersion of the component velocities, the interaction of the nonlinear effects with the velocity dispersion opposes the broadening of the pulse and can actually cause it to narrow.

If the initial energy and shape of the pulse are controlled within certain limits that are fixed by the properties of the fiber, the pulse will evolve in shape as it propagates until the nonlinear effects are exactly balanced by the effects of the velocity dispersion. The pulse that results is a remarkably stable configuration of light energy, in the sense that the area under the pulse does not change as the pulse moves along. Such a pulse is called a soliton, or solitary wave.

Soliton pulses are ideally suited to optical communications. Extremely narrow solitons can be generated, and so the transmission rate can be high. Energy loss along the fiber causes the amplitude of the soliton pulse to diminish, but the area under the pulse remains the same. Energy supplied to the fiber by Raman light can restore the soliton to its original amplitude, but because the area under the pulse must remain constant, the energy also restores the soliton to its original shape. The transmission of solitons enhanced periodically by Raman light amplification could lead to transmission rates on the order of 10 to 100 billion bits per second.

Oncogenes in vivo

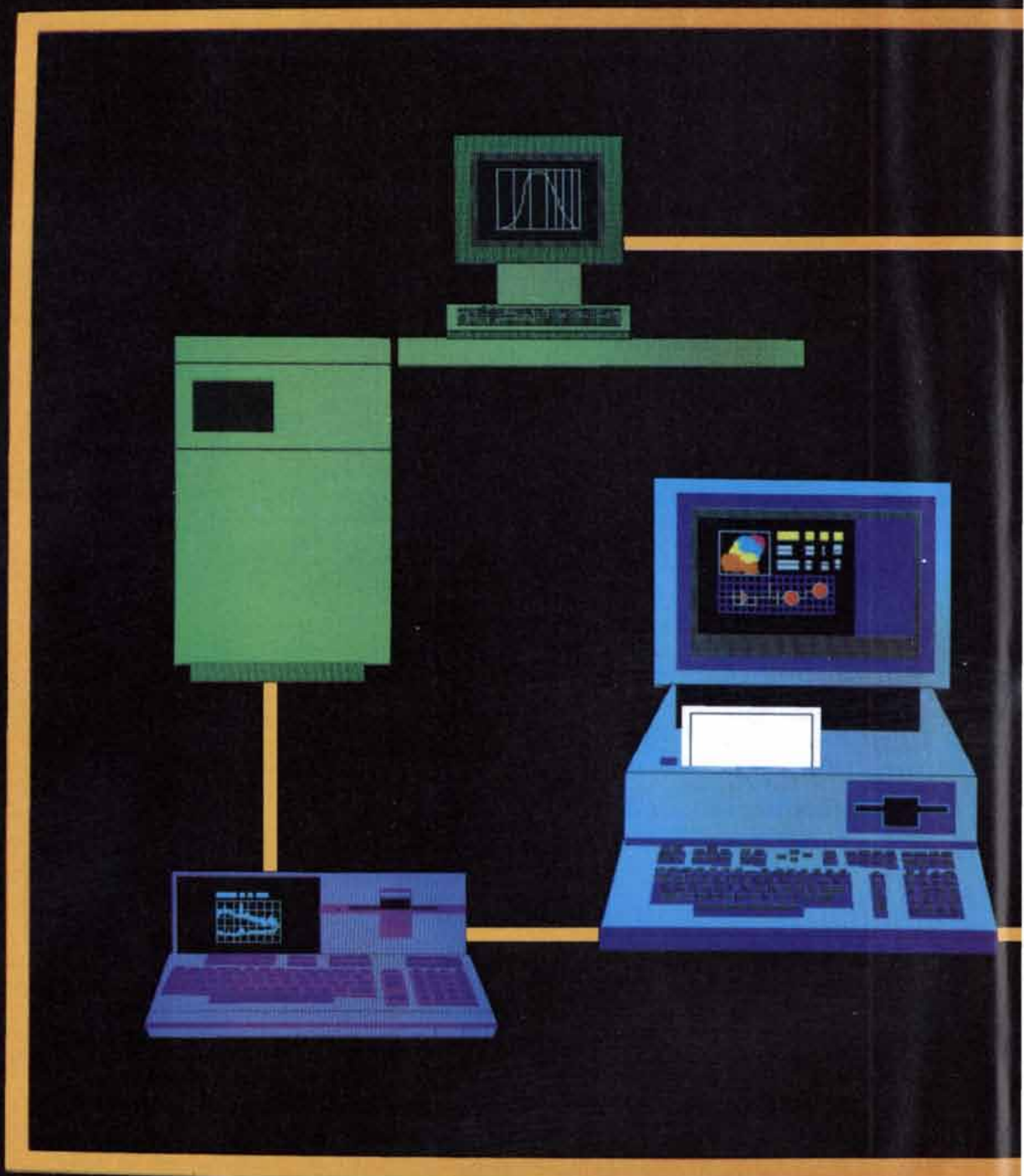
Over the past five years strong evidence has accumulated indicating that specific genes are responsible at least in part for the development of human cancers. The evidence is, however, circumstantial. What is well established is that certain ordinarily benign "proto-oncogenes" in normal cells can be activated to become oncogenes, which are found in animal and human tumor-cell lines and tumor tissue. When oncogenes are transferred into mouse cells that have been maintained in laboratory cultures, the cells are transformed into the cancerous state. It takes a combination of two different oncogenes or of an oncogene and a carcinogenic agent to transform more nearly normal cells, in keeping with the observation that the development of cancer is a multistep process. In either case injection of the transformed cells into laboratory animals gives rise to tumors. What has not been directly demonstrated is a direct role for oncogenes in initiating the original tumor. Both to establish that role and to investigate the multistep cancer process at the molecular level it would be good to have an animal model in which a particular oncogene could be followed from its activation by a carcinogen to the growth of a tumor. Steps have been taken toward developing such a model.

A first step was reported in *Nature* some months ago by Allan Balmain and Ian B. Pragnell of the Beatson Institute for Cancer Research in Scotland. They induced squamous-cell skin carcinomas by treating mice with two carcinogens, first an "initiator" substance and then a "promoter," and transplanted the tumors to increase their malignancy. DNA from the transplanted tumors transformed cultured mouse cells. The transformed cells were shown to contain an oncogene, of the *ras* family of genes, that is found in a number of human tumors. The implication was that the oncogene had been activated by a carcinogen and had given rise to the skin cancer.

Now a more precisely defined experimental system and clearer identification of the mechanism whereby a *ras* oncogene is activated in it have been reported in *Nature* by Saraswati Sukumar, Vicente Notario, Dionisio Martin-Zanca and Mariano Barbacid, working in Barbacid's laboratory at the National Cancer Institute. Sukumar and her colleagues induced mammary carcinomas by treating female rats with the carcinogen nitroso-methyl-urea (NMU), an initiator. DNA from nine such tumors transformed cultured mouse cells; DNA from normal breast tissue did not. Mouse cells transformed by DNA from each of the nine rat tumors were shown to contain a particular activated *ras* oncogene.

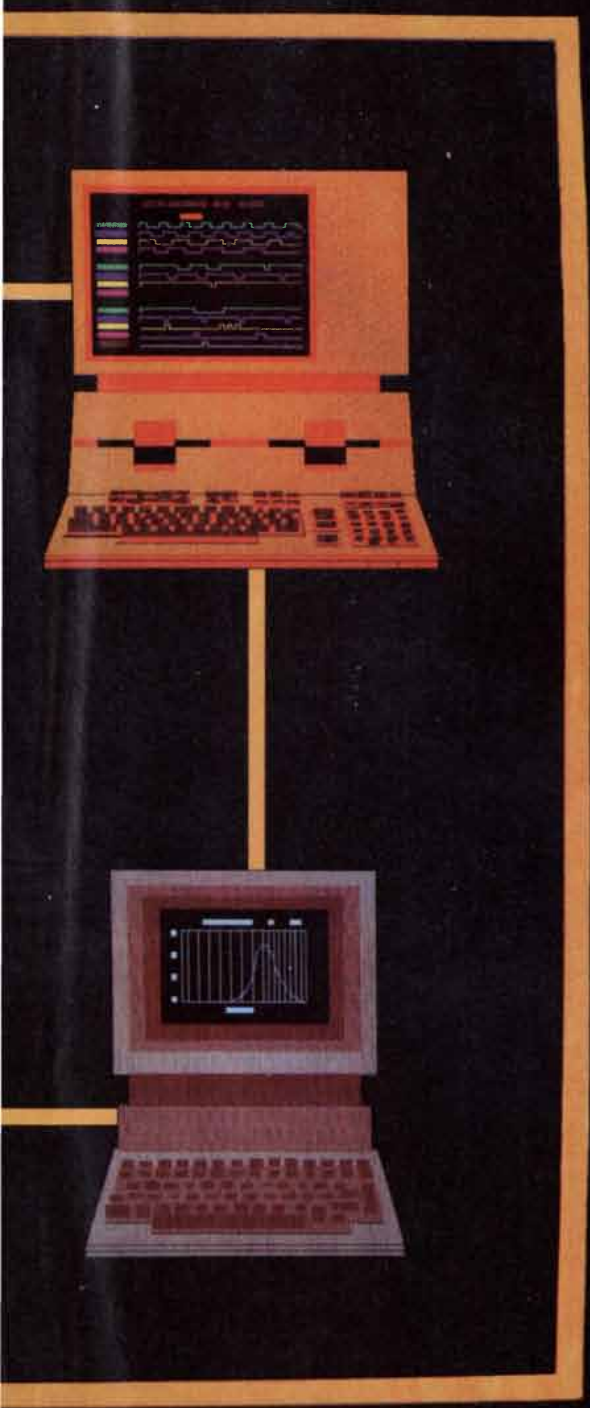
The next step was to identify the mechanism of activation. In 1982 Barbacid's group and groups headed by Robert A. Weinberg of the Massachusetts Institute of Technology and Michael Wigler of the Cold Spring Harbor Laboratory had shown that the human *ras* oncogene acquires its malignant properties as the result of a single point mutation. The mutation changes the codon (the sequence of three nucleotides in DNA that codes for a specific ami-

The Engineering Pr



With HP's broad range of workstations, you can choose the one that best fits your application and budget. From personal computers for engineers to desktop mainframes, with beginning prices ranging from $3,950 to $28,250.

oductivity Network.



Presenting your broadest choice of engineering workstations anywhere. From a low-cost personal computer to a powerful 32-bit system.

No doubt about it: when you put a computer on every engineer's desk, it helps them be more productive.

In fact, at Hewlett-Packard, the workstation approach has worked so well for our own engineers that we've taken it to an entirely new dimension.

It's called the Engineering Productivity Network. As you can see, it includes the broadest range of workstations available anywhere. From the recently-introduced 32-bit HP 9000 desktop mainframe to our personal computer designed just for engineers.

With such a wide range to choose from, you can match the computing power to the individual task. And give every engineer the tool it takes to do the best possible job.

To help you tailor every workstation to a task even more closely, we have over 200 specific application software packages. And since our workstations use HP-UX—our enhanced version of the UNIX* operating system—we'll be able to bring you a lot more in the future.

Of course, when you tie all these individually-optimized workstations together into a network, you can look for an even higher level of performance. Whether you choose our HP-UX networking, Ethernet local area network, or Shared Resource Manager, your engineers can share data and peripherals, communicate much better, and get more done than ever before.

The HP Engineering Productivity Network. You can start with one workstation, or with dozens. But no matter where you start, you're actually making your whole department more productive. One step at a time.

To find out more, just call the nearest HP office listed in the white pages of your telephone book, and ask for a Technical Computer Representative. Or write to Pat Welch, Dept. 41189, Hewlett-Packard, 19447 Pruneridge Avenue, Cupertino, CA 95014. In Europe, write Henk van Lammeren, Hewlett-Packard, Dept. 41189, P.O. Box 529, 1180 AM, Amstelveen, the Netherlands.

We'll be happy to show you how far you can really go with a good idea.

*UNIX is a trademark of Bell Laboratories.

HEWLETT PACKARD

BD-02316

no acid in a protein) for the 12th amino acid of the protein encoded by *ras*. Where the proto-oncogene codon specifies the amino acid glycine, the oncogene codon specifies valine; it has since been shown that the substitution of any other amino acid for this glycine would have the same activating effect.

Now Barbacid's group isolated the *ras* gene derived from one of the mammary tumors and determined the nucleotide sequence of the first part of the gene. Comparing it with the sequence of a normal *ras* gene from the same rat strain, they again found a single-nucleotide difference that changes the 12th codon so that it no longer specifies glycine; in this case it encodes glutamic acid. The investigators also showed that the *ras* oncogenes in the other eight mammary tumors had undergone mutation in the same location.

The high degree of reproducibility in oncogene activation by NMU opens the door to a reexamination of the role of chemical carcinogenesis in the development of cancer. Barbacid's group noted that the particular mutation identified in the study is one to be expected if NMU was in fact responsible for the activation. If the link can be firmly established, the NMU-induced mammary carcinoma in particular and chemical carcinogenesis in general should provide a model system for studying the malignant activation of an oncogene and the mechanism by which it contributes to the development of a tumor. Because the mode of activation is the same in the mouse carcinoma and in some human tumors, what is learned from such a system should be directly applicable to human cancer.

Message from the Moon

Since it was first realized in the early 1970's that the Antarctic ice tends to preserve meteorites that would have weathered away in other parts of the globe, annual expeditions to the area have added thousands of specimens to the world's collections of extraterrestrial rocks. One of them, ALHA 81005, has caused particular excitement in meteoritic circles since its discovery in the Allan Hills region in January, 1982. The papers in a special issue of *Geophysical Research Letters* reveal why, namely that investigators are certain this modest rock, weighing just 31 grams and measuring three centimeters across, is the first known meteorite from the moon.

How do they know? To begin with, ALHA 81005 *looks* like the Apollo moon rocks: its fine-grained texture and its glassy brown constituents identify it as a breccia (a rock composed of fragments) from the lunar regolith (the surface layer worked over by meteoritic impacts). In mineral composition it is also like lunar samples but unlike any other meteorite, consisting of approximately 75 percent plagioclase, a calcium aluminum silicate. Then there is a plethora of evidence from chemical and physical analysis. The ratio of iron to manganese, the relative abundances of three isotopes of oxygen and large concentrations of trapped noble gases (such as helium and argon) all serve to distinguish ALHA 81005 from other meteorites and to suggest its lunar origin. The noble gases demonstrate that the rock is not terrestrial; they can only have been implanted by the wind of particles from the sun, which do not reach the earth's surface because they are blocked by the atmosphere.

The rock is not, however, the same as other moon rocks, which makes it all the more interesting. It seems to have come from an area on the moon not visited by any of the Apollo (manned American) or Luna (unmanned Russian) missions, which were concentrated close to the center of the near side. The main reason for believing this is the relative absence in the meteorite of the KREEP elements (standing for potassium, the rare-earth elements and phosphorus), which were found as contaminants in most Apollo and Luna rocks. According to Carle M. Pieters of Brown University and her colleagues, remote-sensing data on the moon's geochemistry (from X-ray and gamma-ray spectrometers mounted on two of the Apollo spacecraft) also suggest that ALHA 81005 came from the far side of the moon or the edge of the near side.

The National Aeronautics and Space Administration authorized a consortium of workers led by Klaus Keil of the University of New Mexico to break off up to 10 grams of the meteorite for more detailed study of its composition. Keil reports the identification in it of pristine fragments, that is, fragments that have not remelted and whose radioactive clocks have not been reset since their minerals first crystallized out of a lunar magma. By radioactive dating of the pristine fragments it may be possible to determine their age. ALHA 81005 could thereby offer a glimpse into the early history of the lunar crust, which is thought to have formed about 4.5 billion years ago. Such glimpses are rare, Keil observes; crystallization ages have so far been determined for only a few lunar samples.

The implications of ALHA 81005 may extend even further to eight puzzling objects known as the SNC meteorites (for shergottites, nakhlites and chassingites). Most other meteorites crystallized 4.5 billion years ago, and they are therefore thought to have come from asteroids, the only bodies small enough to have cooled so soon after the formation of the solar system. In contrast, the SNC's are no more than 1.3 billion years old. Therefore they must be from a body sufficiently large and well insulated to have had volcanic activity that recently. The most popular theory, based on indirect geochemical evidence, is that a powerful meteorite impact expelled them from the surface of Mars. The main objection to the hypothesis is that no one has been able to demonstrate convincingly how a rock could be accelerated to the Martian escape velocity of five kilometers per second without being pulverized.

Until the discovery of ALHA 81005 the same argument had been advanced to explain why no meteorites found on the earth could have come from the moon. The lack of meteorites from the moon, in turn, supplied a strong motive for dismissing the possibility that rocks could reach the earth from Mars. The petrologists and geochemists who believe in the Martian origin of the SNC's now argue that ALHA 81005 has buttressed their case. Not only was it not pulverized in escaping the moon but also it exhibits no evidence (such as the melting of some of its components to glass) that it was shocked in the process. Since it is widely agreed that the only event that could knock a rock off the moon or Mars is a giant meteorite impact, this is surprising.

H. J. Melosh of the University of Arizona offers a possible explanation. When a meteorite buries itself in the lunar surface (to a depth of several times its diameter), the shock travels upward as a compressional wave and is reflected back off the surface as a tensional wave. The total pressure at any point below the surface is given by the sum of the two wave amplitudes. At the surface, however, the two waves cancel and the pressure is zero. According to Melosh, a thin surface layer could therefore be "spalled" off the moon without being shocked. The spalling hypothesis fits the observation that ALHA 81005 is a surface rock.

There are many reasons, however, why the hypothesis cannot explain how the SNC meteorites might have got off Mars. Most important, some of the SNC's are highly shocked and are clearly igneous rocks from below the surface. Furthermore, the escape velocity of the moon is 2.4 kilometers per second, half that of Mars. Also some of the SNC meteorites are about 10 times larger than ALHA 81005, and attempts to model their ejection from Mars theoretically require them to have started off larger still—at least several meters across. According to Melosh, the dynamical objections to the hypothesis of Martian origin are still strong. Nevertheless, says Donald Bogard of the Johnson Space Center of NASA, a majority of geochemists now accept the hypothesis and think the onus is on dynamicists to come up with a suitable physical basis.

10 GOOD REASONS WHY EAGLE IS WORTH MORE THAN ANY ORDINARY WAGON.

The American Eagle is one luxury wagon whose value increases with every inch of snow, every drop of rain, and every mile of bad road or highway. In fact, the uglier the driving, the better Eagle looks. Because when ordinary wagons are slipping, sliding, or just plain stuck, the Eagle can switch into the sure-footed traction of 4-wheel drive to transport you confidently through winter's worst storms.

But a feeling of security is by no means Eagle's only luxury. The interior room and comfort is outstanding. And with standard features like power steering and brakes, plush 18 oz. carpeting, woodgrain instrument panel overlay and individual reclining seats of genuine leather, the Eagle Limited makes 4-wheel drive extremely attractive.

It's a great looking wagon, but more importantly, it's rugged, eager and very reliable. No wonder the National Ski Patrol has chosen the Eagle as its official vehicle. They *have* to get through.

NATIONAL SKI PATROL

In addition to impressing you with its 4-wheel drive performance, Eagle can switch back to 2-wheel drive to impress you with its outstanding gas mileage, (24) EPA EST. MPG and 30 HWY. EST.*

The Eagle is the perfect all-season, all-weather family wagon. Because when other luxury wagons are grounded, the Eagle flies. SAFETY IS A SNAP

Now available for purchase or lease at participating dealers.

THE 2-WHEEL/4-WHEEL DRIVE

Eagle
AMERICAN MOTORS

* 4-cylinder/4-speed. Use these figures for comparison. Your results may differ due to driving speed, weather conditions and trip length. Actual highway figures are probably lower.

The Structure of Proteins in Biological Membranes

Membrane proteins are embedded in lipid, but they extend into water. The configurations that adapt them to contact with two environments are beginning to be revealed by novel electron-microscope techniques

by Nigel Unwin and Richard Henderson

Cells are bounded, and organelles within them are enveloped, by an extremely thin but tenaciously stable film of lipid and protein molecules: the biological membrane. The membrane functions basically as a permeability barrier that establishes discrete compartments and prevents the random mixing of the contents of one compartment with those of another, enabling the chemistry of life to proceed in an orderly way. Biological membranes are not, however, just containers; they serve as highly specific mediators between the cell or organelle and its environment. The diverse functions of different cells and different organelles are accountable in large part to the design of their individual membrane and the properties of the membrane's proteins.

The membrane proteins have a wide range of vital functions. The permeability of membranes is highly selective, and the reason is that particular proteins form channels or pumps allowing specific ions or molecules to pass through the membrane or driving them through it against a concentration gradient. The response of a target cell to hormones, neurotransmitters or foreign antigens depends on a receptor protein that detects the presence of these molecules and transmits their specific signal through the membrane. Still other proteins, such as enzymes or antigens, are simply anchored in the membrane to localize a particular reaction there.

To understand the critical roles of membrane proteins one needs to know their three-dimensional structure and how it changes in response to specific stimuli. We have been analyzing the structure of membrane proteins with the electron microscope, and some of the details are beginning to come clear.

Lipids and Proteins

The ground substance of a biological membrane is a bimolecular sheet of lipid molecules. A lipid molecule is amphipathic, that is, it has a hydrophilic portion and a hydrophobic one. The bulky head of a lipid molecule is hydrophilic: it reacts easily with water, commonly by forming hydrogen bonds. The two long hydrocarbon tails of the molecule are hydrophobic: they cannot form hydrogen bonds and they avoid water. In an aqueous physiological environment such molecules tend to arrange themselves into a two-layer membrane. The hydrophilic heads line up to face outward, interacting with water molecules in the intercellular space, the cytoplasm of a cell or the interior of an organelle; the hydrophobic tails line up pointed away from the water, interacting with one another in the interior of the bilayer. The result is a membrane some 45 angstrom units (ten-millionths of a millimeter) thick. The lipid molecules are usually in thermal motion and are free to diffuse in the plane of the membrane, so that the bilayer is a two-dimensional film of liquid only slightly more viscous than water. For all its thinness and fluidity the bilayer is stable because the amphipathic lipid molecules are in their preferred orientation.

Membrane proteins are densely intercalated among the lipid molecules, dispersed over the membrane or grouped in patches. What distinguishes the diverse proteins of membranes from other, water-soluble proteins is that they make contact with two environments, or are partitioned between two phases: the lipid film in which they are embedded and the water into which they protrude, often on both sides of the membrane. They are tethered firmly (although they are free to diffuse in the plane of the membrane) because they too are amphipathic: they have hydrophilic portions that interact with the lipid heads at the membrane's surface and with water and they have hydrophobic portions that interact with the lipid tails in the interior of the membrane. Membrane proteins can also interact with water-soluble proteins, with one another and with small molecules soluble in lipids.

We reasoned that the special environment of the lipid bilayer must impose special constraints on the structure of the embedded proteins and that the fundamental principles of protein structure should apply. Amino acids, the constituents of proteins, are linked in a linear sequence by peptide bonds to form a long polypeptide chain. A polypeptide chain folds on itself (or several chains fold together) in a specific way determined by the sequence of amino acids. The sequence dictates the folding because the side chains that characterize each of the 20 amino acids differ in size, shape, charge, hydrogen-bonding capacity and chemical reactivity; evolution has selected sequences that give rise to stable, functional structures.

Alpha Helix, Beta Sheet

Most of the water-soluble proteins studied to date consist of more or less regular conformations of the polypeptide chain—the alpha helix and the beta sheet—connected by less regular regions. The alpha helix is a rodlike structure: the backbone of the chain is coiled tightly into a helix, forming a cylinder from which the amino acid side chains project. In the beta sheet, on the other hand, two or more backbones, almost fully extended, are arrayed in parallel and the amino acid side chains extend perpendicularly from opposite surfaces of the sheet. Both the alpha helix and the beta sheet are stabilized by hydrogen bonding between the hydrophilic NH and CO groups of the backbone chain. In the alpha helix this bonding is between groups at successive turns of the helix; in the beta sheet it is between groups on adjacent strands.

In less regular regions of the protein (for example where a polypeptide chain links two successive alpha helixes) these

STRUCTURE OF BACTERIORHODOPSIN, a protein that powers an unusual form of photosynthesis in the membrane of salt-loving bacteria, is seen in a computer display generated by one of the authors (Henderson) and Arthur M. Lesk. The protein has 248 amino acids in a single polypeptide chain, which is linked to one light-absorbing retinal group. Seven segments of the chain, accounting for most of its length, are coiled into the regular alpha-helix conformation. It is the seven alpha helixes that are seen in this display. They are presumably linked to one another by six nonhelical stretches of the chain, but neither these links nor the retinal is visible at a resolution of seven angstrom units, the resolution of the contour maps of electron density (*see bottom illustration on page 91*) on which the display is based.

stabilizing bonds may not form; there are free hydrophilic groups in the backbone, and they are normally hydrogen-bonded to water molecules. A structure in which the hydrophilic groups are unable to interact with one another or with a water molecule would be unstable. Soluble proteins achieve stability by folding so that the self-bonded alpha and beta structures are largely on the inside, with less regular turns and bends of the chain situated at the molecule's surface, where water is available.

Similar principles must surely apply, perhaps with even greater force, to a protein embedded in the hydrophobic lipid-tail environment. The less regular turns and bends with free hydrogen-bonding groups should be mainly in parts of the protein exposed to the hydrophilic heads of lipids or to the aqueous environment. The alpha helixes and the beta sheets, whose hydrophilic backbone groups are bonded to one another, should be sequestered within the membrane in the hydrophobic environment of the lipids' hydrocarbon chains, and so should the more hydrophobic of the amino acid side chains. One could imagine a variety of possible configurations for membrane proteins, for example bundles of alpha helixes, beta sheets formed into barrels or combinations of the two. Such configurations could constitute functioning proteins or could be subunits of more complex assemblies.

Electron Analysis

To test these hypotheses we set out some 10 years ago to analyze the structure of membrane proteins with the electron microscope, adapting techniques developed by X-ray crystallographers for the analysis of crystals of soluble proteins. A crystal is made up of many identical molecules, which are assembled in positions and orientations that depend precisely on the individual molecule's structure; the crystal's structure is in effect the sum of that of each of the repeating units. By recording the way in which the atoms of a crystal diffract X rays one obtains information about the positions of atoms within the individual molecules of the crystal. Most of what is known about the structure of soluble proteins, including the details of the alpha and beta configurations, has been learned through X-ray-diffraction analysis of large three-dimensional crystals. Such analysis was not possible for the membrane proteins because they could not be induced to form three-dimensional crystals.

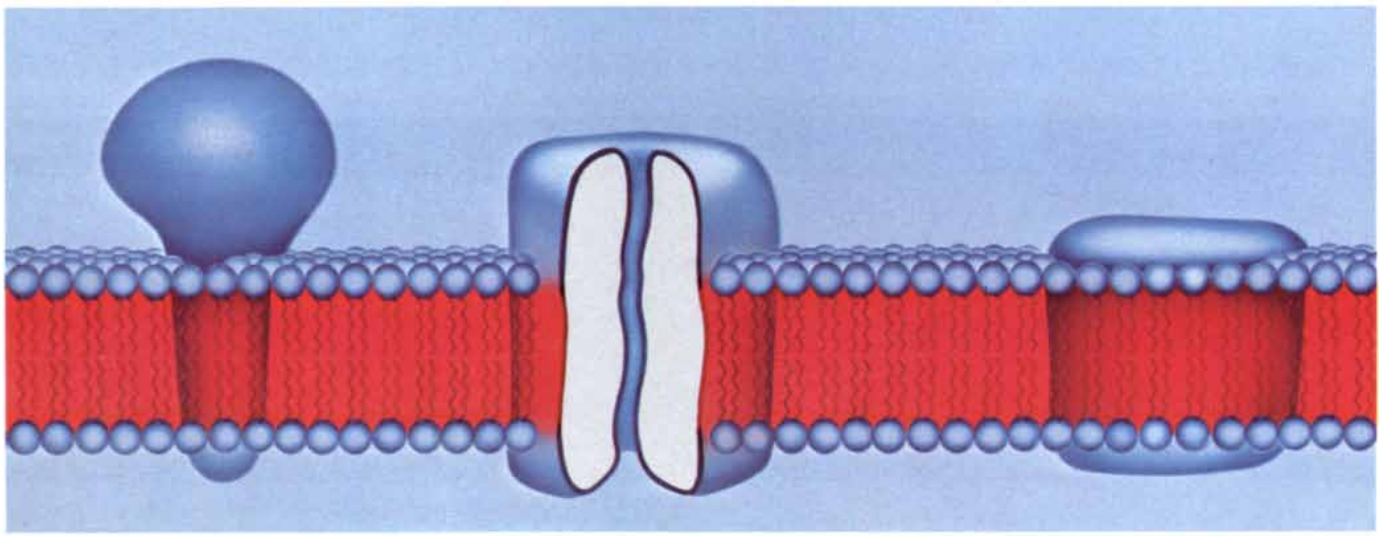
These proteins are adapted to floating in the lipid bilayer, a two-dimensional world. They tend to associate more easily in two dimensions than in three. Some of them tend to aggregate into two-dimensional lattices within the membrane. Others can be induced to form such arrays, which are in effect two-dimensional crystals, if they are extracted from the membrane with a detergent, purified and then reconstituted with a suitable proportion of lipid. X rays would hardly be diffracted at all in passing through such a thin crystal, but electrons, which interact with the atomic nuclei, are scattered much more strongly than X rays and yield a diffraction pattern. Moreover, electrons are charged particles, and so they can be focused to form an image: an electron micrograph. The diffraction pattern yields some information about the structure of an individual molecule; more information can be extracted from the micrograph.

For conventional electron microscopy a membrane must be negatively stained: embedded in the salt of a heavy metal that is opaque to the electron beam. The stain outlines the external, hydrophilic surfaces of the specimen and so reveals the shape of surface features; details of structure within the lipid bilayer and within the protein molecules remain obscure. To observe these details it is necessary to eliminate the stain and visualize the specimen directly. That is not easy to do because biological matter, unlike the stain, is rapidly damaged by the electron beam. Such features as alpha helixes and beta sheets are destroyed by an electron dose far smaller than the dose required to define them in a micrograph of a single molecule.

The solution is to make a micrograph of a crystalline array of several thousand molecules with an electron dose too weak to destroy them, and then to apply a summation procedure that reinforces genuine detail associated with the repetitive molecular structure while canceling out the large amount of noise: the random electron impacts associated with the weak electron dose. In this way a complete and accurate picture of the molecule is built up from fragments of information provided by each of the repeating units.

Fourier Transforms

One way to carry out this reinforcement, or averaging, would be to take successive repeating units from the micrograph and superpose them on one another, thus summing the information reflecting genuine, repetitive detail. That might not work, however, because it is often hard to determine the precise positions of repeating units. A more elegant and powerful approach would be



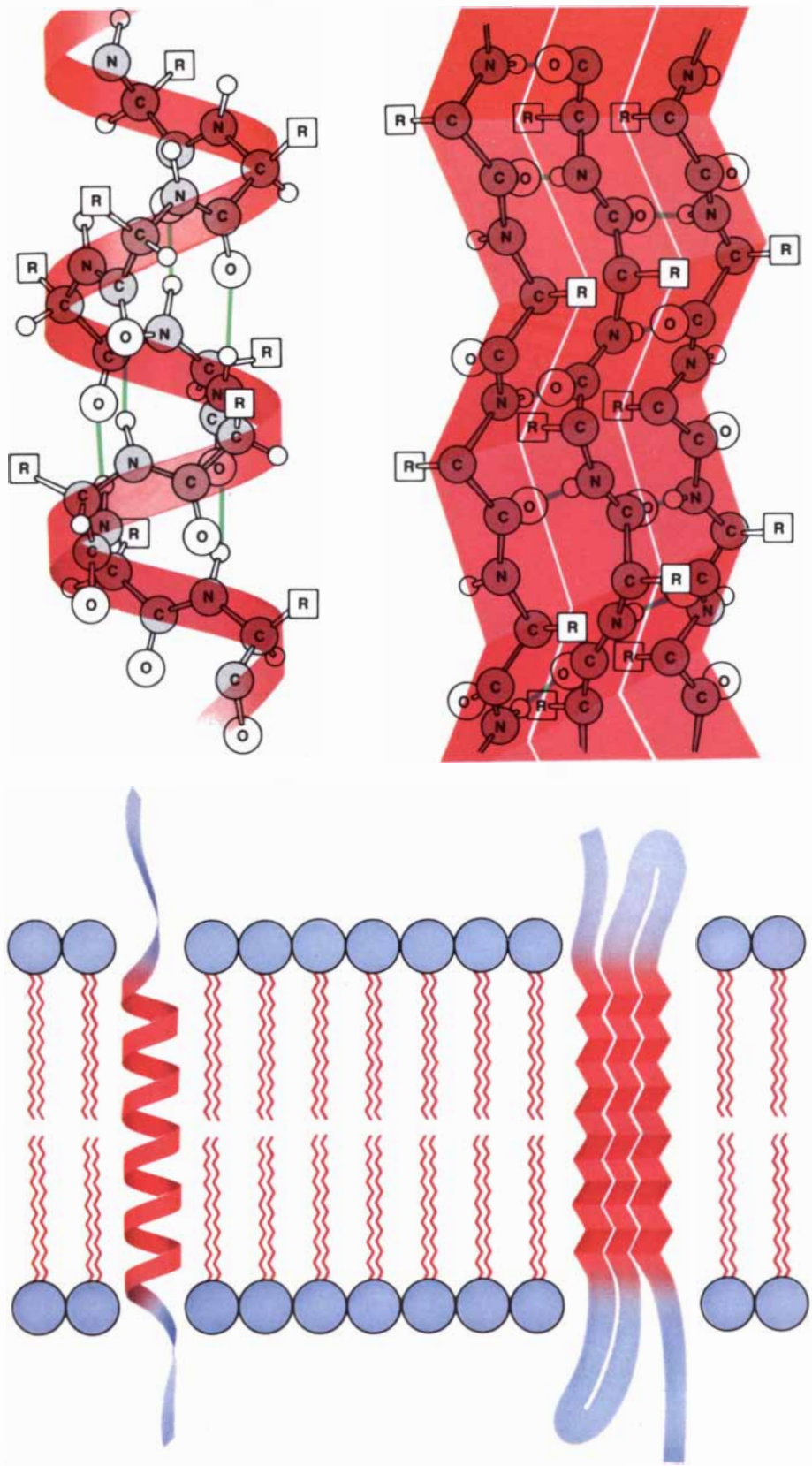
BIOLOGICAL MEMBRANE is made up of proteins set into a bimolecular film of closely packed lipid molecules. The hydrophilic heads (*blue spheres*) of the lipids are oriented toward the aqueous extracellular or intracellular medium (*light blue*) and their hydrophobic hydrocarbon tails (*red*) are oriented toward the inside of the bilayer. The exposed portions of proteins embedded in the bilayer are also either hydrophilic or hydrophobic, depending on whether they interact with lipid heads and water molecules or with hydrophobic tails.

the one developed some years ago by David J. DeRosier and Aaron Klug of the Medical Research Council Laboratory of Molecular Biology in Cambridge for analyzing electron micrographs of ordered molecular aggregates such as virus particles. The averaging is done in steps. The first step is to scan a micrograph of a crystalline object with a microdensitometer, thereby generating a two-dimensional array of numbers representing the photographic density at every point in the picture. The second step is to compute the Fourier transform of the array, in effect the numerical analogue of its diffraction pattern. The third step is to select from the Fourier transform only the information derived from the regular, repeating features of the object. The final step is to reconstruct the image of the object by combining just this information in a second Fourier-transform calculation: a Fourier synthesis, which reconstructs the original image without its obscuring noise.

Except for the third step the entire process is analogous to the formation of an image in a simple two-lens optical microscope. The first lens creates a diffraction pattern of the object, equivalent to the first Fourier transform, and the second lens projects an image, equivalent to the second transform, or Fourier synthesis. Indeed, we often inspect the micrographs with an optical diffractometer that displays their diffraction pattern (the first transform) in order to check their quality before going on to process them.

To understand the third step, which isolates the genuine information associated with repetitive details from the random fluctuations associated with noise, one needs to consider some properties of Fourier transforms. An image can be conceived of as being built up of sine waves of various wavelengths that differ in their amplitude, orientation and displacement from a common origin. In the Fourier transform of the image each sine wave, or Fourier term, appears as two sharp peaks positioned symmetrically on each side of the origin, at a distance from it that is inversely proportional to the wavelength; each peak has an amplitude and a phase, which respectively indicate the strength and the relative displacement of the wave. The pattern of peaks in the transform therefore reflects the regularity and symmetry in the image.

Because the densities of the atoms constituting the crystal are arranged on a repeating lattice, the first Fourier transform reveals a regular pattern of strong peaks derived from those few sine-wave components that fall on the lattice. It also has peaks derived from noise, but since the random density fluctuations of noise contain all possible sine-wave components, those peaks are present everywhere at random. The first



ALPHA HELIX AND BETA SHEET are two regular conformations of a protein's polypeptide chain. In the alpha helix (*top left*) the backbone of the chain (*gray*) is coiled into a helix, forming a rodlike structure from which the side chains (*R*) of the constituent amino acids project. In the beta sheet (*top right*) two or more backbones are arrayed side by side. In both cases the structure is stabilized by hydrogen bonds (*green*) between NH and CO groups of the backbone. These groups are hydrophilic: in the absence of such bonding to each other they need to be able to form hydrogen bonds with water molecules. One would therefore expect the parts of membrane proteins buried in the hydrophobic interior of the membrane to be largely in the form of alpha helixes or beta sheets (*bottom*); less regular parts of the chain, whose NH and CO groups may not be "protected" by bonds to each other, should be at a hydrophilic surface of the membrane or in the aqueous environment, where the groups can form hydrogen bonds.

If it's worth remembering, it's worth 3M data recording products.

3M computer media products remember everything, every time. Their reputation for reliability is backed up by something very rare in the magnetic media industry: over thirty years of manufacturing experience. That experience, plus the fact that 3M controls the entire manufacturing process, assures that the same high quality is built into every computer media product we make. So you can have complete confidence in every 3M data cartridge, diskette, and roll of computer tape you buy.

Insist on the best for all your DP needs. Insist on information processing products from 3M. Inventor of the data cartridge and the world's largest manufacturer of flexible magnetic media for the home, office and computer room.

Look in the Yellow Pages under computer supplies and parts for the 3M distributor nearest you. In Canada, write 3M Canada, Inc., London, Ontario. If it's worth remembering, it's worth 3M data recording products.

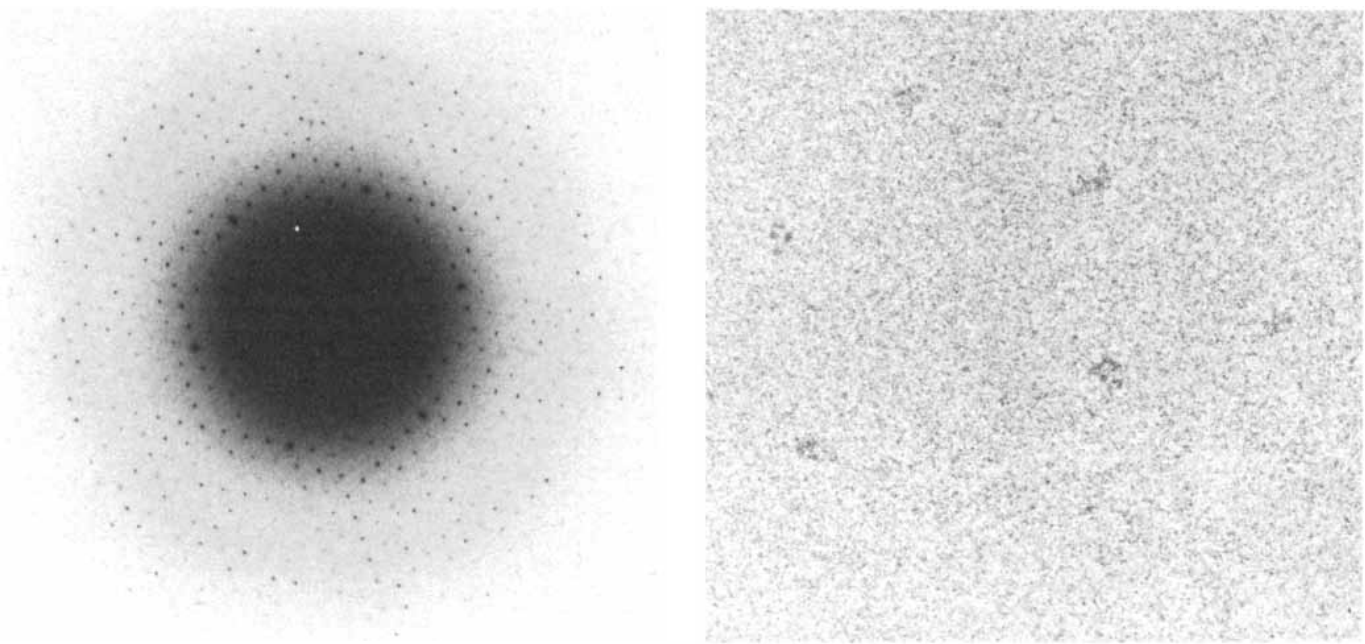
3M hears you...

3M

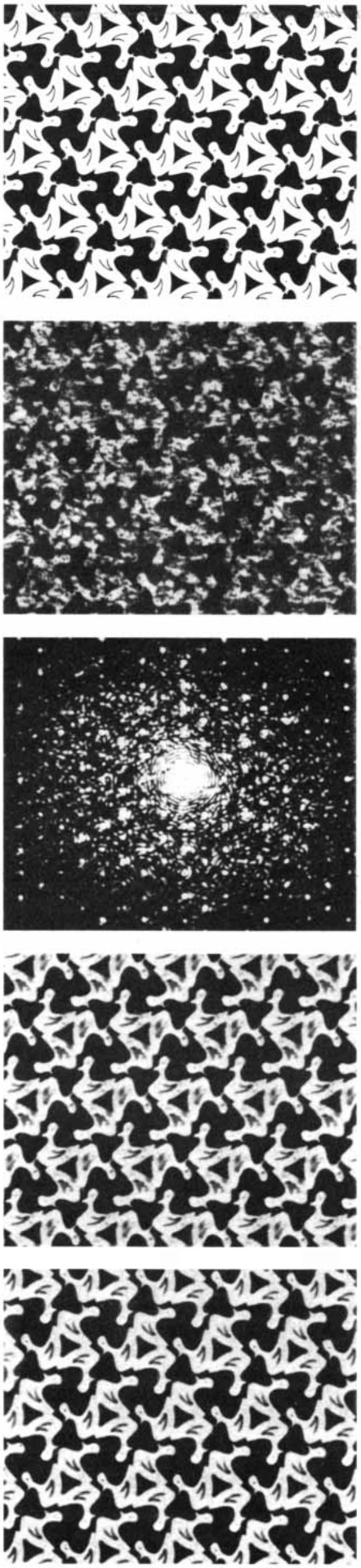
Fourier transform therefore consists of a lattice of strong peaks superimposed on a random background. What happens in the third step is that a computer program selects only those peaks forming a regular pattern and masks out the intervening peaks derived from noise. Only information reflecting genuine detail is included in the final Fourier synthesis that reconstructs the image. In the final step it is also possible to identify and sum just those peaks related by crystal symmetry, thus reinforcing the image of the molecule's true structure.

The idea that one might learn the structure not only of a molecular aggregate but even of a single molecule by analyzing an image made with a weak electron dose had been advanced before the technology was sufficiently advanced to accomplish the task. Robert M. Glaeser of the University of California at Berkeley had done computer simulations that demonstrated the proce-

CRYSTALLINE ARRAYS of bacteriorhodopsin are enlarged some 100,000 diameters in an electron micrograph made by one of the authors (Henderson) and David M. Shotton. To make the micrograph a stack of membranes was prepared by the freeze-fracture method: frozen, split at a shallow angle, shadowed with platinum and replicated in carbon. Membranes treated in this way tend to split along the central plane of the bilayer. Rough-textured regions of the micrograph show the cytoplasmic fracture face, that is, the cleaved surface of the inner layer of the membranes. The surface is studded with a hexagonal lattice of bacteriorhodopsin molecules, which adhered to the inner layer when the membranes were split. Smooth regions show the extracellular fracture face: the cleaved surface of the outer layer.



ELECTRON MICROSCOPE yields both a diffraction pattern and an image of an unstained crystalline array of bacteriorhodopsin. The diffraction pattern (*left*) can be analyzed to give the amplitude of the sine-wave components of the structure; their relative displacement can be learned only from analysis of the image on an electron micrograph. A micrograph (*right*) of an unstained membrane is very different, however, from a picture of a stained membrane, such as the one in the upper illustration: the signal-to-noise ratio is so low that no regular features can be discerned. To extract structural information from such a micrograph it is necessary to turn to Fourier analysis.



STRUCTURAL INFORMATION is separated from noise by Fourier averaging. To illustrate the process a regular lattice of ducks (*top*) was overlaid with a grainy negative and photographed to make a simulated noisy electron micrograph (*second from top*). The photograph was scanned with a densitometer and the array of density values was processed to make a Fourier transform, whose visual analogue is a diffraction pattern (*third from top*). Now the signal of the ducks (*strong peaks on a regular lattice*) is separate from the noise (*random peaks*), whereas the two kinds of information were jumbled together in the photograph. Only the signal information was selected for a second transform (a Fourier synthesis), from which the image was reconstructed (*fourth from top*); noise is suppressed and the duck pattern greatly enhanced. Still clearer pattern (*bottom*) was generated by averaging information in peaks related by lattice symmetry.

dure's potential. It became feasible only after the advent of computers powerful enough to calculate the large Fourier transforms (hundreds of thousands of terms) within a reasonable time and the development of methods for preserving the integrity of unstained membrane specimens in the vacuum of the electron microscope. We found that one way to protect the specimen is to immerse it in a glucose solution. When the water evaporates in the vacuum, some liquid glucose remains, providing a hydrophilic medium that substitutes for the membrane's usual aqueous environment. Another way, developed by Kenneth A. Taylor and Glaeser at Berkeley, is to freeze the specimen in a film of water or salt solution and examine it at a temperature so low that the ice does not sublime away.

Bacteriorhodopsin

We first took advantage of these technical developments to analyze the structure of bacteriorhodopsin, a simple protein found in the cell membrane of the halobacteria. Such bacteria thrive only in water with a very high concentration of sodium chloride: high-salinity lakes and evaporation ponds. The protein consists of a chain of 248 amino acids linked to a single group called retinal. Retinal is the light-absorbing element of visual pigments in animals, including rhodopsin, the "visual purple" of human rod cells. In the halobacterial membrane bacteriorhodopsin functions as a light-driven proton pump that powers an unusual form of photosynthesis [see "The Purple Membrane of Salt-loving Bacteria," by Walther Stoeckenius; SCIENTIFIC AMERICAN, June, 1976]. Bacteriorhodopsin has the unusual property of aggregating within the membrane, forming crystalline patches composed of tens of thousands of protein molecules and hundreds of thousands of lipid molecules packed tightly together in a hexagonal lattice. It is not hard to isolate the patches and prepare them in glucose for direct examination.

Our plan was first to record the electron-diffraction pattern of the crystalline array and then to make images of the array for processing as Fourier transforms. To record the diffraction pattern one simply interrupts the usual series of operations performed by an electron microscope: electrons deflected by atoms in the sample are allowed to fall directly on a photographic film instead of being focused by the microscope's magnetic lens to form an image. The purple membrane proved to be so sensitive to radiation damage that it was impossible to follow the usual procedure of first viewing the diffraction pattern on a fluorescent screen and then recording it; the diffraction pattern vanished before our eyes. We learned to avoid such pre-irradiation and got some clear hexagonal diffraction patterns. The intensities of the spots gave us the strength of the sine waves composing the array of molecules, that is, gave us the amplitudes of the Fourier terms.

To get information on the phase of the sine waves, and thus on their displacement from a common origin, we went on to make an actual image of the molecular array. An unstained patch of purple membrane recorded with a low electron dose yields a featureless micrograph that would dismay an electron microscopist. Viewed in an optical diffractometer, however, such a micrograph yields a sharp crystalline diffraction pattern: the weak periodic detail, spread out in the picture over thousands of repeating units, is concentrated into a much smaller number of sharp points.

Areas of the micrographs calculated to include at least several thousand bacteriorhodopsin molecules were scanned with a microdensitometer. The Fourier transforms of the density values, together with amplitudes derived from the diffraction patterns, were combined to synthesize an averaged representation of the molecular array: a contour map of atomic densities showing the structure of the purple membrane in projection (a vertical view in two dimensions). It was limited by the microscope's imaging properties to a resolution of seven angstroms, that is, features separated by less than that distance were hard to distinguish from each other.

The map showed a striking pattern of peaks, such as would be generated by a set of rods seen end on. The peaks were spaced 10.5 angstroms apart. In other proteins such rodlike patterns seen in contour maps at low resolution had turned out to represent alpha helixes. It looked as if the principal feature of the bacteriorhodopsin molecule must be a set of alpha helixes running roughly perpendicular to the lipid bilayer. Just such an alpha-helical packing of the peptide

IT'S GETTING TOUGHER FOR NANCY RAY TO BE A MOTHER.

Nancy Ray has multiple sclerosis.

Every day it becomes harder for her to see. Harder for her to walk. Harder to do the things she wants to do for her family.

Four years ago, Nancy Ray was young and healthy. She was a mother with a four-year-old son, a teacher, a writer for educational TV.

Then, like about 200 young adults every week, she learned she had MS.

Multiple sclerosis destroys the myelin covering that surrounds nerve fibers in the brain and spinal cord. Messages to and from the brain get lost as they travel through the damaged areas. So your eyes, arms, legs don't do what you want them to do.

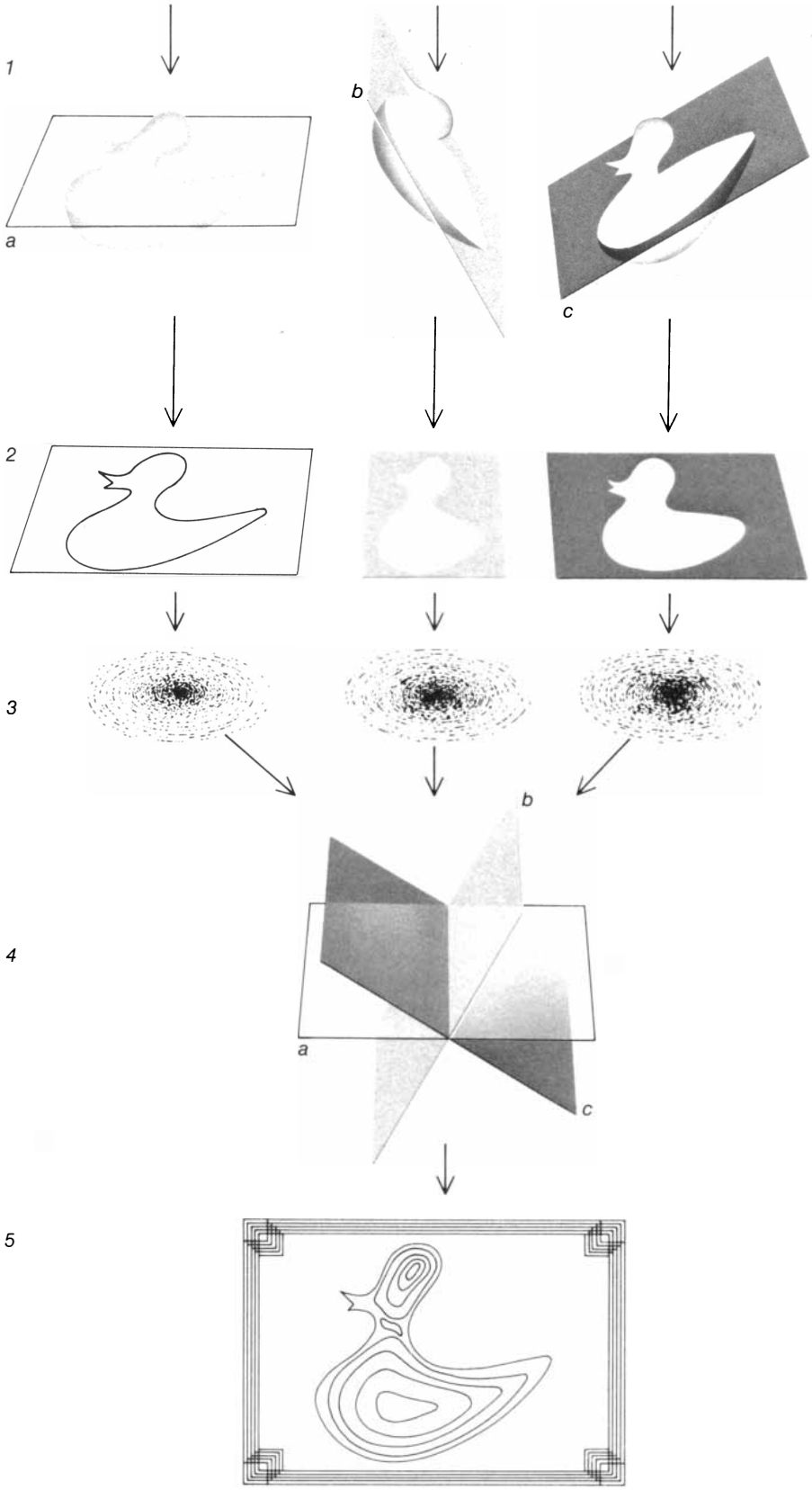
MS can be mild. Or it can be severe. Once you've got it, you've got it for life. Most days, Nancy Ray can get around with a walker. With a lot of time and effort, she can even prepare meals for her family. Some people with MS aren't that fortunate.

The National Multiple Sclerosis Society is helping people with MS live with MS. And funding research for a cure, all over the world. We can do it, but only with your help.

Please give generously. It's tough enough to be a mother without MS.

NATIONAL MULTIPLE SCLEROSIS SOCIETY
205 East 42nd Street
New York, New York 10017

JUST WHEN YOU'RE STARTING TO LIVE, MS CAN STRIKE.



FURTHER PROCESSING is necessary to get a three-dimensional view. Specimens are tilted through a range of angles to the incident beam (*1*). The resulting micrographs constitute a set of different projection views (*2*). From the micrographs Fourier transforms (*3*) are computed, each equivalent to a different plane through the origin of a three-dimensional Fourier transform. From a large number of such planes (only three are indicated in this illustration) the complete three-dimensional transform is computed (*4*). From it a number of contour maps are plotted on transparent plastic sheets, each map showing the density values at a different level in the membrane. When maps are stacked, complete three-dimensional structure is revealed (*5*).

chain had been predicted by one of us (Henderson) and Allen E. Blaurock of King's College London on the basis of an examination of X-ray-diffraction patterns recorded from pellets of purple membrane. The apparent confirmation was pleasing, but some tantalizing questions remained. We could not tell to what extent the helical rods penetrate the bilayer, or how large a part of the molecule they represent, because in addition to the distinct peaks the projection map of each molecule showed a continuous strip of density that defied interpretation. It was clear the analysis had to be carried into three dimensions.

Three-dimensional Structure

DeRosier and Klug had formulated general principles for solving the three-dimensional structure of biological molecules by combining the information contained in a number of different two-dimensional views. In an optical stereomicroscope each eye looks at the specimen from a different angle of tilt and the brain combines the two flat images to create a single three-dimensional impression. Similarly, with an electron microscope one can record two micrographs of an object at different angles and then look at them through a stereo viewer, but this provides too little detail for accurate reconstruction of the object in three dimensions. Fourier theory indicates that such reconstruction is possible, however, if one records not just two views of the object but an entire series of them at slightly different angles. Each view gives a distinct projection of the object, which is then decomposed into its transform so that the Fourier terms of the signal can be separated from the noise. The terms from all the projections are then combined to construct a complete three-dimensional set of terms: a Fourier synthesis that is a three-dimensional representation of the object.

To obtain a perfect three-dimensional representation we would have had to record electron micrographs of the purple membrane at angles of tilt ranging from zero to 90 degrees. This was not possible, but we did manage to record 18 micrographs at angles between zero and 57 degrees, moving the specimen so that each micrograph was of a different patch of membrane, undamaged by the electron beam. The Fourier transforms of these micrographs gave us the amplitudes and phases of the set of Fourier terms. To get better values for the amplitudes we also recorded 15 electron-diffraction patterns at different angles of tilt. Then we combined the information gained from the two methods in a Fourier-synthesis calculation and plotted the results on 10 contour maps, each one representing a different level in the membrane.

The bacteriorhodopsin molecule was

Radio Shack Presents the Transportable TRS-80®

The New Model 4P Computer—From the World's Largest Computer Retailer

The Computer for Today's Upwardly Mobile Society

If you're headed for success, our newest TRS-80 can help you get there in record time. The Model 4P is a compact, disk-based desktop computer with a big difference: it has a handle. So wherever your work may take you, you can get a handle on scheduling problems, ever-changing sales projections, and last minute reports. It's like having your office with you on trips, at home—anywhere!

A Completely Portable, Self-Contained System

We don't sacrifice features for portability, either. The Model 4P comes standard with two 184,000-character disk drives and 64K internal memory. You also get a full 80-character by 24-line 9″ display, an electric typewriter-quality keyboard with numeric keypad, and printer and communications interfaces. And you can add an extra 64K of memory, an internal modem or even external hard disk drives!

"Off the Shelf" Software

In addition to running the entire selection of TRS-80 Model III/4 disk software, you can add the optional CP/M Plus™ operating system to use thousands of additional ready-to-run programs.

Easy to Use, Easy to Tote

The Model 4P works anywhere there's AC. And when you're ready to move on, slip the keyboard into the hi-impact case, snap on the protective cover and go. It weighs just 26 pounds.

Like Every TRS-80, We Stand Behind the "4P" 100%

Tandy Corporation is a two-and-one-half billion dollar company with over 6,600 stores nationwide. That means software, service, leasing plans and friendly faces are always just around the corner.

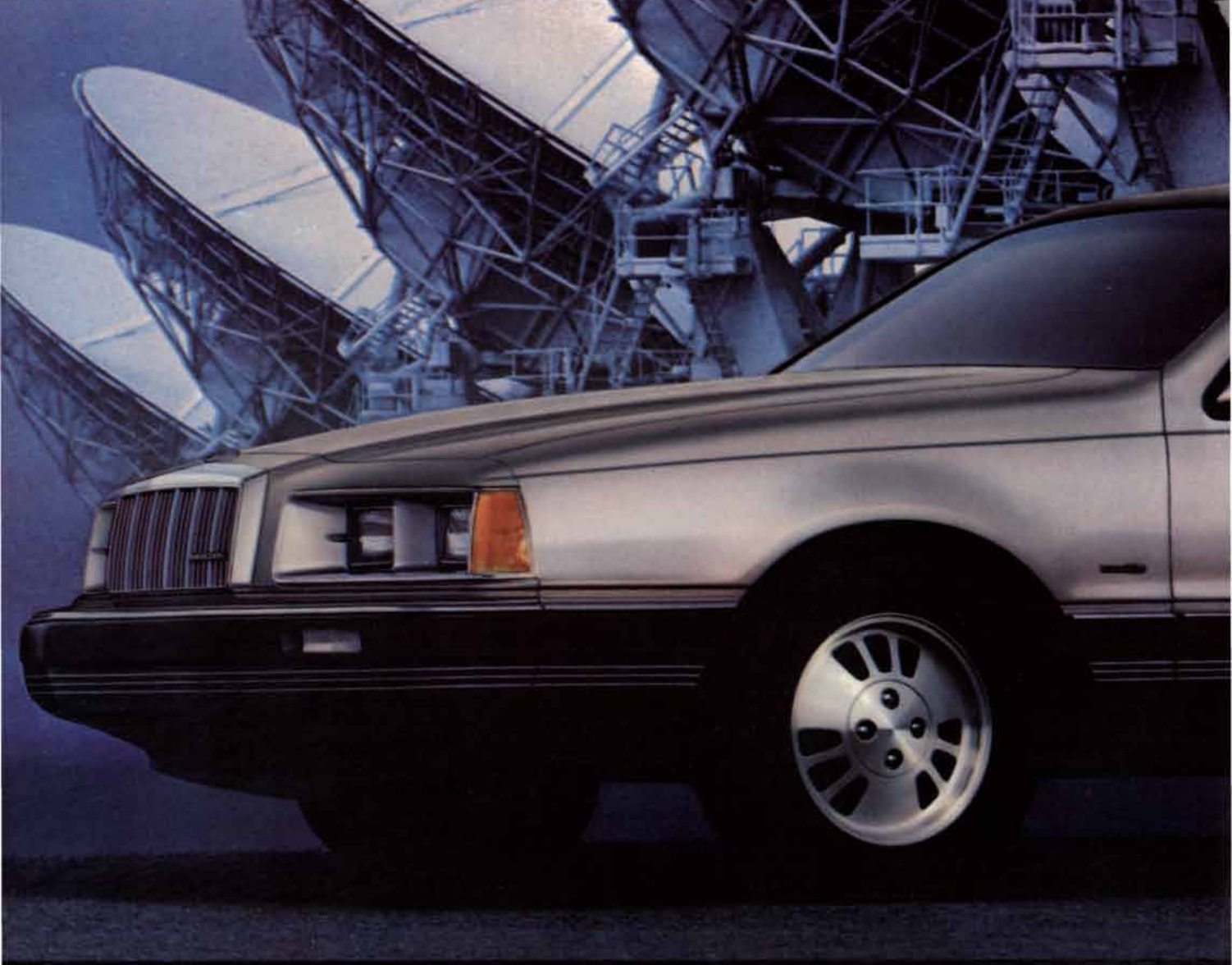
Available Nationwide

Stop by your nearest Radio Shack Computer Center, participating store or dealer today and find out where Model 4P will take you.

Radio Shack®

The biggest name in little computers®

A DIVISION OF TANDY CORPORATION



ONE CAR THAT ASKS TO BE JUDGED ON MORE THAN MERELY ITS ABILITY TO ACCELERATE.

There is a temptation, in talking about a turbo-charged, fuel-injected automobile, to dwell on how well it does against a stopwatch.

But in trying to explain the Cougar XR-7, one realizes immediately how much more there is to be said.

Because this is not a car that offers merely the thrill of acceleration.

Rather, it is a car built with the idea that sophisticated performance has more to do with the impression the car makes on the driver than the one it leaves at a stoplight.

For this reason, we draw your attention to the 60-percent boost in horsepower* the XR-7's turbo delivers and then to the way the XR-7's advanced electronic engine-control system intelligently complements that

Get it together—buckle up.

power, giving you the fuel efficiency** of a 4-cylinder engine most of the time, and the power of a V8 when you want it.

To the advanced suspension system that utilizes 6 gas filled shocks to translate the turbo's exhilarating power into a comfortable, controlled driving experience.

To the choice offered of either an automatic transmission for driving convenience or a 5-speed manual transmission, which involves the driver in every shift decision.

To the carefully detailed interior that boasts such comforting features as functional sport seats that adjust 5 ways to your hips, legs and back.

And to the driver's space that was designed by using studies of the human body to place controls exactly where you can most easily reach them.

In every detail of the Cougar XR-7, the responsive power of its turbocharged engine has been matched with a sophistication of style, handling and ride that you would expect from a very special Cougar.

We invite you to test-drive one.

We think you'll find its performance can be measured with more than just a stopwatch.

Call us toll-free at 1-800-MERCFAX for a copy of our 1984 Cougar catalog.

*As measured by SAE standard J-1349
**28 EST. HWY. [redacted] EPA EST. MPG. Compare to the estimated MPG of other cars. Your mileage may differ depending on speed, weather and trip length. Actual highway mileage less.

MERCURY. A MORE ENLIGHTENED APPROACH.

LINCOLN-MERCURY DIVISION Ford

MERCURY COUGAR XR-7 TURBO

The U.S. Navy will save $24 million on the F/A-18 program in the next nine years as a result of a redesigned electrical equipment rack for the strike fighter's radar. The new rack is far simpler to make and helps improve the aircraft's center of gravity distribution. Bulkheads made of bonded honeycomb panels were replaced with machined aluminum panels, and wraparound honeycomb panels with a riveted aluminum assembly. The rack rail pan's riveted sheet assembly, with its numerous parts, has been replaced by a cast assembly. The modifications reduce costs by $24,000 per unit. Hughes Aircraft Company builds the AN/APG-65 radar under contract to McDonnell Douglas for the Navy and Marine Corps aircraft.

Tests of a prototype ducted rocket engine hold promise of increased range and velocity for future tactical missiles. In milestone demonstrations for the U.S. Air Force, Hughes solid-propellant ducted rocket engines were fired successfully in a wind tunnel simulating supersonic missile speeds at a variety of altitudes. Whereas conventional air-launched motors contain all the fuel and oxidizer they need for combustion, the ducted rocket obtains a large portion of its oxygen from the atmosphere. An important benefit is that the ducted rocket motor can contain more fuel for a given weight.

Space shuttle crews will soon be able to rendezvous with satellites in low orbit in order to repair or recover them, thanks to a new integrated radar and communications subsystem that passed its first tests on the shuttle last June and August. The Hughes subsystem, also called a Ku band radar, has an antenna dish at the front of the cargo bay. It can pinpoint objects as small as 1 square yard from up to 14 miles away. If the object is equipped with an electronic signal enhancer, the range increases to 345 miles.

NATO early-warning aircraft are being equipped with a communications system that uses four primary encoding techniques to hamper enemy eavesdropping or jamming. The Joint Tactical Information Distribution System (JTIDS) provides E3A AWACS aircraft and NATO ground command centers with secure voice and digital communications. One JTIDS encryption technique is spread spectrum, in which a signal is expanded over a large bandwidth. With frequency hopping, a second method, frequencies are changed many times a second. Another technique, time division multiple access, assigns certain users to specific time slots no longer than a fraction of a second. Finally, to verify messages, JTIDS repeats messages automatically. Hughes is supplying JTIDS to NATO and the U.S. Air Force.

A system that provides three-dimensional positioning information to military field commanders is in production. The Position Location Reporting System (PLRS) gives commanders continuously updated information on the location of troops, vehicles, helicopters, fixed-wing aircraft, and weapons. The system links small portable radios into a network controlled by microprocessors. Users can locate their positions, navigate to predetermined positions, and be informed when nearing or crossing boundaries. Join our high-tech team. Send your resume to Hughes Ground Systems Group, Employment Dept. SE, P.O. Box 4275, Fullerton, CA 92634.

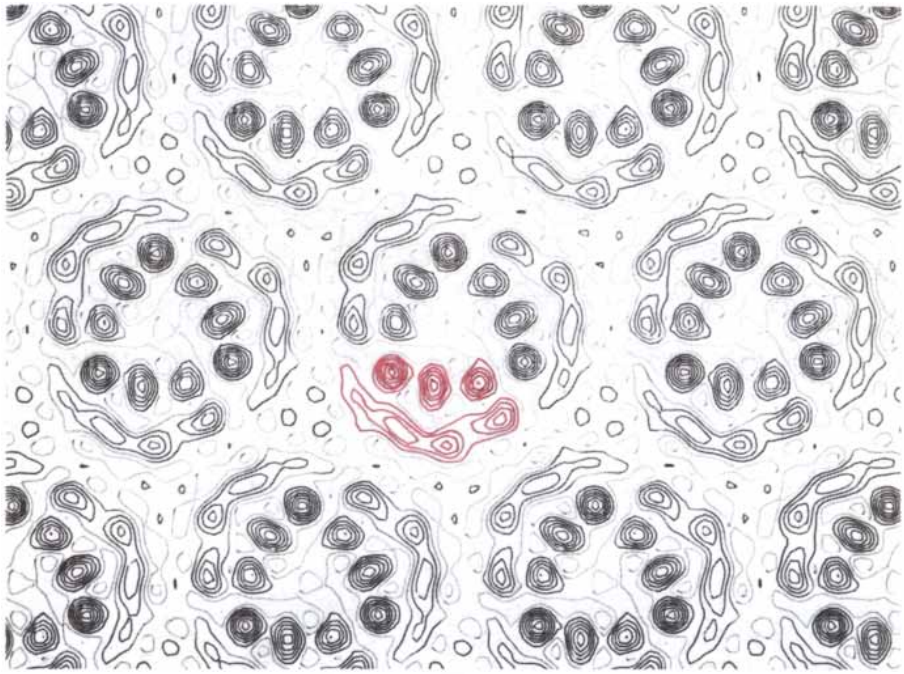
seen to be composed of seven rods of density, presumably alpha-helical segments of the polypeptide chain, extending all the way through the membrane. Three of the rods are indeed perpendicular to the plane of the bilayer, as we had suspected from the projection view. The other four rods are tilted slightly, so that various cross sections of their density had been partly superimposed in the projection view and the rods had therefore been indistinctly resolved. The direction in which the rods are tilted is consistent with theoretical arguments for the side-to-side packing of alpha helixes that were put forward as long ago as 1953 by F. H. C. Crick.

The stretches of the polypeptide chain that connect the seven rods are not visible at a resolution of seven angstroms. We could, however, estimate the total length of the nonhelical segments by subtracting the number of amino acids necessary to account for the seven helixes from the 248 amino acids known to constitute the complete protein. The difference, some 70 amino acids, suggested that the nonhelical stretches are short links connecting the helixes at the two surfaces of the membrane. This impression was confirmed when Yu. A. Ovchinnikov of the M. M. Shemyakin Institute of Bio-organic Chemistry in Moscow and H. Gobind Khorana's group at the Massachusetts Institute of Technology determined the amino acid sequence of bacteriorhodopsin. They found seven segments of generally hydrophobic residues separated by short hydrophilic segments, a sequence that is likely to correspond to seven alpha-helical rods set into the membrane with connecting links at the two surfaces.

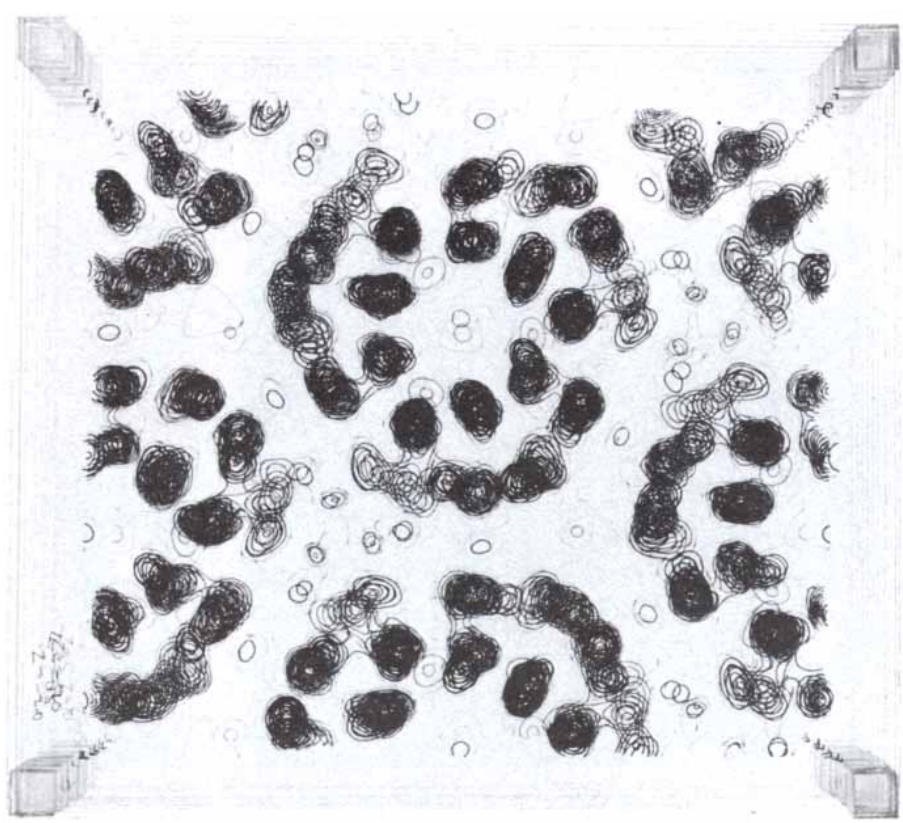
The available chemical and structural data fall short of providing a complete description of the bacteriorhodopsin molecule and of the mechanism by which it pumps protons. For that one needs to know the exact positions of the individual amino acids and of the retinal group. Our structural findings do, however, give confidence that the general principles derived from X-ray studies of water-soluble proteins apply also to the lipid-soluble proteins residing in membranes.

Connexons

Adjacent cells need to communicate with one another if they are to coordinate their metabolism and otherwise work together in a tissue. They communicate by means of short intercellular "pipes," which are arrayed in disk-shaped patches at what are called gap junctions. Each pipe is composed of two roughly cylindrical proteins called connexons. Each connexon is embedded in one of the apposed cell membranes and extends into the intercellular space, where it is linked with an abutting con-



TWO-DIMENSIONAL contour map derived by Fourier averaging shows the crystalline array of bacteriorhodopsin and lipid molecules in projection. A single protein molecule is delineated in color. The contours connect points of equal electron density; since the values come from a single overhead view of the sample, the density values at every level in the sample are summed. As a result the only readily identifiable structures are those that are perpendicular to the membrane: three distinct peaks of density (*circular contours*) in each bacteriorhodopsin molecule are suggestive of three alpha helixes seen end on. The molecule also includes a continuous strip of density (*ridgelike contours*), however, that cannot be interpreted straightforwardly.

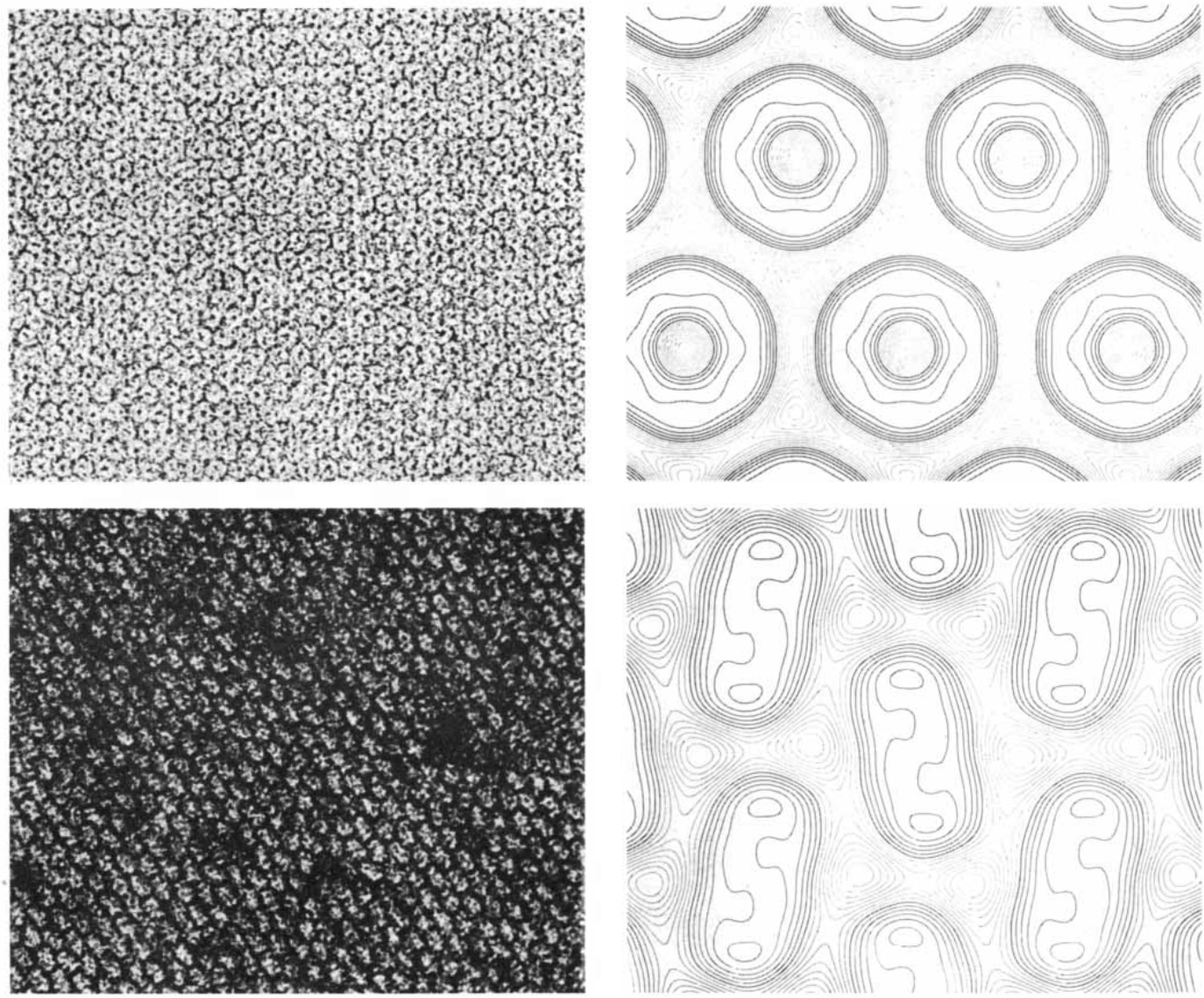


THREE-DIMENSIONAL VIEW of the array of bacteriorhodopsin and lipid molecules is revealed by a stack of contour maps representing density levels at different heights. Now the ridgelike contours of the projection view are resolved into four separable sets of roughly circular contours. They represent four more alpha helixes, tilted from the perpendicular. Each bacteriorhodopsin molecule, then, has seven alpha helixes. The computer drawing on page 79 is a smooth-surface representation based on the contours that delineate a single density level.

nexon to form a continuous pathway between the cells' interiors. A connexon is made up of subunits, each of which has about the same molecular weight as a single molecule of bacteriorhodopsin. The subunits are arranged to form a central channel: the lumen of the intercellular pipe. By microinjecting fluorescent dyes of various sizes into a single cell and observing their penetration into nearby cells, Werner R. Loewenstein and his colleagues at the University of Miami School of Medicine determined that the diameter of the channel is from 15 to 20 angstroms, large enough for the passage not only of small ions (charged atoms) but also of certain nutrient molecules and such messenger molecules as cyclic AMP.

By isolating gap junctions in a detergent solution one can obtain crystalline hexagonal arrays of connexons in two apposed sheets of membrane, with abutting connexons still linked to each other. These arrays have fewer repeating units than the purple-membrane patches do, so that they can be analyzed by electron microscopy only at a rather low resolution. We visualized specimens with two complementary techniques. One technique was negative staining, which outlines only the external, hydrophilic surfaces. The other was the method, mentioned above, in which the membrane is frozen in a thin film of water. Differences in the density of protein, lipid and ice provide enough contrast to reveal the internal details of the gap-junction membrane, and the density difference between protein and water (which is large compared with the difference between protein and glucose) reveals the parts of the protein that protrude from the bilayer. For gap junctions the advantages of a truly aqueous environment outweigh the difficulty of keeping the specimen stably frozen in the electron microscope's vacuum.

In pictures of negatively stained gap junctions the connexon looks like a ring surrounding a central region in which the stain is concentrated: the channel. To examine the arrangement of the subunits around the channel Guido A. Zampighi of the Duke University Medical Center and one of us (Unwin) applied the image-analysis and reconstruction procedures described above. In a projected view at a resolution of 18 angstroms the contour lines defining the



TWO LARGER PROTEINS studied by the authors and their colleagues are the connexon (*top*) and cytochrome *c* oxidase (*bottom*). Their micrographs (*left*) and projection maps (*right*) are shown here. The micrographs, of negatively stained samples, have more contrast than the micrograph of unstained bacteriorhodopsin, but because there are fewer repeating units of these larger proteins in a given area, the resolution of their maps is lower: only about 18 angstroms in the case of the connexon and about 25 angstroms for the cytochrome *c* oxidase. The heavy contours in the connexon map suggest a six-lobed ring surrounding a central region of lesser density (*gray contours*): a channel through the membrane. The map of cytochrome *c* oxidase suggests that the protein is present as a dimer, with the two molecules related by a twofold axis of symmetry. Three-dimensional contour maps of the two proteins revealed further details of their structure.

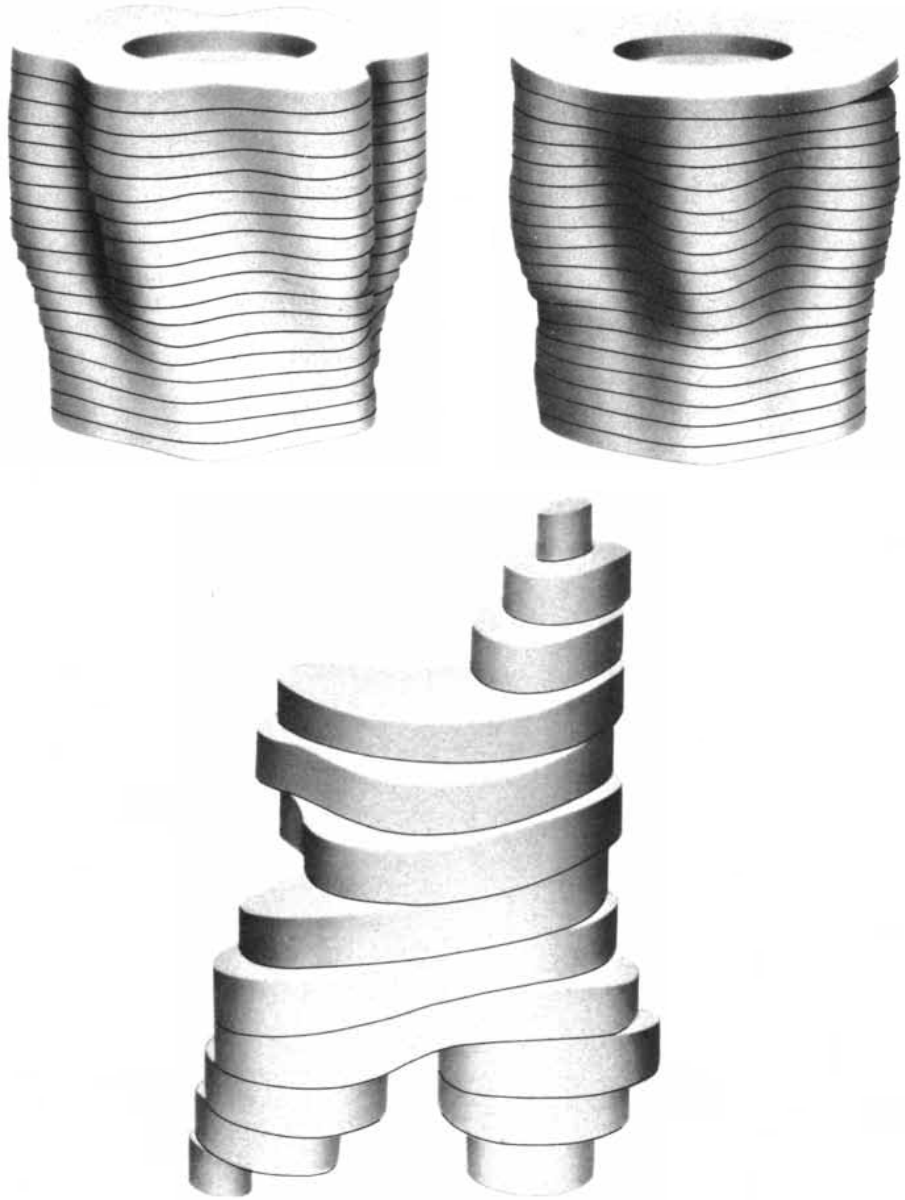
ring were seen to have six lobe-shaped modulations suggestive of a channel surrounded by six protein subunits. The projected view was complicated, however, by the superpositioning of connexons embedded in the two apposed membranes. Once again a three-dimensional map was necessary to reveal the structure's configuration.

The combination of the information from many images made at different angles showed that there are indeed six subunits. They protrude from both surfaces of the membrane, extending less than 10 angstroms into the cytoplasm of the cell and between 15 and 20 angstroms outside the cell, where they meet the directly apposed connexon. The subunits tilt in opposite directions on opposite sides of the channel, indicating that they are skewed at an angle to its axis. The direction of tilt is the same as that of the alpha-helical elements of bacteriorhodopsin. The channel itself is seen to taper somewhat: it is perhaps 20 angstroms wide near the extracellular end and narrow enough to be incompletely resolved near the cytoplasmic end.

The Gating Effect

In live tissue the gap-junction channels are known to open and close in response to changes in intracellular conditions. This gating effect is crucial to the life of tissues; among other things, it seals off cells that die, preventing the loss of vital nutrients from adjacent healthy cells. The molecular mechanism of the opening and closing was suggested by some of our gap-junction images, which showed what appeared to be a form of connexon different from the one we had first analyzed: the central channel seemed to be less clearly defined. The three-dimensional map derived from these images revealed a change in the configuration of the subunits. They were less tilted and came closer together at the cytoplasmic end, constricting the channel there. The change seemed to be the result of small tilting and sliding motions of the subunits along their lines of contact. It occurred to us that such displacements, somewhat like those that open and close the iris diaphragm behind a camera lens, might be responsible for the opening and closing of the channels in live tissue.

At the Stanford University School of Medicine, Peter D. Ennis and one of us (Unwin) recently put that notion to the test. The opening and closing of gap junctions has been shown to depend on the intracellular level of calcium ions. We applied the freezing technique to reveal the protein's structure both in the absence of calcium ions and in their presence. The effect of the calcium was indeed to straighten the subunits and constrict the channel, confirming that just such a rearrangement quite likely does regulate the passage of molecules between the interiors of live cells.



SCHEMATIC DRAWINGS of two connexons (*top*) and of a single cytochrome *c* oxidase molecule (*bottom*) are based on models derived from their three-dimensional maps. The connexon was studied both in a calcium-free medium (*left*) and in the presence of calcium (*right*). The calcium causes the six tilted subunits to rotate and become more nearly perpendicular to the plane of the membrane, thereby apparently constricting the central channel. The cytochrome *c* oxidase monomer is a large *Y*-shaped molecule. The stem of the *Y* protrudes from the mitochondrial inner membrane and is responsible for the enzyme's binding to cytochrome *c*.

We still do not know much about the internal detail of a connexon assembly. Some findings have been reported, however, on the basis of X-ray studies. Lee Makowski, Donald L. D. Caspar, Daniel A. Goodenough and Walter C. Phillips of Brandeis University have recorded X-ray-diffraction patterns (from pellets of gap junctions) suggesting that, unlike bacteriorhodopsin, the connexon's polypeptide chains have some beta-sheet regions rather than being nearly all alpha-helical. We hope to make crystalline arrays that have a larger number of repeating units than the samples we have so far analyzed. This would enable us to improve the resolution of our maps while reducing the exposure of the specimen to the electron beam and thereby overcoming the problem of radiation damage. Maps with higher resolution should reveal the configuration of the chains in greater detail.

Cytochrome *c* Oxidase

The third membrane protein whose three-dimensional structure is beginning to be understood is cytochrome *c* oxidase, an enzyme that consists of at least seven polypeptide chains and is about

six times as large as bacteriorhodopsin. It resides not in the cell membrane but in the inner membrane of the mitochondrion, the energy-transducing organelle of the cell, where it catalyzes the oxidation of cytochrome *c*. In the reaction four protons from the interior of the mitochondrion (the matrix) combine with four electrons from cytochrome *c* molecules (on the outside of the inner membrane) and one oxygen molecule to yield two molecules of water. The removal of protons from the interior of the mitochondrion creates a proton gradient across the inner membrane, and that gradient powers the production of adenosine triphosphate (ATP), which supplies energy for metabolism.

The membranes of living mitochondria do not contain crystalline arrays of cytochrome *c* oxidase, but such arrays can be made by treating mitochondria with detergents that extract some of the components selectively and concentrate other components. Roderick A. Capaldi of the University of Oregon found that two such arrays contain only cytochrome *c* oxidase. James F. Deatherage and one of us (Henderson), together with Capaldi and Stephen D. Fuller, have analyzed the two arrays. To look at the parts of the proteins extending outside the membrane we treated the specimens with negative stain; to see the parts embedded in the membrane we applied a glucose film.

The seven polypeptide chains of cytochrome *c* oxidase are different in size and have different amino acid sequences, and the protein's chemical asymmetry is reflected in its shape. The part of the structure that is largely embedded in the lipid bilayer is divided into two clearly resolved domains, which protrude less than 20 angstroms into the matrix. These domains are connected to a larger third domain, which protrudes about 55 angstroms from the opposite (exterior) surface of the inner membrane, the same side that reacts with cytochrome *c*. The overall structure is that of a lopsided *Y*, with the arms mostly embedded in the lipid bilayer and the stem protruding into the intermembrane space. We were fortunate in that the protein molecules were arranged differently in our two crystalline arrays, enabling us to trace the boundaries between neighboring molecules. In one array, which is thought to reflect the state of the enzyme in its normal environment, the protein is present as a dimer, with the two molecules of the dimer related by a twofold axis of symmetry.

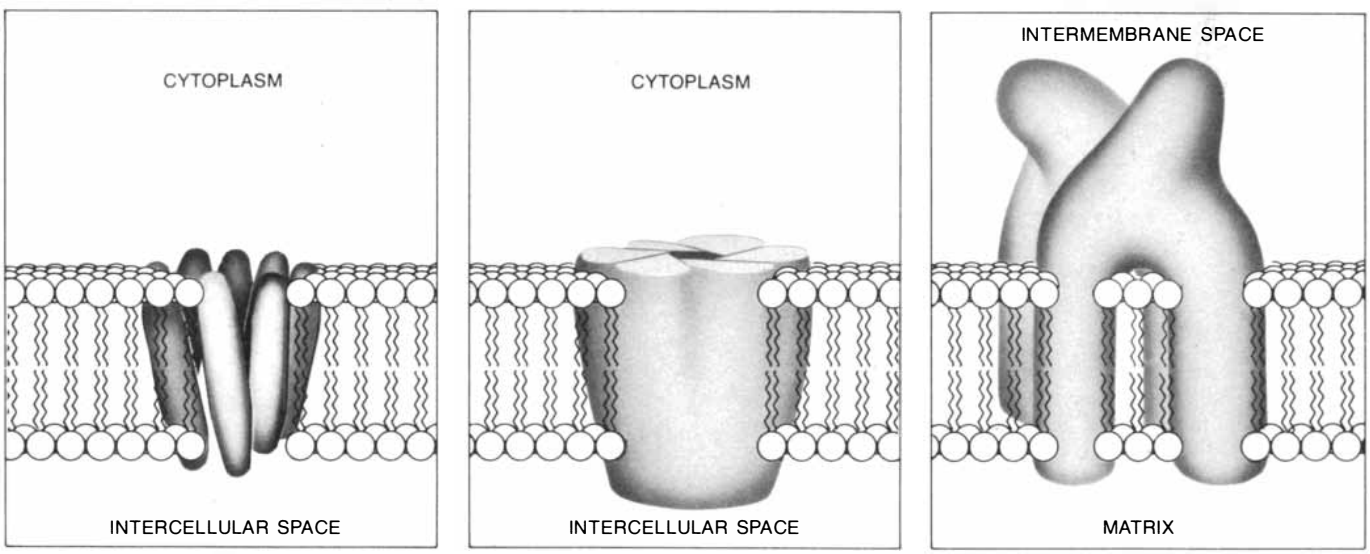
As in the case of the gap-junction crystals, the resolution of our picture of cytochrome *c* oxidase is limited by the relatively small number of repeating units in the two-dimensional crystal. The oxidase is a particularly large protein: a dimeric complex of seven different polypeptide chains with a total molecular weight of 330,000. It therefore seems likely that the high-resolution structure of the two smaller proteins will be determined before that of cytochrome *c* oxidase.

Membrane-Protein Structure

What do these preliminary three-dimensional studies tell one about the structure of membrane proteins in general? They indicate that the folding of such proteins' polypeptide chains and other aspects of their design are similar to those of water-soluble proteins. The chains probably fold to form regular groups of alpha helixes and beta sheets in the parts of the proteins that traverse the hydrophobic interior of the bilayer, and they probably form irregular bends and turns outside the membrane. The groups of helixes and sheets probably coalesce into compact, globular molecules that vary in size, shape and number of polypeptide chains. Several such molecules may associate within the plane of the membrane to create composite structures. Such structures would be comparable to certain water-soluble enzymes, consisting of several subunits, that undergo small "cooperative" rearrangements in response to specific stimuli.

Such structural features will have to be described in detail if one is to understand how membrane proteins work. Improvements in electron-micrographic techniques, combined with amino acid sequence data, may soon make it possible to map at least some of the proteins in atomic detail, which is to say at a resolution of perhaps two angstroms.

There has also been progress in the X-ray diffraction of membrane proteins. Hartmut Michel of the Max Planck Institute for Biochemistry in Munich has developed a new method for crystallizing such proteins in three dimensions by treating membranes with detergents of low molecular weight; Jürg Rosenbusch and Michael Garavito of the Biozentrum in Basel have devised a related method. The ability to prepare large three-dimensional crystals of membrane proteins means that the powerful established techniques of X-ray diffraction can be applied to analyze the molecules' structure at atomic resolution. The investigators in Munich and Basel, working with two quite different membrane proteins, have already made excellent progress toward that goal.



MEMBRANE PROTEINS whose three-dimensional configuration is best understood are shown at the same scale. Bacteriorhodopsin (*left*) is a monomer; the connexon (*center*) is a hexamer of identical subunits, and cytochrome *c* oxidase (*right*) is a dimeric complex each half of which has seven polypeptide chains. The first two proteins are in cell membranes; third one is in inner membrane of mitochondria.

Today's Chevrolet

Move into more. New Chevrolet Celebrity.

Four or five years ago, your smaller car seemed enough car. But now you're ready to move into more.

More room. More comfort. More engine. And more than just a little more prestige. For which we introduce our smooth, quiet, front-drive Chevrolet Celebrity.

Move into Celebrity comfort, Celebrity convenience. Comfort that begins with more passenger room than 91 out of 95 import cars. More passenger room, in fact, than Chrysler LeBaron. And more trunk space than Mercury Marquis.

More? Turn the key. Celebrity's standard electronically fuel-injected engine gives you more standard cubic inches than any front-drive Datsun, Toyota, Honda, Mazda, Ford, Mercury or Dodge.

More excitement. And front drive. Exemplified by our new Celebrity Eurosport shown here. All the virtues of Celebrity, wrapped in head-turning European styling, mounted atop a Special F41 Sport Suspension. All at a price that will make Europe's best blush with envy. And you with pride.

Isn't it time you moved into more car? Especially when it doesn't cost that much more. In fact, try to find any other front-drive car with more passenger room at a lower sticker price.

You'll see what we mean by Chevrolet–Taking Charge.

OFFICIAL U.S. CARS AND TRUCKS OF THE XIV OLYMPIC WINTER GAMES

Sarajevo '84

Some Chevrolets are equipped with engines produced by other GM divisions, subsidiaries, or affiliated companies worldwide. See your dealer for details.

1984 Celebrity Sedan shown with available Eurosport Equipment Package, Aluminum Wheels and Eagle GT tires.

The Spectroscopy of Supercooled Gases

A new technique for cooling a gas far below the normal temperature of condensation can dramatically simplify its spectrum. The result gives a highly resolved picture of complex molecular energy states

by Donald H. Levy

Chemicals and their reactions have traditionally been studied by making bulk measurements that describe the average properties of a relatively large amount of material. Such quantities as the total amount of heat liberated in a reaction, the amount of a given chemical that is generated and the time required for the reaction to yield a given amount of the chemical are all properties of macroscopic aggregates of matter. Although the determination of these data remains important, modern physical chemistry is increasingly concerned with what is going on at the molecular level. Many experiments are now designed to study the structure of the individual molecule, the mechanism by which its energy is stored and the probability that it will react with another molecule when the two molecules collide. Such investigations probe both the static and the dynamic properties of the molecule.

One of the most powerful tools for examining the microscopic properties of a substance is molecular spectroscopy. Matter does not uniformly absorb light of all frequencies; on the contrary, certain frequencies are strongly absorbed and others are not absorbed at all. Similarly, matter in an excited state can fluoresce, or emit light of certain frequencies. The amount of light absorbed or emitted by a collection of molecules, plotted against the frequency, or the color, of the light, is the molecular spectrum of the substance. By measuring, analyzing and interpreting the frequencies present in a molecular spectrum it is possible to infer a great deal about the physical properties of the molecule responsible for the spectrum.

In principle the more complex the molecular spectrum is, the more information about the molecule it contains. In practice, however, the spectrum of a compound can be so complex that analysis is impossible. For example, the absorption spectrum at visible-light frequencies is a collection of dark lines spaced at various intervals against the familiar rainbow of colors from red to violet. The dark lines represent the absorbed frequencies of light, and the rest of the spectrum is made up of transmitted light, which passes through the compound without being absorbed. For many substances the dark spectral lines are so dense that they cannot be separated even with instruments of the highest resolving power available; indeed, some lines are so close together that their resolution is theoretically precluded by the uncertainty principle. Hence any experimental technique that can reduce the number of spectral lines and thereby simplify molecular spectra can be extremely useful. Although it sounds contradictory, what has been sought is a means of decreasing the information in the spectrum so that the remaining information can be understood.

For the past several years my colleagues and I at the University of Chicago have been developing a technique, based on the properties of gases at very low temperatures, whereby the number of spectral lines absorbed or emitted by gaseous substances can be systematically reduced. Our strategy is simple. Every dark line in the absorption spectrum represents the absorption by the substance of at least one photon, or quantum of light; the absorption of a photon is accompanied by a transition of the substance from its initial energy state to a higher, more energetic state. Similarly, in fluorescence the emission of a photon marks the transition of the substance from some initial energy state to a lower, less energetic one. The laws of quantum mechanics require that the frequency of the photon, or in other words the position of its line in a spectrum, be proportional to the difference in energy between the initial energy state and the final one.

If the molecules that make up a substance are found in many different initial energy states, the number of possible energy transitions, and therefore the number of lines in the spectrum, can be extremely large. On the other hand, if the number of initial energy states is limited, the number of possible transitions is reduced and the spectrum that results can be analytically tractable. The most practical way to limit the number of initial energy states is to lower the temperature of the substance, because at low temperatures only the states of lowest energy are populated. Hence our strategy for reducing spectral complexity is to measure the spectrum of the substance at temperatures within one degree of absolute zero.

If the only requirement for high-resolution spectroscopy were a conventional low-temperature refrigerator, there would be no need for our new technique. At low temperatures, however, molecules tend to condense into the liquid or the solid state, and the interactions of the molecules with one another bring about significant changes in the spectrum. There can be little doubt, given the remarkable advances in electronics brought about by an understanding of the solid state, that the spectroscopy of condensed (liquid or solid) matter is an investigative technique of great value. Nevertheless, it is also important to study the properties of individual molecules, free from interactions with their surroundings. Such study is best carried out on the gas phase. The spectral lines identified with the absorption or emission of light by a low-pressure gas arise, to a good approximation, solely from transitions in the energy states of the individual molecules of the gas. The problem, therefore, is to make a gas cold enough to limit the number of initial energy states without allowing the gas to condense and form a liquid or a solid.

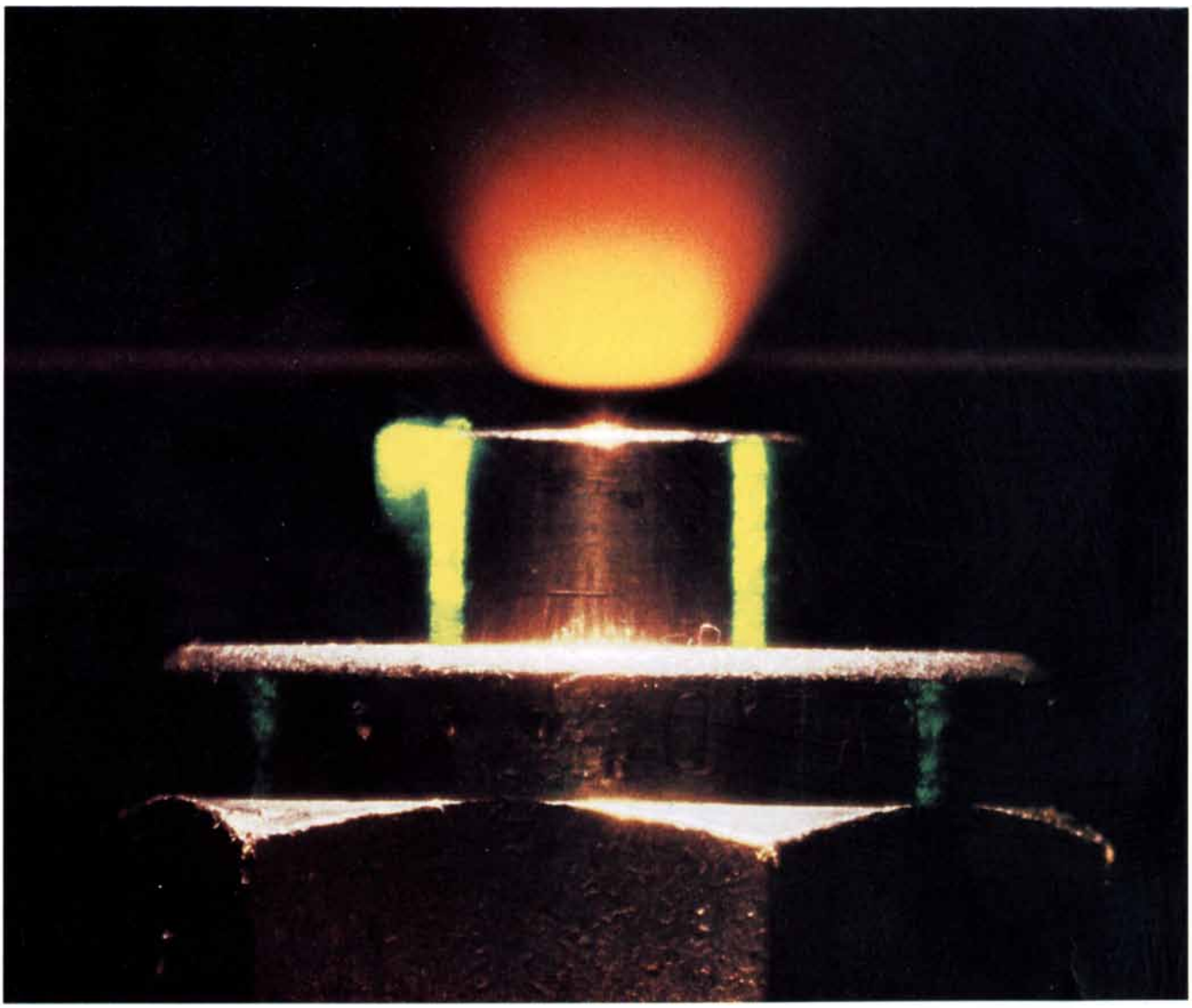
At this point it will be useful to consider the physical properties that are responsible for the spectral lines. Accord-

ing to the laws of quantum mechanics, energy cannot be added to a molecule in arbitrary amounts. If energy is to be added to or taken from the molecule, it must be sufficient to bring the molecule to its next higher or its next lower energy state; otherwise the initial energy state cannot be changed at all.

Electromagnetic radiation is absorbed or emitted in the discrete, indivisible unit called the photon. The photon is best described in some contexts as a particle and in others as a wave; as a wave it has both a frequency and a wavelength. Remember that the energy of the photon is directly proportional to its frequency, and so the absorption or emission of a photon signals the transition of a molecule from one energy level to another. A molecular spectrum has an enormous range of frequencies. Low-energy transitions between closely spaced energy levels are signaled by radio-frequency photons, transitions between levels of intermediate spacing by infrared photons and transitions between widely spaced levels by visible-light, ultraviolet and X-ray photons. It is possible to associate various kinds of molecular motion or degrees of freedom with absorption or emission in particular regions of the spectrum.

The motions of the heavy, slow-moving atomic nuclei and the motions of the light, fast-moving electrons that make up a molecule can be considered independently. The nuclear motions are so slow that the electrons can almost always adjust to a changing configuration of nuclei, and so the electronic orbits about moving nuclei are essentially the same as they would be if the nuclei were at rest. Furthermore, the nuclear motions can be divided into two distinct types: the end-over-end rotation of the molecule as a whole and the small vibrations of the nuclei about their equilibrium positions. These two motions also are largely independent. The rotational motion is close to what it would be if the molecule were rigid, and the vibrational motion is approximately the same as it would be in a nonrotating molecule.

The energy transitions associated with change in the rotational motion of a molecule can be quite small, and so the transitions between adjacent rotational energy levels are induced by photons of very low frequency, in the radio or mi-

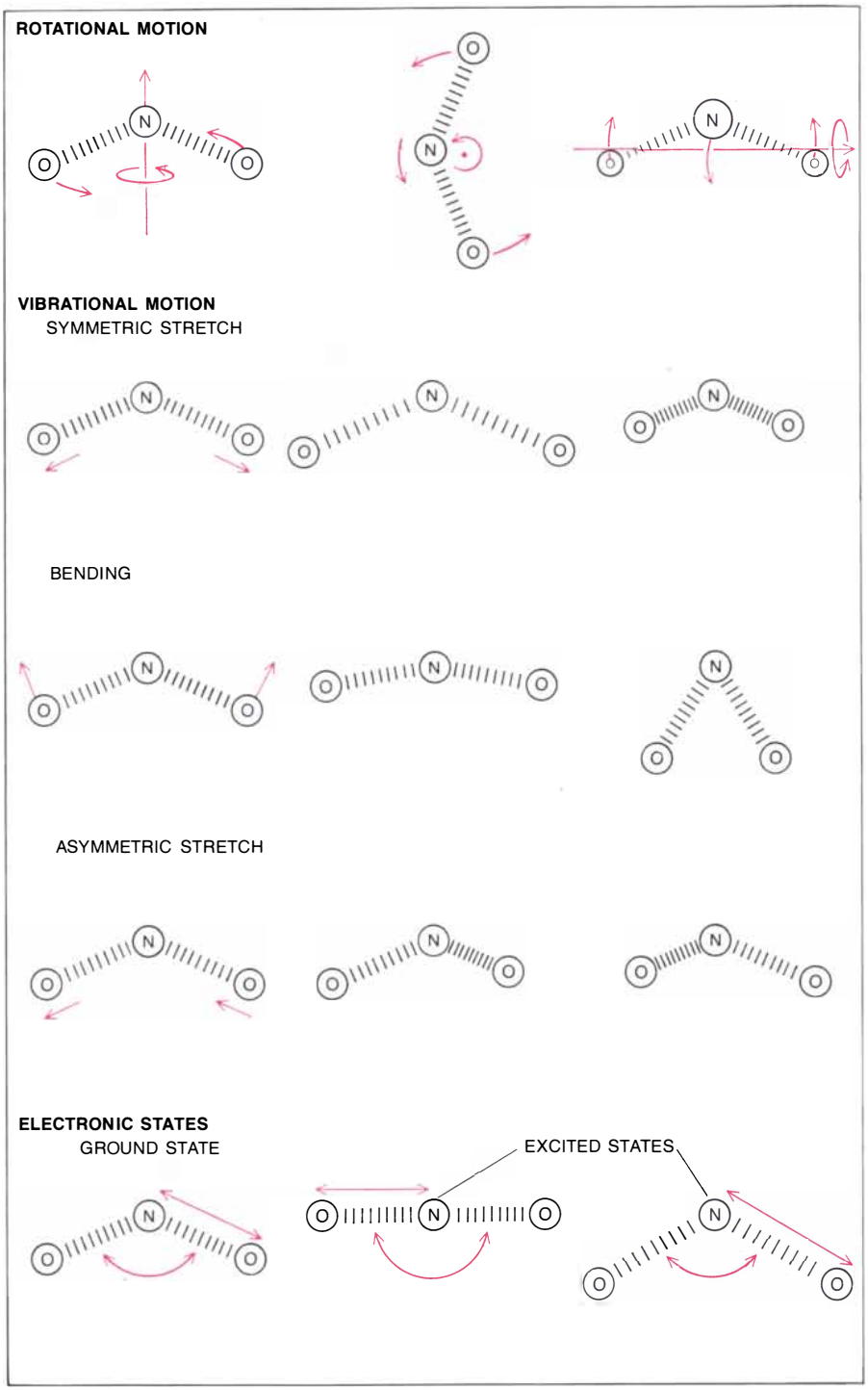


SUPERSONIC FREE JET of helium gas seeded with a small amount of molecular iodine (I_2) fluoresces a bright yellow-orange during exposure to the energizing radiation of a laser. The path of the laser beam is indicated by the thin line of red fluorescent light created by the beam as it passes through the residual gas surrounding the jet. The gas in the jet emerges from a high-pressure container through a small hole about .05 millimeter in diameter and expands at supersonic speed into an evacuated chamber, where it is exposed to the laser beam. The rapid expansion of the jet cools the gas and simplifies the spectrum of the fluorescent-light emission. Laser-induced fluorescence spectroscopy carried out on supersonic free jets of gas has become a deep and powerful probe of molecular structure and dynamics. The photograph was made by Fritz Goro in the author's laboratory at the James Franck Institute of the University of Chicago.

crowave regions of the spectrum. The spacing of the rotational energy levels is determined by the moments of inertia of the molecule, and each moment of inertia is in turn determined by the configuration of the atomic nuclei in the molecule. The analysis of the pattern of rotational energy levels can therefore lead to a determination of the bond lengths and the angles between the molecular bonds.

The vibrational motion of a molecule can be quite complex, and the larger the molecule, the greater the complexity in the energy-level pattern. The spacing of the vibrational energy levels is usually greater than the spacing of the rotational energy levels, and the photon frequencies associated with vibrational transitions lie in the infrared region of the spectrum. Infrared spectroscopy has therefore been the traditional tool for unraveling the structure of vibrational energy levels. The vibrational frequencies are determined by the strength, or stiffness, of the restoring forces that hold the atomic nuclei in place, and so the vibrational spectrum provides information about the forces responsible for chemical bonds.

The lightest particle associated with transitions in molecular energy is the electron, and the separation between adjacent electronic energy levels is greater than the separations associated with the rotational or the vibrational motions of the nuclei. Accordingly the transitions between electronic energy states are induced by photons in the visible-light or ultraviolet regions of the spectrum. Chemical properties of a molecule such as the configuration of its atomic nuclei, the strength of the bonds it can form and its chemical reactivity are determined by the distribution of the electrons, particularly the valence electrons: the outer electrons in the molecule. Hence there is a sense in which different electronic states of the same molecule can be considered different chemical species.



MOLECULAR ENERGY is stored in the rotational or the vibrational motion of the molecule as well as in the configuration of the orbiting electrons. Here the three rotational modes and the three vibrational modes of the nitrogen dioxide molecule (NO_2) are shown. The amount of stored rotational energy depends on the rotational velocity and on the moments of inertia of the molecule about the principal axes of rotation. The electronic energy structure can also alter the configuration of the nuclei. The orbitals of the electrons have been omitted for clarity.

It is possible for a single photon to cause changes in more than one kind of energy state. For example, an infrared photon can change both the rotational and the vibrational state of a molecule: most of the energy is expended in the vibrational excitation, but a small fraction of it causes rotational excitation as well. Thus each vibrational state is split into a large number of energy levels representing the possible rotational states of the molecule for the given vibrational state. An energy transition from another vibrational state to one of these levels corresponds, in general, to a vibrational transition and a rotational transition. Similarly, a visible-light photon or an ultraviolet photon can cause electronic, vibrational and rotational transitions all at once. An electronic energy state is made up of a series of vibrational energy states, each of which is further split into rotational energy levels.

The energy transitions accessible to a molecule in a given energy state can thereby give rise to tens or hundreds of spectral lines, but it is not this complexity the spectroscopist wants to avoid. Indeed, such a spectrum could be both resolved and analyzed. The trouble is that the spectrum is never that of a single molecule; it is the spectrum of a large collection of molecules such as those of a gas in an experimental glass bulb. The molecules in the bulb occupy many different energy states, and each state gives rise to a different spectrum. The observed spectrum is the sum of all the spectra associated with the large num-

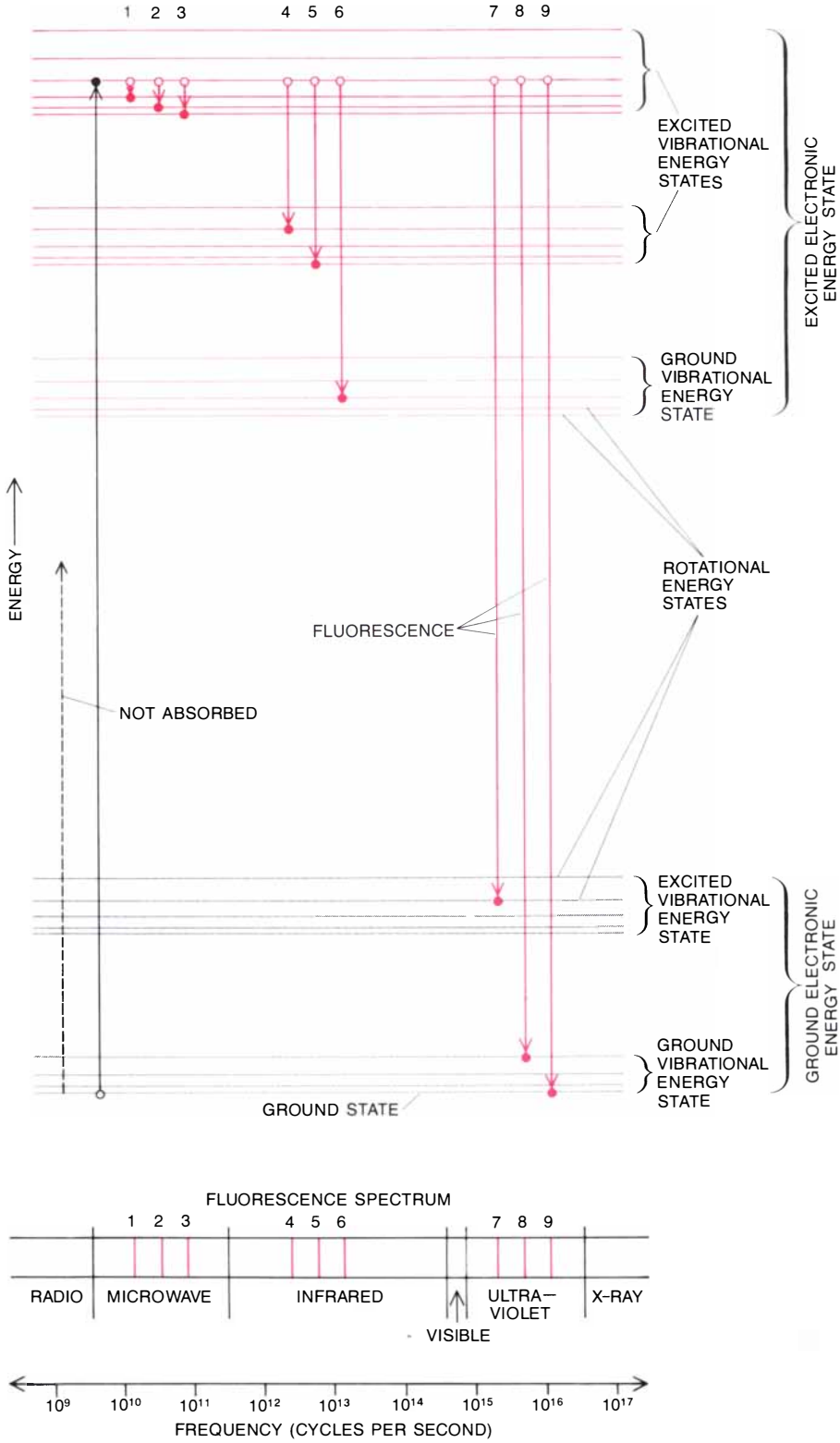
ber of initial energy states in the bulb.

When the molecules in a gas can freely exchange energy with one another, they are in thermal equilibrium, and the laws of statistical mechanics make it possible to estimate the distribution of the molecules among the available energy states. For any given absolute temperature the fraction of the total number of molecules in the gas that are in any single energy state can be calculated from the statistical law named the Boltzmann distribution, after the Austrian physicist Ludwig Eduard Boltzmann. In order to calculate the population of a given energy state one must find the difference between the energy of the given state and the energy of the ground state, or lowest energy state possible. This difference E is then divided by a constant k, called the Boltzmann constant, which is equal to approximately 1.4×10^{-16} erg per degree Kelvin. Since the Boltzmann constant is expressed in units of energy per unit of temperature, the energy separation between the given state and the ground state is thereby expressed in units of temperature.

The Boltzmann distribution gives the ratio of the number of molecules in a gas that occupy the given excited energy state to the number of molecules that occupy the ground state. Equivalently, the ratio is the probability that a single molecule will occupy the excited state, divided by the probability that it will occupy the ground state. This ratio is equal to e, the base of the natural logarithms, raised to the power $-(E/k)/T$, where T is the absolute temperature of the gas. When the energy separation is equal to the gas temperature, the equation shows that the excited state is about .37 as likely to be populated as the ground state. When the energy separation is about 2.3 times the gas temperature, the number of molecules in the excited state is about one-tenth the number of molecules in the ground state. If the population of the excited state is at least one-tenth as great as the population of the ground state, let us arbitrarily call the excited state populated.

The number of populated states depends on both the energy separation between states and the absolute temperature. At absolute zero all energy levels except the ground state are depopulated. The energy separation between rotational energy states depends on the size of the molecule: for small diatomic molecules it is about one degree K., for medium-size, polyatomic molecules it is roughly 10^{-2} degree and for the largest molecules in the gas phase it is on the order of 10^{-3} degree. Hence at room temperature, which is about 295 degrees K., the number of populated rotational energy states can be enormous.

The Boltzmann distribution makes it clear why molecular spectra can be sim-



QUANTIZATION OF ENERGY in a molecule can be represented as a collection of discrete horizontal lines, each of which corresponds to a level of energy allowed by quantum mechanics. The energy levels typically exhibit a three-layered hierarchy: many rotational energy states are allowed for each vibrational state, and many vibrational states are allowed for each electronic state. Hence each electronic energy state is split into separate clusters of lines; each cluster corresponds to a state of vibrational energy, and each line within a cluster represents a discrete level of rotational energy. Energy transitions of the molecule can take place only between two energy levels; if the energy does not match the energy difference between two levels, the rules of quantum mechanics forbid the transition (*broken arrow at left*). Transitions are accompanied either by the absorption (*black arrow*) or the emission (*colored arrows*) of a photon. The frequency of the photon varies directly with its energy, and so the emitted photons give rise to a spectrum of bright lines from which certain physical properties of the molecule can be deduced. The numbered energy transitions correspond to the numbered lines on the schematic spectrum at the bottom. The relative heights of the transitions are not drawn to scale.

executive health

the report that briefs you on what to watch

Volume XX, Number 3 • December, 1983 • P.O. Box 589, Rancho Santa Fe, Calif. 92067 • Area 619:756-2600

Walter M. Bortz II, M.D.:

ON DISEASE . . . AGING . . . AND DISUSE

Much of illness . . . and even many biologic changes thought due to aging . . . may really be due to simple disuse

PUBLISHER'S NOTE: *Dr. Walter M. Bortz II of Stanford University School of Medicine and the Palo Alto Medical Clinic, in many medical journal reports, is urging physicians to recognize the striking evidence: That human frailty has a third dimension which is neither disease nor time. And that overcoming the disuse dimension by natural means can sustain health and slow aging-related changes as no drug can. We have asked him to brief you on this vital subject.*

"Use it or lose it" seems like a trite, trivial aphroism. But it is not. Contained within those five small words is immense wisdom.

All natural processes link vitality with purposeful activity. When purpose is lost, nature has little tolerance. Each cell, tissue, and organ requires active stimulation to function optimally. When denied maximal usage, deterioration in function and structure is certain to follow.

Until recently, we have had no scientific basis for fully understanding *the many effects of disuse—and the strikingly beneficial multiple effects of activity*—on the human body.

We now have mounting evidence of both. Disuse on the one hand . . . and activity on the other—*touch virtually every area of the body, virtually all body systems, and, hardly least of all, the brain and its functioning.*

Medicine and the individual today are—or should be—faced with a potentially greatly rewarding new challenge: recognizing how much of all illness and slowing down result from disuse rather than disease or aging. And, of course, *the coronary challenge: how much can be done by something inherent, natural—becoming active, fully, purposefully active.*

The "disease theory"—not enough

Traditional medical "disease theory" holds that ill-

How Long Will You Live?

Discover what research scientists now know about how to take the best care you can of the only body you will ever have!

Have you ever asked yourself, "What keeps well people well?" Medical men have long concentrated upon sick people and how to get them well . . . not upon well people and how to keep them well. Now many top research scientists are concentrating their efforts on *preventive* medicine . . . how to keep well people well.

May we suggest that you get the benefits of this new research for yourself? Do as so many thousands of executives do. Subscribe to Executive Health Report. The members of our ***Editorial Advisory Board*** are among the world's most distinguished authorities on *preventive medicine*. Their wise advice can help you not only live longer but enjoy those extra years! (Note their high qualifications.)

Executive Health Report is not sold on newsstands but only by private subscription at $30 (U.S. Funds) a year in the U.S.A. and its possessions, Canada and Mexico. Individual reports (back issues) $3.00 per copy. All other countries $36 (air mail only). (Only International Money Order or check on U.S. bank acceptable.)

Subscribe now under this unusual introductory offer:

(1) Your choice of any three of the reports listed below ($3.00 each) FREE!

(2) If you are not satisfied with your first issue, your money will be *promptly* refunded.

Please study the reports listed here and circle your three choices:

•

Walter M. Bortz II, M.D.: ON DISEASE . . . AGING . . . AND DISUSE. Much of illness . . . and even many biologic changes thought due to aging . . . may really be due to simple disuse.

•

If you sit at a desk all day . . . How to avoid a POT-BELLY and DOUBLE-CHIN or get rid of them for good! (updated)

•

Lot B. Page, M.D.: ON MAKING SENSE OF SALT AND YOUR BLOOD PRESSURE. An authoritative report on how much salt may matter to you if excessive . . . and how much good moderation can do.

Herbert A. deVries, Ph.D.: ON EXERCISE FOR RELIEVING ANXIETY AND TENSION. Authoritative new insights into how physical activity — surprisingly modest activity — can be at least as effective as tranquilizers and far safer.

•

Richard B. Shekelle, Ph.D. and C. C. Tangney, Ph.D.: ON DIETARY VITAMIN A . . . Beta-Carotene . . . and Lung Cancer Prevention. An authoritative report on our deadliest malignancy . . . the possible anti-lung cancer effect of some foods . . . and what, wisely, to put on your plate.

•

Linus Pauling, Ph.D.: ON GOOD NUTRITION FOR THE GOOD LIFE . . . in which the two-time Nobel Laureate informs you of his personal approach . . . and his own intake of many vitamins . . . and why — for your own possible guidance — that intake differs so markedly from conventional recommendations.

•

Hans Selye, M.D., Ph.D., D.Sc., F.R.S. (C): "IT'S NOT THE AMOUNT OF STRESS YOU HAVE, IT'S HOW YOU RESPOND TO IT . . . Adaptability is probably the most distinctive characteristic of life."

•

Joseph T. Coyle, M.D.: ON SENILITY . . . ALZHEIMER'S DISEASE . . . AND THE CHOLINERGIC LINK. New insights into what is NOT a natural accompaniment of aging . . . the causes of mental impairment . . . and ALL that now can be done to combat it.

•

ON THE B VITAMINS. Are they doing well by you . . . can they do more for you?

•

Linus Pauling, Ph.D.: ON VITAMIN C and infectious diseases.

Roger J. Williams, Ph.D., D.Sc.: ON YOUR STARTLING BIOCHEMICAL INDIVIDUALITY. Some amazing facts about your body you need to know if you want to understand yourself (and other people) better.

•

Robert J. Krane, M.D.: ON THE BRIGHT NEW OUTLOOK FOR IMPOTENCE. How major advances both in pinpointing causes — and treating them — now can benefit close to 90% of the affected.

•

Joan Borysenko, Ph.D. and Myrin Borysenko, Ph.D.: ON PSYCHONEUROIMMUNOLOGY: HOW THE MIND INFLUENCES HEALTH AND DISEASE . . . AND HOW TO MAKE THE INFLUENCE BENEFICIAL. Insights from a burgeoning new medical field.

•

Roy L. Walford, M.D.: ON RETARDING AGING BY NUTRITION. Can it be done? The evidence now is yes— even starting in middle age.

•

The **TRUTH** about the **RIGHT DIET** and **EXERCISE** to lose weight safely and stay slender for life.

•

New research findings ON WALKING . . . nature's own amazing "anti-age antibiotic!" Not running, not jogging, but walking is your most efficient form of exercise and the only one you can safely follow all the years of your life!

•

Bruce Proctor, M.D., F.A.C.S. and Conrad Proctor, M.D., M.S., F.A.C.S.: ON A DIET TO TREAT MENIERE'S. Relieving the incapacitating inner ear disorder . . . with its prostrating vertigo, nausea, ear ringing, and hearing loss.

Please use the coupon below under our special money-back guarantee.

EXECUTIVE HEALTH, Pickfair Bldg., Box 589, Rancho Santa Fe, CA 92067

Gentlemen: Enclosed is my check for $_____ for a year's subscription to Executive Health to start with this month's issue. I have circled the three $3.00 reports I am to receive free. It is understood that if I am not satisfied with my first issue, my money will be refunded. In addition, send me a list of some 100 other reports of yours still in print. There may be some from which I might greatly benefit and would otherwise miss.

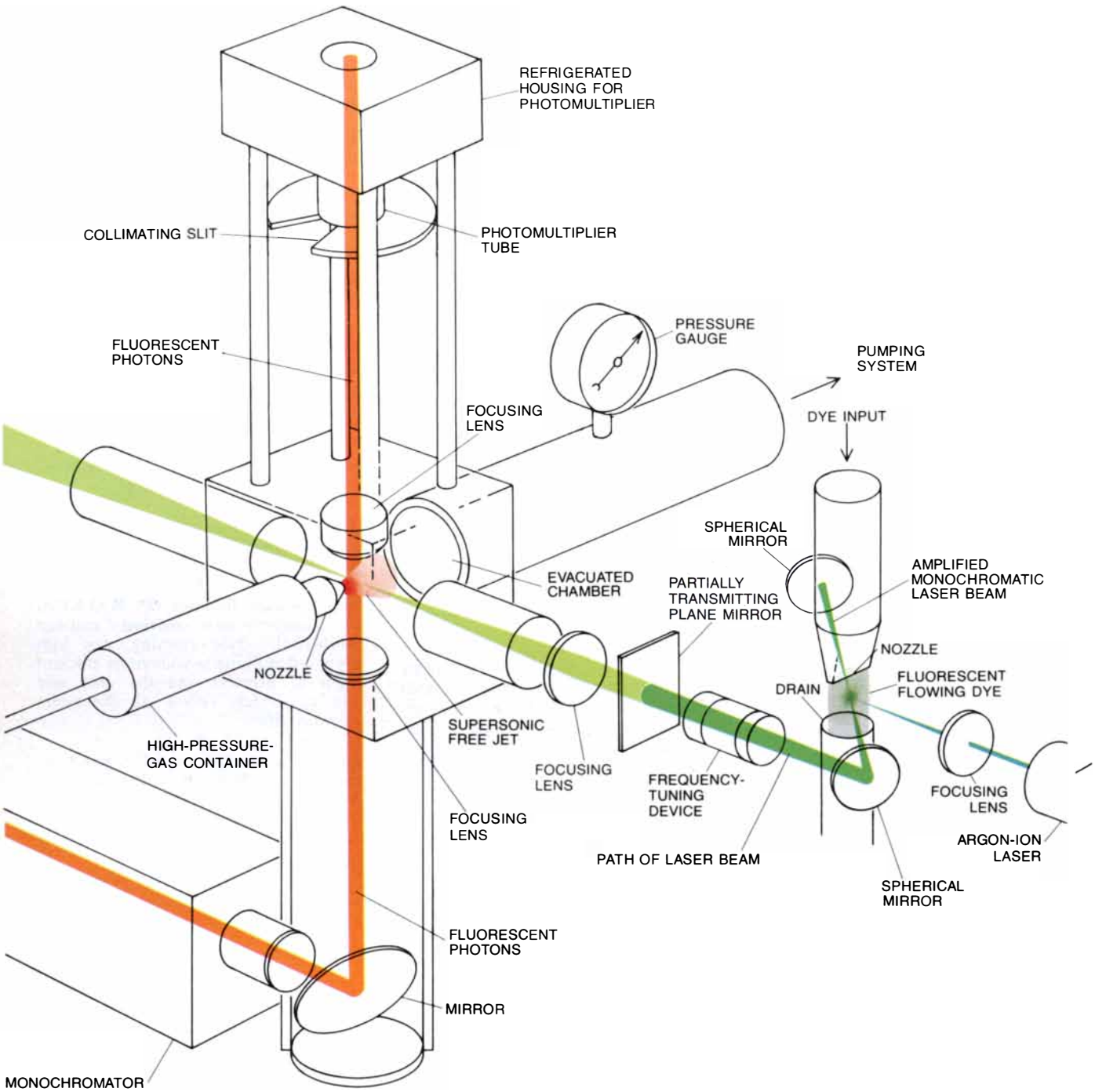
Firm/Name (Please print clearly) _____

Street & No./Route _____ **Suite or Apt.** _____

City, State, Country _____ **Zip** _____ SC-1

plified by cooling the sample. At sufficiently low temperatures the separations between energy levels even for the largest molecules are large compared with the temperature, and most levels become depopulated. The lines they contribute to the spectrum disappear, and the spectrum of the few populated levels that remain becomes strong and simple. The observed small separations between molecular energy states imply that unless the sample is cooled to well below one degree K. the necessary spectral simplification cannot be achieved.

In order to study molecules in the gas phase at such temperatures my colleagues and I rely on a novel technique of refrigeration that chills the gas molecules very quickly and then thins them out before they can condense. The refrigerant is an uncollimated, or free, jet of gas ejected into an evacuated chamber under conditions where the flow is many times faster than a sound wave can travel in that medium; the technique is therefore called supersonic-free-jet spectroscopy. Imagine a container filled with an atomic gas such as helium, pressurized to perhaps 100 atmospheres,



EXPERIMENTAL APPARATUS for observing the fluorescence spectrum of a supersonic free jet has three main parts: the apparatus for generating the jet, the laser system for exciting the gas molecules in the jet and the detector for the fluorescent photons. In this schematic diagram the jet expands through a small hole in the high-pressure-gas container into the evacuated chamber. The source of the laser beam is a dye laser, which is pumped, or energized, by a second laser, the argon-ion laser. The frequency of the light emitted by the argon-ion laser is fixed, but the frequency of the dye laser can be varied. Two detectors are shown, both of which count the photons emitted by the molecules in the jet. The photomultiplier tube counts photons over a range of frequencies, and the monochromator counts photons of a predetermined frequency. Other kinds of detector such as a charged-particle detector (not shown) can also be employed.

Why this New York ad man leaves his $300 attaché case in the closet and carries our $39.50 Square Rigger.

Dick Anderson, an Executive Vice President of Needham, Harper and Steers, explains it this way:

"I'm always trying to jam more things into an attaché than it was ever meant to hold. That plays hell with stiff-backed cases. So I find myself leaving my $300 leather case in the closet, and carrying my canvas Square Rigger. It swallows overloads without complaint, has inside pockets for my calculator and appointment book, even a snap hook for keys."

And more.

That's not even the whole story of our Square Rigger. It's made of tough yet soft 18 oz. Square Rigger cotton canvas. Has comfortable padded handles. Smooth-operating YKK zippers. Comes in six businesslike colors.

Pretty impressive for $39.50. And we'll add your monogram for just $5 more.

We make the Square Rigger and over a hundred other Lands' End items at our duffle works in Boscobel, Wisconsin. That way we can guarantee the best possible materials and workmanship. And sell our soft luggage to you without middleman mark-ups.

First quality, then price.

What we have to offer goes beyond soft luggage, even beyond our great sportswear and accessories. The Lands' End tradition affects everything we offer you.

First, *quality.* Then, *price.* And always, always *service.*

That's why we offer a guarantee that would put lesser merchants out of business. Simply: "If you are not completely satisfied with any item you buy from us, at any time during your use of it, return it and we will refund your full purchase price."

If we're new to you, we don't ask that you trust us. Just try us. By phone, you can reach us toll-free 24 hours a day at 800-356-4444 (except Alaska and Hawaii call 608-935-2788). Or fill in the coupon. Order a Square Rigger or not, as you like. But let us send you a free copy of our Lands' End catalog.

LANDS' END
DIRECT MERCHANTS

of fine wool and cotton sweaters, Oxford buttondown shirts, traditional dress clothing, snow wear, deck wear, original Lands' End soft luggage and a multitude of other quality goods from around the world.

☐ **Please send free catalog.**

Lands' End Dept. Q-02
Dodgeville, WI 53595

Ship____ Attachés. $39.50 ea.
plus $2.50 shpg. ☐ Tan ☐ Navy ☐ Green
☐ Brown ☐ Gray ☐ Burgundy

$5.00 to monogram 3 initials: ☐☐☐

☐ Check enclosed ☐ American Express
☐ Visa ☐ Master Card

Card No._____

Expiration Date_____

Name_____

Address_____

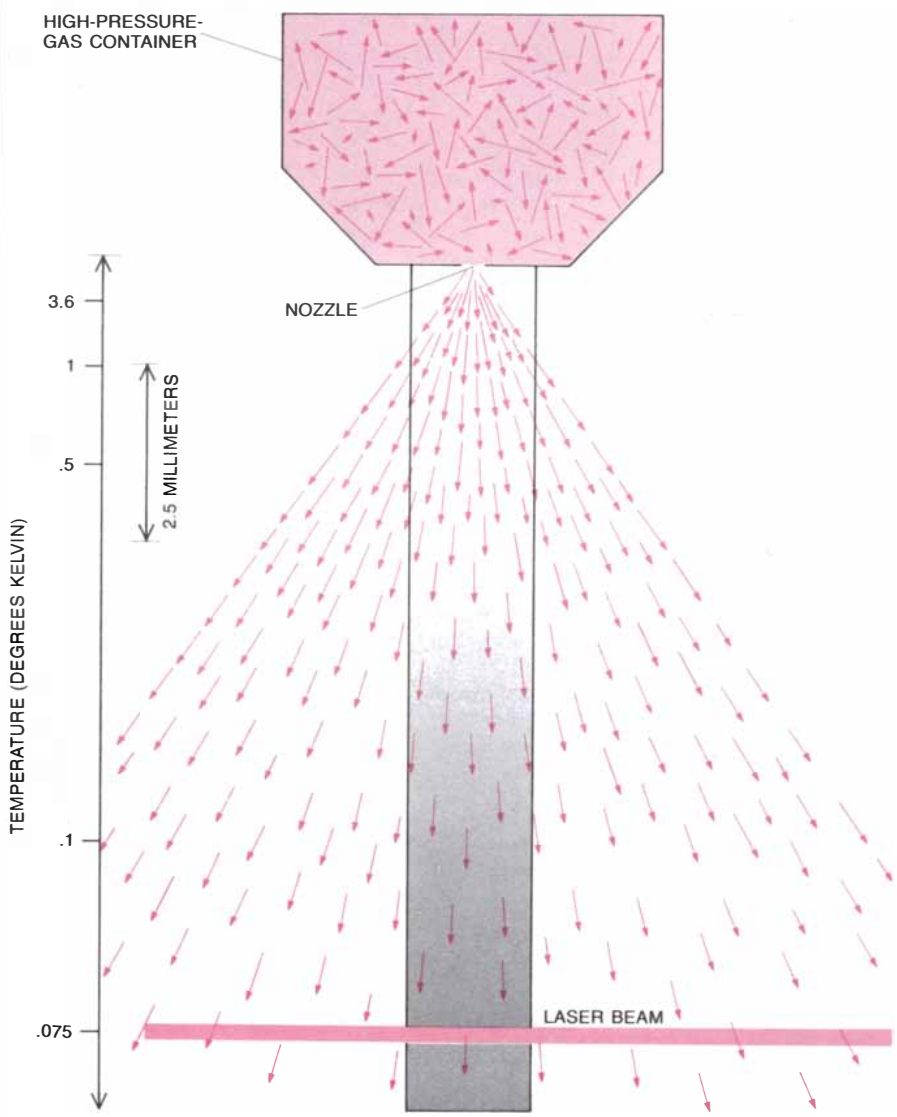
City_____

State_____ Zip_____

Or call Toll-free:
800-356-4444
(Except Alaska and Hawaii call 608-935-2788)

and assume the container is enclosed in an evacuated chamber. If a hole is punched in the container, gas will rush out of the hole into the vacuum. If the pressure in the container is high enough or the hole is large enough, atoms in the container will tend to collide with the atoms escaping from it, and so a jet of gas is pushed out of the hole.

The laws of thermodynamics state that under appropriate conditions the expanding gas in the jet must cool. On a microscopic scale the cooling of a gas implies that the energy of the gas associated with the random motion of the particles in the gas is reduced. The atoms of a gas inside a bulb move about at many different velocities. For simplicity consider motion in only one direction. Since the motion is random, an atom is as likely to be moving at a certain velocity to the right as it is to be moving at that same velocity to the left. Hence the velocities of the particles in the left-right direction are distributed symmetrically about the most probable velocity, which is zero. The temperature of the gas is reduced as an increasingly large fraction of the particles move at low velocity with respect to the most probable veloc-



MOLECULAR VELOCITY DISTRIBUTION downstream from the nozzle determines the temperature of the gas in the jet. The direction of each molecule is represented by an arrow, and the magnitude of the velocity is represented by the length of the arrow. In the high-pressure-gas container the density of the molecules is high and their velocities are random in both direction and magnitude. Moreover, because the high-pressure gas is relatively hot, the range, or dispersion, of the velocities assumed by the molecules is relatively large. Outside the nozzle the density of the molecules drops precipitously, and it continues to drop with the distance from the nozzle. Collisions among the molecules in the jet narrow the range of molecular speeds along any radial direction downstream from the nozzle, and so all the molecules in any small region of space begin to move at nearly the same velocity. Because the width of the velocity distribution is a measure of temperature, the gas in the jet can be cooled to within a few hundredths of a degree of absolute zero. The cooling of the central region of the supersonic jet from which the spectrum is observed is indicated by the progressively darker shading.

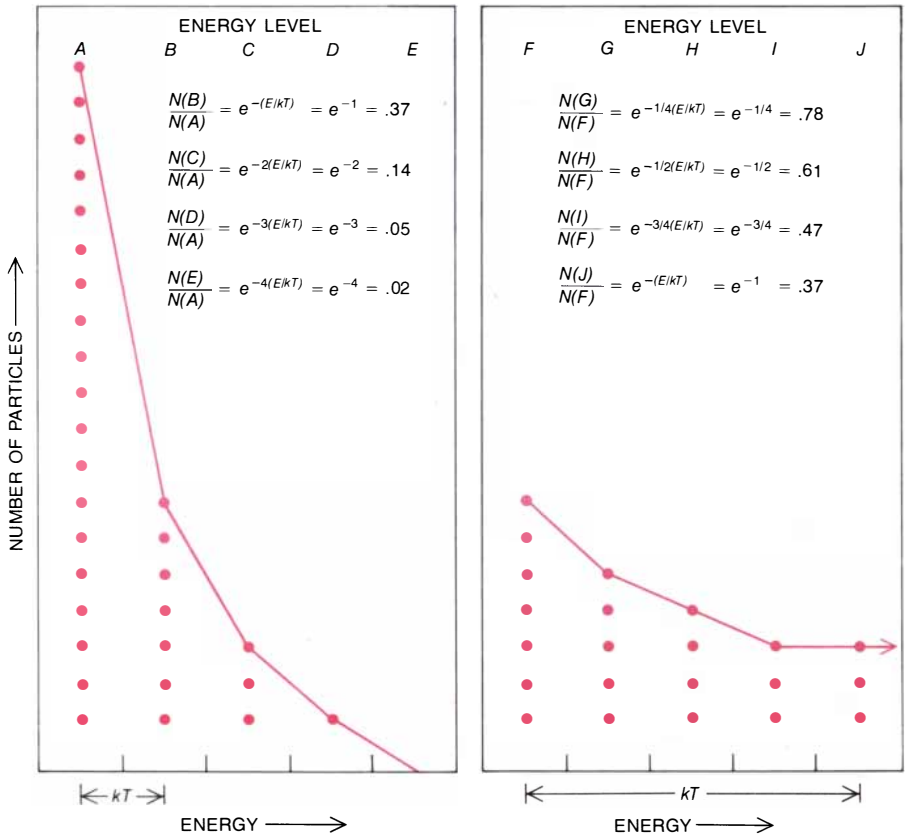
ity. In other words, the temperature is measured by the width of the velocity distribution about the most probable velocity. At low temperatures almost all the atoms move at velocities close to zero, whereas at high temperatures a larger fraction of the atoms move at higher velocities to the left or right.

In a supersonic free jet the temperature is also measured by the width of the velocity distribution. In the jet the most probable velocity in one direction is not zero, since the gas is moving away from the hole at some nonzero flow velocity. The temperature is determined by the width of the velocity distribution centered about the flow velocity; the temperature is lowered as an increasing number of atoms move at approximately the same velocity, even though the entire mass of gas is moving forward at the flow velocity. The definition of temperature is independent of the flow velocity because chemical and physical properties of a parcel of gas are determined by what is going on in its immediate vicinity. These properties are no more altered by the high flow velocity of the jet than they would be if the parcel were carried aboard an airplane moving at a high but constant velocity.

The narrowing of the velocity distribution that takes place in a supersonic expansion is called translational cooling. Consider the transverse velocity component of each atom, or the component of the velocity that is perpendicular to the direction of the flow. If there were no collisions between atoms, the atoms with small transverse velocities would remain near the center of the expansion and the atoms with large transverse velocities would move away from the center line. Because of the spatial segregation of the atoms according to their transverse velocities, the atoms in any small parcel of gas relatively far downstream from the nozzle must have similar transverse velocities. In other words, the transverse velocity distribution within any small parcel is narrowed and so the transverse velocity is cooled.

Near the nozzle, however, where the pressure of the gas in the jet is relatively high, the atoms in the jet collide with one another continually. These collisions allow full or partial equilibration between the transverse velocity components and the other degrees of freedom in the motion of the atoms. The cooling of the transverse velocity thereby leads to a cooling of the longitudinal velocity, which is the velocity component parallel to the direction of the flow of the jet.

In order to cool a molecular gas of spectroscopic interest a small amount of the gas is mixed with a large quantity of a monatomic carrier gas such as helium. When the mixture expands through a hole or nozzle, it is translationally cooled. As long as the density of the



DISTRIBUTION OF MOLECULES in random thermal motion among various energy levels is given by the Boltzmann distribution. The number of molecules at any given energy level in a gas depends on the absolute temperature of the gas and on the difference in energy between the given energy level and the ground state, or lowest energy level possible. The ratio of the number of molecules in the given level to the number of molecules in the ground state is e, the base of the natural logarithms, raised to the power $-E/kT$, where E is the energy separation above the ground state, k is the constant known as Boltzmann's constant and T is the absolute temperature of the gas. Because k has units of energy per degree Kelvin, the quantity kT has units of energy. Hence if the separation E between energy levels is roughly equal to k times the absolute temperature, only the lowest energy levels are populated (*left*). On the other hand, if the energy separation is less than kT, higher energy levels are populated as well (*right*). In the equations the symbol $N(X)$ designates the number of particles that occupy energy level X.

mixture is high enough the molecules collide at low velocity with the cold atoms, and the energy of rotational and vibrational motion of the molecules is transferred to the translationally cold bath of monatomic gas. As the gas flows downstream from the nozzle, however, the average distance between the atoms and molecules increases, and so the number of collisions per second between the particles decreases. Thus the cooling of the molecules caused by the exchange of energy in collisions is a self-limiting process. The cooling is rapid in the upstream region of the jet, but it decreases sharply as the collisions in the jet become rare.

The approximate independence of the rotational, vibrational and electronic energy states of a molecule implies that cooling can be considered independently for each kind of motion. For these internal degrees of freedom cooling is equivalent to limiting the motions of most of the molecules in a gas to their lowest energy states. The exchange of energy between molecular rotational motion and atomic translational motion is extremely fast, and so the molecular rotations are cooled to a temperature that is close to the ordinary, translational temperature.

Molecular vibrations exchange energy with translational motion somewhat more slowly; they are cooled to a temperature significantly below their original temperature but above the final translational temperature. Both cooling processes, however, are much faster than condensation. Indeed, the main advantage of the supersonic free jet as a refrigerant is that the molecular motions can be cooled for spectroscopic purposes before the molecules condense to form a liquid or a solid.

Condensation is ultimately checked by the same mechanism that limits the cooling, namely the rarefaction of the gas as it moves downstream in the jet. Even during the translational cooling, however, condensation is slow because its initiation requires simultaneous collisions among three or more particles. Unlike cooling, which requires only that energy be exchanged between two colliding particles, condensation requires that at least two particles stick together. If a bond is to form, however, a third particle must collide with the first two during the time of the two-body collision so that the third particle can carry off the binding energy of the pair as translational energy. Three-body collisions require much higher densities in the gas than two-body collisions do, and the cooling is complete before a significant number of three-body collisions can take place. In effect the molecular gas is rapidly supercooled, or cooled below its ordinary condensation temperature, and then allowed to expand until its low density precludes the formation of clusters of molecules on which other molecules could condense.

In principle any kind of molecular spectroscopy that can be done on gases in a bulb can also be done on a supersonic free jet. In practice the spectroscopic technique must be tailored for the jet because the sample of molecular gas is small. In a typical experiment, such as the study of the free-jet spectrum of molecular iodine, the sample of gas exposed to light at any one moment is made up of 10^8 molecules and weighs 3×10^{-14} gram. With a sample that small a spectroscopic method of great sensitivity is necessary.

In addition to the sample molecular spectroscopy requires a source of light for exciting the molecules out of their initial states and a detector for recording the resulting spectral lines. One highly sensitive method we have employed is fluorescence spectroscopy. When the molecules in the jet are exposed to light of a frequency at which the molecules absorb, each molecule can absorb a photon from the light beam. After 10^{-9} to 10^{-6} second the excited molecule can return to some lower energy state by emitting another photon; in other words, the molecule can fluoresce. The detection of the fluorescent photon makes it possible to infer that a molecule in the sample has absorbed a photon at the known frequency of the light. The resulting spectrum, in which the intensity of the fluorescence is plotted against the frequency of the incident light, displays information similar to that carried by the absorption spectrum. Fluorescence detectors can count individual photons with high accuracy, so that the method is an extremely sensitive way to obtain the absorption spectrum.

If fluorescence spectroscopy is understood as a chemical reaction in which two reactants (the incident photons and the molecules) are combined to form a product (the fluorescent photons), then in spite of the low concentration of one reactant the reaction rate can be kept at a tolerable level by forcing an excess

concentration of the other reactant. In short, because the molecular sample is small, the light source must be bright. The brightness of a light source is its power per unit frequency per unit of solid angle; brightness can therefore be achieved not only by increasing power but also by collimating the light beam and limiting the photons in the beam to a narrow frequency range.

Among the brightest sources known is the laser, whose intense, monochromatic and highly collimated output beam can deliver photons of the proper frequency at a high rate to the small volume occupied by the sample. Indeed, supersonic-jet spectroscopy would have been either impossible or at best a curiosity applicable only in a few favorable cases if the frequency-tunable laser had not been available. The growth of supersonic-jet spectroscopy as a tool of general utility has been made possible largely by the development of more reliable and more powerful lasers that can be tuned to produce photons from the infrared to the ultraviolet.

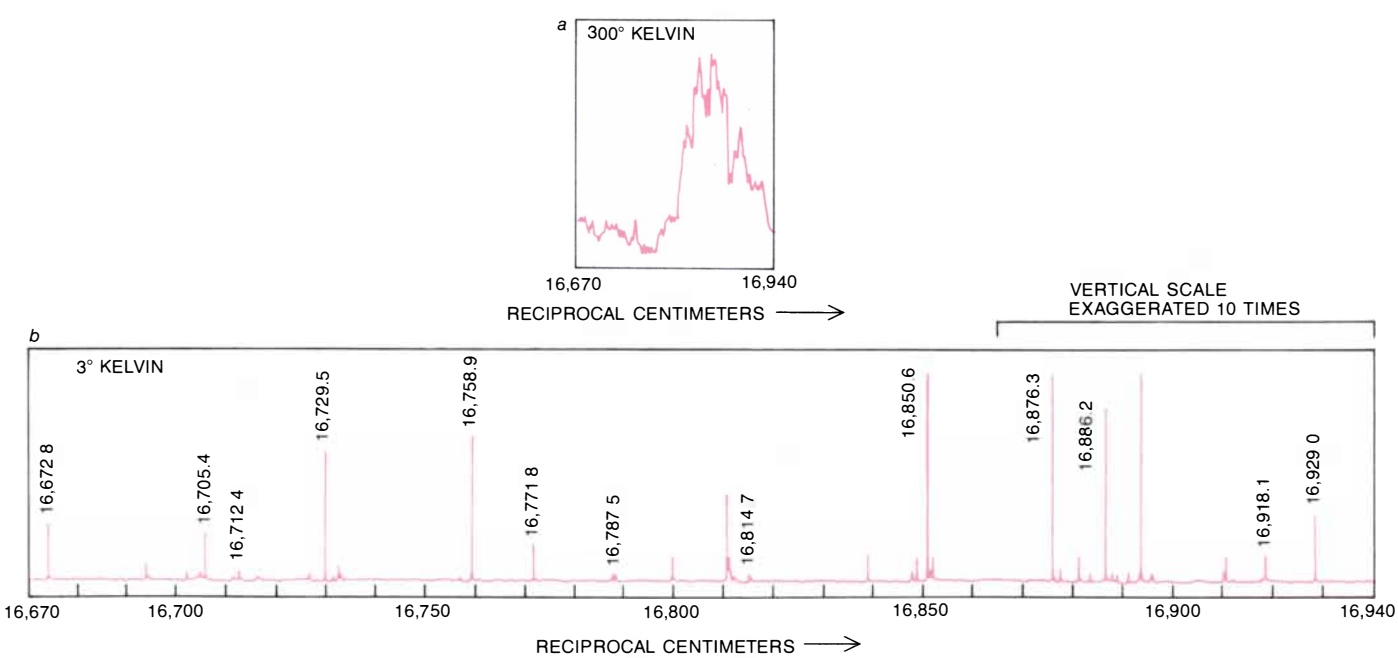
In our apparatus the carrier gas, which is usually helium but can on occasion be argon, is passed over or bubbled through the molecular compound of interest. The carrier gas thereby picks up a small amount of molecular vapor, and the resulting mixture is expanded through a nozzle, which is simply a hole about .05 millimeter in diameter in a thin piece of metal. When the vapor pressure is not high enough at room temperature, the molecular sample and the nozzle can be heated before the gas is passed into the evacuated chamber.

The largest part of the apparatus is the pumping system necessary to maintain the vacuum into which the free jet expands. The final temperature along the center of the jet can be lowered either by increasing the diameter of the nozzle or by increasing the pressure behind it, but both these modifications increase the amount of gas that must be removed by the pumps. Hence the size of the pumping system ultimately limits the cooling performance of the jet. We have been able to attain a final translational temperature of about .05 degree K., which chills the molecular rotational motion to a few tenths of a degree and cools the vibrational motion to between 10 and 50 degrees. (Most molecules are in the electronic ground state even at room temperature.)

In order to generate the spectrum an ultraviolet or visible-light laser beam, whose radiation can be varied in frequency, is crossed with the supersonic free jet at right angles. The laser photons are energetic enough to raise a significant proportion of the molecules to an excited electronic energy state. The detector is a photomultiplier tube, which converts the photon signal into a pulse of electric current. We can distinguish the fluorescent photons from the photons emitted by the laser in one of three ways. First, the fluorescent photons do not necessarily have the same frequency as the laser photons. If the electronically excited molecule does not fall back to the same energy level from which it was excited, the frequency of the fluorescent photon does not match the frequency of the laser. For such frequencies a simple colored-glass filter or a monochromator can transmit the fluorescent photons to the photomultiplier tube and block the laser light.

A second means of distinguishing the fluorescent photons from the laser photons is simply to place the photodetector outside the path of the laser beam. Since the laser beam is collimated, the only photons counted by the detector are the fluorescent ones, which are emitted in all directions. Finally, it is sometimes possible to filter the laser photons temporally. If the laser is pulsed, or turned rapidly on and off, the molecule can emit fluorescent photons for some time after the laser is off. The laser photons can then be eliminated if the photodetector is turned off while the laser is on and turned on while the laser is off and the molecule is still fluorescing.

Fluorescence detection is not the only technique for obtaining the spectrum of a supersonic jet. Another method has recently been developed that detects electrically charged particles instead of photons. The molecule in the sample is first excited electronically by the absorption of a laser photon. A second photon is then absorbed either from the same laser or from a second laser; the energy of the second photon is sufficient to ionize a molecule, or remove an electron from it, only if the molecule has previously been put in an excited state by the first photon. Since the molecular ion is a positively charged particle with a deficit of one or more electrons, the presence of a charged particle in the sample is a signal that the molecule had



RESOLUTION OF THE SPECTRUM of nitrogen dioxide is dramatically improved when the spectrum is generated by a supersonic jet seeded with nitrogen dioxide molecules. The upper trace is a spectrum generated by static nitrogen dioxide gas; its complexity is beyond the scope of analysis. The lower trace is the supersonic-free-jet spectrum in the same frequency region, shown with an expanded frequency scale. Individual spectral lines are resolved and information about the structure of the nitrogen dioxide molecule can be deduced.

Business. 200 corporations know that dance is important to the people important to them. That's why they are investing in seven of America's greatest dance companies through The National Corporate Fund for Dance. Don't let your corporation sit this one out. Contact William S. Woodside, Chairman, American Can Company c/o The National Corporate Fund for Dance, Inc., 130 West 56th Street, New York, N.Y. 10019.

THE NATIONAL CORPORATE FUND FOR DANCE

originally absorbed a photon of known frequency when it was first electronically excited. Moreover, a charged-particle detector can include a mass spectrometer, so that the mass of the absorbing molecule can be determined.

Supersonic-jet spectroscopy has now been applied to the study of a wide variety of molecules that range in complexity from diatomic molecules to the large organic molecules called porphyrins. The first molecule we investigated with the technique, nitrogen dioxide (NO_2), provides an excellent demonstration of the power of supersonic-jet spectroscopy to resolve densely packed spectral lines. The electronic spectrum of nitrogen dioxide extends from the infrared through almost all the visible frequencies; the densely packed lines give rise to the brown color of nitrogen dioxide that is seen as air pollution over many cities.

Spectral lines are often analyzed by grouping them into bands that correspond to many rotational energy levels and a single vibrational level. In many small molecules the vibrational bands are well separated from one another, but in nitrogen dioxide the spectral lines are so dense that there is no obvious separation; as a result the spectrum had resisted analysis in spite of substantial effort for many years. In a supersonic free jet the nitrogen dioxide can be so cold that essentially all the molecules in the sample occupy the lowest vibrational and rotational energy level. When the molecule is electronically excited, the spectrum is made up of one rotational line for each vibrational level of the excited energy state, and the separation of bands is straightforward. In fact, we found it useful to back away from the most extreme conditions we could achieve so that a few rotational energy levels were initially populated. More rotational structure was thereby present in the spectrum, and it yielded additional information about the molecule.

One of the most promising applications of supersonic-jet spectroscopy is the study of the dynamic processes in molecules, such as energy transfer and decomposition by light. We have been particularly interested in the flow of energy within a molecule from one vibrational mode to another. It is useful in the study of such phenomena to have a means of detecting the transfer of small amounts of energy; if such means are available, the molecule can be studied near its vibrational ground state, where its structure and properties are well known. This requirement is met by a class of molecules that are held together by the weak intermolecular forces called van der Waals forces. In these molecules the transfer of a small amount of energy from a vibrational mode to the van der Waals bond breaks the bond; the breaking of the bond signals that the energy transfer has taken place. In van der Waals molecules the threshold of energy required to break the bond and signal the energy transfer is much smaller than it is in molecules held together by ordinary chemical bonds; hence it is possible to investigate photochemical reactions whose excitation energy is extremely low.

Unfortunately the van der Waals force is too weak to form stable bonds under ordinary laboratory conditions. At the low temperatures in the supersonic free jet, however, van der Waals bonds can form between molecules such as molecular iodine (I_2) and atoms of a noble gas such as helium or neon. If these molecules are then electronically excited and slightly vibrationally excited by the absorption of a photon, the two iodine atoms begin to oscillate along the bond between them. After a short time this energy is transferred to the van der Waals bond, causing the rupture of the bond and the production of two particles: an electronically excited but vibrationally deexcited iodine molecule and a free atom of the noble gas. The iodine molecule can then fall back to its electronic ground state and emit a photon, signaling that the van der Waals molecule had previously absorbed a photon and that a photochemical reaction had taken place.

The spectral line registered by the absorbed photon is smeared, or broadened, but the broadening itself is informative: from the width of the line one can calculate the duration of the excited energy state. The calculation depends on the uncertainty principle, which states that the accuracy with which the time of the excited state is known must be offset by an uncertainty about its energy. Since the frequency, and so the position, of the spectral line is proportional to the energy of the photon, the line must be broadened in frequency to reflect the uncertainty in its energy. The width of the observed line shows that several tens or hundreds of picoseconds (trillionths of a second) elapse before the energy is transferred to the van der Waals bond; the vibration of the iodine atoms must therefore be maintained throughout hundreds of oscillations.

Such studies, which provide high resolution in time as well as in space, are helping to make supersonic-jet spectroscopy a standard method for studying the spectra of molecules in the gas phase. The major limits to the technique are the brightness and reliability of the laser light source, and the degree to which it can be tuned to a given emission frequency. The rapid development of laser technology is likely to make supersonic-jet spectroscopy an increasingly powerful tool.

Numbers and Measures in the Earliest Written Records

As early as the end of the fourth millennium B.C. proto-Sumerian and proto-Elamite scribes had well-developed systems of numbers and measures. They included precursors of our own decimal system

by Jöran Friberg

Among the world's earliest written records are inscriptions on clay tablets unearthed in Iraq and Iran, in particular at the sites of two great ancient cities: the early Sumerian city Uruk and the early Elamite city Susa. The inscriptions, mainly accounts and receipts of various kinds, were written toward the end of the fourth millennium B.C. and soon afterward. After more than 100 years of scholarly effort all the systems of numbers and measures in these "proto-literate" texts have now been identified. They turn out to include precursors of the later Sumero-Babylonian sexagesimal number system (counting in 10's and 60's) and of our own decimal system (counting only in 10's). In addition they include a previously unrecognized system of capacity measures, used in all accounts dealing with barley, which in this early period was both the basic food grain and the currency.

The reader who would enjoy knowing about the proto-literate systems of numbers and measures will have to join me in a two-directional journey. We shall travel backward in time with respect to the historical record and forward from the past to the present with respect to the scholars who have studied the ancient tablets. The reason we must do so is that the oldest tablets were buried the deepest and were therefore the last to be excavated and made available for study. The oldest tablets were also the most difficult to interpret.

Let us take as our point of departure the Greek coastal island of Cos, some 20 miles northwest of Rhodes. There in about 340 B.C. the founder of a school of astrology, a Babylonian named Berossos, wrote a history of his homeland. In it he told his Greek readers that the numbers *sossos* (60), *neros* (600) and *saros* (3,600) occupied a special place in Babylonian arithmetic and astronomy. Practically nothing more was known about Babylonian numbers and measures for the next 2,200 years. Then in 1855 Sir Henry Rawlinson, one of the pioneers in the decipherment of cuneiform script, published a summary of the cuneiform numbers inscribed on a small clay tablet found at the site of the ancient Mesopotamian city Larsa. Rawlinson realized, among other things, that the last two lines of the tablet stated in effect that "58 1 is the square of 59" and "1 is the square of 1." He concluded that the tablet was the final part of an incomplete table of square roots, beginning with the square of 49 (equal to 2,401, or $40 \times 60 + 1$) and ending with the square of 60 (equal to 3,600, or 60×60). His interpretation, of course, was possible only if it was assumed that the numbers 60 and 60×60 were both represented by the same symbol, namely the symbol for the number 1.

Rawlinson drew the conclusion that the Babylonians had worked with a sexagesimal number notation of a quasi-positional nature, in other words, a number notation in which the symbol for 1 also stood for the powers of 60 and for 10 times the powers of 60. He further concluded that the Babylonians did not have any special sign to represent zero.

Here it is necessary to briefly consider the relative merits of number systems with different bases. Let us begin with the so-called metric system, which is actually a family of interrelated systems of units for several kinds of measures. The metric system owes its current acceptance to its structural simplicity and to the fact that it is constructed to match the base-10, or decimal, system used today for all kinds of routine computations. Since its conception in France in the aftermath of the French Revolution it has gradually spread across the world.

The very length of time it has taken the metric system to gain general acceptance is proof of how difficult it is to suppress other "customary" systems of weights and measures. English examples of such systems include the sequences "mile," "furlong," "chain," "rod," "yard," "foot" and "inch" for measures of length, "barrel," "bushel," "peck," "quart" and "pint" for measures of dry capacity and "ton," "hundredweight," "pound" and "ounce" for measures of weight. For that matter even the metric system has come to incorporate nondecimal systems: the 12-month year, the 24-hour day, the 60-minute hour and the 60-second minute as units of time, and the 360-degree circle, with its subdivisions the 60-minute degree and the 60-second minute as units of angle. These customary measures can be traced back to classical Greek astronomy and beyond that to the general use of sexagesimal numbers for computation in Babylon and Sumer. Many other customary systems of weights and measures, however, were doomed to be replaced by the metric system because they were insufficiently matched to the widely adopted decimal number system.

Still, the survival of some customary systems has been partially the fault of the decimal number system itself. The decimal system has the weakness that its base of 10 is really too small. This will become more apparent when I give additional examples of computations within the framework of the sexagesimal system with its larger base: $60 = 3 \times 4 \times 5$ (as opposed to $10 = 2 \times 5$). As will be seen, the base-60 system made it possible for the Sumerians' proto-literate predecessors to construct a family of nicely interrelated measure systems, with sequences of naturally occurring standard units that were easy to deal with in computations.

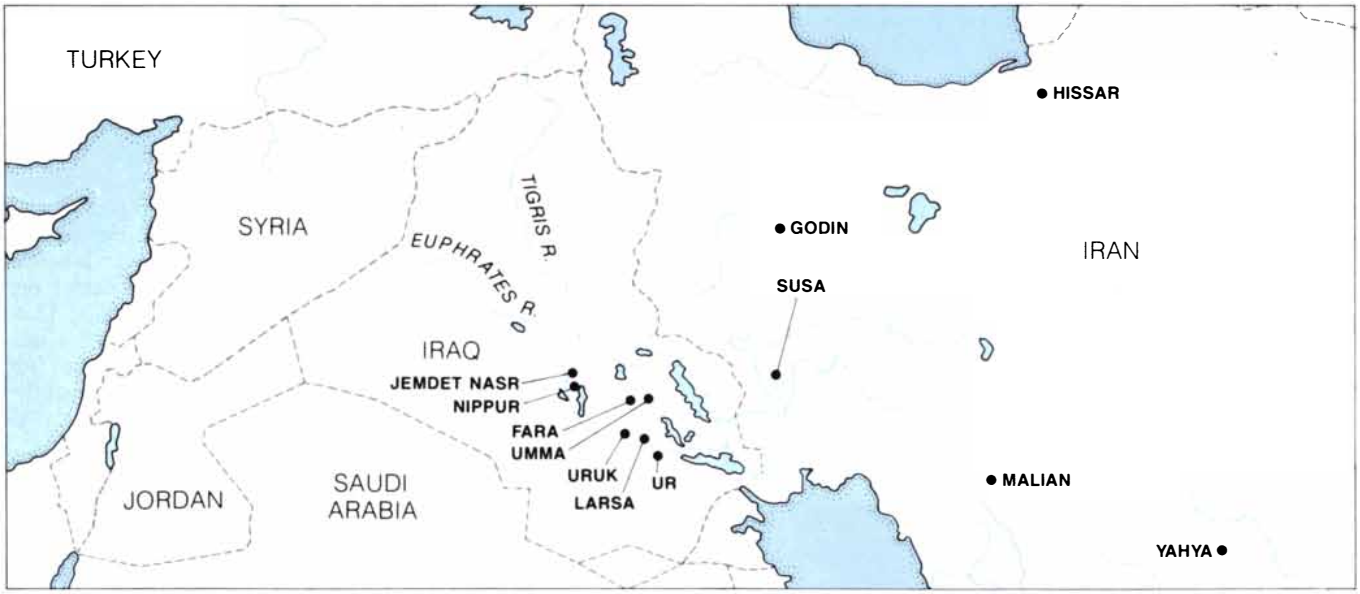
To find the correct interpretation of the system of cuneiform notations serving to represent sexagesimal (or base-60) numbers was relatively easy. It proved to be much harder to understand how the various systems of measures

that appear clearly in many cuneiform inscriptions were constructed. Some decisive clues were offered by certain tablets known to scholars as school texts.

The copying of standard texts was an essential part of the school curriculum in Old Babylonian times (1900 to 1500 B.C.). Many of the texts were lists and tables: lists of geographic names, lists of the names of birds and fishes, lists of words in two languages, grammatical tables for the study of the difficult Sumerian language and so on. Also copied were mathematical tables and lists explaining the structure of the Babylonian systems of measures and their representation in cuneiform script. By doing this kind of copying a student trained himself in cuneiform writing and at the same time accumulated a small personal library of tablets.

The first example of a table of measures to be described in a scholarly publication was a fragmentary tablet also uncovered at Larsa. The table was discussed by George Smith, a prominent student of cuneiform, in 1872, but its meaning was not fully understood until much later. On the left side of each column of the tablet is a systematically arranged sequence of linear measurements, expressed in standard units. The units are, from the smallest to the largest, the *she* (a grain), the *shu-si* (a finger), the *kush* (a cubit) and so on up to the *beru,* equal to $30 \times 60 \times 12$ (6×60^2), or 21,600, cubits. On the right side of each column are the same linear measurements expressed as multiples of cubits in sexagesimal notation. For example, the line at the lower right labeled *b* in the reproduction of the tablet on page 113 reads "Two *beru* [equals] 12." It should be noted that *beru* is how the Babylonians pronounced the symbol for the Sumerian word *danna* (normally written as *kas-gid,* meaning "long way").

PROTO-SUMERIAN TABLET from Jemdet Nasr in Iraq records the rations allotted to a total of 40 men in the course of a five-day week. The signs farthest to the left in the top three rows indicate "Day 1," "Day 2" and "Day 3." The text on the reverse of the tablet indicates that each of the men received rations equivalent in value to two minor units of barley per day. Barley was the currency of the period. The triangular sign near the right end of the fourth row probably means "workers." The tablet is in the British Museum.



IRAQ AND IRAN are the regions of southwest Asia where early forms of writing arose before 3000 B.C., primarily for the purpose of recording numbers and commodities. The seven sites in Iraq have yielded numerous cuneiform tablets. Two of them, Jemdet Nasr and Uruk, are the sources of tablets bearing the earliest Mesopotamian script, proto-Sumerian. The five sites in Iran, Susa in particular, are the sources of a linguistically unrelated script, proto-Elamite, that is about as old as proto-Sumerian. Both scripts relied on essentially identical special notations to represent numbers and measures, so that the tablets bearing numerical records and accounts are decipherable.

The 12, however, represents not 12 cubits but the much larger sum of 12×60^2 cubits. With a cubit being equal to about half a meter in length, the length of the *beru* was more than 20 kilometers.

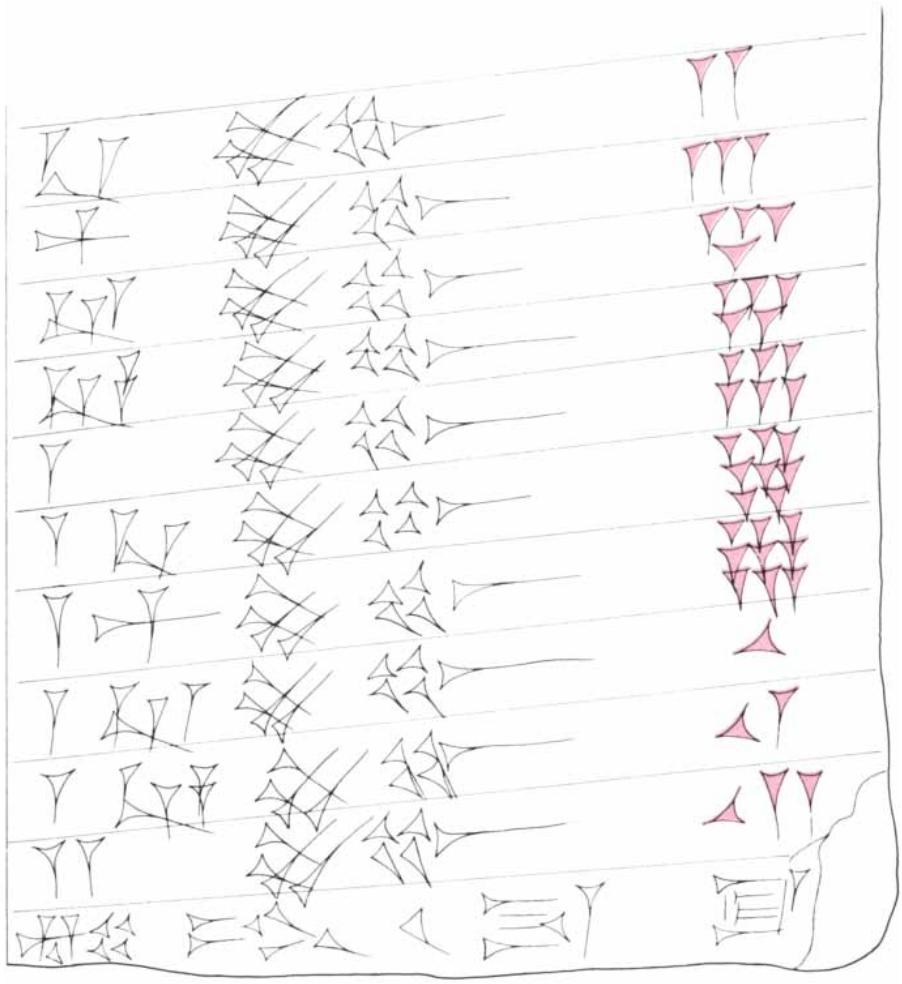
When another fragment of the same tablet was identified soon after the first fragment was found, it proved to contain an additional metrological table of the same kind as the first, except that here the right side of each column was concerned with multiples of a *nindan* (equal to 12 cubits) in sexagesimal notation. Only much later did the study of Babylonian mathematical texts dealing with the computation of volumes make it clear that whereas the cubit was the basic unit for vertical measurements, the *nindan* was the unit for horizontal measurements. Hence the small-

CUNEIFORM TABLET FROM LARSA, described in 1855 by Sir Henry Rawlinson, is reproduced in part. The script is Old Babylonian cuneiform; the text is a table of square roots. Next to the top two and bottom two lines are transliterations into English script and Arabic numerals (*black*). Below the transliterations are free translations (*color*). Many such tablets were study copies made by novice scribes.

est Sumero-Babylonian unit of area, the *shar,* was one square *nindan.* By the same token the smallest unit of volume, also called a *shar,* was the space enclosed by a bottom area of one square *nindan* that had sides one cubit high. This seemingly peculiar choice of units was actually quite practical because it usually excluded the need to count with small fractions of a volume unit.

These two metrological tables are eloquent witness to how well adapted the Sumero-Babylonian system of linear measurements was to the sexagesimal number system. Consider the conversion rules for units of the system of linear measures. Six *she* is equal to one *shu-si,* 30 *shu-si* is equal to one *kush,* 12 *kush* is equal to one *nindan,* 60 *nindan* is equal to one *USH* and 30 *USH* is equal to one *kas-gid* (or *beru*). The information contained in this sequence of conversion rules can be condensed as follows: the "conversion factors" for the Babylonian linear system are 6, 30, 12, 60 and 30. Note that each one of these factors is also a numerical factor of the sexagesimal number system. By way of comparison, the Anglo-Saxon sequence from the inch to the mile involves the following conversion factors: 12, 3, $5\frac{1}{2}$, 4, 10 and 8. Whatever the origins of the factors for this customary system, they are clearly in no way adapted to our decimal number system.

Excavations in Mesopotamia have yielded not only mathematically interesting texts such as those from Larsa but also "problem texts" that are even more substantial. As early as 1900 the British Museum reproduced two large Old Babylonian tablets that displayed mixed mathematical problems. Their mathematical terminology was difficult and unusual, and so some 30 years passed before the majority of the problems were interpreted and fully understood. Today a large number of Old Babylonian mathematical problem texts are available for study. To say more about them here would not, however, be within the scope of this exploration of origins. Let me therefore proceed to some other texts published at the beginning of this century.



ANOTHER TABLET FROM LARSA, part of a conversion table for linear measurements, was also probably a novice scribe's practice exercise. At the right, from the top to the 10th row, is a single column of numbers from 2 to 12 (*color*), omitting 7. The last intact line, from left to right, reads "Two *beru* [equals] 12." Because in the Old Babylonian script the numbers 12, 12×60 and 12×60^2 were all written in the same way, what this actually means is that two *beru* are equal to a distance of 43,200 cubits (12×60^2 cubits). The missing 7 may be a mistake.

Between 1889 and 1900 an American expedition conducted a major excavation at Nippur, one of the largest and most important sites in Mesopotamia. By 1906 Herman V. Hilprecht of the University of Pennsylvania had reported some of the expedition's results in a volume that included reproductions of several important Old Babylonian mathematical and metrological texts and a single problem text written in Sumerian. (As recently as a year ago I was able to show that this problem text bears a series of three-dimensional geometric problems solved by reduction to cubic equations and extractions of cube roots.)

The Old Babylonian metrological texts Hilprecht published included sexagesimal conversion tables for various kinds of measures and lists of measures from the smallest to the largest. It can be supposed that such lists served to teach both the structure of the Sumero-Babylonian systems of measures and the form of the number signs and other signs belonging to each system.

Hilprecht's publication demonstrated an obvious similarity between on the one hand the Old Babylonian tables and lists and on the other the Sumero-Babylonian tables and lists. Over the years since then quite a number of additional Sumero-Babylonian texts have been unearthed. The oldest of them are a few proto-literate nonmathematical texts that date to the end of the fourth millennium B.C. The oldest mathematical tables are half a millennium younger.

If we now continue our journey backward in time to the beginning of the third millennium B.C. we come to the Sumerian period known as Ur III (2050–1950 B.C.), which immediately preceded the Old Babylonian period. Ur III is known from an enormous number of cuneiform texts. At that time the quasi-positional notation for sexagesimal numbers the Babylonians used most frequently had not yet become common. In the Sumerians' nonpositional notation separate signs stood for 1, 60 and 60^2, for 10, 10×60 and 10×60^2, and so on. Clearly with this kind of nonpositional notation a special sign for zero was not necessary.

Only a few of the many Ur III texts are of any interest because of their mathematical or metrological content. Here I shall mention just a small group of them, concerned with calculations of the amount of seed grain needed to sow fields of a given size when the spacing of the furrows was given as N furrows per *nindan.* Preserved copies of a kind of Sumerian *Farmer's Almanac* indicate that grains regularly were dropped in a furrow at intervals of two fingers (*shu-si*), that is, at a rate of 180 grains (equal

to a unit of capacity known as a shekel) per *nindan*. When, as was quite commonly the case, *N* was equal to 10, that rate could be expressed as precisely one *gur* (the largest unit of capacity) per *bur* (the largest unit of area). This example, in which two grains per finger is equal to one shekel per *nindan* or one *gur* per *bur*, shows in what a nice way the various Sumero-Babylonian units of measure were interrelated in spite of their seeming complexity.

Sumerian domination in Mesopotamia during the third millennium was broken by a brief Semitic interlude, beginning with the rule of Sargon of Akkad (2350–2300 B.C.). The existence of mathematical activity in the period of Sargon and his successors, which is hardly to be doubted, is confirmed by a handful of small tablets recording simple but far from trivial geometric exercises. The Sargonic period is exemplified here by a tablet from the city of Umma that gives an account of daily and monthly religious offerings of beer. Its text contains interesting metrological computations. It also clearly shows that number symbols could be written in two ways: either as cuneiform signs, inscribed with the wedge-shaped end of a stylus, or as circular signs made with the blunt end of a stylus.

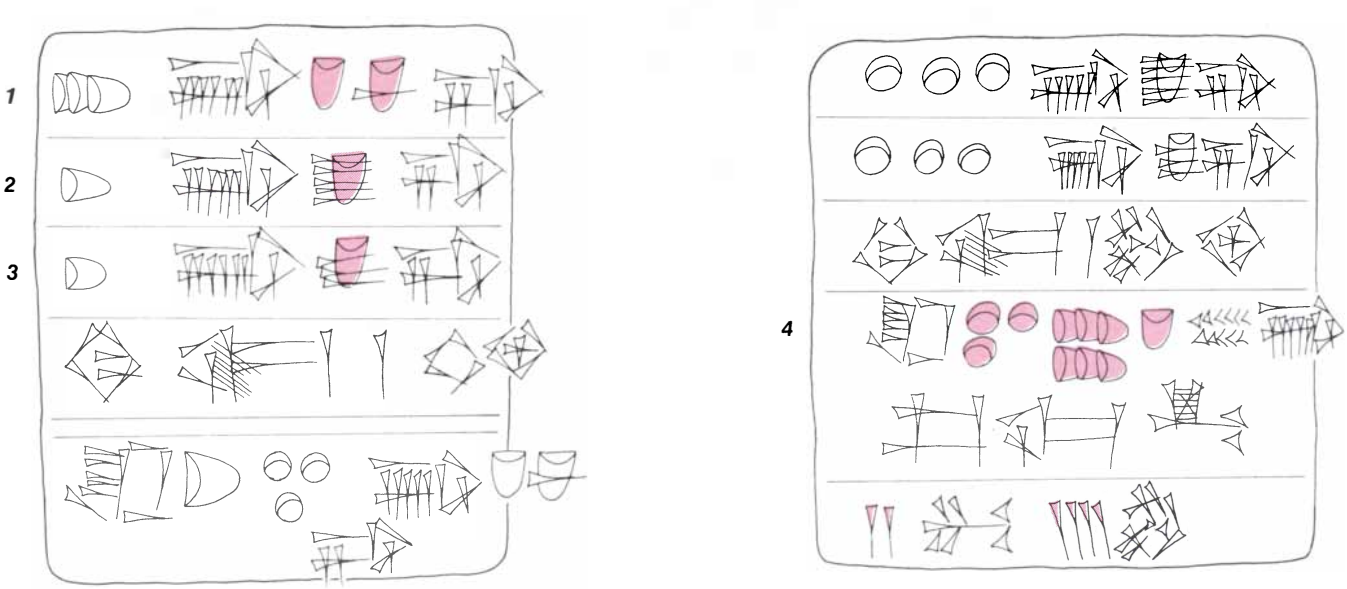
A few further mathematical texts are known from the Sumerian Presargonic period and from the preceding Early Sumerian Fara period in the middle of the third millennium B.C. The oldest metrological text from the Fara period is of particular interest. It is a table listing the areas of large square fields up to and including a field of $(10 \times 60\ nindan)^2$. I recently had the opportunity to identify another Fara text as a related geometric exercise. Two other mathematical texts from the Fara period are well known. Both deal with the same division problem involving large sexagesimal numbers. Although such a small sample of texts does not allow any far-reaching conclusions, it seems clear that the early Sumerian teachers, whose work would appear to have been instruction in what might be called applied elementary mathematics, actually excelled in working with fairly abstract problems involving very large or very small numbers or measures, with algorithms for multiplication or division, with area computations based on tables of square areas and so forth. As we shall see, the predilection for working with unrealistically large and small numbers can be traced back as far as the proto-literate period.

How did Sumerian mathematical and metrological systems evolve? Presenting the principal evidence calls for another step backward toward the beginning of the third millennium B.C. In 1928 Stephen H. Langdon of the University of Oxford published a collection of some 200 texts from tablets and tablet fragments that had been excavated a few years earlier at the small site of Jemdet Nasr in Iraq. (The Jemdet Nasr collection is now divided between the Ashmolean Museum in Oxford and the Iraqi Museum in Baghdad.) The new inscriptions were written in an archaic pictographic script that was recognizably a precursor of Sumerian cuneiform. Many of the signs in the archaic texts, however, were no longer in use by the time the cuneiform texts of the Early Sumerian Fara period were written.

For that reason even today the correct reading of many signs in the "proto-Sumerian" inscriptions from Jemdet Nasr is still not known and the texts remain more or less unintelligible. It is not even completely clear that the language of the Jemdet Nasr tablets is Sumerian, which is why I refer to the inscriptions as proto-Sumerian. Nevertheless, the numbers that appear in these texts, as a rule made with the round end of styluses, are always easy to identify. It was also immediately recognized that the number notations on the Jemdet Nasr tablets were closely related to those on even more enigmatic tablets from Iran. These tablets are for the same reason called proto-Elamite, and they are of about the same age. We shall be returning to them.

In 1936, less than a decade after Langdon's publication of the Jemdet Nasr tablets, the German scholar Adam Falkenstein published another collection of proto-Sumerian tablets, clearly older and somewhat more primitive than those from Jemdet Nasr. Falkenstein's collection included 600 inscriptions, but this was only a small part of the total excavated by a German expedition that had worked at the site of the important Mesopotamian city of Uruk in the early 1930's. The majority of the published tablets came from Level IV of the site. The balance of the collection (now in Berlin) is scheduled for publication in the near future.

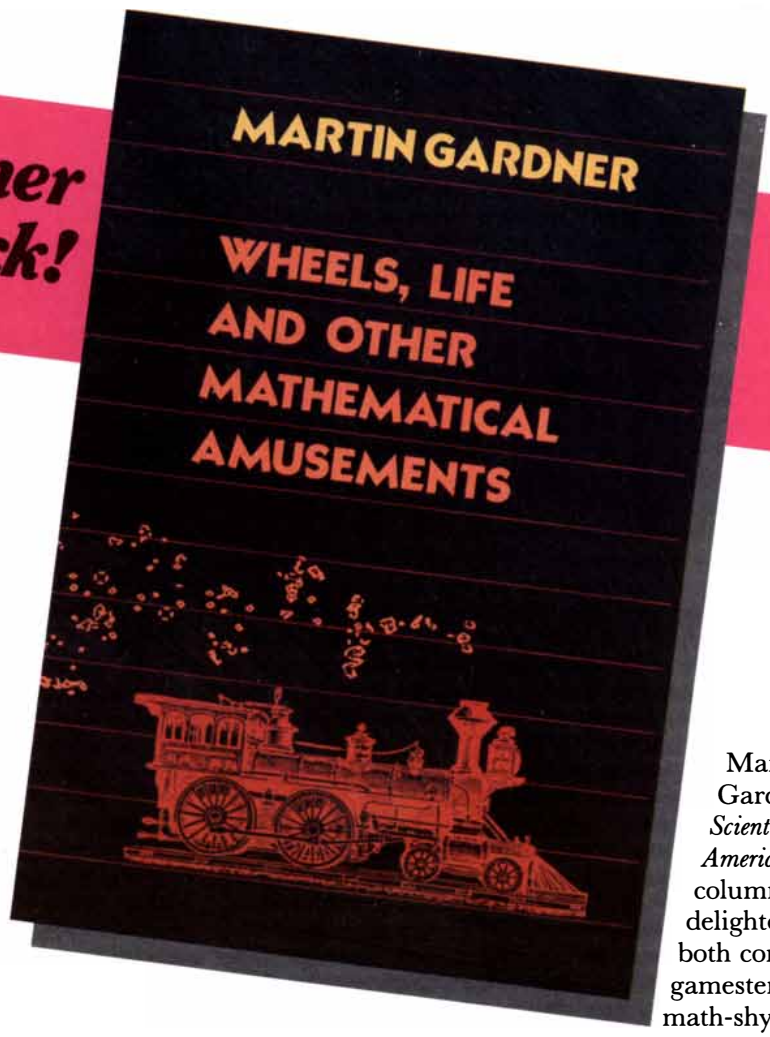
Three texts from Jemdet Nasr dealing



RELIGIOUS OFFERINGS OF BEER are recorded on the obverse side (*left*) and reverse side (*right*) of this Sargonic-period tablet from Umma in Iraq. The signs at the left in the top three rows of the obverse side respectively indicate three jars of beer, one jar and one jar. The colored symbols to the right in the first line are units of capacity, from left to right a *bariga* and a *ban* (six *ban* make one *bariga*). The unit below that is five *ban* and below that three *ban*. The sense of the three lines is that a daily offering of beer is to be made, consisting of three one-*ban* jars with a value in barley of one *bariga*, of one jar with a value of five *ban* and of one additional jar valued at only three *ban*. The wide line (*4*) of the reverse side summarizes the total value of the offering over a month of 30 days: from left to right (*color*) 3 times 10 *gur* plus 6 *gur* plus 1 *bariga*. (Four *bariga* is equal to one *gur*.) The last line of the reverse side (*5*) records the date "Year 2, Month 4."

Wheels, Life, and Other Mathematical Amusements

Martin Gardner is back!



"In WHEELS the prestidigitating master is better than ever."
—Isaac Asimov

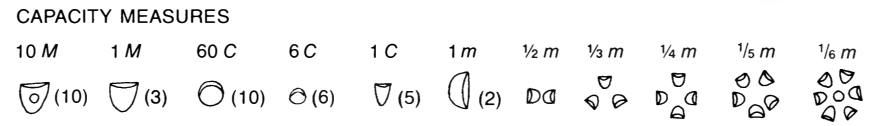
Four years have passed since the last volume in this highly successful series drawn from Martin Gardner's popular "Mathematical Games" column in *Scientific American*. This new collection of problems and paradoxes, teasers and tricks is guaranteed to enhance his already considerable reputation as a master conjurer.

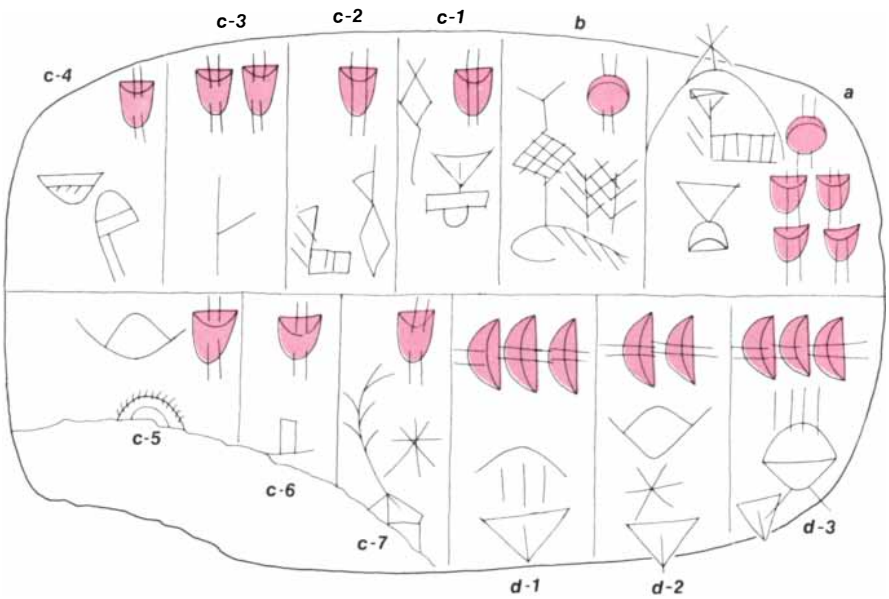
Ranging from classical paradoxes to geometric fallacies and games, WHEELS, LIFE, AND OTHER MATHEMATICAL AMUSEMENTS includes entertainments as easy as pi to wonderfully challenging excursions into the intellectual beyond. With anecdotes about famous (and not-so-famous) mathematicians sprinkled among the games and problems, WHEELS, LIFE, AND OTHER MATHEMATICAL AMUSEMENTS will provide hours of enjoyment for recreational math devotees, games buffs, or those who simply want to sharpen their wits.

Martin Gardner's *Scientific American* column delighted both confirmed gamesters and math-shy novices for almost 20 years. This new treasure-trove is sure to win him even more fans.
350 pages, 172 illustrations; $15.95 (cloth)
ISBN 0-7167-1588-0

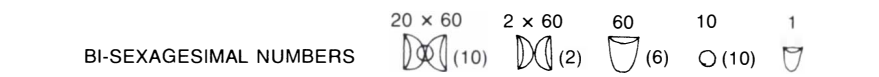
For your copy, fill out the accompanying order form, or send payment (including $1.50 for shipping and handling; California, New York, and Utah residents add appropriate sales tax) to:

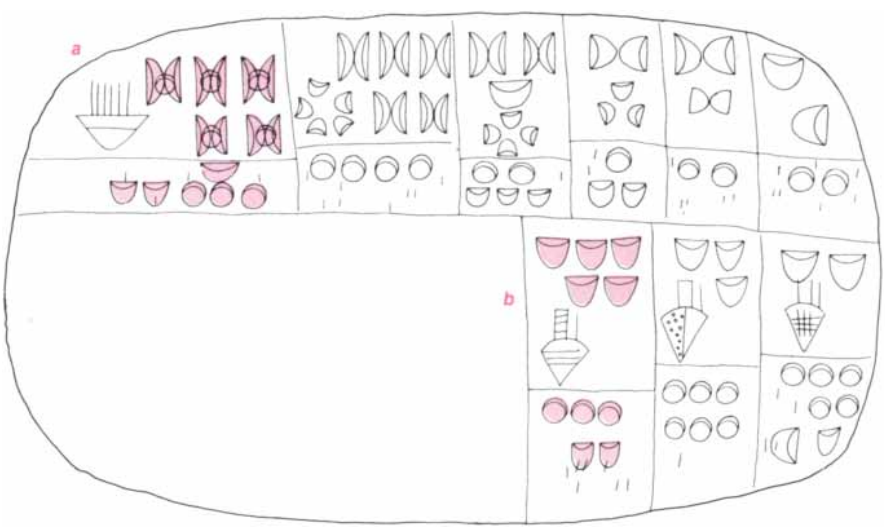
Order Department, W. H. Freeman and Company, 4419 West 1980 South, Salt Lake City, Utah 84104.





PROTO-SUMERIAN TEXT FROM URUK records the amount of emmer wheat given to a number of men and women. The system of capacity notations in archaic texts such as this one proceeds from a major unit (*M* in the factor diagram above the text) via a principal unit (*C*) to a minor unit (*m*) and fractions of a minor unit. The double lines that appear with the units of measure indicate the grain is emmer. (A different symbol, an ear of grain, represents barley.) The text reads from right to left and units are in color. Two supervisors, *a* and *b*, respectively receive 64 *C* and 6 *C*. Six of seven men (*c*-1, *c*-2, *c*-4, *c*-5, *c*-6 and *c*-7) receive 1 *C* each; one man (*c*-3) receives 2 *C*'s. Three women (*d*-1, *d*-2 and *d*-3) respectively get 3 *m*'s, 2 *m*'s and 3 *m*'s.





ANOTHER PROTO-SUMERIAN TEXT, an exercise dealing with large quantities of bread and beer, uses a "bi-sexagesimal" number system shown in the factor diagram above the text. The largest number is 20 × 60; the smallest is 1. Boxes farthest to the left in the top row (*a*) state that 6,000 loaves of bread (valued at a fractional minor unit of barley each) would have a total cost of 1 *M* + 3 × 6 *C* + 2 *C* (equal to 200 *C*). Boxes farthest to left in the second row (*b*) state that 5 (×60 is understood) jars of beer would have a cost of 3 × 6 *C* + 2 *C*, or 15 jars per *C*.

Three texts from Jemdet Nasr dealing with area computations were interpreted in 1930 by the French scholar François-Maurice Allotte de la Fuÿe, who was able to prove that the sexagesimal number system and the Sumerian system of area measures were also used in these proto-Sumerian inscriptions. With this exception, however, proto-Sumerian notations for numbers and measures continued to be little understood. It was generally assumed, however, that along with the Sumerian sexagesimal system the early scribes resorted to a decimal (or "centesimal") number system, possibly borrowed from the proto-Elamite. It was also assumed on loose grounds that in the proto-Sumerian grain texts (characterized by the sign *she,* meaning grain) the main capacity unit was a *gur* consisting of 300 *sila* (or 30 *ban*), just as in the grain texts of the later period Ur III.

In 1978 I was able to demonstrate the falsity of these assumptions by proving that in both kinds of proto-literate grain texts, proto-Sumerian and proto-Elamite, the number sign invariably read by scholars as 10 is just as likely to have the value 6. As a matter of fact, even in a single text the same number sign can be read either as 10 or as 6, depending on the context.

This new insight at once made it possible to understand the computations in hundreds of proto-Sumerian and proto-Elamite capacity-measure texts that before had been impossible to analyze. By drawing parallels between certain classes of proto-literate texts and their Sumerian counterparts I was further able to determine, at least approximately, the absolute size of the main proto-literate unit of capacity, which turned out to be about one Sumerian *ban* (equivalent to 10 liters) rather than the assumed 30 *ban*.

As a result of these advances it is now possible to begin to correctly assess the scale and character of the proto-Sumerian economies at Jemdet Nasr and Uruk and of the proto-Elamite economy at the important ancient Iranian city of Susa. Thus, for instance, a certain "bread and beer" text belonging to the Jemdet Nasr period but found at Uruk that contains computations with large numbers and small fractions has proved to be of particular interest. The text lacks any kind of date or signature. This fact alone suggests that it is no ordinary administrative record but a school text: an exercise in mathematics and metrology. As is indicated by the bottom illustration on this page, the upper part of the tablet is devoted to the computation of the amount of grain expended in baking given numbers of loaves of bread of various sizes (indicated by fractions of a minor proto-Sumerian capacity unit). The lower part of the tablet does a similar calculation for the quantity of grain

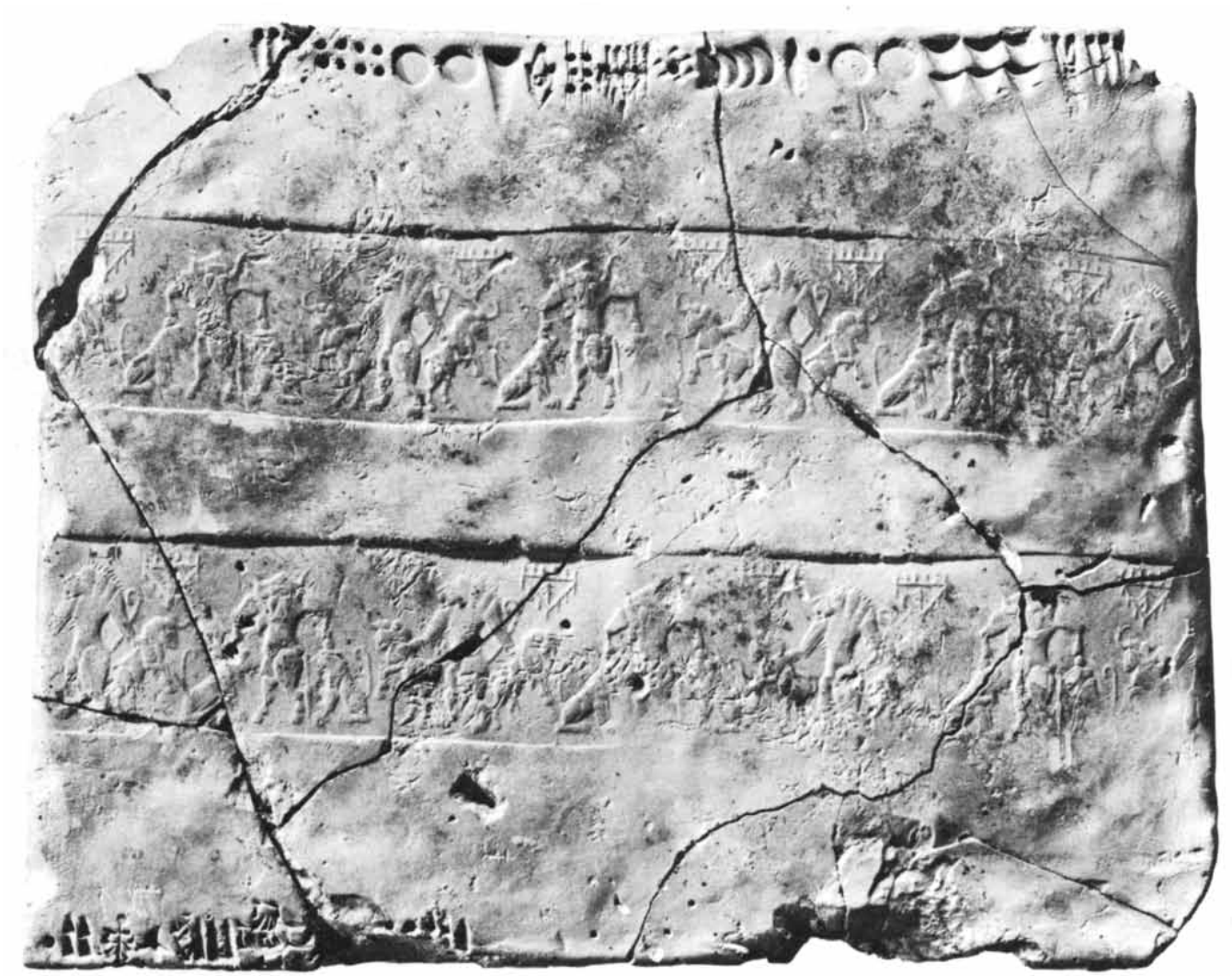
expended in brewing two jars of strong beer, three jars of medium-strength beer and five jars of weak beer. The different strengths of the three batches of beer are evidenced by the different quantities of grain per jar in the three instances.

This particular example of a proto-Sumerian text of the Jemdet Nasr type is important for several reasons. First, it establishes beyond doubt the *relative* values of quite a few units of the proto-Sumerian system that served for measuring capacity, including a number of fractional units. Second, it demonstrates the use of both the sexagesimal number system and a special "bi-sexagesimal" system. Equally important, the text can be used to find the *absolute* values of the units of this archaic system of measuring capacity.

Let us turn now to the second kind of proto-literate inscription, the proto-Elamite texts from various sites in Iran. Elam was the name given by the Babylonians and the Assyrians to the western region of Iran extending to the borders of Mesopotamia. French expeditions have been excavating at the site of ancient Susa, a capital city of Elam, since late in the 19th century. At the time of the Old Babylonian period, almost 4,000 years ago, the Elamites were using the Sumero-Babylonian cuneiform system of writing. Excavations at Susa and in other parts of Iran have revealed, however, the existence of an earlier Iranian civilization that wrote a noncuneiform script and flourished for a brief period a millennium before the time of the Old Babylonians and the Elamites. This proto-Elamite civilization was vigorous enough to extend its influence beyond Elam itself northward and eastward to distant corners of the Iranian plateau.

Hundreds of archaic clay tablets from Susa (they constitute the greater part of all known proto-Elamite texts) were published by the French scholar Vincent Scheil between 1900 and 1935. Scheil was also the first scholar to note one parallel between the Sumerian and the proto-Elamite scripts: the method of producing numerical notations. Otherwise, however, the proto-Elamite script shows no relation to any other known script. Its repertory of signs, representing the words or syllables of an unknown language, is also unhelpfully abstract and nonpictographic, a circumstance that would seem to make interpretation of the non-numerical parts of the proto-Elamite texts virtually impossible.

Scheil's first attempts to understand the nature of the number systems in his proto-Elamite texts were unfortunately not very successful. He failed to recognize that one sign—a small circular impression—need not always have the numerical value of 10. He was thus



PROTO-ELAMITE TABLET FROM SUSA shows lines of number signs only along its top and bottom edges. Much of the remaining surface of the tablet is marked by the impression of a cylinder seal that shows lions dominating bulls and vice versa. The signs to the right at the top of the tablet summarize many entries on the reverse side of the tablet: slightly more than 1,200 principal units of capacity. To the left is a smaller number, about 360 principal units. It is probably a tax or fee imposed on the transaction by the authorities.

for a time led to the conclusion that the proto-Elamites, like the ancient Egyptians, worked with a decimal number system in all their accounts and computations. Later, however, he found that some proto-Elamite texts actually feature a special sign for 60. Scheil also identified a series of fractional notations and went on to correctly identify the proto-Elamite ideograph for "grain" or perhaps "grain measure."

A proto-Elamite mathematical-metrological exercise text, published by Scheil in 1935, deals with the addition of a long series of numbers, made up of many digits and representing measures of capacity. This particular text yields excellent confirmation of my own analysis of the structure of the proto-Elamite units for measuring capacity. The system was clearly constructed to match the sexagesimal (or bi-sexagesimal) number system; it has the following sequence of conversion factors from one unit to the next: 6, 10, 3, 10, 6, 5, 2, 3, 2, 2. The corresponding series of conversion factors for the units of the proto-Sumerian system, omitting the first factor and the last four, is 10, 3, 10, 6, 5. Hence the two systems differ from each other only in their respective ways of representing fractions of the small capacity unit (a measure that corresponds to the liter in the metric system).

I wrote an account of this analysis of the two systems' units of capacity measure in a 1978 research report. The same report included another surprising discovery, namely that the proto-Elamites (but not the proto-Sumerians) used the sexagesimal number system only when they were counting people or inanimate objects such as loaves of bread or clay vessels. When they were counting animals, they worked with a decimal number system!

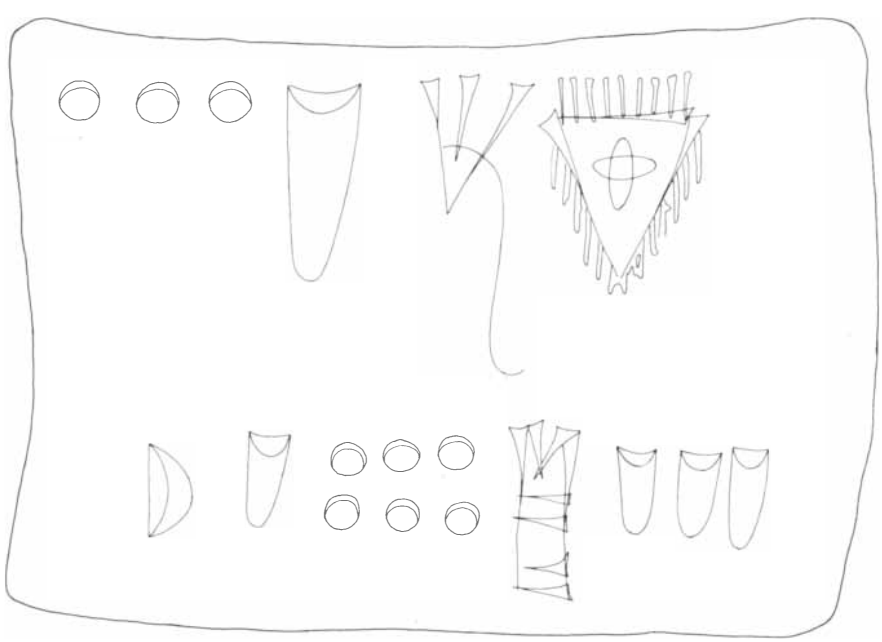
The use of the decimal system for counting animals is confirmed by a famous proto-Elamite text published by Scheil in 1923. The text is uncharacteristic: it includes unmistakable pictographs rather than the customary obscure ideographs. The heads of horselike animals are displayed in four different categories, probably divided according to sex and age. There can be no doubt, however, that the "horses" were counted with decimal numbers. To add to the general confusion resulting from the proto-Elamite use of two different number systems for counting, several of the proto-Elamite signs for numbers (not only the infamous small circular impression) have different values depending on the context. For example, the sign for 1,000 is the same as the one for 2×60.

To summarize, as a result of my identification of the systems used in the proto-literate texts for numbers and measures, it is now possible for the first time to divide the archaic proto-Sumerian and proto-Elamite numerical texts so that they fall into a comparatively small number of categories according to their content. The content of a text is indicated in particular by the kind of numbers and measures used in it. For example, among the proto-Elamite texts the main categories are enumerations of people (sexagesimal counting numbers), computations of grain expended on rations for people (capacity numbers and sexagesimal numbers) or for animals (capacity numbers and decimal numbers), bread-and-beer texts (capacity numbers and sexagesimal numbers), accounts of groups of animals (decimal numbers), accounts of loaves of bread or clay vessels (sexagesimal numbers) and accounts of large quantities of grain delivered to or dispatched from storehouses, in one case during a five-day "week" (capacity numbers). Additional categories are to be found among the proto-Sumerian texts: area computations (area numbers and sexagesimal numbers) and in a single instance a seed-grain text (area numbers and capacity numbers).

Much work remains to be done before the existing collections of proto-literate texts, published or still unpublished, can be persuaded to give up all the information they contain. Nevertheless, so much progress has been made, particularly in the past decade, that it is now at last feasible to sketch a plausible outline of the development of numeration and metrology in southwest Asia from prehistoric to Late Babylonian times.

In particular it is now known, from the work of Denise Schmandt-Besserat of the University of Texas at Austin, that a variety of clay "tokens" served in the region continually from the ninth millennium B.C. to the end of the fourth millennium to designate numbers, measures and perhaps categories of objects. It is also known that the use of loose tokens was complemented late in the fourth millennium by the device of enclosing selected symbols in protective envelopes of clay, many of them bearing surface impressions resembling the enclosed tokens. The envelopes in turn seem to have inspired the invention of "impressed tablets." (The term refers to clay tablets bearing number notations and usually cylinder-seal impressions but lacking any abstract or pictographic word signs.)

The similarity of the form of the number signs on the impressed tablets and the shape of the earlier clay tokens show a continuity of representations of numbers and measures from the time of the rise of city-states in southwest Asia backward in time to the early Neolithic, as long as 10,000 years ago. Moreover, the fact that notations for units belonging to various proto-literate measure systems also clearly appear on some impressed tablets shows a continuity in representations of numbers and measures forward in time from late in the fourth millennium B.C. all the way to the Old Babylonian period, which was some 2,000 years later. And, as we have seen, our own decimal system is a reflection of the same continuity down to the present day.



GRAIN-RATION RECORD FROM PROTO-ELAMITE SUSA is read from right to left. The text begins with a "hairy triangle," thought to be the issuing authority. Next is a "plow" sign that is probably a field worker. The number signs that follow continue at the right of the bottom line: $1 \times 60 + 3 \times 10 + 3$. Next is a sign like a sheaf of grain; it means barley. Last is a capacity number: $6 \times 6\,C + 1\,C + 1\,m$. The sense is that 93 workers receive two minor units each.

Solitary Bees

Most species of bees are not social like the honeybee. They are solitary, meaning that the female builds her own nest. They play a profoundly important role in pollinating crops and wild plants

by Suzanne W. T. Batra

For most people the beehive and the intricate social organization of the honeybees that populate it are the hallmark of bee life. In reality more than 85 percent of the some 20,000 bee species are not social but solitary. Each female independently mates, makes her own nest of about 10 brood cells, stocks the cells with food for the young, lays an egg in each cell and dies before the next generation emerges. The solitary bees play immensely important roles in ecological systems, particularly in the pollination of crops and wild plants. Indeed, farmers in some parts of the world are beginning to make specific arrangements to accommodate solitary bees for the pollination of crops (such as alfalfa) that are not pollinated effectively by honeybees.

Bees apparently evolved from predaceous ancestors such as the solitary mud-dauber wasps in about the middle of the Cretaceous period (100 million years ago), when flowering plants became the dominant vegetation on the earth. Now bees live everywhere, showing the greatest abundance and diversity of species in semiarid, warm temperate climates. Many solitary bees do not resemble honeybees in general appearance but instead look more like wasps, flies or bumblebees. They range in length from 1.5 to 46 millimeters. Some species are hairless and shiny, others are densely furry. Although many of them are drab, some (such as the emerald green orchid bees) are quite beautiful. The hair or the spots and stripes may include all the colors of the rainbow.

Solitary bees are distributed among nine families: the Colletidae, or membrane bees, which are most numerous and diverse in the Southern Hemisphere; the Andrenidae, or digger bees, which are mainly in the Northern Hemisphere; the Halictidae, or sweat bees (so named because they are attracted by human sweat), of worldwide distribution; the Megachilidae, or leafcutter and mason bees, also worldwide; the Anthophoridae, or carpenter and miner bees, predominantly tropical; the Melittidae, Oxaeidae and Fideliidae, which are small families with a limited distribution and no common names, and the Apidae, or honeybees, stingless bees and orchid bees, all of which are mostly tropical, and bumblebees, which are mainly in the Northern Hemisphere and are the most able of all the bees to tolerate low temperatures.

The nests and brood cells of most solitary bees are made underground. Bare, dry, light soil exposed to the morning sun is preferred. In such habitats many nests of their relatives, the ants and the solitary wasps, may also be found. In a favorable area thousands of nests of ground-dwelling solitary bees may form dense aggregations. The females do not, however, cooperate in nest-building chores.

The largest recorded aggregation of nests occupied about 360,000 square meters extending for seven kilometers along a bank of the Barysh River in the U.S.S.R. It consisted of an estimated 12 million nests made by two species of sweat bee and a species of miner bee. In about mid-May every year I receive at my U.S. Department of Agriculture office in Beltsville, Md., many agitated telephone calls about swarms of solitary bees nesting on lawns in the Washington area. These are usually the membrane bee *Colletes thoracicus,* which is active as an adult for about six weeks. The largest aggregation of this honeybee-size insect included about 104,000 nests in a suburban backyard of 1,187 square meters.

Aggregations of nests may persist for many years. For example, a species of miner bee has occupied an adobe wall near Baltimore for at least 40 years. The aggregations are probably maintained because the young females on their first flight memorize visual landmarks in the area, just as new honeybee workers remember the location of their hive. Pheromones (chemical signals) or odors released by other females in the aggregation may also be attractive. After mating, a young bee returns to make her own nest among the other nests in an environment shown to be satisfactory by the survival of her ancestors.

Bees that nest in the ground usually have plenty of room for nest sites. They face, however, the problem of protecting for almost a year the nutritious honey-and-pollen provision they have laid up for their brood (and the delicate, helpless, grublike brood itself) from soil moisture and a wide variety of soil organisms such as bacteria, yeasts, fungi, nematodes and mites. The large majority of ground-dwelling bees line their underground brood cells with special waterproof secretions before provisioning them. These bees possess an abdominal gland (Dufour's gland) that opens near the sting. In some species it occupies half of the abdominal cavity. The gland contains an oily liquid of a fragrant-to-musky odor. (A few species collect plant resin or cut pieces of leaf to line their brood cells. The Fideliidae and other bees that live only in deserts do not line their cells at all.)

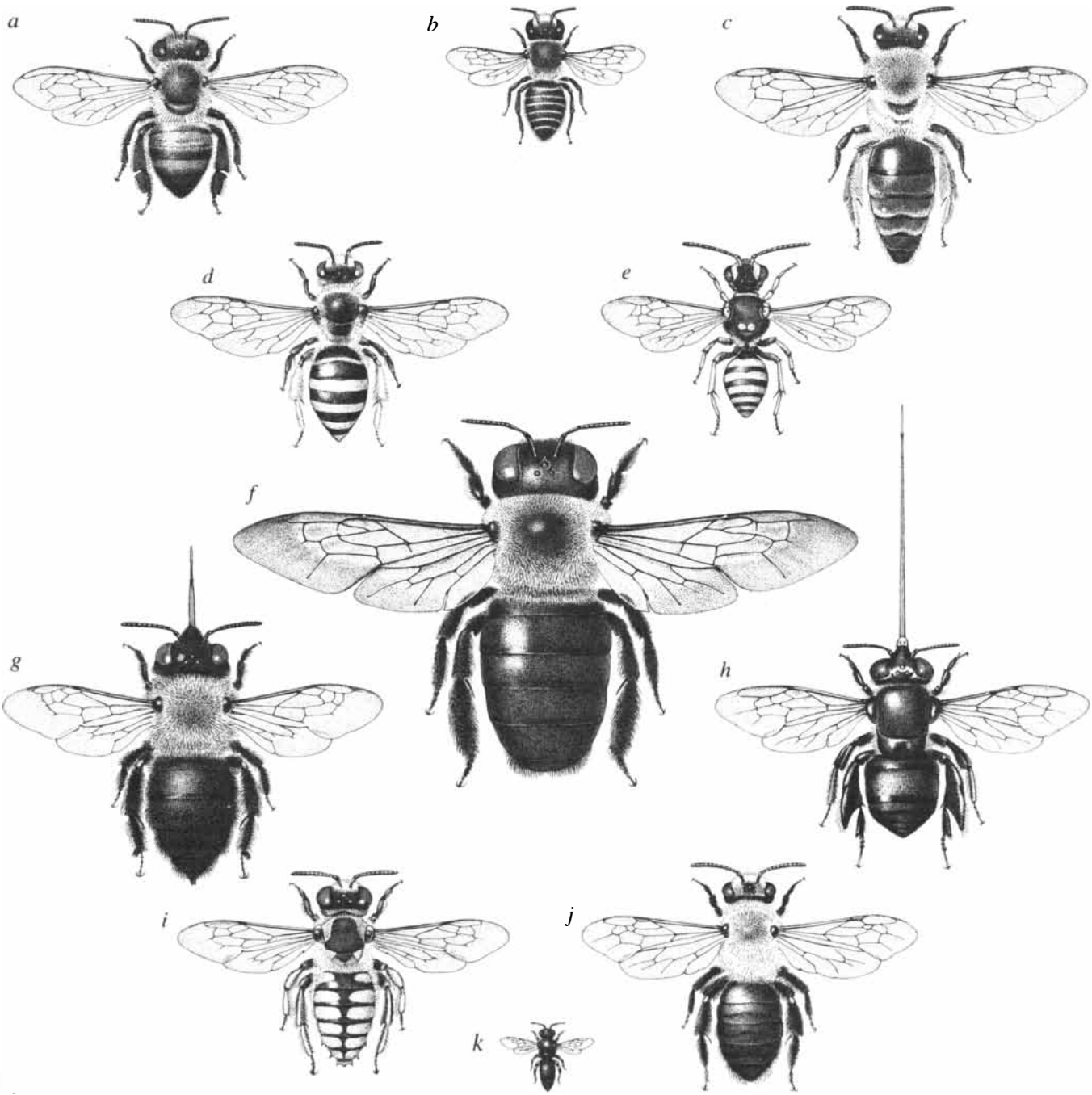
By inducing bees to nest in narrow, glass-sided chambers filled with soil I was able to observe their subterranean nest-making behavior. After a chamber for a cell is excavated by the bee and often polished to a shiny smoothness by the pygidial plate at the end of the abdomen, dollops of Dufour's-gland secretion are deposited in it and are repeatedly smeared around by rhythmic sweeps of the bee's tongue, by special velvet-like pads on the abdomen or by hairy brushes on the legs. Most solitary bees fly about during the day collecting provisions, lay an egg in the afternoon, seal the cell, begin digging a new cell chamber in the evening and apply the Dufour's-gland secretion at night, so that the new cell is dry and ready for provisioning the next morning. Such bees work day and night with little respite.

Henry M. Fales and Abraham Hefetz of the National Heart, Lung, and Blood Institute and I analyzed the secretions of *Colletes* membrane bees. We found that the Dufour's gland contains fragrant, musky-smelling compounds (macrocyclic lactones). When they are deposit-

ed in the brood cell, they polymerize to form a flexible, transparent, waterproof membrane composed of a natural polyester (mainly 18-hydroxyoctadecanoic acid and 20-hydroxyeicosanoic acid). The membrane forms a miniature plastic bag in each cell. The bag even has a "zip lock" closure: a flap that is lifted and sealed by the female bee after she has laid an egg. The membrane is quite durable, remaining intact in soil for a year or more.

For comparison we also analyzed the Dufour's-gland secretion and brood-cell lining of *Anthophora abrupta,* a miner bee. This bee's large Dufour's gland holds a mixture of transparent, slightly fragrant oily triglycerides. When they are secreted into the brood cell, they are converted, possibly by salivary enzymes, into white, waxy diglycerides with a rancid odor. This secretion is also added to the provision for the larvae. The developing larvae eat it and even rasp away and eat the cell lining. Until the miner-bee larvae's feeding habits were discovered the honeybee was the only bee known to feed its larvae with a secretion (royal jelly, secreted orally and consisting mainly of protein).

Bees that do not live underground need not protect their young so carefully from excess moisture, and probably for this reason they have relatively small Dufour's glands. Most ma-



SOLITARY BEES are compared with the honeybee, *Apis mellifera* (*a*), in drawings that are done to a common scale. The solitary bees are the alfalfa leafcutter, *Megachile rotundata* (*b*); the membrane bee, *Colletes thoracicus* (*c*); the alkali bee, *Nomia melanderi* (*d*); a cuckoo bee, *Nomada luteoloides* (*e*), so named because it lays its eggs in the nests of other solitary bees; the carpenter bee, *Xylocopa virginica* (*f*); the miner bee, *Anthophora abrupta* (*g*); the orchid bee, *Euglossa ignita* (*h*), shown with its unusually long tongue extended; a mason bee of the genus *Anthidium* (*i*); a digger bee, *Andrena carlini* (*j*), and a silk-making membrane bee of the genus *Hylaeus* (*k*). In life *Hylaeus* is four millimeters long. The bees portrayed represent only a few of the many solitary-bee species, of which there are some 17,000.

son bees and leafcutter bees use existing aboveground holes as nests, building brood cells of mud, resin and plant fiber or of neat layers of ovals and circles precisely cut out of leaves and petals. Leafcutters may damage plants. In an extreme example the yield of two hectares of soybeans was reduced 30 percent by the removal of leaf tissue by bees nesting nearby. Some mason bees build clustered brood cells of mud or resin on branches or walls. One common species of solitary bee is notable for defacing Egyptian and Greek monuments with its numerous large, lumpy nests of dried mud.

Carpenter bees drill neat round holes with their powerful mandibles into wood, bamboo or pithy stems. In their galleries they construct series of brood cells separated by partitions of agglutinated sawdust. They sometimes do significant structural damage to the beams or eaves of buildings.

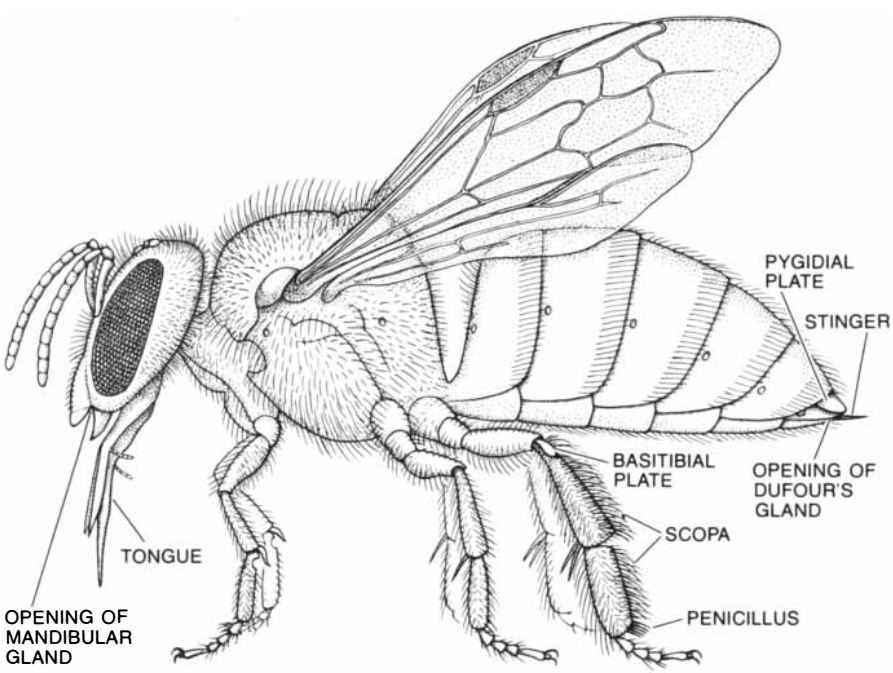
Membrane bees of the genus *Hylaeus* are wasplike in appearance and carry provisions internally as wasps do instead of on special hairs in the manner of other bees. *Hylaeus* bees are unique in making brood cells of true silk inside existing holes. The silk is secreted orally and applied by the female bee's broad tongue. Small predatory wasps belonging to the family Pemphredoninae also (uniquely among wasps) make cells out of silk. It is possible that bees have evolved, through primitive forms such as *Hylaeus,* from this group of wasps. The secretion of silk by adult insects is quite rare.

Most female solitary bees, like the honeybees and the bumblebees, can defend themselves individually by stinging, but their venom is usually mild. The nests of solitary bees, unless they are densely aggregated, seldom attract vertebrate predators to their meager provisions and brood. Therefore the bees have not developed the aggressive mass-stinging counterattack provoked when honeybee and bumblebee nests are disturbed. Solitary bees' nests, however, are particularly vulnerable to attack by other invertebrates and by microorganisms, and the bees have evolved ways to repel them.

Fungi are the primary enemies of solitary bees that nest underground. The bee must work quickly to finish the preparation of the provision and to lay an egg so that the larva can consume the food before it spoils. The brood-cell lining isolates the contents from free water in the soil. Water vapor enters many cells, however, along with the gases needed for the respiration of the larvae.

The provisions of most bees consist of a mixture of honey and pollen. The sugars of the honey are hygroscopic (they attract moisture from the air), and so they may gradually become more dilute. This may account for the fact that bee larvae may weigh more than their provision originally did. All is well with this system unless populations of yeasts and bacteria in the provision explode. These microorganisms, usually present in the nectar from flowers, are normally deposited along with the provision. Their growth is prevented, however, by the high sugar concentration; with solitary bees, as with honeybees, the excess water in the honey evaporates away as the flower nectar is converted into honey by enzymatic reactions. If yeasts grow unchecked, the bee larva gets indigestion and becomes weak or dies. The fermented provision is then overrun by filamentous fungi.

Working at Utah State University with two of my colleagues in the Department of Agriculture (George E. Bohart, an entomologist, and my husband Lekh R. Batra, a mycologist), I found that bee adults and broods are attacked by at least 124 species of fungi. The solitary alkali bee has developed an interesting method of protecting its brood from some of the 49 species of fungi that afflict it. In common with most other solitary bees the alkali bee seals each cell after provisioning it and laying an egg. It has long been assumed that such "mass provisioning" bees have no further contact with their young. When the contents of a cell become infested with a fungus, however, the mother bee opens the cell and packs soil into it. This procedure reduces the amount of air available for the growth of the fungus and so prevents the fungus from spreading to adjacent cells. Although some soil-dwelling bees may annually lose as much as half of their brood to fungi, the species that nest in drier sites such as sound wood, stems and adobe walls are usually less severely affected.



FEATURES OF THE ANATOMY of typical solitary bees include the pygidial plate, which the female uses to smooth the inside of a nest she has excavated in the ground, and Dufour's gland (within the abdomen), with which she applies a secretion that waterproofs the inside of the nest. Mandibular-gland secretions serve in communication and self-defense. The scopa, which in some species is on the abdomen, is an array of hairs for carrying pollen. A solitary bee walks on its knees (the basitibial plates) when it is excavating or provisioning its burrow.

Specialized solitary bees, known as cuckoo bees or inquilines, are common parasites of other solitary bees. Like its bird namesake, a cuckoo bee does not make its own nest but instead lays its eggs in the nests of other bees. About 15 percent of all bee species are cuckoo bees. These parasites may be quite host-specific, each attacking a single host or a group of related host species. Most families of bees include some cuckoo species, which may attack hosts belonging to their own family or to other families. The young larvae of many cuckoo bees have long mandibles that serve to crush the larva of the host bee; older cuckoo-bee larvae, feeding after the host has been eliminated, have normal small mandibles. Usually the most easily observed cuckoo bees are the slender, wasplike, rust-colored or yellow-and-black species of *Nomada* (Anthophoridae), which flit about in early spring near the ground among the woodland flowers in search of temporarily unoccupied nests of their *Andrena* hosts.

A solitary bee must spend several hours daily away from its nest foraging for nectar and pollen. During this time the helpless brood is particularly vulner-

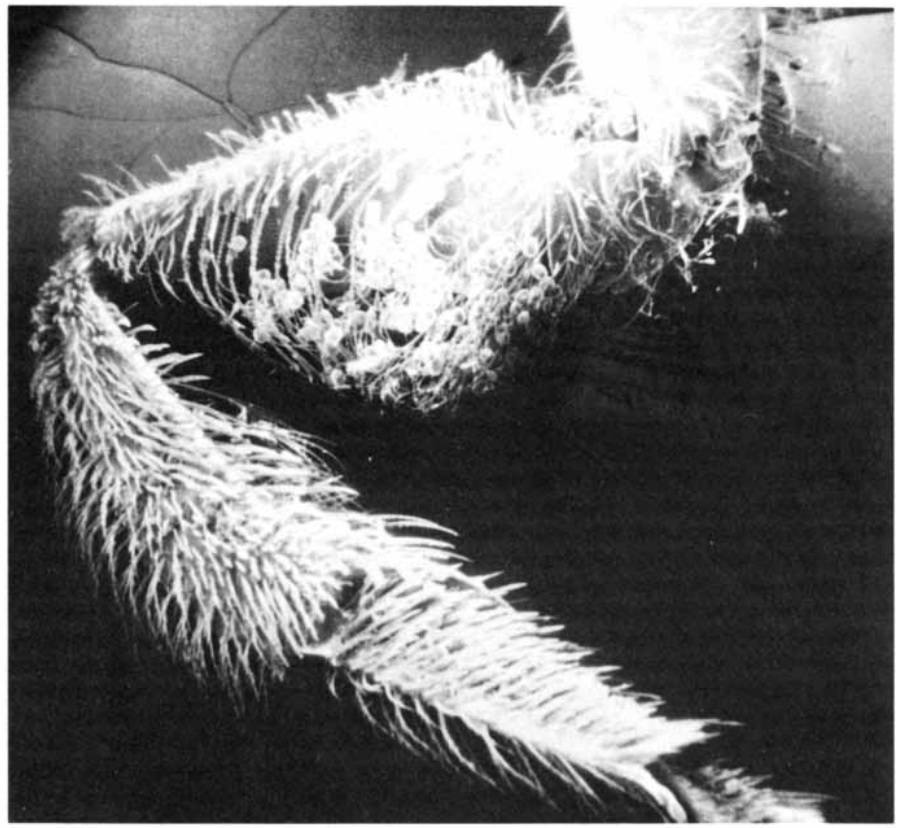
able. The brood cells of some species are tightly sealed with a thick cap of resin, secretion or clay, but the cells of many solitary bees have a permeable cap of lightly packed soil, leaves or wood dust that is readily penetrated by insect parasites and predators. The mother bee may try to protect the nest in her absence by positioning it so that the entrance is hidden, by temporarily filling the burrow with soil, by heaping soil in a small tumulus over the entrance or by building a compact turret of earth that may deter crawling insects. Nevertheless, solitary-bee broods suffer losses to a variety of predatory and parasitic wasps, ants, several kinds of flies and beetles, strepsipterons, mites and nematodes. Some of the parasites deposit their eggs or larvae on flowers, where they are accidentally picked up by the foraging bee, brought home to the nest and included in the brood-cell provision, where they lie ready to attack the bee larva.

The colorful, symmetrical and fragrant blossoms of higher plants evolved to attract pollen-dispersing animal life. Bees are the most efficient and effective of all pollinators for several reasons, including their sheer abundance, their rapid flight, their tendency to visit several flowers of the same species in succession, their need for large quantities of nectar and pollen and their specialized hairs, which can trap and hold as many as 15,000 pollen grains per bee.

An individual bee may possess several types of hair, which are grouped on the body and legs according to their function. The dense hairs on the hind legs of most solitary bees form a pollen-holding brush called the scopa. (In leafcutter and mason bees it is below the abdomen.) Honeybees and their relatives have shiny bare patches on their hind legs known as pollen baskets or corbiculae, which carry nectar-moistened lumps of pollen or nesting material. The hairs of bees are modified according to the type of pollen collected by the particular species. For holding fine, dry pollen grains a cluster of down-like, branched, plumed hairs is best. Large or sticky pollen grains call for stiffer hairs. Such hairs often have shorter branches or are formed with spiral ridges resembling a corkscrew. They may also have spatulate or hooked tips.

Combs of bristly hairs on the legs and mouthparts of bees serve for grooming and for transferrring pollen to the scopa. The Dufour's-gland secretion is applied with special straight hairs in tufts resembling a paintbrush. To collect floral perfumes or plant oils some solitary bees have velvety hairs that form sponges and flattened hairs arranged in combs.

The value of bees for the pollination of wild plants and crops is incalculable. No price can be placed on wild plants, which are pollinated primarily



POLLEN-CARRYING APPARATUS is the scopa, the group of branched hairs near the top of this electron micrograph of the hind leg of a sweat bee of the genus *Lasioglossum*. Several spherical grains of pollen are visible in the scopa. The enlargement is about 62 diameters.

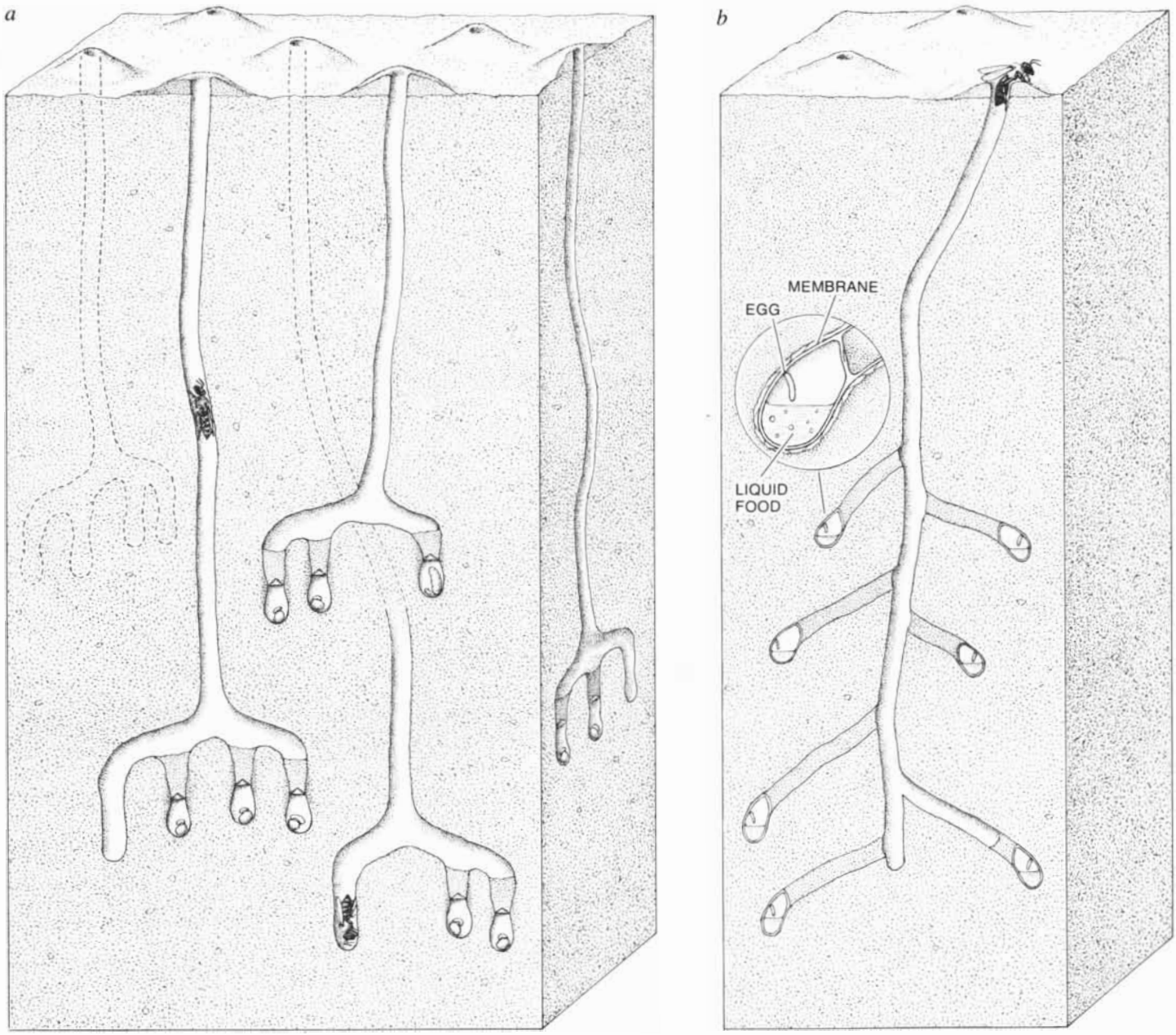
INNER SURFACE of the sweat bee's hind leg shows (*left*) the spiral-ridged hairs that serve to manipulate pollen and (*lower right*) the penicillus, used somewhat like a paintbrush for spreading a waterproof secretion in the nest. The enlargement of the hairs is 300 diameters.

by solitary bees. It has been estimated that in the U.S. the value of some 90 crops that depend on insect pollination for maximum yield is $17 billion per year. If seeds of insect-pollinated forage plants for livestock are included, about a third of the nation's food supply depends directly or indirectly on pollination by insects, mainly bees. It is reasonable to assume that solitary bees account for at least half of this crop-pollination activity in the U.S. and perhaps more in areas where beekeeping is less prevalent.

Unfortunately populations of solitary bees decline under the conditions of large-scale modern agriculture and urbanization. When large fields of single crops are planted as monocultures, where all plants flower simultaneously for a short time, solitary-bee populations may be insufficient to handle the pollination, and when the flowers finish blooming, nothing is left for the bees to eat. Cultivation and construction eliminate supplementary floral hosts and nesting areas. Irrigation may destroy nests. It also increases the damage done to the bee brood caused by fungi. The careless application of pesticides kills many bees. The sulfur dioxide in polluted air can cause bees to reduce their flying time.

These effects can be observed by comparing, say, the bee populations on flowering apples in the center of a large, well-sprayed commercial orchard with the bee populations on flowering trees in abandoned orchards or at the edge of a forest. In order to achieve adequate pollination in the absence of enough native bees, growers have to rent hives of honeybees from beekeepers at a cost of about $20 per hive. Hundreds or even thousands of hives may be needed. The value of pollination services far exceeds the value of the honey and wax obtained from honeybees.

Honeybees pollinate many types of plant, but some plants are pollinated only ineffectively. For example, the flower of alfalfa has a little spring in it that is triggered when a bee pushes her head into the flower. The stamens pop



NESTING ARRANGEMENTS of six species of solitary bee are depicted. Each nest has several individual cells, each of which the female supplies with provisions for the larva before she lays an egg. The cells are usually made at a rate of one a day (in good weather) in the period of several weeks when the adult female is active. The nest of the alkali bee, *Nomia melanderi* (*a*), is made in moist soil. The membrane bee, *Colletes thoracicus* (*b*), which also nests in soil, gets its common name from the fact that it lines the inside of each cell with a poly-

up and rap the bee on the head, dusting her with pollen. Honeybees apparently do not like being rapped on the head, and since they are fairly intelligent insects, they learn to poke their long tongue into the side of the flower, avoiding the spring. Thus they take nectar but accomplish no pollination. Because alfalfa is one of the major forage legumes, the supply of seed (a $100-million-per-year industry in the U.S.) is of critical importance.

In the northwestern U.S. and Canada two species of solitary bee—the alfalfa-leafcutter bee and the alkali bee—are intensively managed to pollinate alfalfa. These persistent insects do not seem to mind being rapped by the flowers, and they collect the pollen for provisioning their nests as it is showered on them. The alfalfa-leafcutter bee, originally European, quickly visits many flowers in succession and is a highly efficient pollinator. It nests in groups. Farmers make suitable holes in "bee boards," which they put in shedlike shelters in the alfalfa fields. The brood cells of this bee are built out of bits of alfalfa leaf.

The alkali bee is a species of sweat bee native to the U.S. It likes to nest on alkaline flats, where the subsoil stays fairly moist. Growers who accommodate this bee construct "bee beds," which consist of plastic-lined trenches in the alfalfa fields. The trenches are filled with specially prepared soil that is maintained at just the right level of moisture to attract and hold dense aggregations of alkali bees. There may be as many as 2,100 nests per square meter.

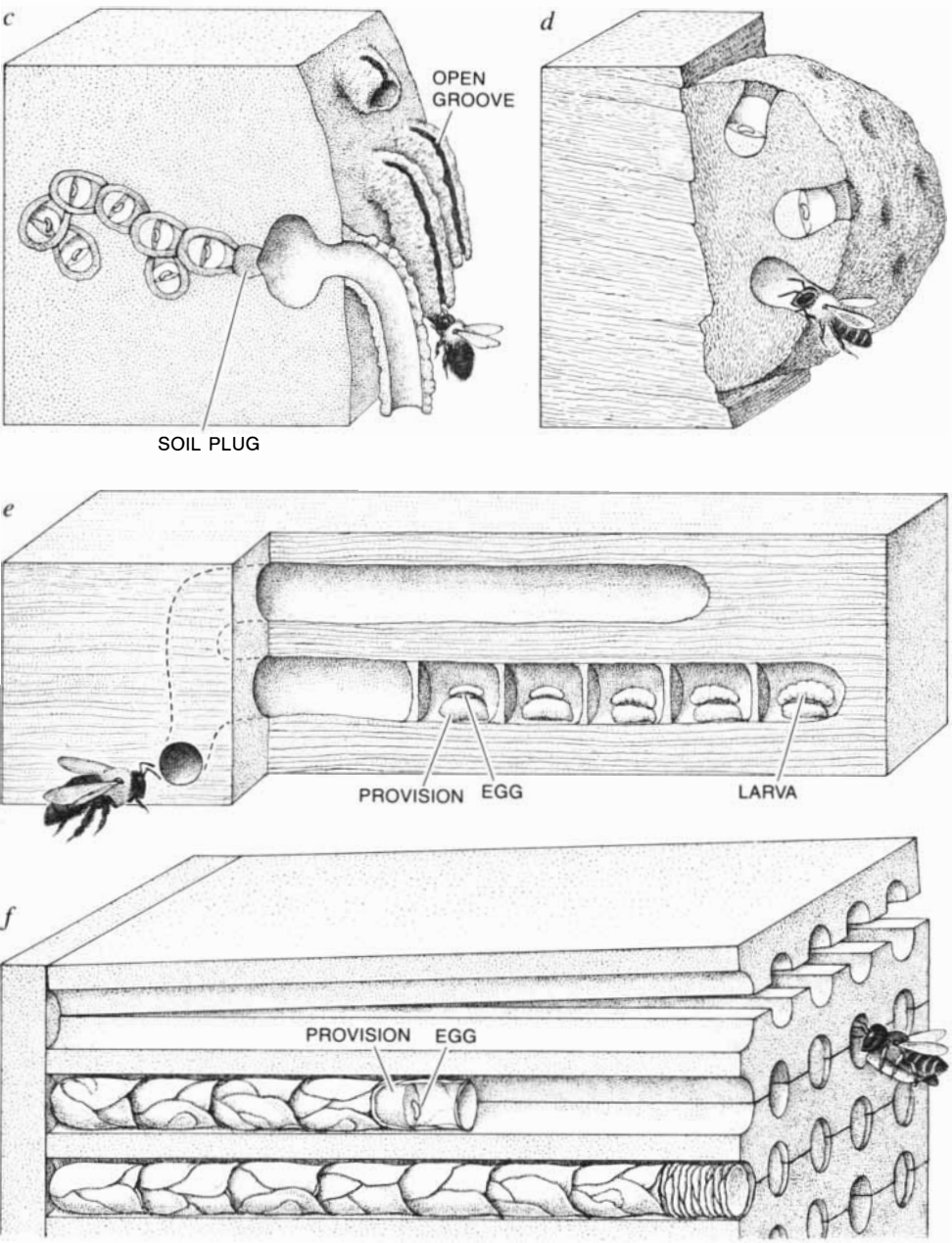
Another solitary bee utilized commercially is a mason bee, the Japanese hornfaced bee. Apple flowers are often not pollinated efficiently by honeybees, which tend to work slowly, to avoid touching the stamens and to fly away to forage on dandelions and other weeds. Some 40 years ago Japanese farmers noticed that the little hornfaced bees nesting in their thatched roofs were energetic and effective pollinators of apples. Over the years the management of this bee has been perfected, and bee shelters containing hollow reeds for nesting are seen throughout the apple-growing districts of Japan. Recently, in collaboration with Yasuo Maeta of Shimane University and other Japanese workers, I have successfully established populations of this Asian pollinator in the eastern U.S., where it is being tested for its adaptability to the climate and its acceptance of the available nesting material.

Many species of bees have life cycles and habits that depend on the availability of specific host plants. Bees that gather pollen from one species or a few closely related plants are termed oligolectic. Bees such as the honeybee, which can gather pollen from a variety of plants because of the hive's long collective period of adult activity, are termed polylectic.

Flowering plants and oligolectic bees have developed some remarkable mutualistic relations. For example, the American native gourds are most effectively pollinated by squash bees, which are several native species of miner bee. Some species of sweat bee pollinate evening primroses and fly at night, which is unusual for bees. This activity is aided by their eyes, which are enlarged for light gathering. Other specialized miner bees pollinate wild morning glories. Perhaps the most highly specialized solitary bees are those visiting flowers that yield oils, which are collected instead of nectar (for mixing with pollen in the provision for the brood); to pick up the oil these bees have hairs arrayed in comblike structures.

Orchids are among the most specialized of flowers. Some of them have evolved unique ways of exploiting the behavior patterns of solitary bees. Some orchids have false nectaries: colored spots yielding no nectar. When a bee probes a false nectary for nectar, a pollinium (pollen packet) is attached to the insect, which carries it along to pollinate the next orchid. This deception works until the bees learn that the nectaries yield nothing.

Other orchids have color patterns and aphrodisiac odors that resemble the ap-



ester membrane of its own making. The miner bee, *Anthophora abrupta* (*c*), employs a different technique in dry soil. The mason bee of the genus *Chalicodoma* (*d*) builds its nest on the surface of wood or rock; the carpenter bee, *Xylocopa virginica* (*e*), drills into wood. Some growers of alfalfa make grooved "bee boards" (*f*) for alfalfa-leafcutter nests to aid pollination.

pearance or the odor of female bees or wasps. Males are attracted. They try to copulate with the flower and in so doing pick up a pollinium, which is deposited at the next enticing orchid. Still other orchids mimic certain male miner bees. When an aggressively territorial male bee sees the orchid flower nodding in the wind, the bee strikes it in an effort to drive the "rival" away. As a result a pollinium is stuck to the bee's head.

The most complex relation is the one between certain South American orchids that yield neither nectar nor pollen and several species of orchid bee. The male bees have swollen hind legs with hair-lined slits leading to internal pockets and have velvety pads on their front feet. The males form leks, or all-male swarms, that attract females for mating. Certain males are more attractive than others because they have collected mixtures of fragrances from the orchids. They brush the flower's odor glands with their front feet and stuff fragrant particles into their hind-leg slits. While the bee is doing so it picks up a pollinium from the orchid. Since these bees can fly as far as 23 kilometers in a day, they are effective pollinators for sparsely distributed plants, such as certain of the orchids.

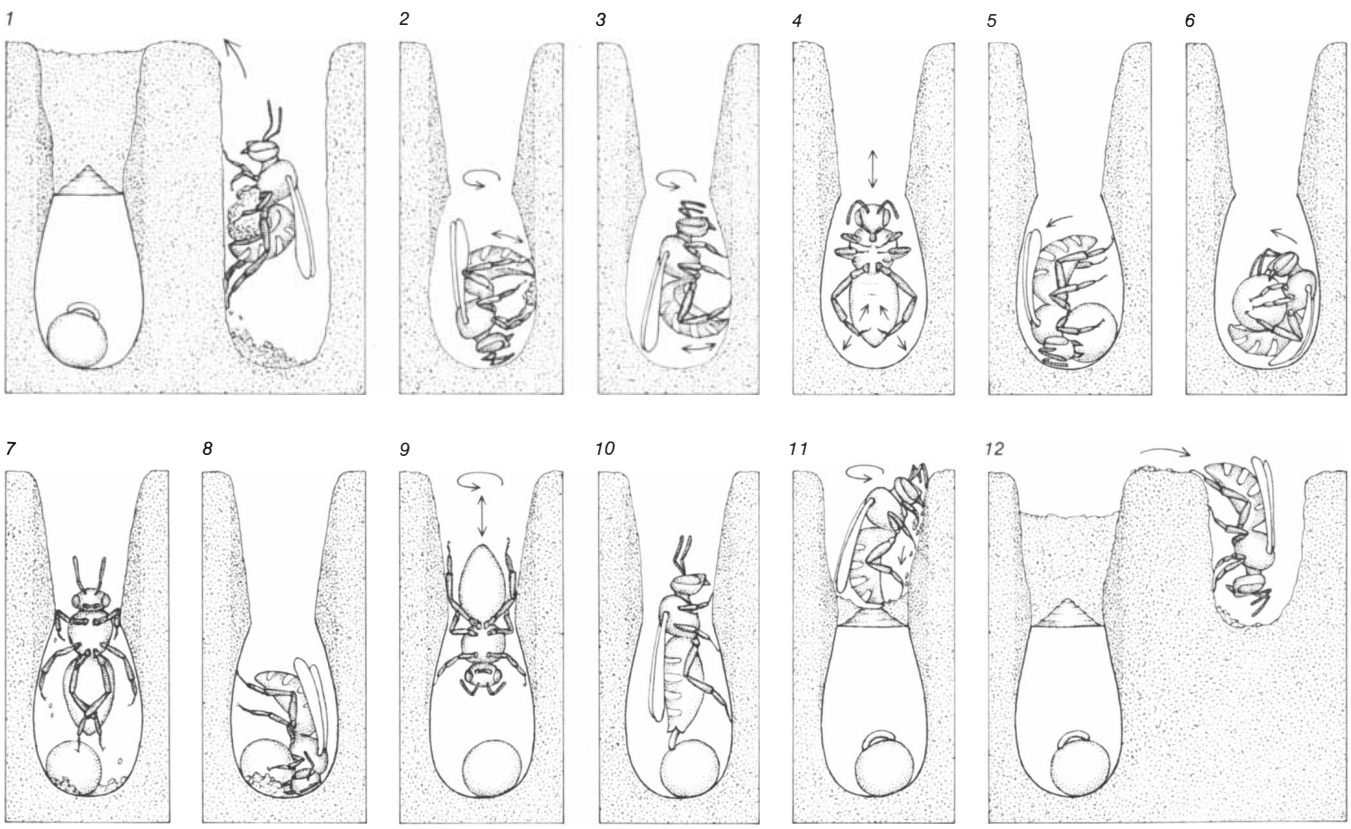
Orchid bees apparently respond to other odors. Males of one species have become a nuisance because they are attracted to houses that have been dusted with DDT for mosquito control. They collect the powder and pack it into their leg slits. The compound does not seem to harm the male bees, but whether or not it has the intended aphrodisiac effect on the females is not known.

Male solitary bees ordinarily emerge from the pupa a few days earlier than the females of their species. They feed on floral nectar and pollen and search for newly emerged virgin females. In some species the males form loose groups that jointly patrol the nesting site, nearby landmarks or flowers that are regularly visited by females. When males see a female bee or an object shaped like one, they pounce on it, but unless a sex pheromone is present they let it go immediately. If a virgin female appears, a pile of writhing males, all attempting to mate, may form rapidly on top of her. Usually the first male to grasp the female holds on to her tightly and succeeds in mating. Males of most species merely compete to reach the females first and are not noticeably aggressive toward one another. In some species of carpenter bee and mason bee the males are aggressive and battle to exclude other male bees from their small home territory.

In ancient India the territorial battles of carpenter bees were called "bee sport." It is said that the homing of such bees was exploited for the carrying of tiny messages. Much of the carpenter-bee activity noticed by homeowners today consists of males zooming after one another and even heading in the direction of passing birds and airplanes. The territorial males stage battles of bluff, zigging and zagging face to face until one of them gives up and leaves.

Some *Anthidium* mason bees are unusual among insects because the males are larger than the females. The reason is probably that the males butt and grapple with each other in their aerial battles; physical strength and weight would make it easier to retain a territory visited by females. These males attack other insects that wander into their territory. They can break the intruder's wings with their strong mandibles and puncture its body with their stout abdominal spines. Curiously, male bees that fight over territory during the day may cluster together peaceably in the evening as they sleep.

In 1802 the naturalist William Kirby



NEST-MAKING ACTIVITY of an alkali bee, *Nomia melanderi*, is followed over a period of 24 hours. At the outset (*1*) the female has finished one cell and has begun excavating another. When she reaches the appropriate depth, she packs soil along the edges (*2*). Then she smooths the inside of the cell with her pygidial plate (*3*) and applies a waterproof lining with a secretion from her Dufour's gland (*4*). Now she is ready to lay in provisions (*5, 6*), consisting of a honey-pollen mixture, which will sustain the larva. After grooming the ball of provisions (*7*) and removing debris from the cell (*8*) she does more polishing of the cell walls (*9*) and lays an egg (*10*). She caps the cell (*11*) and adds a plug of dirt that comes from the early stages of her excavation of what will be the third cell in the nest-building cycle (*12*).

wrote: "I have often thought that if gentlemen, who amuse themselves with chemical experiments, should direct their attention to insects, it might lead to the discovery of some powerful medicines. The variety of strong scents, which these little creatures emit, is wonderful.... Many Melittae have a strong scent, in some approaching to that of garlick or onion." An interesting if somewhat risky pastime is to carefully sniff captive live bees and wasps, which indeed produce a spectrum of odors that evidently serve in communication and defense. Small glands at the base of the mandibles release when the mandibles are opened various combinations of terpenes, aldehydes, ketones, esters, pyrazines and other pungent or fragrant compounds.

Some male bees apply such secretions to vegetation to delineate their territory. In other species the secretions serve as attractants for one or both sexes. For example, if a large number of bees such as species of *Colletes, Nomia* or *Anthophora* are netted at an aggregation of nests, they bite the net in an effort to escape, releasing mandibular-gland secretions as they do so. Other bees downwind are attracted toward the netful of captive bees. That it is the mandibular-gland secretions and not something else that is attracting the other bees has been demonstrated by the fact that the bees are also attracted by excised mandibular glands or by the appropriate synthetic compounds mixed in the right ratio. The small carpenter bees *Ceratina* and *Pithitis* produce pungent mandibular-gland secretions when they are annoyed. They smear them over their bodies and the intruder. These secretions are repellent to nest-invading ants.

Social behavior among the bees probably arose because the survival of the brood is enhanced if a guard bee is present while the forager is away. Another benefit is that two or more bees can share the work of nest construction and provisioning.

Such behavior could have arisen in two ways. The first is that the mother bee came to have a relatively long life span, so that she could help her daughters and be helped by them; this pattern is seen in some sweat bees and carpenter bees. The second way is that bees of the same generation came to share nests and cooperate in nest making and provisioning; this pattern is seen in a few species of sweat bees, digger bees, mason bees and orchid bees. Truly social behavior, with a division of labor in which some bees are queens, dominating the others and laying the eggs, and other bees are nonreproductive workers, is relatively rare among bees, being known only in some sweat bees and some small carpenter bees in addition to the more familiar social stingless bees, bumblebees and honeybees.

The Skill of Typing

How can skilled typists type as fast as they do? A century of study has not produced a definitive answer. A leading hypothesis is that a typist learning to type well learns to make mental processes overlap

by Timothy A. Salthouse

From almost the time the first typewriter came into use, the skill of typing has intrigued experimental psychologists, and it intrigues them even today, because the rate at which typists (even average ones) perform exceeds by far the rate that laboratory tests quite common in psychology would lead a psychologist to predict. Consider a typical secretarial task, the retyping of a document. In essence it is the transcription of a sequence of symbols (letters, numbers and so on) that come under the typist's gaze into a sequence of motor acts: the pressing of typewriter keys. Thus it closely resembles a series of choice-reaction-time tasks, in which a subject is presented with a single visual stimulus from a set of two or more stimuli after being instructed to rapidly press a particular button for each of the possible stimuli. Under optimal conditions (highly practiced people and a minimal number of stimulus-response alternatives) the average latency, or delay between the presentation of the stimulus and the pressing of a button, is approximately 250 milliseconds. The paradox of typing is that a latency of 250 milliseconds yields a typing rate of 48 words per minute (assuming five keystrokes per word). Yet speeds of twice that rate are fairly common.

How is it possible? What have skilled typists learned that enables them to overcome what appear to be fundamental limitations? Everyone has a minimum reaction time, which should set limits on the maximum typing rate, and yet skilled typists have developed a means of overcoming their perceptual and motor restrictions. Their achievement may have significance far beyond the skill of typing. After all, the detailed understanding of any skill, in the form of knowing how skilled people differ from less skilled people, should have implications for the selection of students and for their training in skills quite generally. Understanding the skill of typing may even have significance for rehabilitation therapy.

Typewriters came into use in the U.S. in the latter part of the 19th century, a period when experimental psychology was also becoming established. Perhaps that is why a number of early psychological studies were done on various aspects of typing. The research uncovered most of the major phenomena of typing: that it goes faster when keystrokes alternate between the hands, when successive keystrokes represent letter pairs with a high frequency of occurrence in the language, when the material being

TYPING SPEED exceeds the speed that psychological tests would lead one to predict. The number below each key on this typewriter keyboard is the average time required for that keystroke by 10 skilled typists the author studied. (Times are in milliseconds.) Most of the times are notably fast: less than 250 milliseconds (*colored numbers*). The same typists were asked to rest each index finger on a certain key and press it whenever the letter *l* or *r* (for "left" or "right") appeared on a screen. Response times were then 500 to 600 milliseconds.

typed is meaningful to the typist and so on. Thus the research provided some of the first scientific evidence pertaining to man-machine interaction. Indeed, the research probed what is sometimes considered one of the most maladaptive man-machine interfaces ever invented: the standard, or QWERTY, typewriter keyboard.

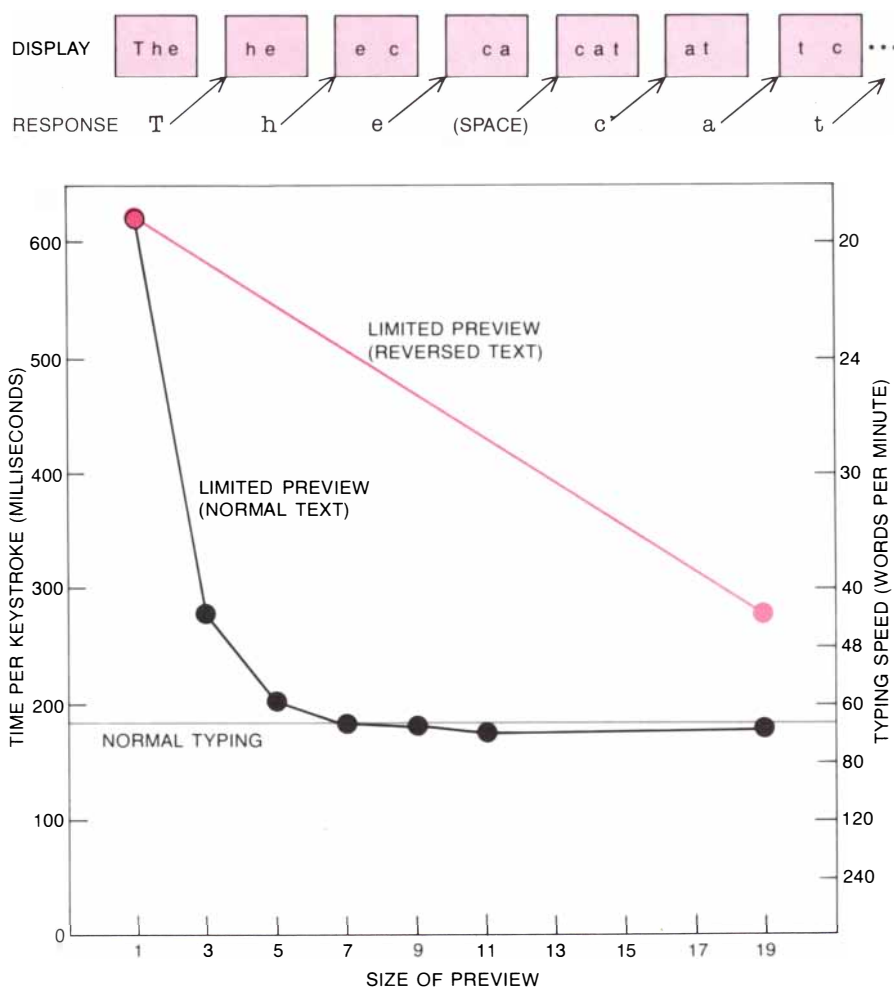
The QWERTY keyboard (named, of course, for the sequence of letters across its top row) had been devised in the 1870's without empirical research. It was a response to a defect of the earlier machines, namely that moderate speeds of typing tended to make keys jam together. QWERTY has been criticized for requiring disproportionate effort by the weakest fingers of each hand, and several alternatives have been proposed. Unfortunately QWERTY is now so thoroughly established that the benefits of a better arrangement are unlikely to offset the cost of retraining present-day typists and scrapping present-day typewriters and other keyboard machines. Clearly specialists in man-machine interaction should participate in the design of virtually all the equipment people use.

One of the earliest studies of typing was reported by William F. Book in 1908. Book had been much influenced by a classic study of telegraphers published 11 years earlier by William Bryan, an experimental psychologist, and Noble Harter, a former telegrapher who was doing graduate work in psychology. These earlier investigators had concluded from interviews, systematic observation and records of performance of telegraphers that a telegrapher in the process of becoming expert focuses attention on successively larger units, beginning with dots and dashes, the fundamental elements of telegraphy, and progressing to characters, syllables, words and even phrases. There is now considerable evidence that many skills are organized in this hierarchical fashion. Moreover, it is clear that skills such as reading would be almost impossibly difficult if the reader always analyzed written language in units of isolated letters. Book was therefore on firm theoretical ground in suggesting that skilled typing was achieved by expanding the size of verbal units from single characters to words and perhaps even phrases.

Fifteen years after Book's proposal the importance of large units of analysis to the skill of typing was empirically demonstrated by John E. Coover. He phrased his finding succinctly: "If copy is presented one letter at a time, so that as soon as the letter is typed another automatically appears, the expert's performance is reduced to a series of reaction times to the letters, and his rate is greatly reduced." This observation is readily verified with the aid of a computer. In the procedure I have implemented on a computer in my laboratory at the University of Missouri at Columbia typists see written material in a "preview display" that can be varied from one character to 39 characters. That is, the material shown to them on a video terminal can range from zero to 38 characters ahead of the character they are typing. They type at a standard keyboard, and the time (in milliseconds) between their successive keystrokes is automatically recorded and stored, along with the character typed.

Many typists have now been tested with this procedure. The results are uniform in that every subject tested typed very slowly when the preview was most limited and typed faster when the preview size increased. The preview size at which the typing rate first equaled its normal rate (its rate with unlimited preview and with a record of the typing that was visible to the typist) ranged from three to seven characters. (The 10 typists in the study varied in age from 20 to 40 and had gross typing rates, that is, rates uncorrected for typing errors, that varied from 50 to 75 words per minute.) If the typing rate is affected by the visual availability of characters up to seven spaces in advance of the character being typed, one must conclude that the typist attends to these characters. Thus there must be a gap or span between the characters receiving the attention of the eyes and the character whose key is being pressed. Similar estimates of such a gap have been reported in analyses of eye movements recorded in typists.

How does preview confer an advantage, so that the limitations of reaction time are circumvented by skilled typists? The reason is still in dispute. The possibility Book would favor is that typists developing their skill move from a mode in which they analyze character by character to one involving larger units such as words and phrases. Thus



LIMITED PREVIEW diminished the typists' speed. Each typist faced a screen on which a text always remained a certain number of characters ahead of the character the typist was typing. (In the example at the top the text is "The cat climbed . . . " and the preview size is three.) The typists' speed approached normal speed only when the preview size approached seven. A similar effect was seen even when the text was randomized by reversing each word letter by letter.

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.

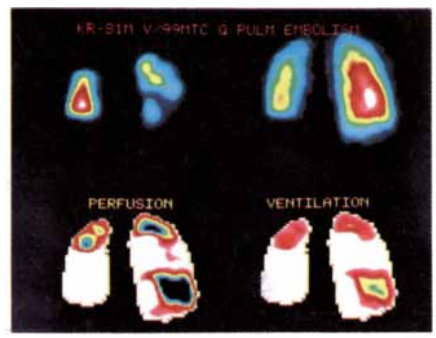
Comprehensive. Leading authorities from Harvard and Stanford cover the full range of medicine for you the seasoned clinician.

Complimentary CME. At no extra cost, we offer a CME program in which you can earn 32 top credits* per year.

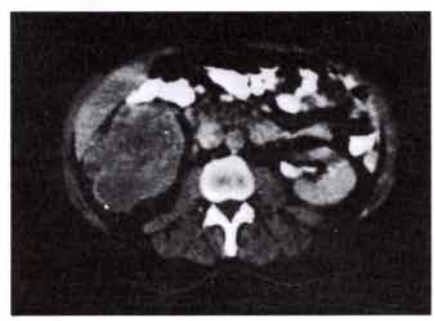
Trial Offer. We invite you to try SCIENTIFIC AMERICAN *Medicine* for two months with our compliments. Send us the coupon and you will receive the two-volume text and two monthly updates. You may also take a CME test for credit. At the end of 60 days, if you decide to continue the subscription, you will owe us $235 for the first 12 months (current renewal is $170); otherwise, return the books; there is no charge.

Please mail the coupon today so that we can put medical progress to work for you.

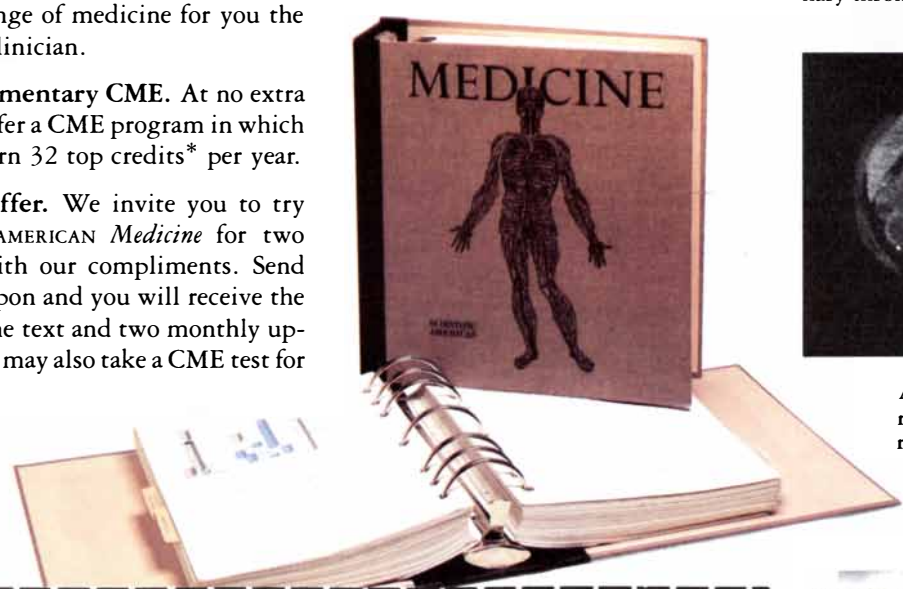
You may also order your trial copy by calling toll-free 1-800-345-8112 (in Pennsylvania, 1-800-662-2444).

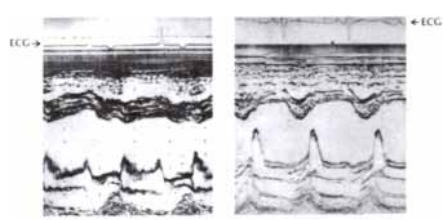


Computerized scintigraphy reveals pulmonary thromboembolism.



Abdominal computed tomogram reveals large renal carcinoma replacing part of right kidney.





Echocardiograms from patients with aortic regurgitation.

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 $235 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 $235*
☐ Bill me ☐ VISA ☐ MasterCard

Expiration Date_____ Account Number_____

*Please add sales tax for California, Illinois, Massachusetts, Michigan and New York

Name_____ MD Specialty_____

Address_____

City_____ State_____ Zip_____

Signature_____

Please allow six to eight weeks for delivery. All payments must be in U.S. dollars. CME is available outside the U.S. and its possessions for an additional $65. Surface routes used unless airmail delivery is prepaid.

7H

*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 has been reviewed and is acceptable for 32 prescribed hours by the American Academy of Family Physicians.

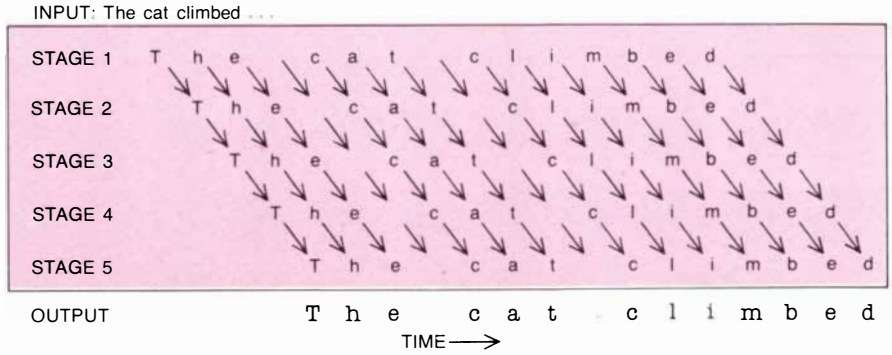
This program has been approved by the American College of Emergency Physicians for 32 hours of ACEP Category 1 credit.

the skilled typist is assumed to impose groupings on both the input (the text to be typed) and the output (the sequence of keystrokes). The input groupings amount to perceptual "chunks." Probably they consist of syllables or common words, evidently having three to seven characters. The output groupings are similar to what contemporary theorists call motor programs and consist of integral sequences of essentially automatic keystrokes. In effect the typing of a given output grouping is taken to be ballistic: it requires only a start signal, after which nothing affects the execution of the movement. According to this hypothesis, maximum typing speed requires a preview of three to seven characters because a smaller preview cripples the ability to group.

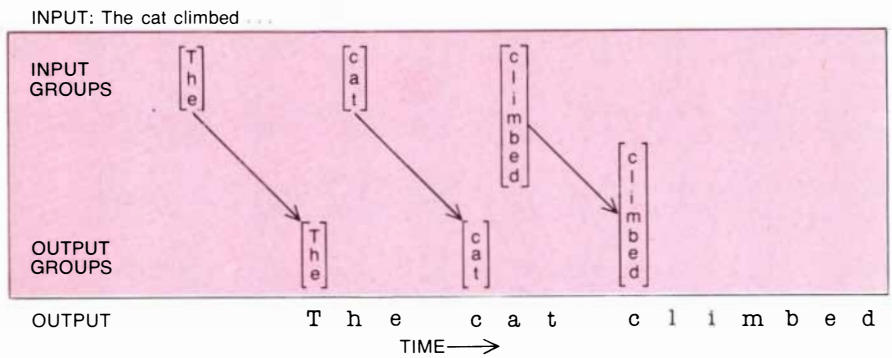
Shortly after Book made his suggestion a number of people proposed the competing idea that typing speeds circumventing the choice reaction time are accomplished by an overlapping of "processing operations." August Dvorak, a strong advocate of this position, was also a critic of QWERTY and the inventor of an alternative keyboard. Dvorak, like Book, asked why typing strokes are faster than isolated choice-reaction-time responses. "The answer," he wrote in 1936, "is that even with a two-letter sequence, while the first finger is stroking, the second finger is starting its play for position and overlapping its stroke with the first."

The processing operations can be given modern labels remarkably similar to the ones incorporated in recent "stage theories" of choice-reaction-time tasks. They might begin with encoding: the perception of each character. Then would come the categorization of the stimulus (say into the class of the left- or right-handed keystrokes), the fixed decision as to which finger must type and finally the preparation and execution of the actual typing movements. On this hypothesis maximum typing speed requires a certain preview because the visual availability of text characters allows the sequence of processing operations to be executed in parallel for a number of characters, the late stages of one sequence accompanying the earlier stages of others instead of the strict succession dictated by the availability of only one character at a time.

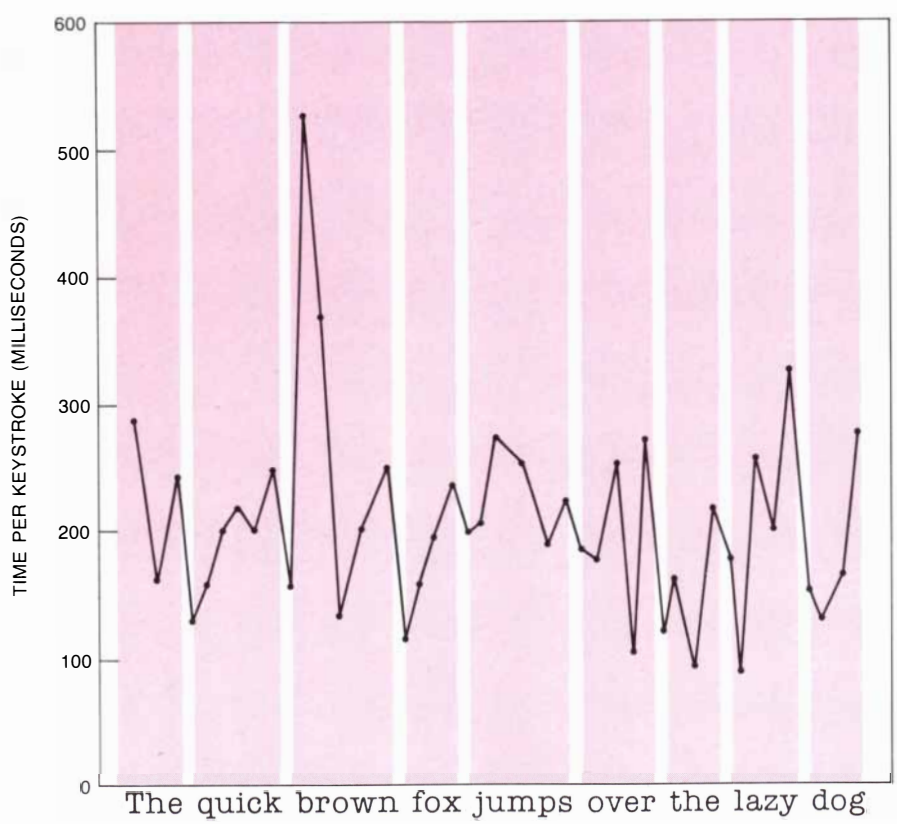
Some support for the "chunking hypothesis" (the one that Book would favor) is available from an analysis of the time intervals between the keystrokes made by a typist who repeatedly types the same text with unlimited preview. Under these circumstances a typist is generally quite consistent in the temporal pattern of keystrokes composing a given word, as though the individual keystrokes were under the control of a single motor program. The consistency cannot simply be attributed to the differ-



CHUNKING HYPOTHESIS is an attempt to account for both the preview effect and the remarkable speed of skilled typists. It proposes that the material to be typed gets mentally "chunked" into multicharacter units (here words) and that the typing of each unit is governed by an autonomous "motor program." If the preview is too limited, chunking is counteracted.



OVERLAPPING OPERATIONS is a competing hypothesis. It proposes that typing of each character proceeds in steps (here five) such as perception of a character, its assignment to a hand and so on, and that the steps overlap. If preview is too limited, overlapping is counteracted.



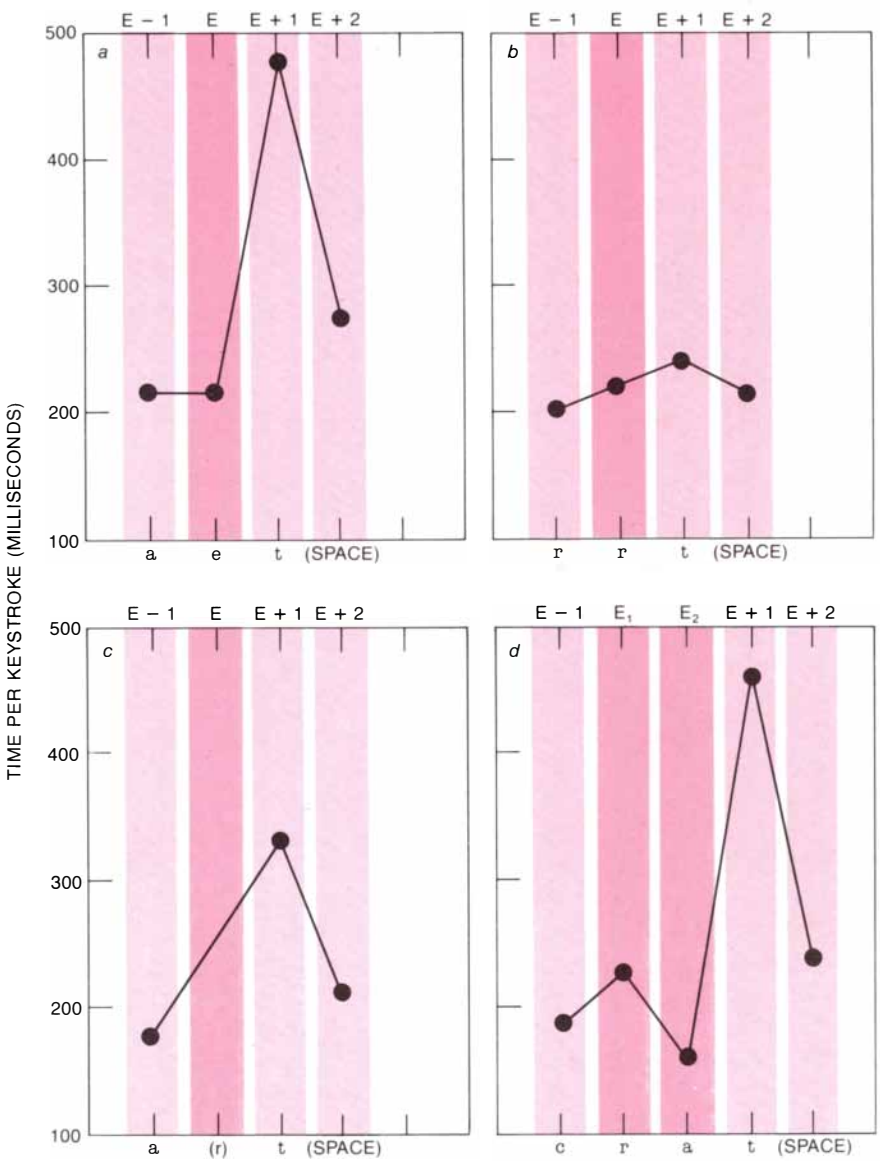
TEMPORAL PATTERN OF KEYSTROKES over 10 typings of "The quick brown fox..." by one typist seems at first to favor the chunking hypothesis. A given letter (such as *o*) is typed at different speeds in different words, as if it were under the control of different motor programs.

ing accessibility of various keys or to variations in the strength or dexterity of individual fingers, because the same character turns out to be typed with different latencies in different contexts. An example is the letter *o* in the familiar sentence "The quick brown fox jumps over the lazy dog." The typists I studied typed it with an average latency of 370 milliseconds in the word *brown,* 160 milliseconds in *fox,* 185 milliseconds in *over* and 130 milliseconds in *dog.*

Further results, however, suggest important qualifications of the chunking hypothesis. If the hypothesis is interpreted literally and the characters of a word are organized into a strictly ballistic motor program, one might predict that errors made at any point in its execution would go undetected until at least the end of the program. The prediction can be examined by an analysis of the errors committed in normal typing. It has been reported (in fact, the first reports came soon after Book's pioneering work) that the keystrokes accompanying the first intimation on the part of the typists that an error has been made are partially inhibited, in that they are weaker or delayed with respect to normal typing. These phenomena might serve, then, as a marker of the instants when errors are detected. The markers do have a drawback: it is difficult to predict a typist's groupings. A preferable strategy is to assume that groupings do not exist. Errors could then be detected immediately. A failure to find evidence of immediate error detection would therefore tend to support the notion that typing is controlled in multicharacter groups.

I determined the median latencies for the keystrokes before and after each of four types of errors: substitution errors (typing *caet* instead of *cart,* for example), intrusion errors (*carrt* instead of *cart*), omission errors (*cat* instead of *cart*) and transposition errors (*crat* instead of *cart*). The interpretation of the omission results is difficult: the latency of the omitted keystroke might be incorporated into that of the next keystroke. In addition some errors are not detected unless the typist proofreads the typed copy. No error-detection phenomena could reasonably be expected from undetected errors. Nevertheless, my results suggest that a large proportion of errors are detected immediately, and not at the end of multicharacter groups. This is not to say output groupings cannot exist. It does, however, indicate that the control of typing is not completely relinquished to multicharacter motor programs.

Additional evidence for the conclusion that keystrokes are controlled in units smaller than words or syllables has recently been provided by Gordon D. Logan of the University of Toronto. Logan employed a technique in which typists were instructed to stop typing whenever they heard a tone. He reasoned that if word-length motor programs controlled typists' keystrokes, the typists would type to the end of a word regardless of both the "position" of the stop signal and the number of letters remaining to be typed. Logan found that regardless of word length the typists hearing the tone actually stopped after an average of only one or two letters.



TYPING ERRORS weaken the chunking hypothesis. Four errors are shown: substitution (*a*), such as *caet* instead of *cart;* intrusion (*b*), such as *carrt* instead of *cart;* omission (*c*), such as *cat* instead of *cart,* and transposition (*d*), such as *crat* instead of *cart.* In each case typists slowed one keystroke after the error (*dark color*) was made. The slowing is taken to signal awareness that something is wrong. The chunking hypothesis implies errors cannot be detected that soon.

A third piece of evidence that weakens the chunking hypothesis emerges from a comparison of the typing of normal and random texts. The chunking hypothesis leads to two predictions about the typing of random texts. First, the process of chunking calls for the matching of grouped input material to preestablished motor programs, and so the process should be disrupted by material that does not have the familiarity of normal English. Hence the rate of typing should be greatly reduced. Second, a random text is not amenable to chunking even with unlimited preview. Hence the rate of typing should not be affected by reducing the size of the preview.

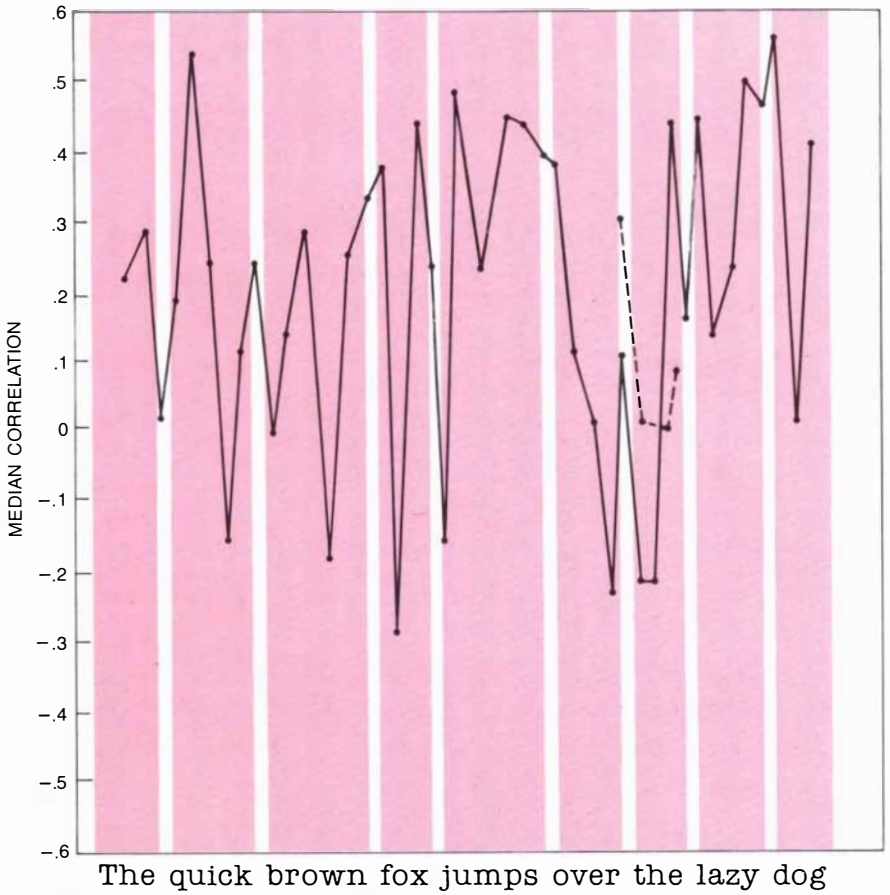
These predictions were examined by having typists type sentences where each word had been reversed letter by letter. The resulting text is not genuinely random, but it does have the advantage of disrupting the familiar letter sequences and word organizations of English while preserving the letter frequencies and the pattern of interword spacing. Preview sizes of 19 characters and one character

were presented, and typists performed under each condition twice. The first prediction proved to be true: typing speed was impaired. The second prediction proved to be mistaken: the viewing of 19 characters resulted in markedly faster typing than the viewing of one character at a time. Thus the preview effect is pronounced. Indeed, increasing the number of visible characters provided nearly as much relative advantage with the meaningless material as it would with normal text. It seems to follow that the benefits of presenting more than one character at a time are not simply attributable to the input chunking and output chunking of meaningful patterns.

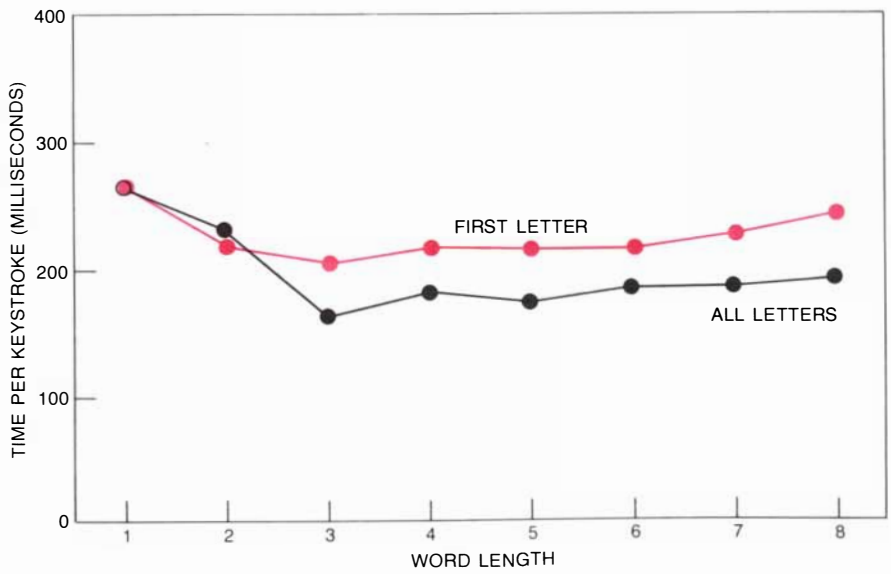
The chunking hypothesis is weakened still further by the finding that the latencies of successive keystrokes have a relatively low correlation. That is, the time it takes a typist to make a given keystroke proves to have little relation to the time it takes the typist to make the next keystroke. If the successive keystrokes were within an output grouping, one might expect to find such a relation. If the first keystroke takes longer, the second one should too. After all, they would both be under the control of a single higher-order process.

Finally, the chunking hypothesis is weakened by the finding that there is no relation between the length of a word and either the latency of the first keystroke in the word or the median interval between successive keystrokes in the word. The reason a word-length effect might have been expected is that the time needed for a motor program to be "retrieved" and "prepared for execution" (to borrow the terminology of computer science) should be related to the number of units (separate keystrokes) governed by the program.

No single item among these lines of evidence is conclusive in itself. Taken together, however, the findings strongly imply that typing speeds circumventing choice-reaction-time rates are unlikely to be accountable solely to a switch from single-character to syllable- or word-length units of analysis. Why, then, does the chunking hypothesis seem so compelling? One explanation might be the strong subjective impression among most skilled typists that they need only think of a word and it seems to be typed automatically. This impression may be due more to the typist's input than to any output processes. In particular, the lifelong habits of reading probably include the perception of written language in multicharacter input units. Thus a manuscript in the process of being typed may get perceptually chunked by syllable or word. After that the preparation and execution of the discrete motor acts that constitute typing



LACK OF SYSTEMATIC CORRELATION between the speeds of keystrokes made in succession by people typing "The quick brown fox..." also weakens the chunking hypothesis. Each dot shows the correlation between the time required for a given keystroke and the time required for the preceding keystroke. The scale ranges from 1 through 0 to −1. A correlation of 1 means that an increase in one of the variables is always accompanied by a proportionate increase in the other. A correlation of −1 means that an increase in one variable is always accompanied by a proportionate decrease in the other. A correlation of 0 means that the variables have no relation. The broad scatter of points suggests that the characters in each word cannot be part of a higher-order chunk. Broken line in the chart shows correlations among keystrokes obtained from the typing of *the* in nonrepetitious material. Again no relation emerged.



LACK OF WORD-LENGTH EFFECT weakens the chunking hypothesis still further. The hypothesis predicts that motor programs for long words should take longer to "retrieve" and "execute." Yet the median time per keystroke needed to type a word (*black curve*) and the time needed to type the first letter in it (*colored curve*) are independent of the length of the word.

If you want reliable results . . . demand reliable software.

BMDP™ Statistical Software Provides "Data Analysis with Confidence"

Professional scientists require the finest instruments and resources to aid in their quest for knowledge, and for over 20 years researchers have called upon BMDP to produce reliable results. As the acknowledged leader in statistical methodology, BMDP has become the standard by which others measure their performance. Dependable software is not developed overnight, it takes time and talent to produce statistical programs of superior quality — BMDP has made that investment in producing the finest statistical software available today.

The 40 programs provided in BMDP feature:

Data Description
Plots and Histograms
Frequency Tables
Missing Values
Regression
Nonlinear Regression
Nonparametric Analysis
Cluster Analysis
Multivariate Analysis
Survival Analysis
Time Series
Analysis of Variance and Covariance

The flexible, English-based control language found in BMDP makes it easy to use. The analyses you select can be as simple or as complex as your data demands.

At BMDP we understand the need for comprehensive documentation and training. We have produced a library of support literature to help our users make effective use of our programs. The BMDP manual provides complete descriptions of input and output options,their uses, and the underlying statistical algorithms on which the calculations are based. In addition, we provide annotated analyses to lead the user through complete applications of complex techniques. Throughout the U.S. and Europe, BMDP also presents two day short courses on the use, application, and interpretation of the BMDP programs.

BMDP operates on most mainframe computers, selected minicomputers, and on our own desktop computer — the StatCat™. Our IBM-PC version is in progress and will be available by mid 1984.

The finest software at an affordable price, BMDP can be obtained under annual license for $2,900/year (discounts available for academic and government organizations). This price includes maintenance, updates, enhancements and training. Since our rates for microcomputer installation are based on the number of programs selected, the cost can be considerably less. Our desktop computer can be ordered to suit the specific computing requirements of your organization. StatCat features a Motorola 68000 UNIX™ based machine with prices beginning at $12,485.

For more information on BMDP Statistical Software and all the services and products we provide, call or write:

BMDP Program Librarian
1964 Westwood Boulevard, Suite 202
Los Angeles, California 90025
(213) 475-5700
Telex: 4992203

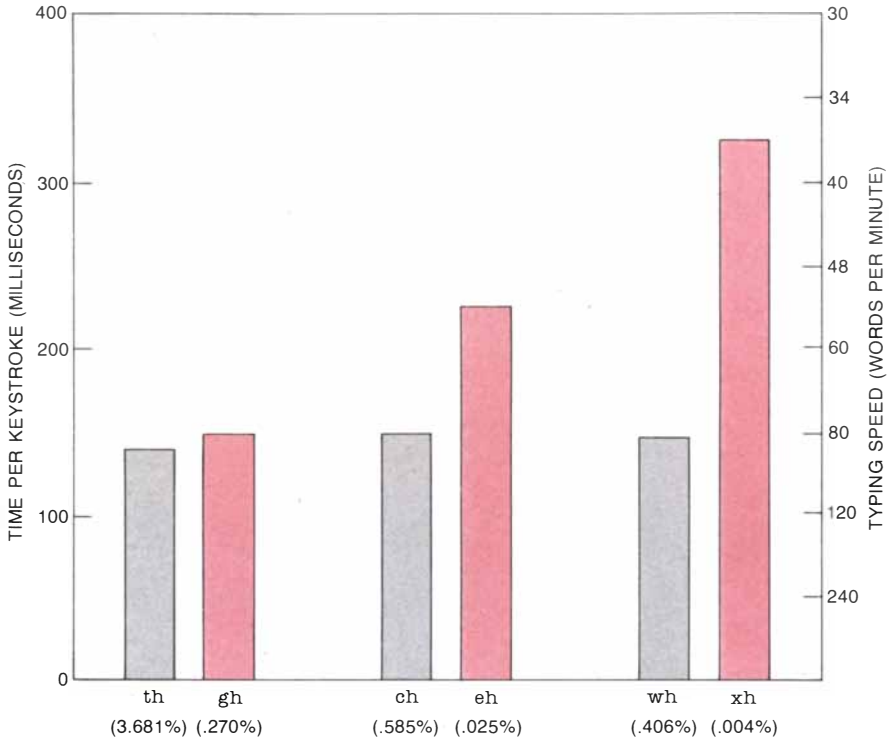
* BMDP is a trademark of BMDP Statistical Software, Inc.
* Statcat is a trademark of BMDP Statistical Software, Inc.
* UNIX is a trademark of Bell Laboratories

may proceed in ways utterly inaccessible to conscious awareness.

How can the evidence that seems to favor the chunking hypothesis (for example the consistency in the temporal patterns of keystrokes making up the words in "The quick brown fox...") be explained if not by a motor program? Specifically, how can the evidence be explained by the competing hypothesis that skilled typists make their processing operations overlap? One possibility is that the time required for a keystroke depends on factors such as the position of the preceding keys and subsequent keys and the direction in which specific fingers must move. The importance of these factors is documented by the finding that the time needed for a keystroke is less when the preceding keystroke was made by a finger of the opposite hand. My own studies of normal typing show that opposite-hand sequences of two keystrokes are executed an average of 40 milliseconds faster than same-hand sequences. Similar effects have been reported by earlier workers investigating sequences involving the same finger or different fingers (the latter are faster) and the same row or different rows of the keyboard. (The results are a bit complicated; for example, the index finger moves inward on a row faster than it moves to a different row.)

One might also expect that the relative frequency with which particular sequences of letters have been typed in the past is important, so that commonly typed sequences are executed faster than less-practiced ones. This possibility can be examined by calculating the correlation between two variables: on the one hand the frequency of occurrence of various two-letter sequences in English and on the other the latency between the two keystrokes in each sequence as the sequences arise in the course of normal typing. The range of occurrences is great: from zero percent for sequences such as *qz* to 3.68 percent for the sequence *th*. Therefore I worked with the logarithms of the frequencies of occurrence. The correlation turned out to be low: only some 4 percent of the variability in the latency between keystrokes could be attributed to the variability in the logarithm of the frequency of occurrence of two-letter combinations. Nevertheless, the average latency decreased by about 52 milliseconds for each log-unit increase in frequency. Thus a frequency effect emerges. I have not done a similar analysis for longer sequences, but it is reasonable to guess that three-letter combinations might also show a frequency effect. It has been estimated that only 104 three-letter sequences (out of 26^3, or 17,576) account for nearly half of the occurrences of three-letter sequences in English.

So far I have emphasized the similarities among typists. I should also mention the differences. One of these is the extent to which typists comprehend what they type. In the course of my studies some typists reported they had virtually no idea of the content of what they typed. Others said they could comprehend the material at least as well as if they had merely read it. Their subjective reports are supported by their scores on posttyping comprehension tests, which ranged from 12.5 percent correct responses to 75 percent.



TYPING OF TWO-LETTER COMBINATIONS suggests one reason typing speeds for a given letter tend to vary from word to word. In each case the letter *h*, typed by the right index finger, was preceded by a letter typed with the opposite hand. In *th* and *gh* (*left*) the *t* and the *g* are typed by the left index finger; *th* is commoner in English. (Frequencies of occurrence are in parentheses.) In *ch* and *eh* (*middle*) the *c* and the *e* are typed by the left middle finger; *ch* is the commoner. In *wh* and *xh* (*right*) *w* and *x* are typed by the left ring finger; *wh* is the commoner. The time required to type *h* is greater for the *h* in each of the less common combinations.

Another way in which typists differed was in the frequency with which they made errors of various types. The least frequent type of error for all typists was transposition, but the most frequent type was either intrusion, omission or substitution. The duration of the keystrokes associated with errors also varied. Some typists made extremely fast keystrokes by pressing two fingers almost simultaneously. The first keystroke would come with a normal latency, the second would come perhaps 10 milliseconds later. The error (often an intrusion error) might be either one of the strokes.

All things considered, it appears that the limits of reaction time are circumvented primarily because skilled typists have learned to overlap the performance of many of the operations involved in making keystrokes. This overlapping is impossible when the material to be typed is displayed only a few characters at a time. What does this conclusion imply for the understanding of skill? First, the results from the study of typing are consistent with results from the study of other perceptual-motor activities in suggesting that a cognitive component is present in nearly all skilled behavior. Attention to characters ahead of the one being typed may be due to the development of multicharacter units of analysis and execution or it may be due to the overlapping of processing operations. Either way, the mechanisms are highly rational adaptations to the limits of processing character by character.

Second, the processes mediating between the input and the output in typing are exquisitely coordinated: the skilled typist does not often hesitate. The intensive practice necessary to achieve high levels of skill appears, then, to result in the elimination of unnecessary operations, in the ability to execute more than one operation at a time and perhaps in a reduction in the attention demanded of the typist by certain operations. These characteristics may be common to a wide variety of activities and so may serve as some of the goals toward which any kind of training should be oriented.

THE AMATEUR SCIENTIST

How to stop a spinning object by humming and perceive curious blue arcs around a light

by Jearl Walker

The next time you are watching a spinning propeller, rotating fan blades or anything else turning rapidly, start humming. By changing your pitch until you reach a certain frequency you can make the rotation (as you perceive it) seem to slow down, stop or move slowly in the opposite direction. Another visual illusion you can test yourself is the appearance of blue arcs when you look under certain conditions at a small light. The humming experiment can be explained but the blue-arc illusion remains something of a puzzle.

It is usually possible to perceive any regular fluctuation if its frequency does not exceed about 40 hertz (cycles per second). At higher frequencies the fluctuations blend together. Imagine a pattern of radial black and white sectors mounted on the turntable of a phonograph and rotating at $33\frac{1}{3}$ revolutions per minute. If the pattern moves across your field of view slowly enough because the rotation is slow or the sectors are wide, the sectors remain visible. At higher speeds the sectors blend into a uniform gray.

In 1967 W. A. H. Rushton of the University of Cambridge reported that an observer can stroboscopically "freeze" such a rotating pattern even when the motion is fast enough to fuse the sectors. All the observer has to do is to hum at the appropriate frequency. At a slightly lower frequency the strobe pattern moves slowly in the same direction as the true rotation. At a slightly higher frequency it reverses direction. The finding was made independently in 1974 by Marlin S. Werner of the Speech and Hearing Center in Oakland, Calif. The cause of the stroboscopic effect was investigated further in 1979 by H. A. K. Mastebroek and J. B. van der Kooi of the Laboratory for General Physics at Groningen in the Netherlands.

The radial pattern employed by Rushton is on a disk often supplied with home audio equipment for checking the speed of the turntable. The check is made by illuminating the disk with a household lamp energized by alternating current. In the U.S. such current reverses direction at a frequency of 60 hertz, causing the lamp to fluctuate in brightness at a rate of 120 hertz. The black and white sectors on the disk are spaced in such a way that when the turntable rotates at the right speed, the fluctuations in the light from the lamp stroboscopically freeze the motion of the sectors.

Suppose that when the lamp is brightest, a white sector is at a certain position in your field of view. The lamp dims and then brightens again, too fast for you to perceive it. By then the motion of the disk brings the next white sector into the same position. Each time the lamp brightens, another white sector is at that position. This matching of the rotation rate of the disk and the fluctuation in the lamp's brightness creates the illusion that a single white sector remains fixed in position.

If the turntable is turning a little too fast, the pattern moves slowly in the direction of the true rotation. On one brightening you find the white sectors in a certain orientation. By the next brightening the turntable has moved them beyond that orientation. Each new brightening reveals the same advance. You misinterpret the motion as a slow rotation of the pattern in the direction of the turntable's movement. On the other hand, if the turntable is rotating a little too slowly, the successive illuminations of the pattern generate the illusion that it is turning slowly backward. In each case the speed of the migration indicates how much the turntable is off speed.

Rushton experimented with a strobe disk intended for use in Britain to check $33\frac{1}{3}$-r.p.m. turntables. Alternating current reverses at a frequency of 50 hertz in Britain, causing the lamp to fluctuate in brightness at a rate of 100 hertz. The corresponding time between bright phases of illumination is .01 second. The strobe disk was laid out with its radial sectors spaced so that the rotation of one white sector into the place previously occupied by another white sector in an observer's field of view took .01 second. When Rushton viewed the disk on a turntable running at the right speed, the pattern seemed to freeze.

As one might expect, stroboscopic effects disappear in steady illumination such as sunlight. Rushton discovered, however, that even in steady light he could freeze the pattern by humming at a frequency of about 100 hertz. If he hummed slightly flat, the pattern migrated slowly in the true direction of rotation. If he hummed slightly sharp, the pattern migrated slowly in the opposite direction. Other sources of sound in the room had no effect.

How does the vibration of humming give rise to the stroboscopic effect? Does it perhaps shake the optic nerve in such a way that the transmission of signals to the brain is periodically diminished or delayed? Rushton devised an experiment to check this possibility. His idea was to force a delay in the transmission of an image to his brain.

For this purpose he resorted to a dark filter. When light falls on the retina, the information is sent immediately to the brain. If the light then gets dimmer, the information is delayed by as much as several milliseconds. Rushton's thought was that if the vibration from his humming was jarring nerve fibers or something else in the pathway leading from the retina to the brain, then dimming his view of the rotating pattern would make a steady stroboscopic pattern migrate because of the imposed delay. He observed no migration, thereby ascertaining that the humming does not act on the pathway leading from the retina. Instead it apparently acts on the eye as a whole. The facial bones conduct the vibrations of humming from the mouth and throat up to the eyes.

Rushton's explanation was the subject of a paper by Mastebroek and van der Kooi. The illustration on page 138 is pertinent to their analysis. The black-and-white pattern on a rotating turntable sweeps across the observer's field of view. In the illustration the motion is shown in relation to a fixed axis in a room. The eye, which is made to oscillate vertically by humming, is in its uppermost position in part *a*. The frequency with which white sectors are replaced in the field of view is matched by the humming frequency and thus by the vibration of the eye.

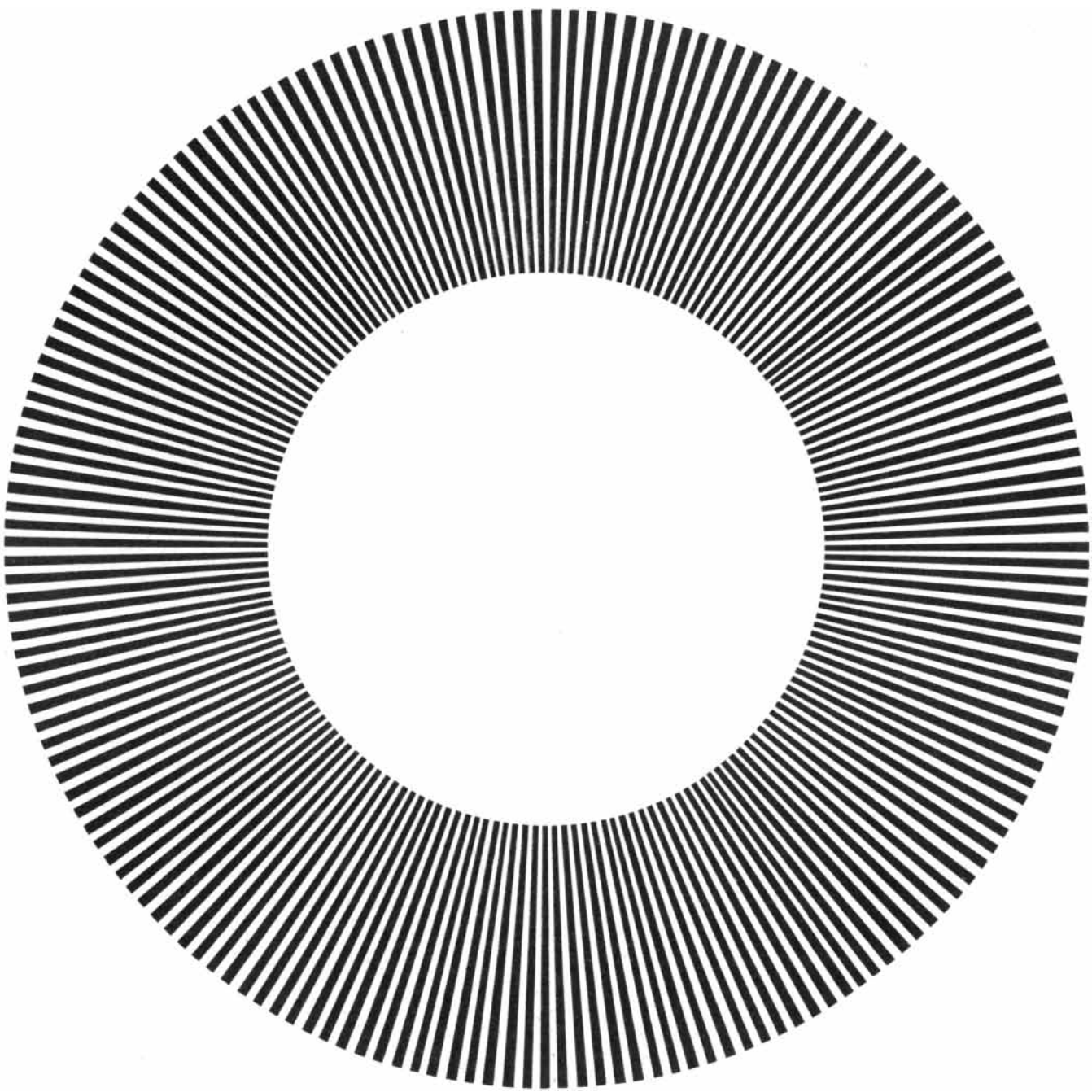
In part *a* of the illustration a white sector lies on the optic axis that runs from the center of the field of view back to the retina. The sector is the central image on the retina. In parts *b* and *c* the disk pattern and the eye move downward in phase, maintaining the white sector as the central image on the retina. In parts *d*, *e* and *f* the disk pattern continues to move downward as vibration

moves the eye upward. The white sector that was previously central on the retina shifts upward. Meanwhile a dark sector is imaged briefly on the center of the retina. In part *f* the cycle begins again.

For most of a single vibration of the eye the center of the retina receives an image of a white sector. Other points spaced periodically along the retina are also mostly illuminated by white sectors. At all such points neither the white images nor the dark ones are perceived individually because of the rapidity of the vibration. Instead the visual system averages the brightness over a full vibration, giving the white-image points a bright average. The points on the retina that receive mostly images of the dark sectors yield a darker average. If the observer is humming at the right frequency, this pattern of averages is reinforced with every oscillation of the eye. Although the pattern is less distinct than the one on the disk (when it is stationary), it is still observable.

Many similar stroboscopic effects induced by humming can be observed. Werner has described how humming can make the image on a television screen unstable. He stood about 30 meters from a television set and hummed through his teeth at a frequency of 60 hertz. Philip C. Williams and Theodore P. Williams of Tallahassee, Fla., had previously reported how humming can alter a television image. In their observations the humming created on the screen an illusion of horizontal dark lines. By controlling the frequency of the hum the hummer could hold a line stationary on the screen or move it up or down. Certain higher frequencies (apparently harmonics of the lowest frequency) yielded multiple lines. On both black-and-white sets and color sets the lines were gray.

Werner also described how humming can stroboscopically freeze spokes on rotating automobile hubcaps and lug bolts on rotating truck wheels. An observer driving past a picket fence can



A pattern for testing the speed of a turntable

freeze the pickets by humming. I have heard of people who can ascertain the frequency of a fluorescent lamp by humming.

The most curious report of a stroboscopic effect was published by J. L. Scott-Scott of Nuneaton, Warwickshire, England. He hit on a way of determining the direction of rotation on an airplane propeller. He made his head vibrate by allowing his tongue to flutter gently against the roof of his mouth as he exhaled. He began with a high frequency of vibration and gradually brought the frequency down until the propellers were stroboscopically frozen. He then decreased the frequency a little more. The slow migration of the strobe image of the propellers revealed their true sense of rotation.

I made a strobe disk by taping a large sheet of white paper onto one side of a long-playing record. I put a reference point on the paper at the center of the hole in the record. Then I marked off radial sectors one degree in width. (All the measurements must be completed before the record is put on the turntable because the spindle breaks through the paper and obliterates the reference point.) I filled every other sector with black ink.

When I rotate the disk at $33\frac{1}{3}$ r.p.m., the turntable is running at a frequency of .555 hertz. Since I had arranged 180 pairs of black-and-white sectors around the disk, the frequency at which a certain arrangement of the pattern is repeated is .555 hertz multiplied by 180 pairs, or about 100 hertz.

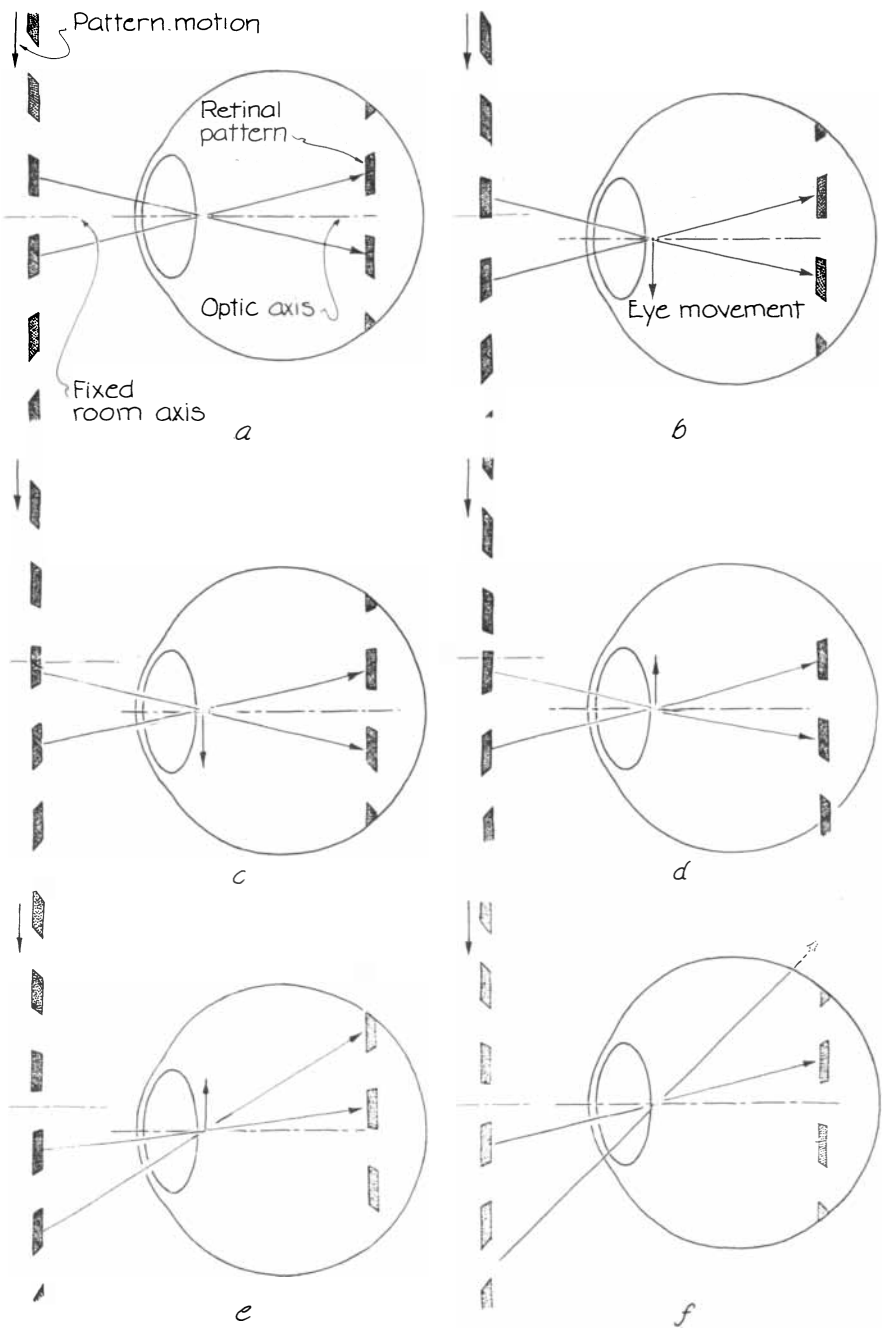
It was tedious to make the disk, and the result was marred by variations in the blackness of the dark sectors. When I rotated the record at $33\frac{1}{3}$ r.p.m., I was easily distracted by the variations in the ink. Other features of the rotation also made me follow the movement of the record instead of fixing my gaze on one area. With experience I learned to ignore most of the distractions and to concentrate on the gray blur.

When I illuminated the rotating pattern with an alternating-current house lamp, I saw only the blur. The light flickered at a rate of 120 hertz, higher than the frequency of 100 hertz needed to freeze the pattern. With the turntable in sunlight I tried to hum at approximately 100 hertz in order to freeze the pattern. Although I could sweep my humming frequency through the proper range, I failed to freeze the pattern, largely because I cannot hold a note of one pitch.

I then turned to an easier method of making my head vibrate. I attached a speaker to an audio amplifier driven by a signal generator. Setting the amplification high, I stuck my chin inside the speaker cone. My head vibrated. With one hand I slowly swept the frequency of the oscillator through a region around 100 hertz as I stared at a section of the rotating pattern. The section I chose was on the right side of the record as I faced it. There the pattern descended in my field of view.

At a frequency of about 100 hertz a faint, stationary strobe pattern was perceptible against the gray blur. As I moved to slightly higher or lower frequencies the strobe pattern migrated in the ways described by other people. Even without looking at the frequency control I could consistently freeze the strobe pattern at the same frequency. When I took my chin out of the speaker cone or moved the frequency well off 100 hertz, the strobe pattern disappeared immediately.

You might like to make your own strobe pattern and continue the experiment. You could also experiment with other vibrating or rotating objects. Can you freeze their motions with an appropriate hum?



How movements of the eye caused by humming can "freeze" a moving pattern

Early in the 19th century Johannes Purkinje, one of the pioneers in physiological optics, discovered that the light from a small piece of glowing tinder created two blue arcs across his field of view. Although the arcs soon faded, he could control their brightness and duration by moving the tinder up and down, sweeping its light across different

RENAULT ALLIANCE WATCHWORD: TECHNOLOGY

If you want to invest in European technology but haven't found a car priced within reach—look at Renault Alliance. Independent suspension with unique twin coaxial torsion bars for a quiet, smooth ride. Sedan comfort for five, with Renault's exclusive pedestal seats for extra rear seat leg room. Power front disc brakes. Electronic fuel injection for an outstanding **52** EST HWY, **38** EPA EST MPG.* Renault Alliance. *Built in America* and priced to bring affordability to European Technology: **$5959.**** It's the one to watch.

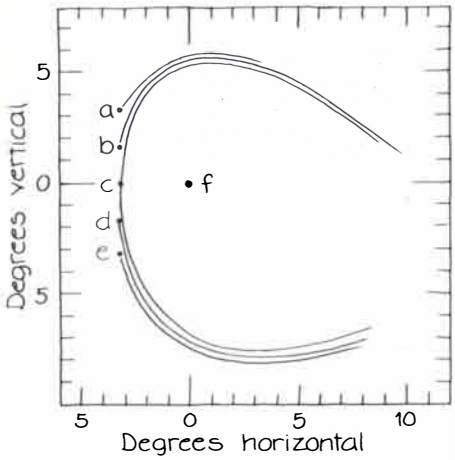
Alliance 2-door

* Mileage may vary with speed, trip length, weather. Actual highway and CA figures lower. ** Manufacturer's suggested retail price for 2-door. Tax, license, destination charges, optional or regional equipment extra.

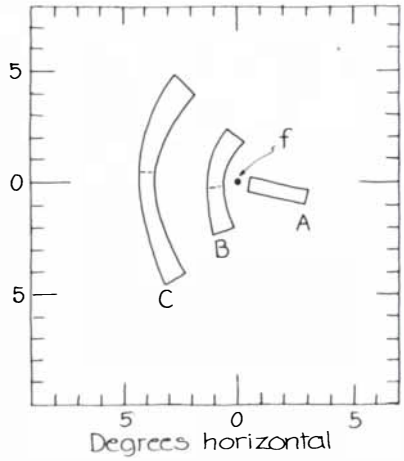
Built by American Motors

RENAULT

THE ONE TO WATCH



Arc locations for the right eye

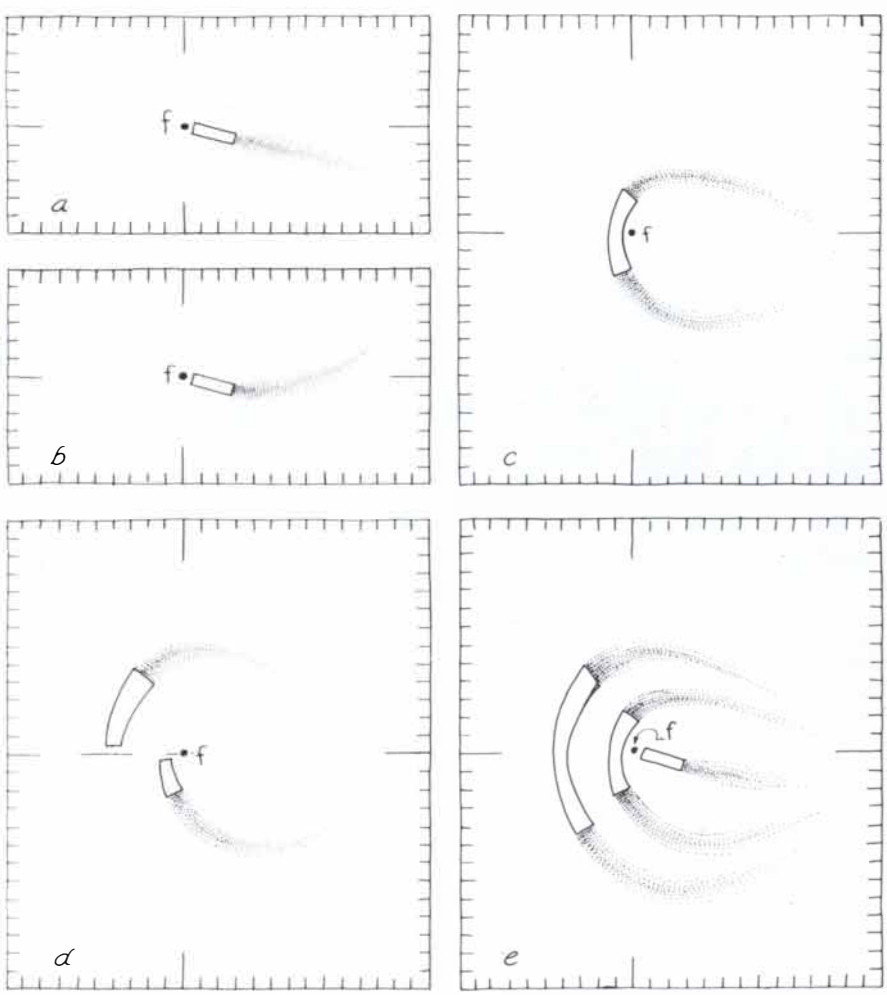


J. D. Moreland's stimulus locations

parts of his retina. If the arcs faded, he could regenerate them by readapting his eyes to darkness for a few minutes.

The cause of the blue arcs, and particularly the reason for their being blue, has been a physiological puzzle ever since. Many investigators have studied the relation between the place on the retina that is illuminated and the shape of the arc that is generated. Recent publications by J. D. Moreland, formerly at the Institute of Ophthalmology in London, provided me with detailed instructions on how to repeat the observations.

Moreland suggests the following procedures for viewing Purkinje's arcs. Al-



Varieties of arc

low one eye to adapt to darkness for at least one minute but not more than three minutes. Switch on a small light. (A yellow or red light in the range of wavelengths between 510 and 620 nanometers is best.) Depending on the location of the light in your field of view, you should see one or more arcs for about a second. For repeated displays turn on the room lights for two or three minutes every five minutes. During the dark period turn on the yellow or red light for half a second, repeating the stimulus every four seconds.

If bright, clear arcs are to be seen, both the orientation and the size of the stimulus light are important. A small rectangular source works well because it can be positioned to excite several overlapping arcs whose brightness can be summed by the visual system. If a rectangular light source is in the part of the field of view toward the nose, its long side should be vertical. If it is in the part of the field of view toward the temple, its long side should be horizontal.

The reason for the preferred orientations has to do with the amount of the retina that is illuminated and with where in the field of view the arcs appear. Every small section of a narrow rectangular source creates a uniquely shaped arc. If the composite arc is to be bright and clear, the spread of the individual arcs should be small.

Several possible arcs for the right eye are shown in the illustration at the top left on the opposite page for a light source that is a vertical slit. The width of the slit is negligible. The height of the slit occupies 15 minutes of angle in the observer's field of view. The point labeled *f* is the point of fixation, that is, the point in the field of view to which you direct your gaze. (The image of this point falls on the fovea, the most sensitive area of the retina.) The graph of the field of view is measured in degrees, with the point of fixation at the origin.

If the slit is at point *c,* the large arc passing through *c* is excited. (Only the center of the arc appears in the illustration; the arc is actually one degree in width.) When the light is moved to point *b,* an arc is excited in the upper part of the field of view. When the light is moved to point *d,* an arc is generated in the lower part of the field of view.

A slit longer than 15 minutes of angle can excite many arcs. What the observer sees is a composite. The visibility of the composite depends in part on how the individual arcs overlap. Suppose the slit spans the field of view from *b* to *c.* The composite is bright because the spread in individual arcs is small enough to allow the visual system to sum their intensities. Similarly, if the slit is long enough to fill the span from *c* to *d,* another bright composite arc is created.

Still longer slits diminish the visibility

Dig here.

"Dynatel" equipment from 3M can scan miles of buried cable and pinpoint a break within inches of where it is located. In minutes, repair crews can also tell the type of fault; before one shovelful of earth is turned.

Not only do telephone calls, electricity and TV programs now flow on underground cables, but so do computer data; even critical health care information. Today thousands of companies count on Dynatel equipment to maintain their cables, reduce downtime and repair costs.

For one company, it cut a repair bill by $500,000.

Answering people's needs has helped 3M pioneer over 600 products for the electronics and electrical manufacturing industries. We now make everything from static eliminators to computer magnets and micro-circuit insulation.

And it all began by listening.

3M hears you...

For your free brochure on 3M's electrical and electronic products and all other 3M products, write: Dept. 030702/3M, P.O. Box 22002, Robbinsdale, MN 55422.

Name_____

Address_____

City_____ State____ Zip_____

Or call toll-free: **1-800-336-4327.**

3M

of the composite arc. For example, if the span of the slit is from *a* to *c*, the overlap of individual arcs yields a composite arc too wide for a complete summation of brightness. This arc is wider than the one generated by a shorter slit but is not proportionately brighter. Perception is even poorer if the slit is oriented horizontally through *c*. With this orientation a large number of arcs are excited to yield a wide composite arc.

The illustration at the top right on page 140 shows three general regions in the right eye's field of view where Moreland considers placing light sources. Good arcs are generated if the light source is in, say, section *C* and has the curved shape depicted in the illustration. Arcs with a different position and curvature are generated by a light stimulus positioned and curved like section *B*. A light stimulus like section *A* is different. A blue spike appears if a narrow light source is inserted between the point of fixation and the blind spot (the spot where the retinal nerve fibers exit to the brain).

Moreland does the following demonstrations of the various arcs. Place a rectangular source of light in section *A*. The source should be about .66 degree wide and about 2.5 degrees long. The probable appearance of the blue spike is shown in part *b* of the bottom illustration on page 140. By rotating your head you can bring the spike in line with the stimulus as in part *a*. If the same arrangement is presented simultaneously to the left eye, you get a striking view of two blue spikes.

Next place a light source in section *B*. For good results the source should be about 4.5 degrees long and about .66 degree wide at its narrowest. The ends are flared. This source generates two arcs that are symmetrical if the source is approximately symmetrical about the line connecting the point of fixation and the blind spot (part *c*). By tilting your head you can complete or break the symmetry of the arcs.

To demonstrate that the shape and the positioning of the arcs depend strongly on the position of the light source, set up two light sources as is shown in part *d*. One source occupies the upper half of section *C*, the other the lower half of section *B*. The resulting arcs are disconnected and have noticeably different curvatures. When stimulus lights are placed in all three positions, the result is two horseshoe-shaped arcs and one spike (part *e*).

In all these demonstrations arcs appear soon after the stimulus light is turned on. Similar but fainter arcs can be seen when a stimulus light is turned off under the right conditions. The on and off times of the room lights remain the same, but now during the dark period the stimulus light is on for 3.5 seconds and off for .5 second. You might have to repeat this sequence 10 times or so before the "off" arcs are visible.

As a light source Moreland suggests a photographic safelight with a 60-watt lamp covered with a red or yellow filter. I found that almost any small light (such as a penlight) covered with a red or yellow filter serves to generate arcs. My filters were inexpensive ones that I obtained from the Edmund Scientific Company (101 East Gloucester Pike, Barrington, N.J. 08007).

To construct sources of light in the forms suggested by Moreland attach to a lamp a sheet of sturdy cardboard with the appropriate form cut into it. Adjust your distance from the lamp until the angular size of the cutout in your field of view is right. You can use a flat lamp of the type normally employed for viewing and sorting slides.

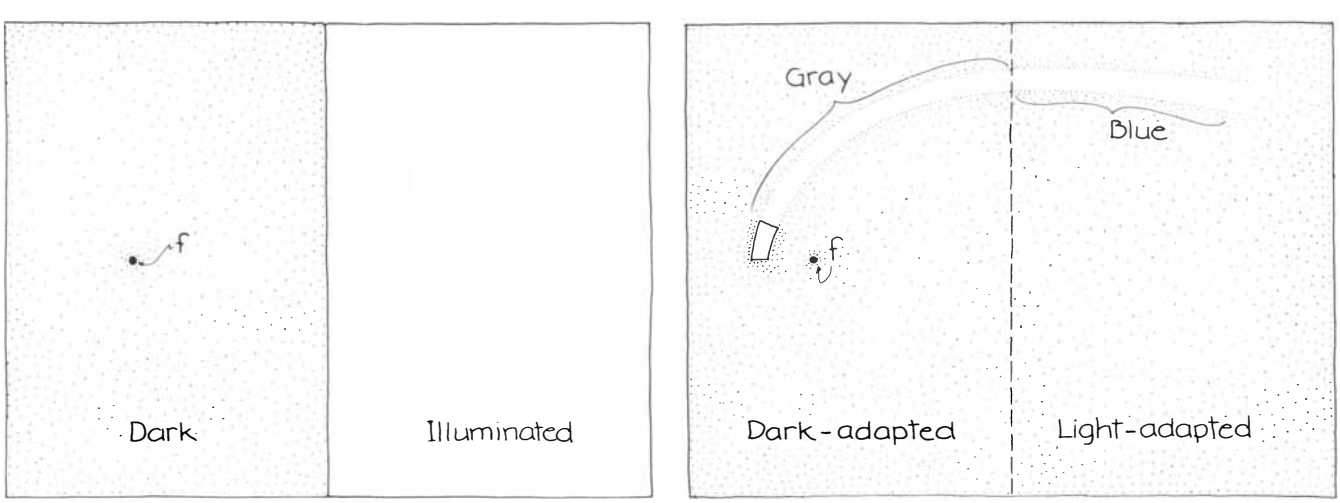
I made most of my observations on my home computer, which has a black-and-white monitor screen. In my first experiments I had a word-processor program in the computer's memory. The blinking cursor on the monitor was the light stimulus. Since the cursor is white, I taped a red filter over the screen. To create a point of fixation I typed in a period near the center of the monitor. With the cursor positioned to the left of the period I fixed the gaze of my right eye by concentrating on the period. After I had determined the best distance between my head and the monitor I held my head in position with books and other objects.

For other shapes and orientations of a stimulus light I wrote a program (in BASIC) that could light various parts of the monitor. The lighted pattern turned on and off at a rate set by FOR–NEXT loops in the program. For example, with a statement of FOR K = 1 to 500: NEXT the pattern remained on the screen until the computer counted to 500. Another loop kept the screen dark (except for the fixation point) for as long as I wanted. (I still operated the room lights by hand.)

Many causes have been proposed for the blue arcs. Purkinje and others thought they result from the scattering of light inside the eye. Other investigators suggested bioluminescence. Their hypothesis was that the excitation of nerve fibers by the stimulus light leads to the emission of ultraviolet radiation, which in turn excites light-receptor cells in the arc regions.

These two hypotheses have been replaced by the hypothesis that the excitation of nerve fibers by the stimulus light gives rise to activity in adjacent nerve pathways. One's interpretation of the additional activity is the perception of blue arcs. The regions that appear to be covered with arcs are actually not illuminated.

The first step in the creation of the perception of the arcs is to have a stimulus light illuminate a small section of the retina, exciting the cone cells there. The nerve fibers connected to the cones are then electrically active. Apparently the



An arrangement for demonstrating the colors of arcs

same fibers also electrically excite adjacent fibers connected to rod cells elsewhere on the retina. Thus the observer receives messages from cones that are excited by light and from the nerve fibers connected to rods that are not illuminated. The brain is misled into thinking the rods are excited by light too. The rods happen to be laid out on the retina in the shape of an arc.

Although this explanation of the arcs is promising, the source of the blueness of the arcs remains unidentified. The rods are thought to be achromatic. How then could the illusion of their excitation deceive the observer into thinking that he is seeing blue? Moreover, the arcs are blue only when the eye is still partially adapted to light. As the eye adjusts fully to darkness the arcs fade to a dull white.

In 1977 Carl R. Ingling, Jr., and Bruce A. Drum of Ohio State University presented an explanation of the color of the arcs that appears to solve the problem. To follow their demonstration you need to arrange a visual field that is dark on one side and illuminated on the other. Adapt your eyes to total darkness for about 10 minutes and then stare at the field. (The illustration on the preceding page suggests a setup that is appropriate for the right eye.) After several minutes replace this field with one that is totally dark except for a small, faint point of fixation and a slowly flashing red light needed to stimulate the arcs. The illustration indicates one type of arc that can be seen.

The part of the arc lying in the side previously adapted to illumination is blue. The part lying in the side that was dark-adapted is dull white. The border between the two colors can be sharp if the eye has been kept steady on the point of fixation.

To repeat this demonstration on my computer I included in the BASIC program some commands that lighted the right half of the screen long enough for that side of my right eye's field of view to become adapted. The fixation point was in the left part of the screen. In the next step of the program the screen went completely dark except for a slowly flashing pattern and the same fixation point. As before, I had taped a red filter over the place where the stimulus pattern appeared.

Ingling and Drum's explanation of the colors of the arcs is based on a previously published model of retinal response called the silent surround. In this model a group of retinal photoreceptors is considered to consist of central cells that respond to light and surrounding cells that can sometimes modify the rate at which the center directs signals to the brain. A silent surround means that when the center is unexcited, excitation of the surround does not alter its firing rate. When the center is excited more than the surround, the surround inhibits the center's signals to the brain.

The silent surround might apply to the generation of Purkinje's arcs if the center includes cone cells that are responsible for the information on yellow and blue sent to the brain. The surround consists of rod cells. Rods do not respond to color, and so it would be impossible for them to directly color the arcs. According to the hypothesis of Ingling and Drum, however, the rods supply the blueness indirectly when they inhibit yellow signals from the yellow-blue cones in the center.

In modern color theory the perception of color is thought to result from two kinds of competition among cones. The cones are specialized in their color response. One type of cone responds to red, another to green and a third to blue. Apparently the visual system sorts the three responses into two pairs of opponent colors: red and green as one pair, blue and yellow as the other.

Purkinje's arcs might arise as follows. The stimulus light illuminates a small section of the retina, exciting cones in it. The cones have nerve fibers running alongside fibers from other parts of the retina, so that the electrical excitation of the first group of fibers excites the second group. Presumably some of the nerve fibers in the second group are connected to rods that are among the silent-surround cells.

When the second group of fibers is excited, what signal do the fibers transmit to the brain? First, they signal the illusion that the retina is illuminated wherever they are connected to rods in the surround. Second, the fibers might also provide an illusion of blue, depending on the state of the centers of which they are a part. If the eye is fully adapted to darkness, the centers are not excited and transmit no color information. Therefore the excitation of the rod fibers sends a colorless signal to the brain. The result is that the observer sees arcs of dull white.

If the eye is not fully adapted to darkness, the cones of the centers are still sending color information to the brain. In particular some of them are sending the signal of yellow. The excitation of the fibers connected to the rods surrounding such cones then inhibits the yellow signal; the effect is the transmission of a blue signal to the brain. Thus the observer is misled into thinking certain regions of the retina are illuminated with blue light.

In my November 1983 column on water striders I incorrectly stated the critical speed below which a water strider does not generate waves. The speed is not .23 centimeter per second but .23 meter per second.

Opportunity. Far beyond the ordinary.

Top career challenge in solid state millimeter/microwave technology is at Hughes.

Here's an exciting opportunity to build a rewarding career in the unique fields of solid-state microwave and millimeter-waves at Hughes' Electron Dynamics Division in Torrance, California.

We are the leader in millimeter/microwave technology. And our rapid growth coupled with an expanding worldwide market, has opened immediate opportunities in a variety of related areas including:

- **Advanced Millimeter-Wave Subsystems Design & Development**
- **Advanced Project Management**
- **Semiconductor Devices & Materials (GaAs, Si) Development**
- **Monolithic Microwave/ Millimeter-Wave Integrated Circuits**
- **FET Technology**
- **Wide Band Amplifiers and Receivers**
- **Microwave/Millimeter-Wave Components and Subsystems Manufacturing Engineering**
- **Phase-Lock Loop & Analog/Digital Circuit Design**
- **Mechanical & Packaging Design**
- **Millimeter-Wave Instrumentation**

Your related microwave experience and/or BS, MS, or PhD degree (AA for research assistants) can be the key to a great future. We offer excellent salary and benefits, big-company stability, and growth opportunity. Your creativity and initiative will be encouraged and rewarded because you will play an important role in a dedicated independent engineering project team — rather than working on a small portion of the development of large systems. That means a much better chance for you to grow, personally and professionally. Also, you'll enjoy a stimulating lifestyle in our modern beach area community, in the heart of Southern California.

If you're ready for a challenge that's far beyond the ordinary, act now. Please send your resume to: Marianne Renz, Hughes Electron Dynamics, Dept. AS-2, P.O. Box 2999, Torrance, CA 90509.

MAKING WAVES IN MILLIMETER/MICROWAVE TECHNOLOGY

Creating a new world with electronics

HUGHES

HUGHES AIRCRAFT COMPANY

ELECTRON DYNAMICS

Equal Opportunity Employer

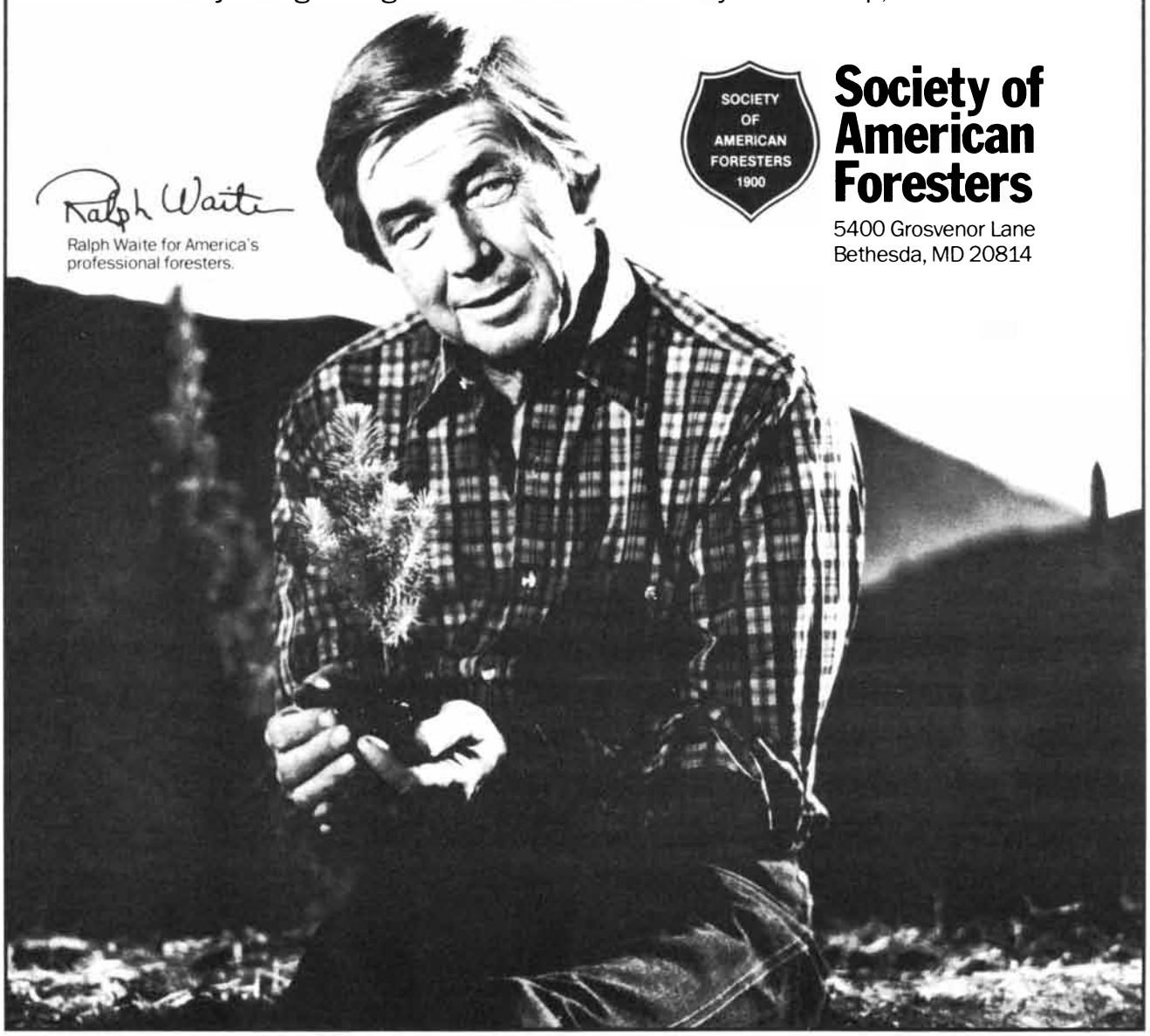
"Imagine where we'd all be today if wood didn't burn."

"We'd all be a little colder—and a lot poorer.

"With plentiful supply, people have turned back to wood to produce dependable inexpensive heat from woodstoves and fireplaces.

"This new demand is coming at a time when we're losing a thousand square miles of forestland each year to urban expansion and other people pressures. So we've got to take extra good care of the forests we have.

"Our job is growing. For information on how you can help, write..."



SOCIETY OF AMERICAN FORESTERS 1900

Society of American Foresters

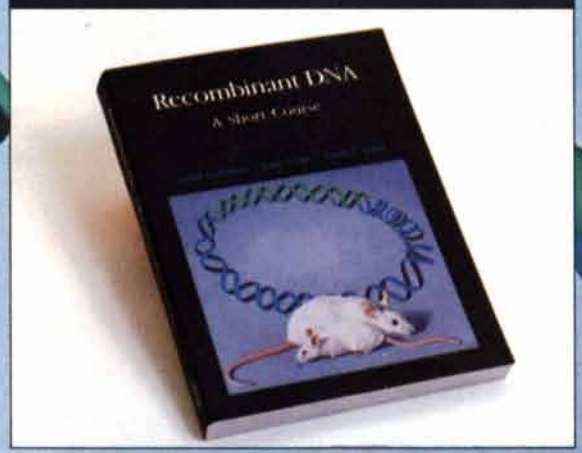
5400 Grosvenor Lane
Bethesda, MD 20814

"An excellent, even extraordinary book" *

An up-to-the-minute overview of the state of recombinant DNA genetics, this eagerly awaited book summarizes the history of the recombinant DNA revolution. Nobel Laureate James Watson and his colleagues survey the recent techniques in recombinant molecular genetics, explaining both the experimental methods now in use and their dramatic results.

Enhanced by 174 exquisitely detailed illustrations, RECOMBINANT DNA: A SHORT COURSE explores such vital topics as tumor viruses, movable genes, the experimentally controlled introduction of DNA into yeast cells, and the genetic engineering of plants. Final chapters are devoted to viral vectors, recombinant DNA and genetic disease, and the science used in the recombinant DNA industry.

RECOMBINANT DNA: A SHORT COURSE



James D. Watson,
Cold Spring Harbor Laboratory

John Tooze,
European Molecular Biology Organization

David T. Kurtz,
Cold Spring Harbor Laboratory

"A masterful blend of 'how-to' methodology with brilliant particulars and conceptual understanding which bears the unmistakable imprint of J. Watson and the additional authority of Drs. Tooze and Kurtz. The three individual styles have somehow managed to become integrated so that by and large the manuscript sounds as if a single author had composed it." —SHERMAN BEYCHOK, *Columbia University**

A Scientific American Book, distributed by W. H. Freeman and Company.
260 pages, 174 illustrations.
Cloth: $27.95 Paper: $17.95

Painting by Marvin Mattelson.

Please send me ____ copy (copies) of RECOMBINANT DNA: A SHORT COURSE at _____ $27.95 (cloth: ISBN 1483-3)
_____ $17.95 (paper: ISBN 1484-1)

☐ I enclose a check made payable to W. H. Freeman and Company (please add $1.50 for postage and handling; New York, California, and Utah residents include appropriate sales tax). Charge my ☐ MasterCard ☐ VISA.

Account # _____ Expires ___ / ___

Signature _____
(Credit card orders must be signed.)

Name _____

Address _____

City _____ State/Zip _____

W. H. FREEMAN AND COMPANY
4419 West 1980 South, Salt Lake City, Utah 84104

0112

BIBLIOGRAPHY

Readers interested in further explanation of the subjects covered by the articles in this issue may find the following lists of publications helpful.

COMPUTER RECREATIONS

MINDSTORMS: CHILDREN, COMPUTERS, AND POWERFUL IDEAS. Seymour Papert. Basic Books, Inc., 1980.

TURTLE GEOMETRY: THE COMPUTER AS A MEDIUM FOR EXPLORING MATHEMATICS. Harold Abelson and Andrea A. diSessa. The MIT Press, 1980.

THE POPULATION OF CHINA

CHINA'S POPULATION: PROBLEMS AND PROSPECTS. Liu Zheng, Song Jian et al. New World Press, Beijing, 1981.

RESEARCH ON THE POPULATION OF CHINA: PROCEEDINGS OF A WORKSHOP. Edited by Robert J. Lapham and Rodolpho A. Bulatao. Committee on Population and Demography, Assembly of Behavioral and Social Sciences, National Research Council, 1981.

POPULATION TRENDS IN CHINA AND INDIA. Ansley J. Coale in *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 80, No. 6, pages 1757–1763; March, 1983.

POPULATION AND BIRTH PLANNING IN THE PEOPLE'S REPUBLIC OF CHINA. In *Population Reports,* Series J, No. 25; Vol. 10, No. 1, pages J-577–J-618; January–February, 1982.

ON THE RESULTS OF THE CHINESE CENSUS. Li Chengrui in *Population and Development Review,* Vol. 9, No. 2, pages 326–344; June, 1983.

THE OLDEST EUKARYOTIC CELLS

THE OLDEST FOSSILS. Elso S. Barghoorn in *Scientific American,* Vol. 224, No. 5, pages 30–42; May, 1971.

THE EVOLUTION OF THE EARLIEST CELLS. J. William Schopf in *Scientific American,* Vol. 239, No. 3, pages 84–102; September, 1978.

SYMBIOSIS IN CELL EVOLUTION. Lynn Margulis. W. H. Freeman and Company, 1981.

PROTEROZOIC PLANKTON. Gonzalo Vidal and A. H. Knoll. Geological Society of America, Memoire No. 30, in press.

THE EARTH'S ORBIT AND THE ICE AGES

VARIATIONS IN THE EARTH'S ORBIT: PACEMAKER OF THE ICE AGES. J. D. Hays, John Imbrie and N. J. Shackleton in *Science,* Vol. 194, No. 4270, pages 1121–1132; December 10, 1976.

ICE AGES: SOLVING THE MYSTERY. John Imbrie and Katherine Palmer Imbrie. Enslow Publishers, 1979.

MILANKOVITCH AND CLIMATE CHANGE. Edited by J. Imbrie and A. Berger. Elsevier Science Publishing Co., Inc., 1984.

THE STRUCTURE OF PROTEINS IN BIOLOGICAL MEMBRANES

STRUCTURAL ANALYSIS OF MACROMOLECULAR ASSEMBLIES BY IMAGE RECONSTRUCTION FROM ELECTRON MICROGRAPHS. R. A. Crowther and A. Klug in *Annual Review of Biochemistry,* Vol. 44, pages 161–182; 1975.

THREE-DIMENSIONAL MODEL OF PURPLE MEMBRANE OBTAINED BY ELECTRON MICROSCOPY. R. Henderson and P. N. T. Unwin in *Nature,* Vol. 257, No. 5521, pages 28–32; September 4, 1975.

STRUCTURE OF THE JUNCTION BETWEEN COMMUNICATING CELLS. P. N. T. Unwin and G. Zampighi in *Nature,* Vol. 283, No. 5747, pages 545–549; February 7, 1980.

MEMBRANE PROTEIN STRUCTURE. R. Henderson in *Membranes and Intercellular Communication; Les Houches Session XXXIII, July 30–August 30, 1979,* edited by Roger Balian, Marc Chabre and Philippe F. Devaux. Amsterdam North-Holland Publishing Co., 1981.

THE SPECTROSCOPY OF SUPERCOOLED GASES

SUPERSONIC NOZZLE BEAMS. James B. Anderson, Ronald P. Andres and John B. Fenn in *Advances in Chemical Physics,* Vol. 10, pages 275–317; 1966.

LASER SPECTROSCOPY IN SUPERSONIC JETS. D. H. Levy, L. Wharton and R. E. Smalley in *Chemical and Biochemical Applications of Lasers,* edited by C. Bradley Moore. Academic Press, 1977.

MOLECULAR OPTICAL SPECTROSCOPY WITH SUPERSONIC BEAMS AND JETS. Richard E. Smalley, Lennard Wharton and Donald H. Levy in *Accounts of Chemical Research,* Vol. 10, No. 4, pages 139–145; April, 1977.

NUMBERS AND MEASURES IN THE EARLIEST WRITTEN RECORDS

A METHOD FOR DECIPHERMENT, THROUGH MATHEMATICAL AND METROLOGICAL ANALYSIS, OF PROTO-SUMERIAN AND PROTO-ELAMITE SEMI-PICTOGRAPHIC INSCRIPTIONS. J. Friberg in *Research Report No. 1978–9.* Department of Mathematics, Chalmers Tekniska Högskola, Göteborg, 1978.

METROLOGICAL RELATIONS IN A GROUP OF SEMI-PICTOGRAPHIC TABLETS OF THE JEMDET NASR TYPE, PROBABLY FROM URUK-WARKA. J. Friberg in *Research Report No. 1979–15.* Department of Mathematics, Chalmers Tekniska Högskola, Göteborg, 1979.

A SURVEY OF PUBLICATIONS ON SUMERO-AKKADIAN MATHEMATICS, METROLOGY AND RELATED MATTERS (1854–1982). J. Friberg in *Research Report No. 1982–17.* Department of Mathematics, Chalmers Tekniska Högskola, Göteborg, 1982.

SOLITARY BEES

GENESIS OF THE HYMENOPTERA AND THE PHASES OF THEIR EVOLUTION. S. I. Malyshev. Methuen and Co., 1966.

INSECT POLLINATION OF CROPS. John B. Free. Academic Press, 1970.

THE SOCIAL BEHAVIOR OF THE BEES: A COMPARATIVE STUDY. C. D. Michener. Harvard University Press, 1974.

EVOLUTION OF INSTINCT: COMPARATIVE ETHOLOGY OF HYMENOPTERA. Kunio Iwata. Published for the Smithsonian Institution and the National Science Foundation by Amerind Publishing Co., New Delhi, available from the U.S. Department of Commerce, National Technical Information Service, 1976.

THE SKILL OF TYPING

THE PSYCHOLOGY OF SKILL, WITH SPECIAL REFERENCE TO ITS ACQUISITION IN TYPEWRITING. William Frederick Book. University of Montana Press, 1908.

A METHOD OF TEACHING TYPEWRITING BASED UPON A PSYCHOLOGICAL ANALYSIS OF EXPERT TYPING. J. E. Coover in *National Education Association; Addresses and Proceedings,* Vol. 61, pages 561–567; 1923.

ON THE ABILITY TO INHIBIT COMPLEX MOVEMENTS: A STOP-SIGNAL STUDY OF TYPEWRITING. G. D. Logan in *Journal of Experimental Psychology: Human Perception and Performance,* Vol. 8, pages 778–792; 1982.

COGNITIVE ASPECTS OF SKILLED TYPEWRITING. Edited by W. E. Cooper. Springer-Verlag, 1983.

THE AMATEUR SCIENTIST

WHY THE BLUE ARCS OF THE RETINA ARE BLUE. Carl R. Ingling, Jr., and Bruce A. Drum in *Vision Research,* Vol. 17, pages 498–500; 1977.

THE EFFECT OF HUMMING ON VISION. H. A. K. Mastebroek and J. B. van der Kooi in *Physics Education,* Vol. 14, pages 253–254; 1979.

BREAKTHROUGH: ONE IF BY LAND. ONE BILLION IF BY LASER.

The use of light to send simple messages is rooted in history. A new nation's destiny once hung on a single gleam.

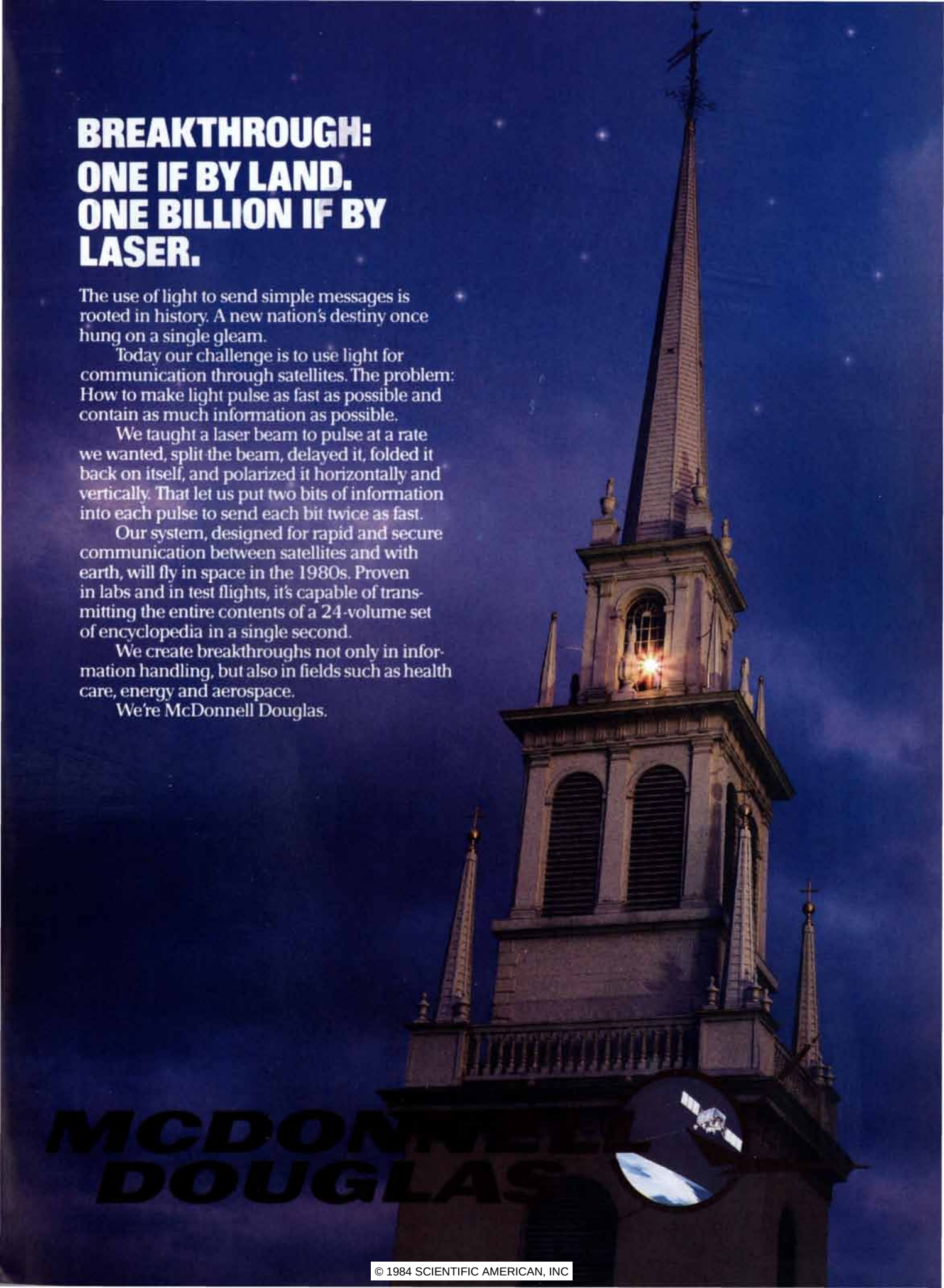
Today our challenge is to use light for communication through satellites. The problem: How to make light pulse as fast as possible and contain as much information as possible.

We taught a laser beam to pulse at a rate we wanted, split the beam, delayed it, folded it back on itself, and polarized it horizontally and vertically. That let us put two bits of information into each pulse to send each bit twice as fast.

Our system, designed for rapid and secure communication between satellites and with earth, will fly in space in the 1980s. Proven in labs and in test flights, it's capable of transmitting the entire contents of a 24-volume set of encyclopedia in a single second.

We create breakthroughs not only in information handling, but also in fields such as health care, energy and aerospace.

We're McDonnell Douglas.

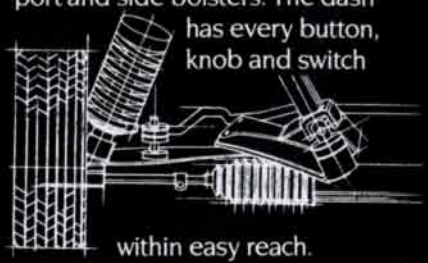


MCDONNELL DOUGLAS

The Toyota Supra. A sleek and powerful driving machine. Smoldering beauty. Scorching performance. An automotive legend in just three years. How does a car earn such a glowing reputation? The answer begins under the hood, the heart of every performance car. For 1984, the 5-speed Supra has been refined, squeezing 10 more horsepower out of its already hot engine. The compression ratio was raised, the intake manifold enlarged for easier engine breathing and the exhaust system streamlined. As a result, the electronically fuel injected 2.8 litre Twin-Cam six propels Supra from 0-60 in less than 8.6 seconds.

In the handling department, standard features like independent rear suspension, MacPherson struts, variable assist power rack-and-pinion steering, and wide 225/60 HR14 radial tires on 14"x7" aluminum alloy wheels, give Supra superb cornering and precise control. A new front air dam was added to channel the wind for less resistance and even better road holding.

Inside the Supra, uncompromising attention has been paid to create the ultimate high performance environment. The Supra's now famous 8-way adjustable sport seat won't take a back seat to anybody else's. In a review of the ten best automotive features, Car and Driver Magazine rated the Supra driver's seat as the best in the business.* The driver's seat has a wide variety of adjustments from a unique pneumatic lumbar adjustment system to movable thigh support and side bolsters. The dash has every button, knob and switch within easy reach.



The 1984 Toyota Supra. Superb performance and stunning beauty. The right stuff to make a car a legend.

* Car and Driver Magazine, January 1983

OH WHAT A FEELING!

TOYOTA

BUCKLE UP—IT'S A GOOD FEELING!

THE NEW 1984 SUPRA. CHARIOT OF FIRE.

HOT!