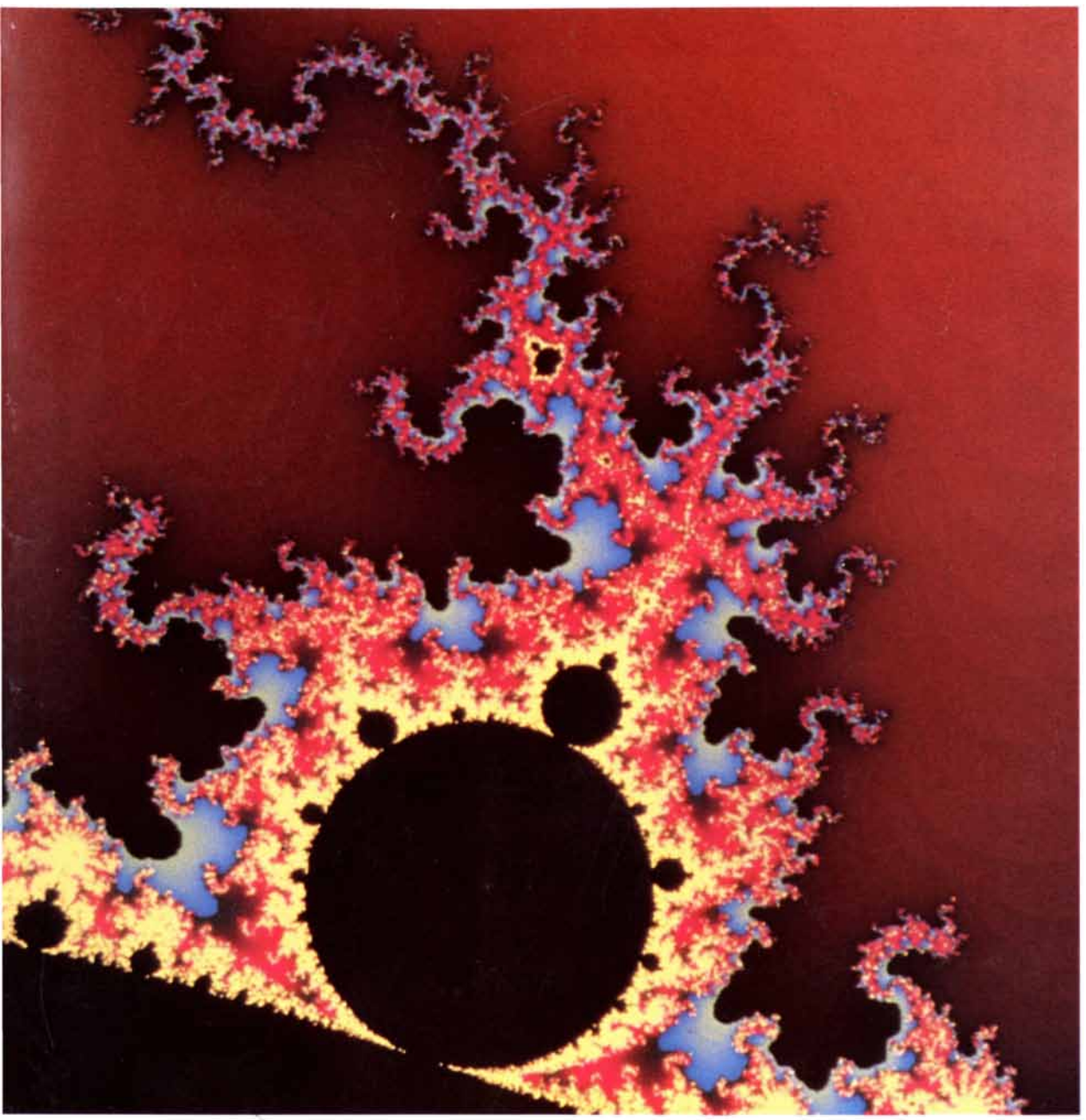
SCIENTIFIC AMERICAN



EXPLORING THE MANDELBROT SET

$2.50

August 1985

How Exxon uses computer our understanding at

At Exxon, theory and mathematics, coupled with the computer, are helping to revolutionize the research and development process. In fact, computer models are fast becoming a new probe for complex systems, coequal with experiment.

The Theoretical and Mathematical Sciences Laboratory: Dealing with complexity

New mathematics helped create the computer. Now the computer is creating an accelerating demand for new mathematics. In the process, linear mathematics is giving way to more complex, nonlinear mathematical ideas and models.

At the Theoretical and Mathematical Sciences Laboratory of Exxon Research and Engineering Company (ER&E), these developments are helping our theoretical scientists and mathematicians gain insight into such basic matters as metallic strength, corrosion and catalysis. Difficult problems are now being solved by combined use of the laboratory and computer modeling.

Catalytic Activity: A new understanding

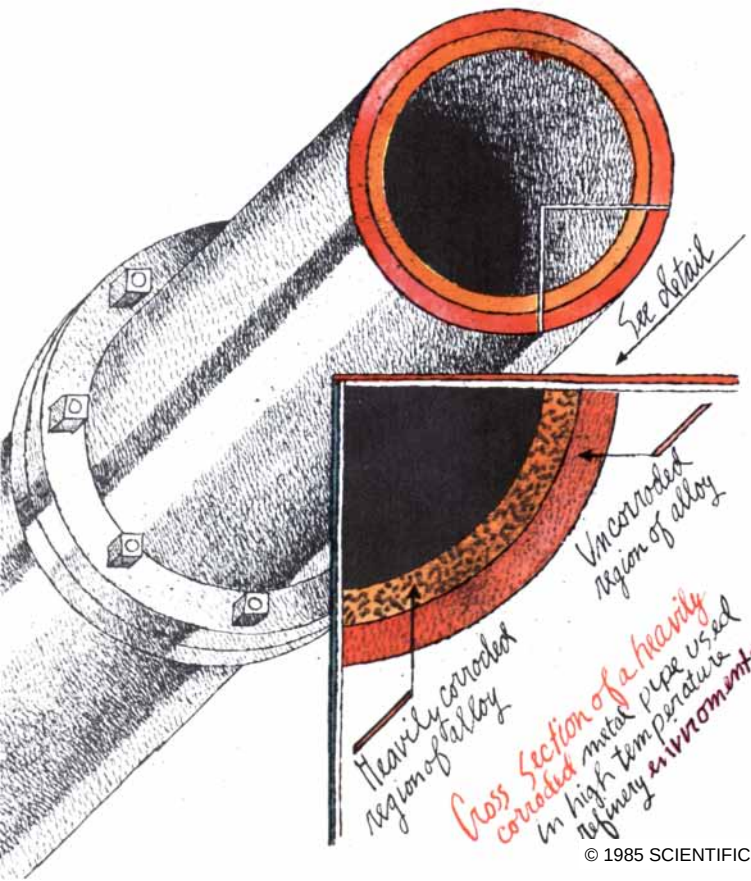
Catalysts based on transition metal sulfides have been widely used for sixty years in processing petroleum, particularly to remove sulfur and nitrogen. Increasing amounts of sulfur and nitrogen in today's feedstocks make these processes particularly important, yet the fundamentals of catalytic activity have been little understood. Now, by numerical solution of the equations of quantum chemistry, a quantitative correlation between catalytic activity and the electronic structure of the catalysts has been demonstrated for the first time.

Data from a model hydrodesulfurization reaction uncovered large differences in activity for various catalysts. Theoretical calculations revealed relationships between this behavior and the electronic structure of the catalysts. Information such as this opens the way to systematically controlling catalyst properties.

Alloy Oxidation Modeli : A new theory

In many industrial processes it is necessary to use complex metal alloys that can survive high-temperature, corrosive environments. These alloys can withstand such harsh environments because of their ability to form protective oxide coatings or scales.

Now a new nonlinear mathematical theory for alloy oxidation has been developed through a collaboration of Exxon metallurgists and mathematicians. Computer-generated solutions have advanced our understanding of how the



modeling to extend the molecular level.

physical processes interact in forming protective oxides. We now see opportunities for controlling an alloy's environment in order to promote these scales.

Computer Simulation of Grain Growth in Metals: A new solution

The physical and chemical properties of materials are determined in part by their microstructure. Grain orientation and size affect yield strength, fracture, surface adsorption phenomena and other properties. In order to effectively tailor the microstructure for specific applications, the understanding of the mechanism and kinetics of grain growth and recrystallizations would provide valuable guidance.

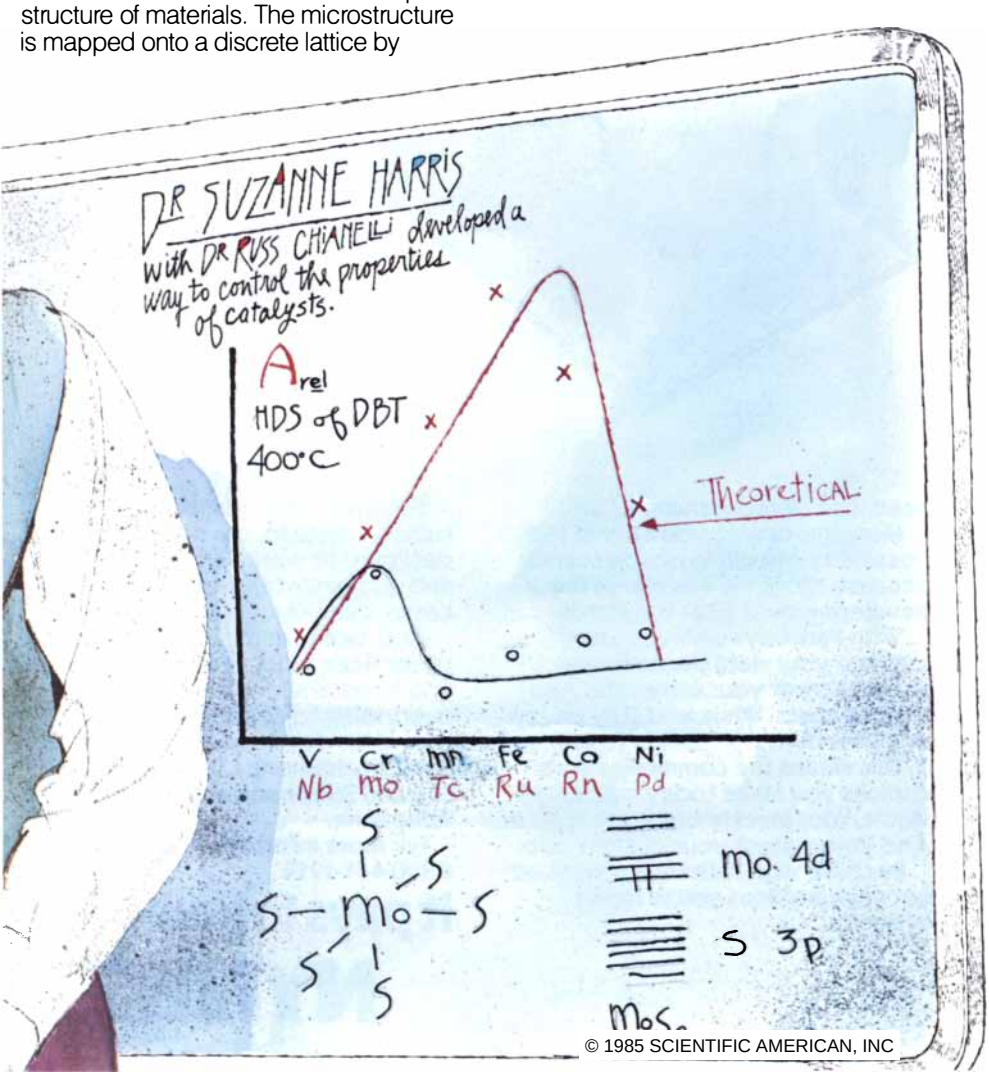
Exxon scientists have developed a method to bridge the gap between the atomistic interactions and macroscopic structure of materials. The microstructure is mapped onto a discrete lattice by dividing the material into small volume elements and placing the centers of these elements on lattice points. This discrete model preserves the topological features of real systems and can be studied by computer simulation using Monte Carlo techniques.

To test this method, the simulation was applied to a long-standing metallurgical problem–grain growth in isotropic polycrystalline metals. For the first time, all features of its microstructure (topology, morphology and kinetics) were correctly predicted. Now, similar modeling is being used to study the development of microstructure when stresses and temperature are applied. Our objective here is to design microstructure starting from atomistic interactions.

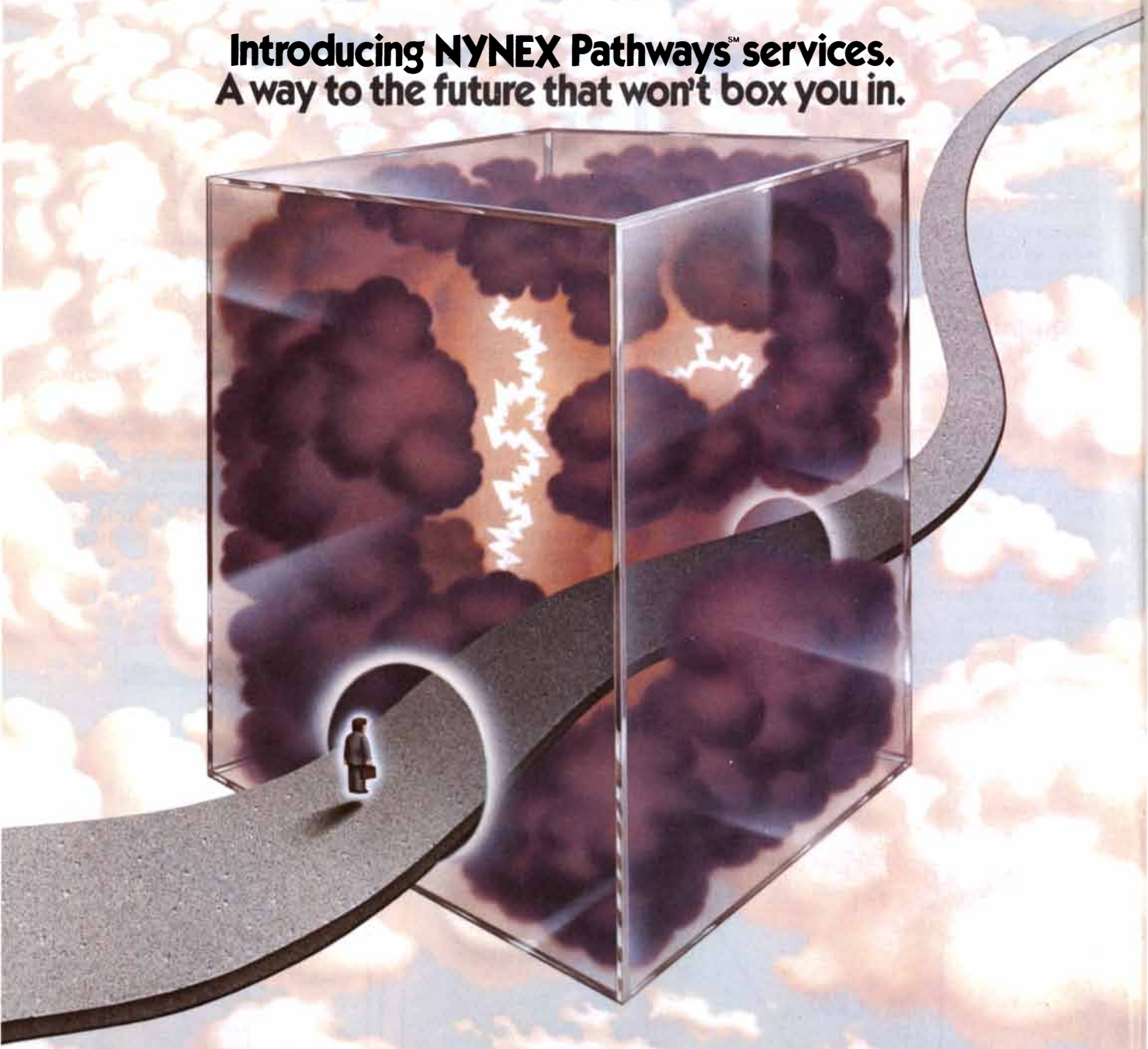
Exxon Research and Engineering Company

Exxon Research and Engineering Company is a wholly owned subsidiary of Exxon Corporation. More than 2,000 scientists and engineers are working at ER&E on petroleum products and processing, synthetic fuels, pioneering science and the engineering required to develop and apply new technology in the manufacture of fuels and other products. For more information, write Ed David, President, Exxon Research and Engineering Company, Room 101, P.O. Box 101, Florham Park, New Jersey 07932.

EXXON



Introducing NYNEX Pathways℠ services.
A way to the future that won't box you in.



Now you can meet your communications needs today. And at the same time guarantee your compatibility with the future.

With Pathways advanced network services, you can take full advantage of evolving digital technology. And you can relax about the future.

Because you'll be prepared for the Integrated Services Digital Network (ISDN). Which will simply and economically connect you to a wide variety of services available through the public switched network.

This means you can avoid the expense and downtime of building your own network. As well as the headache of maintenance.

More important, you'll avoid the possibility of built-in obsolescence. Because NYNEX is a leader in the development of ISDN standards.

With Pathways services you'll improve your data transmission speed. Lower your error rate. And control costs. While you stay on top of all the digital advances to come.

This means the communications choices you make today will remain viable. Your investment is protected. And you protect your options, too.

Because with Pathways advanced network services you're never boxed in.

You aren't locked into one supplier. Instead, our services are designed to work with services and equipment from a variety of carriers and manufacturers.

And, because it's a family of services, you choose only what you need. You don't have to take everything to get one thing.

And it's delivered to you by the NYNEX telephone companies, New England Telephone and New York Telephone.

For more information, call 1 800 441-1919.

It pays to know us.

NYNEX

Established 1845 **SCIENTIFIC AMERICAN** August 1985 Volume 253 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 84104. 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 photograph on the cover is a computer-generated enlargement of part of the boundary of the Mandelbrot set, a mathematical construct that has been called the most complicated object in mathematics (see "Computer Recreations," by A. K. Dewdney, page 16). The set is named after Benoit B. Mandelbrot of the IBM Thomas J. Watson Research Center, who was the first to investigate its properties. The members of the Mandelbrot set are complex numbers, or numbers with a real and an imaginary part, which can be graphed as points in the complex plane. The points in the set are shown in black; the colored regions of the image represent points that, in a specially defined mathematical sense, are fleeing from the boundary of the set. The region of the Mandelbrot set from which this enlargement is made is indicated by the white outline labeled *g* in the illustration on page 17. This photograph, as well as the ones that accompany this month's column, was made in the Graphiklabor Dynamische Systeme by Heinz-Otto Peitgen, Peter H. Richter and Dietmar Saupe at the University of Bremen. Copyright MAPART.

THE ILLUSTRATIONS

Cover photograph courtesy of Graphiklabor Dynamische Systeme

Page	Source
17–19	Heinz-Otto Peitgen, Peter H. Richter, Dietmar Saupe
20	John H. Hubbard
26–27	Canadian Hydrographic Service, © 1984
34–39	Alan D. Iselin
41	Wolfgang Schmidt
42	Mayo Medical Library
43	Rex E. Coupland
44	Bunji Tagawa (*top*), Reiner Fischer-Colbrie (*bottom*)
45–46	Bunji Tagawa
47	Ruth E. Siegel (*top*), Richard Ornberg and Richard D. Leapman (*bottom*)
48	Odile Grynszpan-Winograd (*left*), Wolfgang Schmidt (*right*)
49	Alexandra Patzak
51	Gerd Binnig and Heinrich Rohrer
52	Gerd Binnig and Heinrich Rohrer (*top*), Ian Worpole (*bottom*)
53	Ian Worpole
54	Gerd Binnig and Heinrich Rohrer
55	Gerd Binnig and Heinrich Rohrer (*top*), Ian Worpole (*bottom*)
56	Heinrich Rohrer
70–71	The Granger Collection
72–78	Gabor Kiss
80–85	Patricia J. Wynne
87	Eric Buffetaut and Rucha Ingavat
89	Mendillo Collection of Astronomical Prints
90	David Green and Steven Gull
91	Frederick D. Seward, Paul Gorenstein and Wallace H. Tucker
92–95	Hank Iken, Walken Graphics
96	Frederick D. Seward, Paul Gorenstein and Wallace H. Tucker
99	U.S. Geological Survey
100–101	Andrew Tomko
102	David J. Drewry (*top*), Gordon de Q. Robin, David J. Drewry and D. T. Meldrum (*bottom*)
103	Andrew Tomko
104	Claude Lorius, University of Grenoble
107	Enid Kotschnig
108–110	Tom Prentiss
111	Enid Kotschnig
112–113	Tom Prentiss
115–117	Michael Goodman

Introducing The Single Solution To Many Problems That Can Tie Up Your Personal Computer.

You need to use your computer for a critical job, but can't—because it's tied up running your printer or modem. Or you're rushing to compose and print a report, when you realize the wrong printer is hooked up, and you have to stop and swap cables.

Fortunately, there's a simple solution to these problems—and many more. It's a remarkable new device called Transet 1000, from Hayes.

Transet 1000 relieves your computer from operating your printer or modem, so you and your computer can go on to another task. As fast as your computer can send it, Transet 1000 receives and stores more than 75 pages of text. Then it feeds that data to your printer or modem, at the speed at which the peripheral can accept it. At the same time, you can be using your computer for an entirely different job.

While performing other functions, Transet 1000 can receive data through a modem. In fact, it operates as an unattended electronic mailbox, receiving messages 24 hours a day. Even if the computer's turned off.

Transet 1000 automatically routes information among peripherals, so you don't have to physically change connecting cables. It also allows two computers to share an expensive peripheral, like a laser printer.

And Transet 1000 works with a wide range of systems and configurations.

But these are just a few of the problems Transet 1000 can solve for you. For more information, see your authorized Hayes dealer. Soon.

Because you should be working with your computer. Not waiting for it.

Hayes Transet 1000™

Innovative products for enterprising people

Hayes Microcomputer Products, Inc.
P.O. Box 105203, Atlanta, GA 30348
404-441-1617

Meeting Japan's Challenge
Twenty-First in a Series

HOW WE SELL IN A WORLDWIDE BUYER'S MARKET.

In electronics, competition has never been more intense.

Global marketing has not only added to the number of world marketplaces, but has multiplied the international companies which compete in those markets.

In this environment, the advantage goes to the organization which thinks more about what its customers need to buy, than about what it has to sell.

Here's a good case in point. Not too long ago, we were locked in a heated competition with four other companies, including a leading electronics company from Japan, for a large contract with a major American manufacturing company. Each contestant was challenged to design a new electronic control module. The climax of the competition was to be a field test of an "emulator," an oversized working model of the suggested design. The winning emulator would then be carried to the next step, the actual chip that would be the brain of the control system for their new product.

Our strategy for winning this competition was based on our understanding of the customer's business and demanding standards. We elected to develop both emulator and the actual chip at the same time, even though this would make our participation far more cost and time-intensive.

When the final day came, each of the competitors presented its emulator, and we showed ours, as well. But we also had the chip. It was finished, working, tested and ready for volume production.

And it worked out as we hoped. All the contestants produced satisfactory emulators, but the business was awarded to us. Our design and the quality of our product solved many problems for the company. And the time we saved them helped push production ahead by six to eight months.

In the final analysis, every business in the world exists, functions and prospers at the pleasure of its customers. When companies lose sight of that, they lose sight of their purpose.

MOTOROLA A World Leader in Electronics.

Quality and productivity through employee participation in management.

LETTERS

To the Editors:

In "Brownian Motion" [SCIENTIFIC AMERICAN, February], Bernard H. Lavenda writes: "Einstein was the first with the physical insight to recognize that atoms would reveal their existence in the motions of particles suspended in a fluid."

We think the following paragraph from Lucretius' *De Rerum Natura* is worth quoting: "Observe what happens when sunbeams are admitted into a building and shed light on its shadowy places. You will see a multitude of tiny particles mingling in a multitude of ways in the empty space within the light of the beam, as though contending in everlasting conflict, rushing into battle rank upon rank with never a moment's pause in a rapid sequence of unions and disunions. From this you may picture what it is for the atoms to be perpetually tossed about in the illimitable void. To some extent a small thing may afford an illustration and an imperfect image of great things. Besides, there is a further reason you should give your mind to these particles that are seen dancing in a sunbeam: *their dancing is an actual indication of underlying movements of matter that are hidden from our sight.* There you will see many particles under the impact of invisible blows changing their course and driven back upon their tracks, this way and that, in all directions. *You must understand that they all derive this restlessness from the atoms. It originates with the atoms, which move of themselves.* Then those small compound bodies that are least removed from the impetus of the atoms are set in motion by the impact of their invisible blows and in turn cannon against slightly larger bodies. So the movement mounts up from the atoms and gradually emerges to the level of our senses . . . " (our italics).

It seems that Lucretius, who like Democritus and Epicurus was a convinced atomist, should also be credited with the insight that the existence of atoms could be deduced from the movements of particles in suspension.

JAIME GABARRO-ARPA

Institut Jacques Monod
Paris

ANTONIO LOPEZ-CAMPILLO

Laboratoire de Photophysique
Moléculaire
Orsay

To the Editors:

If we are interested in a philosophical rather than a scientific history of Brownian motion, Drs. Gabarro-Arpa and Lopez-Campillo are correct in tracing the concept back to the atomic philosophers of ancient Greece, whose beliefs are reflected in *De Rerum Natura.* Undoubtedly the notion that matter is made up of tiny particles is one of the oldest in recorded history. Yet we must be careful to distinguish between the implications of a philosophical doctrine and a physical theory that can be tested experimentally.

Nothing can be so poignant as the alarm (and even fear) of 19th-century philosophers at the possibility that science might provide a mechanical explanation of physical phenomena in terms of the motion of atoms. To 19th-century chemists and physicists atoms provided a convenient way of reasoning about the observed properties of matter and expressing relations among them. To philosophers, however, Lucretius' hypothesis that atoms "deviate from their courses at quite uncertain times and places" (a quotation from a talk on molecules presented by James Clerk Maxwell at a meeting of the British Association for the Advancement of Science at Bradford in 1873) seemed to impute to atoms the same free will that was thought to be responsible for the human power of voluntary action. As long as atoms remained unobservable, there was really no conflict between the two conceptions. It was Albert Einstein who first transformed the hypothesis into an experimentally verifiable physical theory.

BERNARD H. LAVENDA

University of Camerino
Italy

To the Editors:

The recent article on the remarkable evasive behavior of the trypanosome ["How the Trypanosome Changes Its Coat," by John E. Donelson and Mervyn J. Turner, SCIENTIFIC AMERICAN, February], accomplished through the rapid replacement of its coat protein antigens, is an excellent example of the power of modern molecular genetics. An ultimate aim of the research, however, is to develop a way of controlling or eliminating trypanosomiasis by means of techniques that will subvert the organism's defense mechanisms. The article suggested that this might be done by a metabolic attack through drugs or by developing resistant strains of cattle.

I should like to have seen some dis-

cussion about the effect of controlling trypanosomiasis on the overall ecology of the affected area of Africa. What will happen, for example, if trypanosome-resistant cattle proliferate greatly and by overfeeding eliminate the grassland and contribute to the rapid expansion of the desert? There have already been reports of the local effects of eliminating the tsetse fly: the desert has been extended and the cattle herds have been greatly reduced, increasing the misery of the human population and resulting in their mass migration into the nearby cities. These are already greatly overcrowded and unable to cope with their own population. In other words, are we really accomplishing a useful end by this premature application of technology? Is it not necessary for us to plan much more thoughtfully? Should we not try to work out a way to prevent the predicted disastrous alteration of the land before we jump in and do "good"?

JEROME GROSS

Harvard Medical School
Boston, Mass.

To the Editors:

Dr. Gross's valid concerns are often discussed among workers in all areas of trypanosomiasis research. Two points can be emphasized here. First, trypanosomiasis is a disease of humans as well as livestock; in recent years an increasing number of human patients have been diagnosed, and the threat of future epidemics is still present. Second, Africa currently has the highest human birthrate in the world. Enormous problems related to nutrition, health, employment and the general quality of life are associated with the population growth. Overcoming these problems requires many different approaches on many different levels. It would therefore be unwise to ignore any potential solution to this serious disease. If a way to control trypanosomiasis is found, it will become one of many considerations in the effort to manage all the land and resources of the world in ways that minimize ecological impact and maximize economic and social benefits.

JOHN E. DONELSON

University of Iowa

MERVYN J. TURNER

Merck Sharp & Dohme
Research Laboratories
Rahway, N.J.

A new look at

Summary:

GTE lighting research operates on many fronts: a space lighting lab to study the motion of gases in a gravity-free environment; the use of various isotopes to enhance the output of fluorescent lamps; the production of light directly from excited molecules.

The science of lights and lighting might seem to be rather mature. Indeed, the standard light bulb has changed very little in at least half a century.

But lighting science is on the brink of revolution. Recent work by GTE points the way to major improvements in every type of lighting.

Lighting research in space.

One of the most powerful and efficient light sources is the high-intensity-discharge (HID) lamp. Its light is derived from gases and ionized vapors which are excited in an electrical arc contained in a quartz arc tube.

The gases circulate by gravity-induced convection, which mixes the radiating species in the arc. This tends to obscure other vital processes such as diffusion, cataphoresis (motion of ions toward the negative electrode), magnetostriction and vapor condensation. Researchers have wanted to observe these processes at leisure, in the absence of convection, for many years.

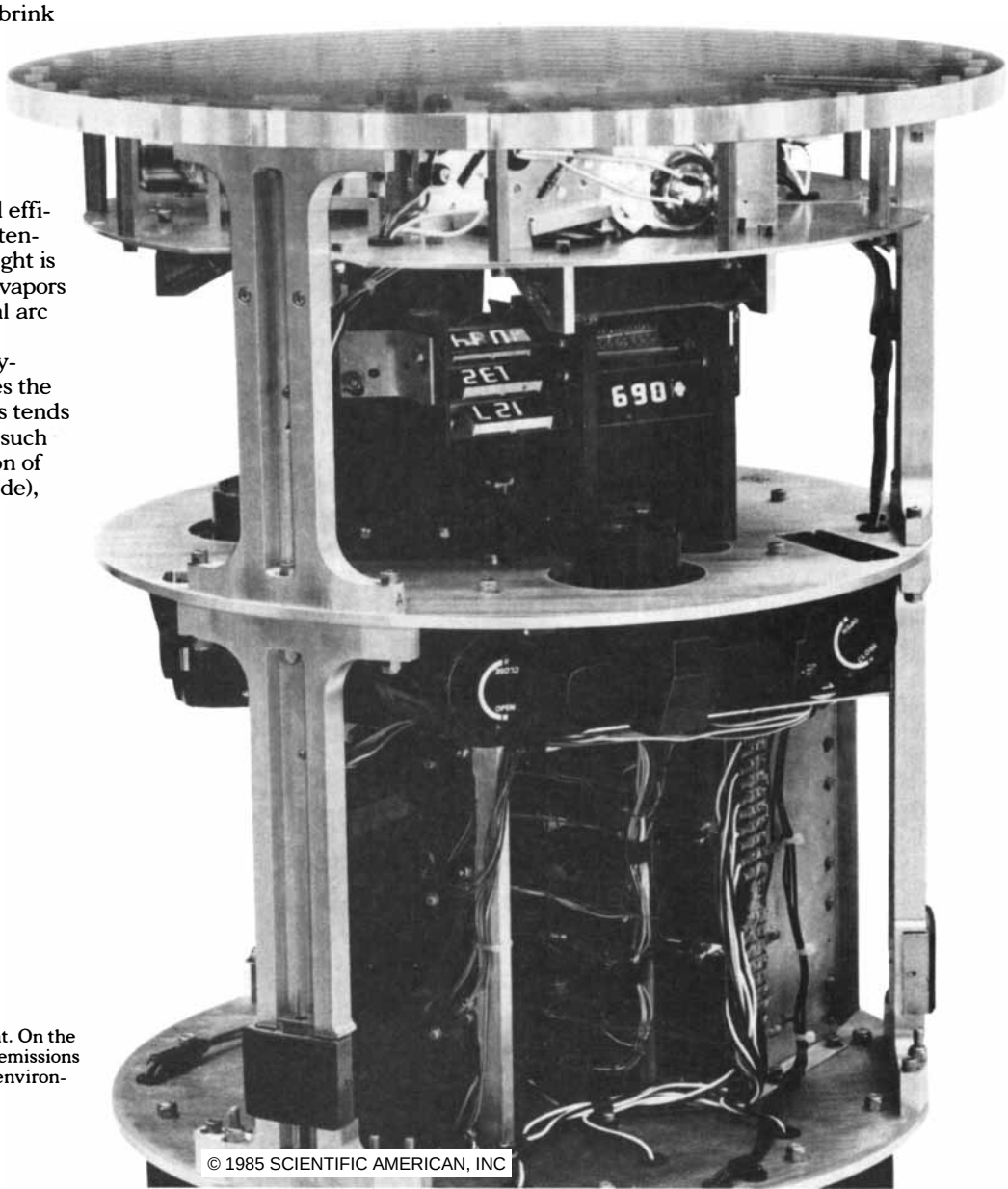
GTE has achieved this goal in a first-of-a-kind experiment aboard the space shuttle. A payload of three metal-halide HID lamps was operated in the microgravity environment of the orbiter. Each lamp was lit for half-hour periods while detailed spectroscopic, light output and electrical measurements were taken.

The results have substantially strengthened the technological underpinning drawn upon for lamp design. GTE scientists now have critical information and new insights that will produce lamps with brighter, whiter light.

Untrapping excited atoms.

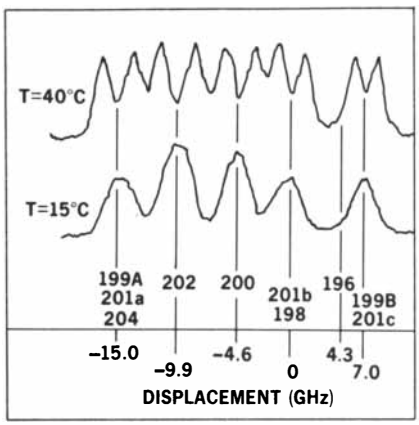
Improvements in fluorescent lamps are on the way, too. As just one example, GTE has discovered how to increase the efficiency of these lamps by about 5%.

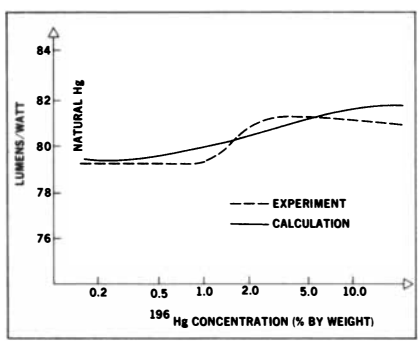
Mercury vapor in the lamp emits



Payload for the GTE space experiment. On the top layer are three HID lamps, whose emissions were investigated in the microgravity environment of the space shuttle.

an old science.





ultraviolet light when it is excited by the electric current. This light is transformed into white when it strikes the phosphors coating the glass tube but some ultraviolet is reabsorbed by the mercury vapor, limiting the lamp's efficiency. GTE researchers have found, however, that by increasing the level of 196 Hg isotope from its naturally-occurring 0.15% to 3.0%, more ultraviolet light escapes to the phosphor. Output improves about 5%.

Light from molecules?

In the future, light may be produced directly from excited molecules in low-pressure lamps. The light spectrum is in broad bands, rather than the narrow-line emission from mercury or sodium atoms.

GTE researchers are investigating ways to produce white light from molecules as the basis for a totally new lamp.

The chemical make-up of the molecules and their behavior in the excited state are undergoing critical studies. In many cases, GTE is applying electrodeless technology with RF power sources as exciters.

This new way of looking at light bulbs promises high-efficiency, long-lived, cool-running light sources with many industrial and residential applications.

The wonderful world of light.

At GTE, we are working on many projects aimed at bringing about the revolution in light. New electrode materials, improved sealants, excimers—these and more are on the GTE research agenda.

The box lists some current papers pertinent to GTE lighting research. For any or all of these, you are invited to write GTE Marketing Services Center, Department TP-L, 70 Empire Drive, West Seneca, NY 14224. Or call 1-800-828-7280 (in N.Y. State 1-800-462-1075).

Pertinent Papers

Convection and Additive Segregation in Metal-Halide Lamp Arcs: Results from a Space Shuttle Experiment
Symposium on Science and Technology of High Temperature Light Sources, Electrochemical Society Meeting, Toronto, 1985

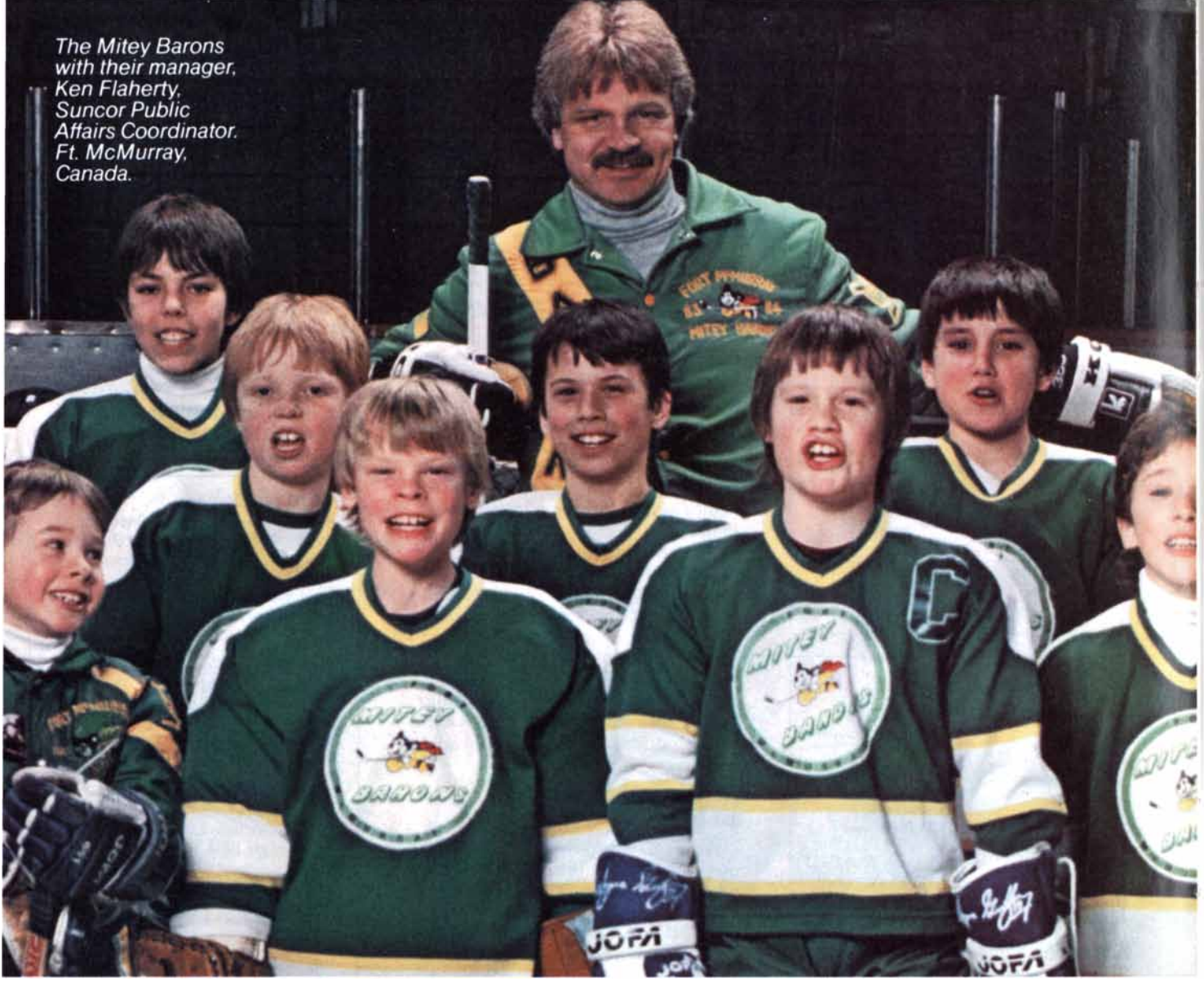
Arc Discharge Convection Studies: A Space Shuttle Experiment
Proceedings of a symposium held at NASA Goddard Space Flight Center, Greenbelt, Maryland, August 1-2, 1984

Energy Conservation Through More Efficient Lighting
Science, Volume 226, pp. 435-436, October 26, 1984

Enhanced HgBr emission at low pressures
Applied Physics Letter 42, May 1, 1983

Bound-free emission in HgBr
Applied Physics Letter 41, November 1, 1982

GTE



The Mitey Barons with their manager, Ken Flaherty, Suncor Public Affairs Coordinator. Ft. McMurray, Canada.

SUN, ENERGY AND THE MITEY BARONS. Twenty years ago Ft. McMurray, Canada, wasn't the town it is today. In fact, it was barely a town at all. That was before Sun Company developed a process to extract oil from the surrounding tar sands. And Sun people like Ken Flaherty can tell you what's developed since then.

"There was a time when this place was mostly tar sands and virtually undeveloped. But Sun brought a lot of energy to Ft. McMurray. We're getting more than fifty thousand barrels of oil a day from these sands. And under the snow around here there's more oil than in Saudi Arabia.

"We've got jobs, a community, and some pretty decent hockey teams. The energy's not just under the snow. It's everywhere."

At Sun we think putting our energy back into a community is just as important as getting it out.

WHERE THERE'S SUN THERE'S ENERGY.

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

AUGUST, 1935: "Flood-control work that has been going forward rapidly on the lower Mississippi River since 1928 has reached such a stage that there is every reason to believe this stupendous project will be completed sometime in the first half of 1936. By then a total of 650,000,000 cubic yards of earth will have been placed in the defensive levees that have been either heightened or built entirely anew to hold within definite bounds the flood waters of the river. This volume of earth is more than twice the amount that had to be excavated in digging the Panama Canal."

"An important biological experiment that may lead to the isolation of the germs that cause colds, influenza and infantile paralysis has been successfully completed at the University of Notre Dame. After six years of work J. A. Reyniers has succeeded in obtaining absolutely germ-free guinea pigs and in raising them without contamination of any kind. The importance of this work lies in the fact that it makes possible a study of any single germ on a living organism."

"You can now acquire a full-size, life-long home simply by ordering from a catalog and telling the dealer where to erect it. When you take possession a month later, the refrigerator will be making ice cubes, the heating plant will be throwing cool, conditioned air through the rooms and, as likely as not, the radio will be playing. This is the pre-fabricated house."

"Police radio was originally limited to one-way operation, that is, transmission from headquarters to police cars. More recently advantages have been seen for two-way operation. Headquarters, thus in touch with any or all members of its mobile unit, can visualize an entire situation and direct maneuvers with full knowledge of how its forces are distributed. The equipment has been developed by the Bell Telephone Laboratories."

"School children of the immediate postwar years were hampered in their studies of geography and the history of Europe because of the many changed boundary lines and new countries. If the present fad for changing well-known place names and names of countries holds, the student will be more than ever bewildered. It has been difficult enough to become familiar with Chosen as the Japanese name for Korea; now Persia insists on use of its ancient name Iran, and Abysinnia wishes to be called Ethiopia. And there is Peiping for Peking, Firenze for Florence and Praha for Prague."

SCIENTIFIC AMERICAN

AUGUST, 1885: "The congress which met in Paris in 1879 to decide on the Panama route and a tide-level canal put the estimated cost at somewhat more than $200,000,000. It is continually becoming apparent that the quantity of excavation to be done has enormously increased. It is to be remembered also that the whole work is not yet surveyed, and the problem of disposing of the waters of the Chagres River is yet to be met. Taking all these items into consideration, and putting off the time of completion at least as far as 1892, the London *Financial News* puts the probable cost of the canal at $530,000,000."

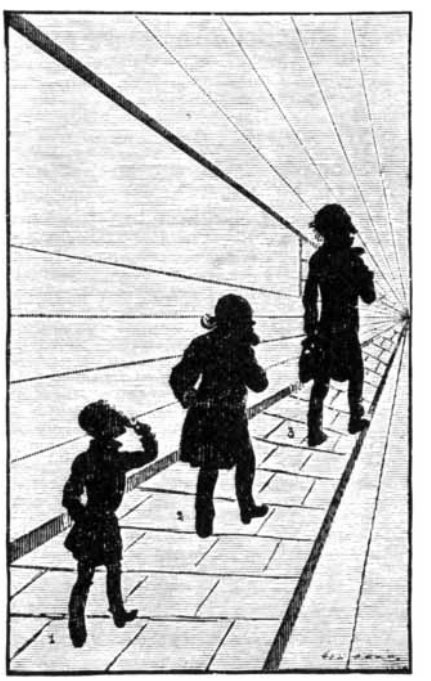
"We publish in this issue charts for great circle sailing. The advantages of great circle sailing have been known for many years, but hitherto the difficulties of calculating and plotting the course have been so great that it never came into more than occasional use. The chart and methods suggested by Prof. Proctor are very simple and may be readily grasped by navigators of even small mathematical knowledge. A great saving of distance is effected by sailing on the arc of a great circle; thus the distance from Melbourne to Cape Town is 6,154 miles on the rhumb course, but is 587 miles shorter on a great circle."

"The *Chronicle* fire tables for 1885 contain a review of the total losses from fire in the United States and Canada during the past ten years, which shows an appalling destruction of values from causes largely preventable. The causes have been classified, and they show an alarming predominance of incendiarism and of defective construction scarcely less criminal."

"The question of burying the telephone, telegraph and electric light wires, instead of having them strung on poles as heretofore, is now attracting a good deal of attention in all our large cities. This goal is being achieved in Paris, the admirable sewerage system affording the ready means."

"The longest bicycle ride ever made has just been completed by H. R. Goodwin of the North Manchester Club. Leaving Land's End on June 1, he journeyed to John o'Groat's, having reached which point in 7½ days he at once turned southward and again arrived at Land's End on the 16th, the double journey of about 1,750 miles, or from one extremity of England to the other, having occupied less than 16 days. From Land's End he rode to London, which was reached on the 19th, the rider thus having completed a journey of 2,050 miles in exactly 19 days, or an average of 108 miles a day. Mr. Goodwin rode a 40-inch 'Facile' safety bicycle and arrived in London in good health."

"Which is the tallest of the three persons figured in the annexed engraving? If we trust to our eyes, we shall certainly say it is No. 3. But if we take a pair of compasses and measure, we shall find that we have been deceived by an optical illusion. It is No. 1 that is the tallest. The explanation of the phenomenon is very simple. Placed in the middle of the well-calculated vanishing lines, the three silhouettes are not in perspective. Our eye is accustomed to see objects diminish in proportion to their distance; seeming to see No. 3 rise, it concludes therefrom that No. 3 is really taller than the two figures in the foreground."



Which figure is the tallest?

THE AUTHORS

WILLIAM EPSTEIN ("A Critical Time for Nuclear Nonproliferation") is a Senior Special Fellow of the United Nations Institute for Training and Research (UNITAR) and an occasional consultant on disarmament to the UN secretary general and the government of Canada. He has been associated with the UN since its founding. For many years he was director of the Disarmament Division, and he represented Canada in the General Assembly from 1978 through 1982. He has also served as technical consultant to the commission that prepared the 1967 Treaty of Tlatelolco, which created a nuclear-free zone in Latin America; as chairman of the group that drafted the 1969 report on chemical and biological weapons, and as a member of the group that prepared the 1980 report on a comprehensive test ban. He represented the UN secretary general at all the negotiations leading to the Nonproliferation Treaty and attended the 1975 and 1980 reviews of the treaty.

STEPHEN W. CARMICHAEL and HANS WINKLER ("The Adrenal Chromaffin Cell") began their intercontinental collaboration on the adrenal medulla during Carmichael's 1979 sabbatical leave in Innsbruck. Carmichael is a consultant in anatomy at the Mayo Clinic and associate professor at the Mayo Medical School, where he teaches gross anatomy, cytology, histology and embryology to medical students and residents. He was graduated in 1967 from Kenyon College and in 1971 from Tulane University with a Ph.D. in anatomy. Winkler is chairman of the department of pharmacology at the University of Innsbruck, where he got his M.D. in 1963. He also has a D.Phil. from the University of Oxford, which he earned in 1968 for work done in the laboratory of Hermann Blaschko. He became professor of pharmacology at Innsbruck in 1975, served as dean of the faculty of medicine from 1977 to 1979 and was made department chairman in 1980.

GERD BINNIG and HEINRICH ROHRER ("The Scanning Tunneling Microscope") are both at the IBM Zurich Research Laboratory. Binnig went to IBM after getting his Ph.D. from the University of Frankfurt in 1978; he has done work in low-temperature physics as well as on the scanning tunneling microscope. His other scientific interests include surface chemistry and electrochemistry. Rohrer received a Ph.D. from the Swiss Federal Institute of Technology in 1960. He did two years of postdoctoral work at Rutgers University, where he studied thermal conductivity, before going to IBM. Rohrer has pursued interests in magnetoresistance and critical phenomena in certain magnetic materials, but he currently devotes all his time to the microscope.

TIMOTHY H. BOYER ("The Classical Vacuum") is professor of physics at the City College of the City University of New York. He was graduated from Yale College with a degree in physics in 1962 and continued his studies at Harvard University, where he got an M.A. in 1963 and a Ph.D. in 1968. He held a postdoctoral appointment at the Center for Theoretical Physics of the University of Maryland at College Park from 1968 to 1970, when he joined the faculty of City College. Boyer writes: "College teaching runs in the family; a great-grandfather, a grandfather and both my parents have been college professors, as are my three brothers."

ERIC BUFFETAUT and RUCHA INGAVAT ("The Mesozoic Vertebrates of Thailand") have collaborated since 1980 in paleontological field work in Thailand. Buffetaut is a paleontologist with the French National Center for Scientific Research (CNRS), working in the Laboratory of Vertebrate Paleontology of the University of Paris. A graduate of the University of Paris, he specializes in the evolution, paleoecology and paleobiogeography of fossil reptiles, notably crocodilians and dinosaurs. He has searched for fossils in France, Canada, Tunisia, Mali, Pakistan, Malaysia and South Korea as well as Thailand. Ingavat is chief of the paleontological section of the Geological Survey of Thailand in Bangkok. She holds a bachelor's degree from Chulalongkorn University in Bangkok and a master's in paleontology from Kyushu University in Japan. Along with her research in fossil vertebrates she also has an interest in foraminifera of the Permian period.

FREDERICK D. SEWARD, PAUL GORENSTEIN and WALLACE H. TUCKER ("Young Supernova Remnants") have worked together for many years in the field of high-energy astronomy. Seward is an astrophysicist at the Smithsonian Astrophysical Observatory in Cambridge, Mass., where he directs the Einstein Observatory Data Bank. He holds a bachelor's degree from Princeton University and a doctorate in nuclear physics from the University of Rochester, which he received in 1958. Since 1976 Seward has been at the Smithsonian Astrophysical Observatory, where he has specialized in X-ray observations of supernova remnants and neutron stars. Gorenstein, an astrophysicist, has been at the Smithsonian Astrophysical Observatory since 1973. A graduate of Cornell University and the Massachusetts Institute of Technology, which granted him a Ph.D. in physics in 1962, he has designed instruments for X-ray astronomy, among them the detectors for the orbiting Einstein Observatory. Tucker divides his time between the Smithsonian Astrophysical Observatory, where he is a senior theoretician, and the University of California at Irvine, where he is a visiting lecturer. He was educated at the University of Oklahoma and the University of California at San Diego, where he got a Ph.D. in physics in 1966. Tucker has written technical and popular books on astrophysics and the history of high-energy astronomy.

UWE RADOK ("The Antarctic Ice") is a senior research associate and a fellow at the Cooperative Institute for Research in Environmental Sciences (CIRES), which is affiliated with the University of Colorado at Boulder and the National Oceanic and Atmospheric Administration. He was educated at the Technical University of Munich and the University of Melbourne, where he earned a Ph.D. in 1953. He remained there and from 1961 to 1977 was chairman of the department of meteorology. He served as president of the International Commission on Snow and Ice from 1975 to 1979; from 1970 to 1983 he was secretary of the International Antarctic Glaciological Project, whose findings are reported in the current article.

JOEL G. KINGSOLVER ("Butterfly Engineering") is assistant professor of population biology at Brown University. He holds bachelor's and master's degrees in zoology from Duke University and the University of Wisconsin at Madison, respectively; he got a Ph.D. in population biology at Stanford University in 1981. He spent several years as a fellow at the Miller Institute for Basic Research in Science at the University of California at Berkeley before joining the faculty at Brown in 1984. The idea behind the current article originated in a course titled "Heat Transfer in Butterflies," which he taught at Stanford together with Robert Moffat of the mechanical engineering department.

COMPUTER RECREATIONS

A computer microscope zooms in for a look at the most complex object in mathematics

by A. K. Dewdney

The Mandelbrot set broods in silent complexity at the center of a vast two-dimensional sheet of numbers called the complex plane. When a certain operation is applied repeatedly to the numbers, the ones outside the set flee to infinity. The numbers inside remain to drift or dance about. Close to the boundary minutely choreographed wanderings mark the onset of the instability. Here is an infinite regress of detail that astonishes us with its variety, its complexity and its strange beauty.

The set is named for Benoit B. Mandelbrot, a research fellow at the IBM Thomas J. Watson Research Center in Yorktown Heights, N.Y. From his work with geometric forms Mandelbrot has developed the field he calls fractal geometry, the mathematical study of forms having a fractional dimension. In particular the boundary of the Mandelbrot set is a fractal, but it is also much more.

With the aid of a relatively simple program a computer can be converted into a kind of microscope for viewing the boundary of the Mandelbrot set. In principle one can zoom in for a closer look at any part of the set at any magnification [*see cover of this issue and illustrations on pages 17–19*]. From a distant vantage the set resembles a squat, wart-covered figure eight lying on its side. The inside of the figure is ominously black. Surrounding it is a halo colored electric white, which gives way to deep blues and blacks in the outer reaches of the plane.

Approaching the Mandelbrot set, one finds that each wart is a tiny figure shaped much like the parent set. Zooming in for a close look at one of the tiny figures, however, opens up an entirely different pattern: a riot of organic-looking tendrils and curlicues sweeps out in whorls and rows. Magnifying a curlicue reveals yet another scene: it is made up of pairs of whorls joined by bridges of filigree. A magnified bridge turns out to have two curlicues sprouting from its center. In the center of this center, so to speak, is a four-way bridge with four more curlicues, and in the center of these curlicues another version of the Mandelbrot set is found.

The magnified version is not quite the same Mandelbrot set. As the zoom continues, such objects seem to reappear, but a closer look always turns up differences. Things go on this way forever, infinitely various and frighteningly lovely.

Here I shall describe two computer programs, both of which explore the effects of iterated operations such as the one that leads to the Mandelbrot set. The first program generated the colored illustrations appearing in this month's column. The program can be adapted to run on personal computers that have the appropriate hardware and software for generating graphics. It will create satisfying images even if one has access only to a monochrome display. The second program is for readers who, like me, need an occasional retreat from infinite complexity to the apparent simplicity of the finite.

The word "complex" as used here has two meanings. The usual meaning is obviously appropriate for describing the Mandelbrot set, but the word has a second and more technical sense. A number is complex when it is made up of two parts, which for historical reasons are called real and imaginary. These terms no longer have any special significance: the two parts of a complex number might as well be called Humpty and Dumpty. Thus $7 + 4i$ is a complex number with real part 7 (Humpty) and imaginary part $4i$ (Dumpty). The italic i next to the 4 shows which part of the complex number is imaginary.

Every complex number can be represented by a point in the plane; the plane of complex numbers is called the complex plane. To find $7 + 4i$ in the complex plane, start at the complex number 0, or $0 + 0i$, and measure seven units east and four units north. The resulting point represents $7 + 4i$. The complex plane is an uncountable infinity of such numbers. Their real parts and their imaginary parts can be either positive or negative and either whole numbers or decimal expansions.

Adding or multiplying two complex numbers is easy. To add $3 - 2i$ and $7 + 4i$, add the parts separately; the sum is $10 + 2i$. Multiplying complex numbers is only slightly more difficult. For example, if the symbol i is treated like the x in high school algebra, the product of $3 - 2i$ and $7 + 4i$ is $21 + 12i - 14i - 8i^2$. At this stage a special property of the symbol i must be brought into play: it happens that i^2 equals -1. Thus the product can be simplified by collecting the real and the imaginary parts: it is $29 - 2i$.

It is now possible to describe the iterative process that generates the Mandelbrot set. Begin with the algebraic expression $z^2 + c$, where z is a complex number that is allowed to vary and c is a certain fixed complex number. Set z initially to be equal to the complex number 0. The square of z is then 0 and the result of adding c to z^2 is just c. Now substitute this result for z in the expression $z^2 + c$. The new sum is $c^2 + c$. Again substitute for z. The next sum is $(c^2 + c)^2 + c$. Continue the process, always making the output of the last step the input for the next one.

Strange things happen when the iterations are carried out for particular values of c. For example, here is what happens when c is $1 + i$:

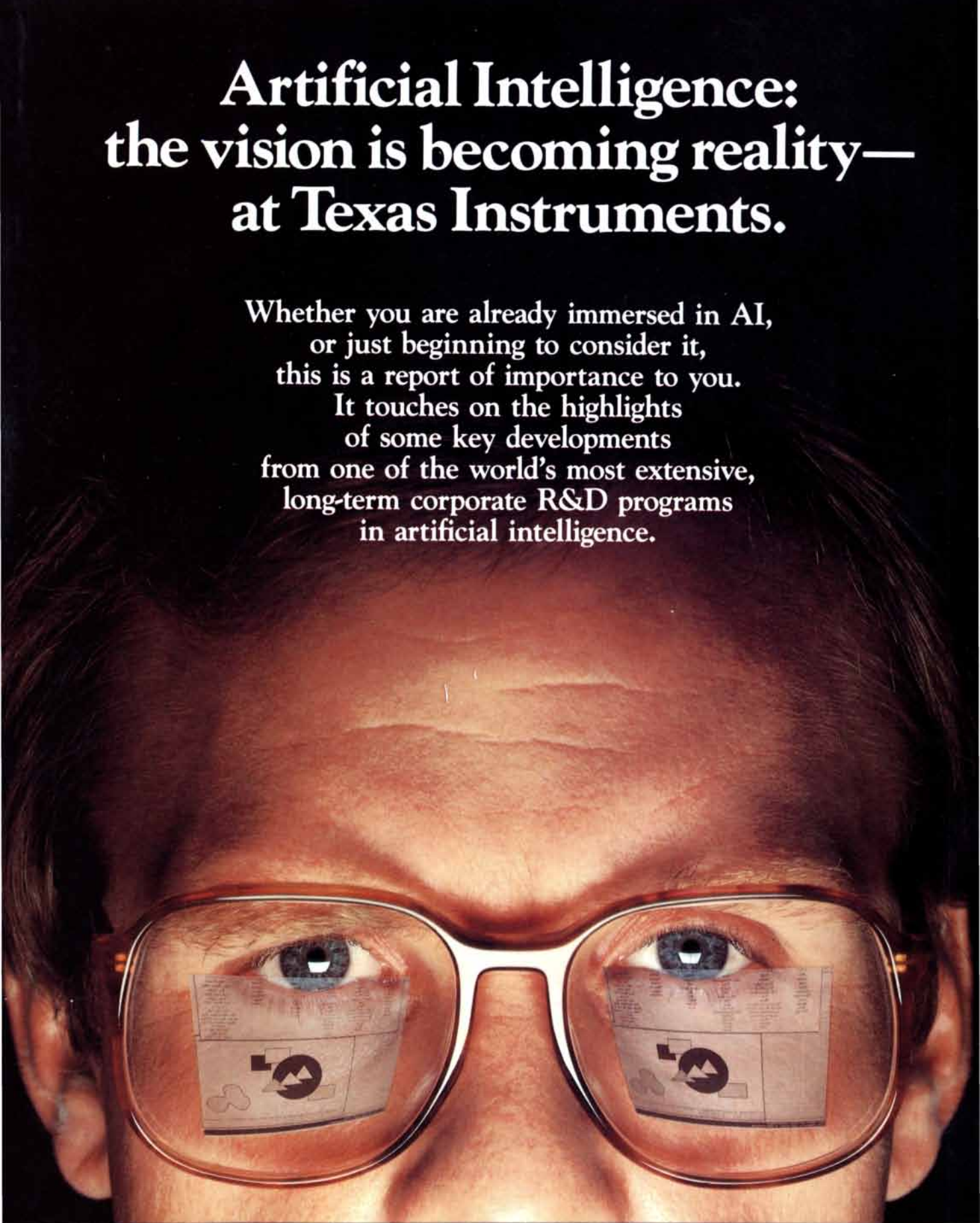
first iteration,	$1 + 3i$
second iteration,	$-7 + 7i$
third iteration,	$1 - 97i$

Note that the real and the imaginary parts may grow, shrink or change sign. If this process of iteration continues, the resulting complex numbers get progressively larger.

What exactly is meant by the size of a complex number? Since complex numbers correspond to points in the plane, ideas of distance apply. The size of a complex number is just its distance from the complex number 0. That distance is the hypotenuse of a right triangle whose sides are the real and the imaginary parts of the complex number. Hence to find the size of the number square each of its parts, add the two squared values and take the square root of the sum. For example, the size of the complex number

Artificial Intelligence: the vision is becoming reality— at Texas Instruments.

Whether you are already immersed in AI,
or just beginning to consider it,
this is a report of importance to you.
It touches on the highlights
of some key developments
from one of the world's most extensive,
long-term corporate R&D programs
in artificial intelligence.



Bringing AI technology to personal computing.

Making technology affordable and accessible to more people is a prime mission at Texas Instruments. In putting AI research to practical use, TI has developed a number of new capabilities for personal computers that are making artificial intelligence applications available to everyone who needs them.

The Personal Consultant™—expert system development tools for personal computers.

The Personal Consultant software package from TI lets you develop and run expert systems on a TI Professional Computer or other personal computer.

Expert systems are computer programs designed to simulate the reasoning processes of human experts in a particular field—in effect, a computerized consultant. It asks a series of questions, and applies rules used by human experts to analyze the answers and make recommendations.

With Personal Consultant software, you can even ask for *help*, inquire *why* you were asked a particular question by the computer, or *how* the system reached its conclusion.

The economies offered by Personal Consultant make adaptation of this advanced technology practical for business and industry today, allowing the expansion of expert systems development and applications to a much greater degree than was affordable up to this time.

Arborist™—powerful decision-analysis software.

Arborist software illustrates the application of AI techniques to solve conventional problems. A combination of LISP processing language, multiple window displays, graphic representation of data, and plain English commands breaks through old barriers to make decision-tree problem solving techniques available to anyone who needs to make better and faster decisions.

Available as a direct result of TI research in artificial intelligence, this new software makes a versatile professional productivity tool easy-to-use and affordable on economical personal computers.

◀ *The Explorer computer system for high-performance symbolic processing.*

Natural language software—helping PCs use plain English.

A leader in the development of natural language interfaces, Texas Instruments developed NaturalLink™ software to allow people to communicate with computers in everyday English sentences, rather than in more traditional computer commands.

Invented in TI's Artificial Intelligence Research Laboratory, NaturalLink packages are available now to provide simplified access to many of today's most popular commercial PC software packages. These include word processing, spreadsheet, database, and graphics programs—programs that impact the majority of personal and business computing needs.

System developers may also build natural language interfaces for their own applications by using the TI NaturalLink Technology Package.

Computers that speak—and respond to the human voice.

Back in 1978, TI brought a major breakthrough in synthesized speech to the fingertips of children in practical, affordable, *talking* learning aids that spawned a whole new generation of electronic products.

With the emergence of personal computers, Texas Instruments has applied advanced speech technologies to PCs with the TI-Speech™ system. This innovation brings together our latest developments in both voice recognition and voice synthesis. The TI-Speech system permits users to verbally enter commands in the computer and provides sophisticated telephone management functions. It makes using a computer easier. And for some hands-busy applications, or for some handicapped individuals, it makes computerization possible.

The Business-Pro™ computer combines advanced personal computing power with capabilities for artificial intelligence applications.

At TI we're using AI today ...for tomorrow.

TI has made an extensive commitment to the application of AI technology for in-house productivity and to solve customer problems. We began our R&D efforts in 1978 and are leading the way in bringing practical, cost-effective AI products to market.

As an established world leader in the research and development of symbolic processing computers, expert systems, and natural language and speech processing, TI is now utilizing the benefits of these efforts. We've discovered new ways to apply AI solutions to improve our own productivity—to speed and improve the development process for advanced VLSI semiconductors, to increase the productivity of automated production systems, to enhance the effectiveness of vital defense systems, to analyze seismic data more accurately for more effective oil exploration, and to develop software that provides better, more timely management information.

The results of these efforts, and our own experience, is bringing thoroughly tested AI products to the marketplace and making them available to an ever increasing number of disciplines.

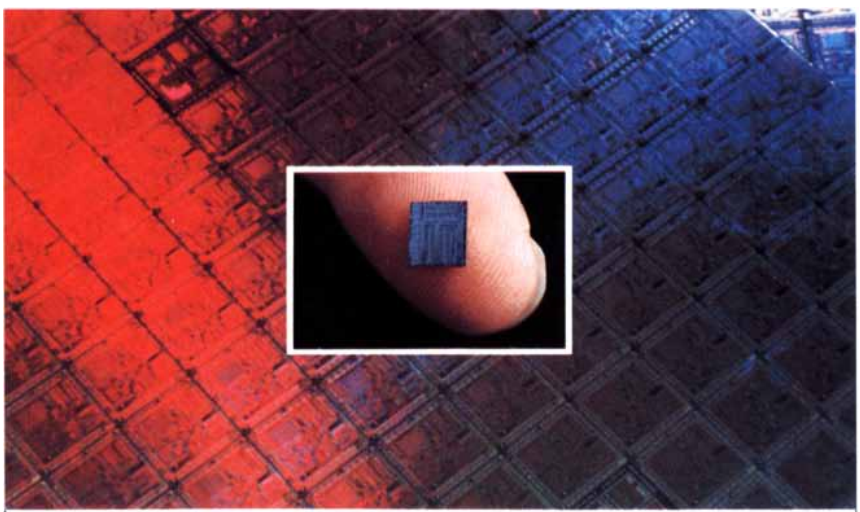
Expert systems, natural language processing, and speech recognition have been applied to fields ranging from genetic engineering to financial management to education. Symbolic processing computers such as Explorer are being used to perform tasks once thought to be solely the province of human experts: planning space missions, diagnosing machinery failures, designing control systems for chemical processing plants.

Texas Instruments dedication to research in artificial intelligence has made significant contributions to this exciting new field. As one of the world's largest technology companies, TI has the resources and commitment necessary to continue its leadership role in the development and production of reliable new products which capitalize on the results of artificial intelligence research.

The next generation of symbolic processing has already begun at TI.

One such program currently underway at Texas Instruments is the development of a LISP processor on a single chip. As part of the Strategic Computing Program of the Defense Advanced Research Projects Agency, this microprocessor is designed to provide up to 10 times the power of symbolic processors previously available.

This new chip will lead the way to many new commercial applications of artificial intelligence. When you link your AI development work to TI, you know you will stay at the leading edge of this fast developing technology—a potentially critical factor in competitive strategies for tomorrow.



A new single-chip LISP processor, being developed by Texas Instruments, will have 10 times the power of previously available symbolic processors.

TI's worldwide support and quality products help you make the most of AI technology.

To make artificial intelligence a practical reality takes more than product innovation. Texas Instruments already has a worldwide support system to provide technical assistance and follow-through. It's a total program of training, documentation, engineering support, and field service designed to provide you with the assistance needed to take maximum advantage of advanced AI technologies.

And for value, TI's exacting design and manufacturing standards enable us to deliver a higher quality product, a more reliable product, with greater functionality and at lower cost than would otherwise be possible.

Explore the promise of AI with TI.

- A proven track record for innovation in new technologies.
- An extensive, long-term R&D effort in artificial intelligence and the resources and commitment to sustain it.
- A leadership role in AI research applied to practical, affordable products—commercially available now—with new products under development.
- Consider TI's one-of-a-kind credentials in artificial intelligence.

Whether you are just beginning to think about this new wave of computing technology or already have a program in place, Texas Instruments can help you lead the way.

For more information.

If you would like to know more about the exciting new developments in AI at TI, write to us at Texas Instruments, P.O. Box 809063, Dept. DEE02, Dallas, TX 75380-9063 and we'll send you additional information.

The next generation of computer technology is coming of age—at Texas Instruments.

Once again, Texas Instruments is pioneering important new technology which promises dramatic increases in human productivity. And, once again, TI brings technology within your reach through practical, problem-solving products at affordable prices.

Artificial intelligence is an emerging set of technologies whose goal is to enable computers to solve problems traditionally thought to require human intelligence or capabilities.

Key to these next generation systems is a significant new approach to computer programming known as symbolic processing. Conventional computer programs require precise sequences of mathematical steps carried out in a prestructured manner. Symbolic processing works with ideas and knowledge rather than numbers, analogous to the way humans reason with knowledge they possess.

Conventional computers experience many shortcomings in areas where symbolic processing systems have proven especially effective: dealing with complex problems, interpreting information, using "rules of thumb" gained by experience, and handling uncertain or incomplete information.

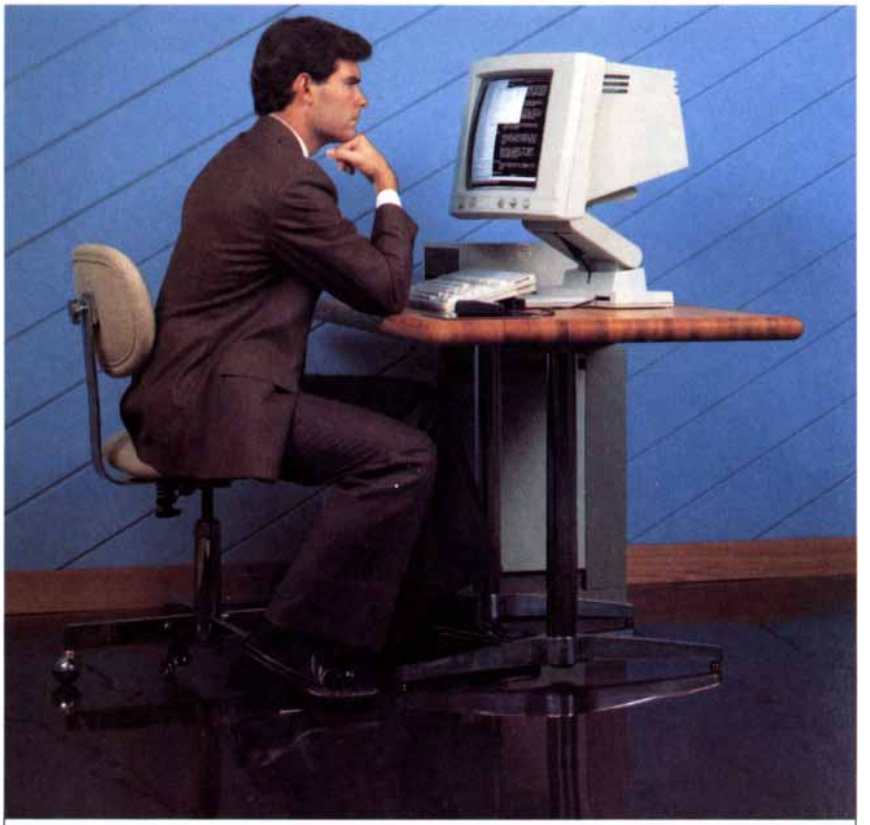
SYMBOLIC vs. ARITHMETIC	
Qualitative	Quantitative
Logical	Numerical
Inferential	Computational

People conceptualize ideas in symbolic terms rather than numbers. Symbolic processing computers manipulate information and knowledge in much the same way.

Symbolic processing, combined with other AI technologies—natural language, speech recognition and synthesis, computer vision—promises to open new dimensions in the way computers serve people. Using AI, computers can now be applied to the broader range of problem solving and decision making that people continually face in the real world.

At TI we're making AI work for you—now.

The transition from basic research to useful products is a commitment at Texas Instruments. We're concentrating not only on fundamental AI research, but on developing and marketing useful products that take full advantage of our experience. Today, TI has a number of products that are a direct result of this commitment.



The Explorer system represents an important advancement in providing exploratory programming and rapid prototyping capabilities for faster development of AI applications.

The TI Explorer™ system. Solving problems beyond the reach of traditional computing.

The Explorer system from Texas Instruments uses a new computer architecture especially designed for AI applications. It executes the language of artificial intelligence, LISP (List Processing). Programmer-friendly ease-of-use features promote rapid productivity for new users.

In contrast to conventional computers, which are designed for numerical data processing, the Explorer system is designed for the efficient processing of symbols and concepts, which represent real-world objects, their properties and relationships, and "chunks" of knowledge.

Until recently, symbolic processing systems were large, expensive computers that required special environments. This limited their use primarily to research laboratories. But, using its advanced semiconductor capability, TI has set a new standard for symbolic processing computers. The Explorer system's compact design, advanced user console, and powerful software packages now make it possible to move artificial intelligence solutions out of the laboratory and into the workplace.

For the company working on AI development, or for the organization just beginning to evaluate it, The Explorer system from Texas Instruments is a "must" consideration.

Discover new dimensions in computing with the Explorer symbolic processing system.

From the company that pioneered the transistor radio, integrated circuit, electronic calculator, integrated-circuit computer, computer-on-a-chip, microprocessor, and synthetic speech.

Once again, Texas Instruments is pioneering an important new technology and bringing it within your reach through a practical, problem-solving product at an affordable price. The Explorer computer system will play a vital role both in extending the boundaries of knowledge and in delivering the products emerging from artificial intelligence research.

A closer look at the Explorer system.

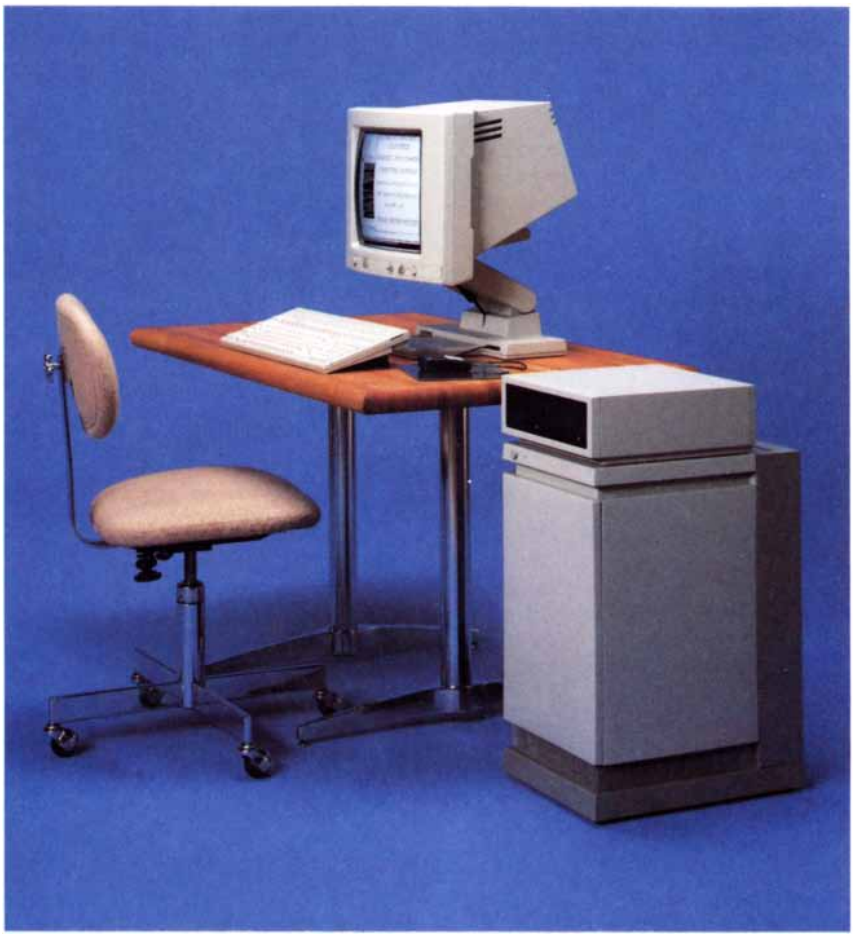
The Explorer system is an advanced computer system designed for high-performance symbolic processing. The Explorer is ideally suited for the development and execution of software that employs artificial intelligence techniques to help solve complex application problems.

The LISP environment.

The system features one of the most productive software development environments available today. The software is based on Common LISP to promote portability and consistency of software applications among different LISP machines. The Explorer also provides high-level extensions to LISP, including Flavors—an object-oriented programming facility.

Unique software.

In addition to the standard LISP machine environment, Texas Instruments provides a number of unique software packages as a standard part of the system software. The Command Interface Toolkit provides standard interfaces to the system. The Suggestions Menu System helps novices rapidly learn the Explorer environment. The Glossary Utility offers online definitions of terms. These packages greatly reduce the time it takes a new user to become productive on the Explorer system.



Advanced hardware.

The Explorer hardware supports high-speed symbolic processing through a number of advanced architectural features. These include a tagged architecture for run-time data typing, bit-field hardware for manipulating complex data structures, hardware assisted memory management (garbage collection), and a 128M byte virtual address space. In addition, physical memory can be expanded to 16M bytes and disk storage can expand to 1120M bytes unformatted (896M bytes formatted).

TEXAS INSTRUMENTS

Creating useful products and services for you.

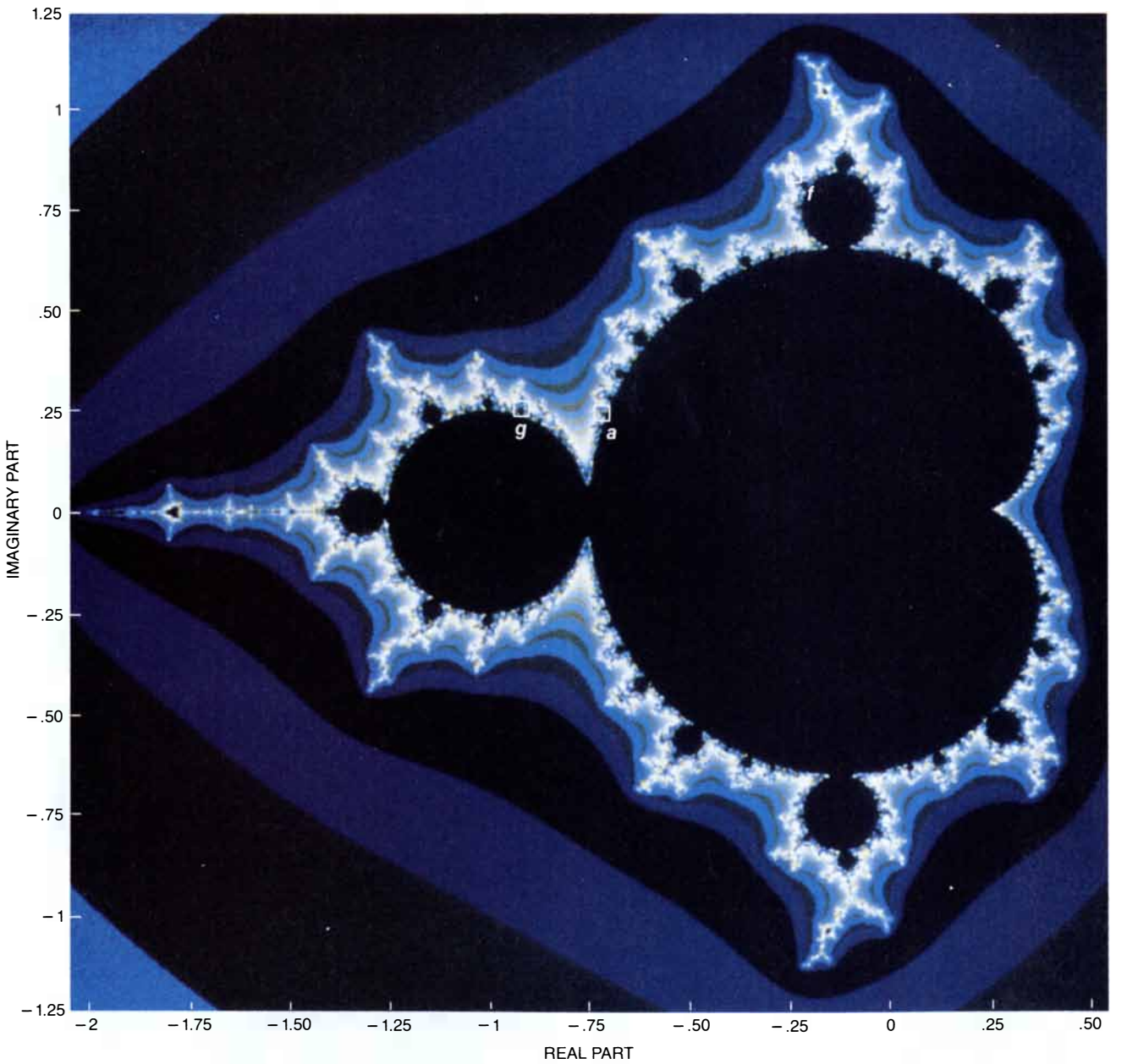
DEAZ040

$7 + 4i$ is the square root of $7^2 + 4^2$, or approximately 8.062. When complex numbers reach a certain size under the iterative process I have just described, they grow very quickly: indeed, after a few more iterations they exceed the capacity of any computer.

Fortunately I can ignore all the complex numbers c that run screaming off to infinity. The Mandelbrot set is the set of all complex numbers c for which the size of $z^2 + c$ is finite even after an indefinitely large number of iterations. The program I am about to describe searches for such numbers. I am indebted in all of this to John H. Hubbard, a mathematician at Cornell University. Hubbard is an authority on the Mandelbrot set, and he was one of the first people to make computer-generated images of it. Most of the images in this article were made by Heinz-Otto Peitgen and his colleagues at the University of Bremen. Peitgen learned the art from Hubbard.

Hubbard's program has inspired a program I call MANDELZOOM. The program sets up an array called *pic,* which is needed for saving pictures. The entries of *pic* are separate picture elements called pixels, which are arranged in a grid pattern. Hubbard's array has 400 columns and 400 rows, and Peitgen's is even larger. Readers who want to adapt MANDELZOOM for personal use must choose an array suited to their equipment and temperament. Larger arrays impose a longer wait for the pictures, but they improve the resolution.

In the first part of MANDELZOOM one may select any square region of the complex plane to be examined. Specify the southwest corner of the square with the complex number to which it corresponds. Two variables in the program, *acorner* and *bcorner,* enable one to enter the real part and the imaginary part of the number respectively. Specify the length of each side of the



The Mandelbrot set and its coordinates in the complex plane. The details shown on the cover and on the next two pages are outlined

square by entering a value for a variable called *side*.

The second part of the program adjusts the array *pic* to match the square of interest by computing the size of a variable called *gap*. *Gap* is the distance within the square between adjacent pixels. To obtain *gap* divide *side* by the number of rows (or columns) in *pic*.

The heart of the program is its third part. Here a search is made for the complex numbers *c* in the Mandelbrot set, and colors are assigned to the numbers that are, in a special sense, nearby. The procedure must be carried out once for every pixel; thus Hubbard's 400-by-400 array requires 160,000 separate computations. Assume the program is currently working on the pixel in row *m* and column *n*; the third part then breaks down into four steps:

1. Calculate one complex number *c* that is assumed to represent the pixel: add $n \times gap$ to *acorner* to obtain the real part *ac* of *c*; add $m \times gap$ to *bcorner* to obtain the imaginary part *bc* of *c*. It is not necessary to include the imaginary number i in the program.

2. Set a complex variable *z* (which has parts *az* and *bz*) equal to $0 + 0i$. Set an integer variable called *count* equal to 0.

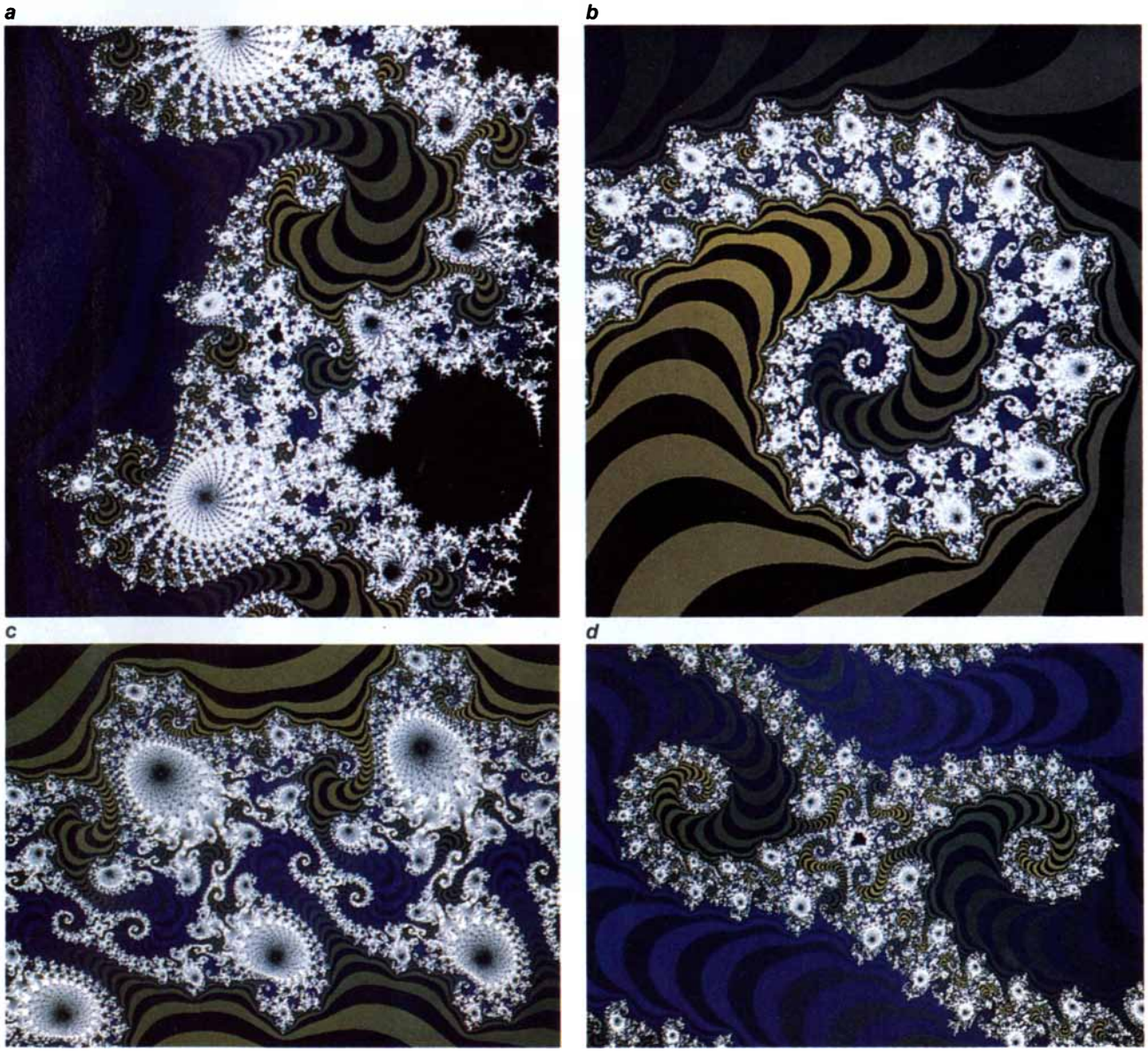
3. Carry out the following three steps repeatedly, until either the size of *z* exceeds 2 or the size of *count* exceeds 1,000, whichever comes first:

$$z \leftarrow z^2 + c$$
$$count \leftarrow count + 1$$
$$size \leftarrow \text{size of } z$$

Why is the number 2 so important? A straightforward result in the theory of complex-number iterations guarantees that the iterations will drive *z* to infinity if and only if at some stage *z* reaches a size of 2 or greater. It turns out that relatively many points with an infinite destiny reach 2 after only a few iterations. Their slower cousins become increasingly rare at higher values of the variable *count*.

4. Assign a color to *pic* (*m*,*n*) according to the value reached by *count* at the end of step 3. Display the color of the corresponding pixel on the screen. Note that the color of a pixel depends on only one complex number within its tiny domain, namely the one at its northeast corner; the behavior of this number then represents the behavior of the entire pixel.

The scheme for assigning colors requires that the range of *count* values attained within the array be grouped into subranges, one subrange for each



Successive enlargements of the "shepherd's crook" in region **a** *of the image on the preceding page*

color. Pixels for which the size of z reaches 2 after only a few iterations are colored red. Pixels for which the size of z reaches 2 after relatively many iterations are colored violet, at the other end of the spectrum. Pixels for which the size of z is less than 2 even after 1,000 iterations are assumed to lie in the Mandelbrot set; they are colored black.

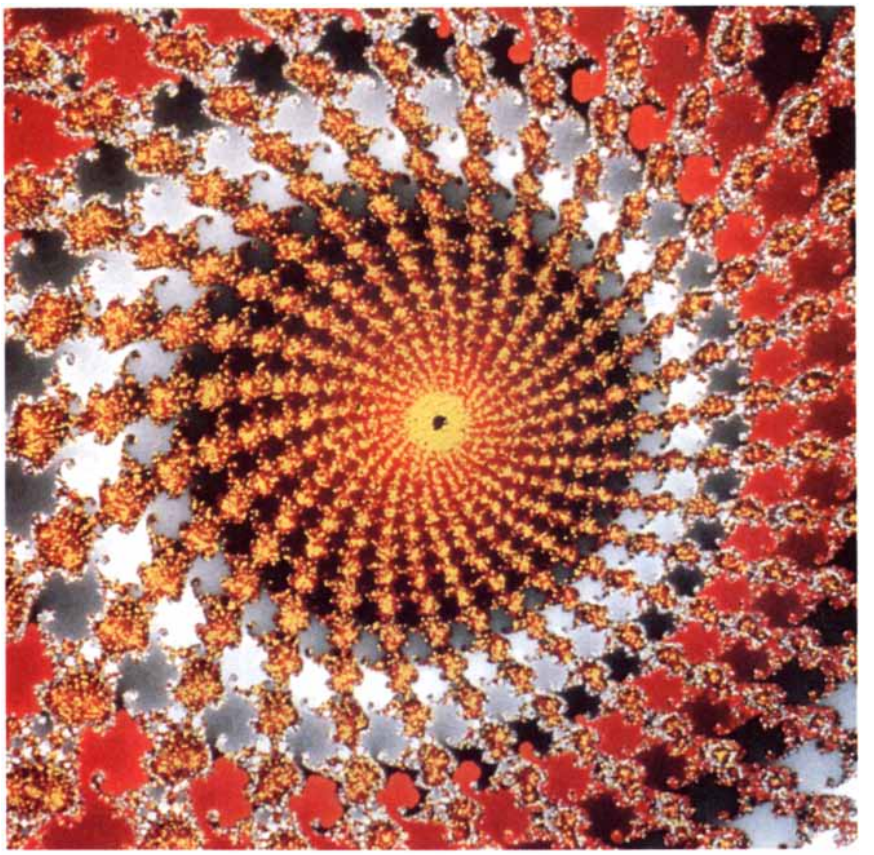
It makes sense to leave the colors unspecified until the range of *count* values in a particular square has been determined. If the range is narrow, the entire color spectrum can then be assigned within that range. Thus Hubbard suggests that in step 4 only the value of *count* be assigned to each array element of *pic*. A separate program can then scan the array, determine the high and low values of *count* and assign the spectrum accordingly. Readers who get this far will certainly find workable schemes.

The reader who does not have a color monitor can still take part in black and white. Complex numbers for which z is larger than 2 after r iterations are colored white. The rest are colored black. Adjust r to taste. To avoid all-night runs the array can be, say, 100 rows by 100 columns. Hubbard also suggests it is perfectly reasonable to reduce the maximum number of iterations per point from 1,000 to 100. The output of such a program is a suggestive, pointillistic image of its colored counterpart [*see illustration on next page*].

How powerful is the "zoom lens" of a personal computer? It depends to some degree on the effective size of the numbers the machine can manipulate. For example, according to Magi (my microcomputer amanuensis at the University of Western Ontario), the IBM PC uses the 8088 microprocessor, a chip manufactured by the Intel Corporation designed to manipulate 16-bit numbers. A facility called double precision makes it possible to increase the length of each number to 32 bits. With such double precision Magi and I calculate that magnifications on the order of 30,000 times can be realized. Higher precision software that in effect strings these numbers together can enhance the numerical precision to hundreds of significant digits. The magnification of the Mandelbrot set theoretically attainable with such precision is far greater than the magnification needed to resolve the nucleus of the atom.

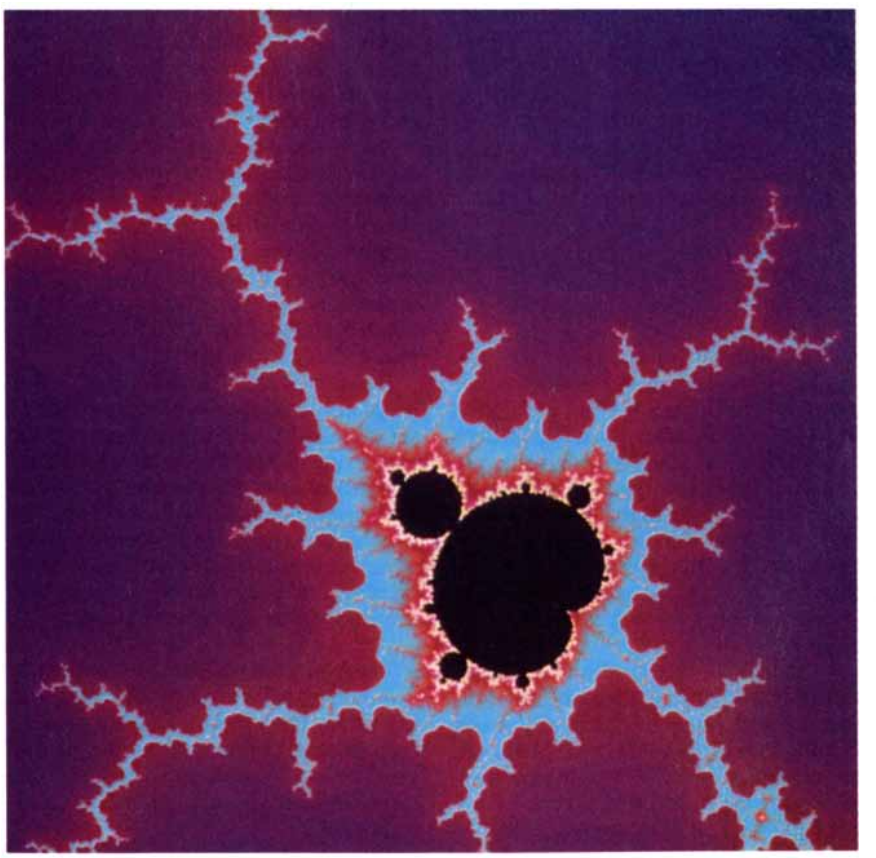
Where should one explore the complex plane? Near the Mandelbrot set, of course, but where precisely? Hubbard says that "there are zillions of beautiful spots." Like a tourist in a

e

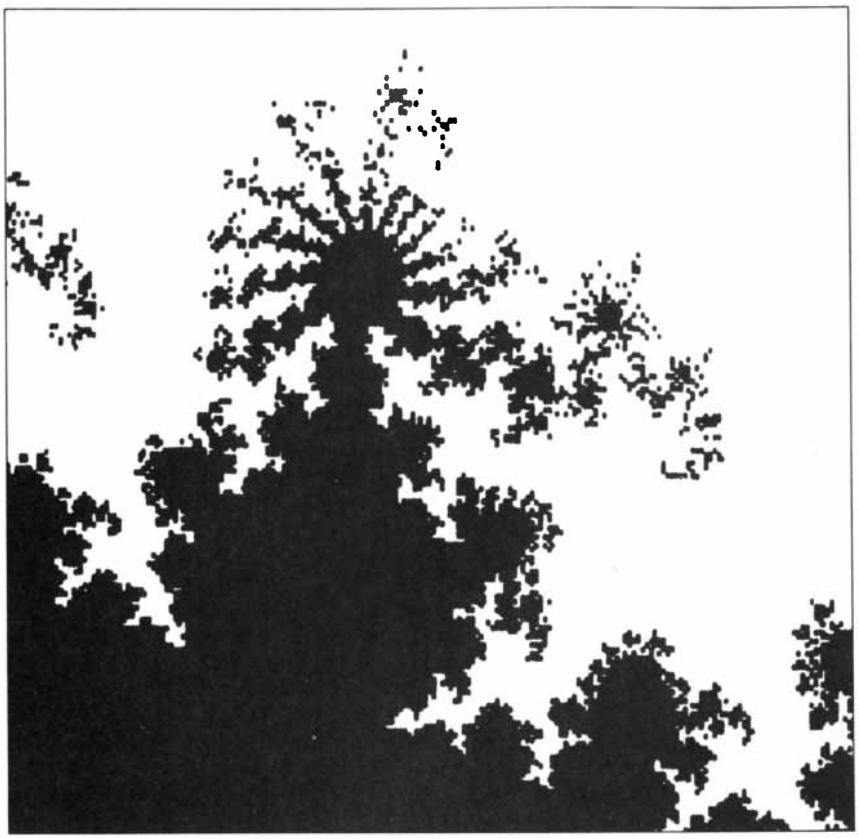


A compound eye peering out from a region of image **d** *on the opposite page*

f



A miniature Mandelbrot in region **f** *on page 17, tethered to the main set by a filament*



Pointillist, miniature Mandelbrot generated by a monochrome monitor

land of infinite beauty, he bubbles with suggestions about places readers may want to explore. They do not have names like Hawaii or Hong Kong: "Try the area with the real part between .26 and .27 and the imaginary part between 0 and .01." He has also suggested two other places:

Real Part	*Imaginary Part*
−.76 to −.74	.01 to .03
−1.26 to −1.24	.01 to .03

The reader who examines the color images accompanying this article should bear in mind that any point having a color other than black does not belong to the Mandelbrot set. Much of the beauty resides in the halo of colors assigned to the fleeing points. Indeed, if one were to view the set in isolation, its image might not be so pleasing: the set is covered all over with filaments and with miniature versions of itself.

In fact none of the miniature Mandelbrots are exact copies of the parent set and none of them are exactly alike. Near the parent set there are even more miniature Mandelbrots, apparently suspended freely in the complex plane. The appearance is deceiving. An amazing theorem proved by Hubbard and a colleague, Adrian Douady of the University of Paris, states that the Mandelbrot set is connected. Hence even the miniature Mandelbrots that seem to be suspended in the plane are attached by filaments to the parent set. The minatures are found almost everywhere near the parent set and they come in all sizes. Every square in the region includes an infinite number of them, of which at most only a few are visible at any given magnification. According to Hubbard, the Mandelbrot set is "the most complicated object in mathematics."

Readers with a simple appetite for more color images of the Mandelbrot set and other mathematical objects can write to Hubbard for a brochure (Department of Mathematics, Cornell University, Ithaca, N.Y. 14853). The brochure includes an order form with which one can buy 16-inch-square color prints that are similar in quality to the Peitgen images shown here.

Confronted with infinite complexity it is comforting to take refuge in the finite. Iterating a squaring process on a finite set of ordinary integers also gives rise to interesting structures. The structures are not geometric but combinatorial.

Pick any number at random from 0 through 99. Square it and extract the last two digits of the result, which must also be a number from 0 through 99. For example, 59^2 is equal to 3,481; the last two digits are 81. Repeat the process and sooner or later you will generate a number you have already encountered. For example, 81 leads to the sequence 61, 21, 41 and 81, and this sequence of four numbers is then repeated indefinitely. It turns out that such loops always arise from iterative processes on finite sets. Indeed, it is easy to see there must be at least one repeated number after 100 operations in a set of 100 numbers; the first repeated number then leads to a loop. There is a beautiful program for detecting the loops that requires almost no memory, but more of this later.

It takes only an hour to diagram the results of the squaring process. Represent each number from 0 through 99 by a separate point on a sheet of paper. If the squaring process leads from one number to a new number, join the corresponding points with an arrow. For example, an arrow should run from point 59 to point 81. The first few connections in the diagram may lead to tangled loops, and so it is a good idea to redraw them from time to time in such a way that no two arrows cross. A nonintersecting iteration diagram is always possible.

One can go even further. Separate subdiagrams often arise, and they can be displayed in a way that highlights some of the symmetries arising from the iterations. For example, the nonintersecting iteration diagram for the squaring process on the integers from 0 through 99 includes six unconnected subdiagrams. The pieces come in identical pairs and each piece is highly symmetrical [*see illustration on opposite page*]. Can the reader explain the symmetry? What would happen if the integers from 0 through 119 were used instead? Is there a relation between the number of unconnected pieces found in the diagram and the largest integer in the sequence?

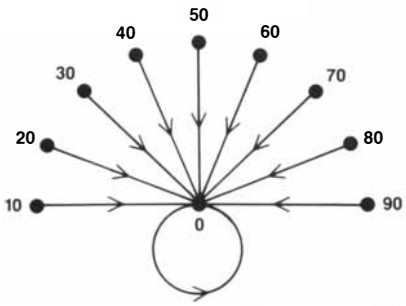
Similar patterns of iteration hold for some of the complex numbers in the Mandelbrot set: for certain values of c repeated iterations of $z^2 + c$ can lead to a finite cycle of complex numbers. For example, the complex number $0 + 1i$ leads to an indefinite oscillation between the two complex numbers $-1 + 1i$ and $0 - 1i$. The cycle may even have only one member. Whether such cycles are found in a finite set or in the infinite Mandelbrot set, they are called attractors.

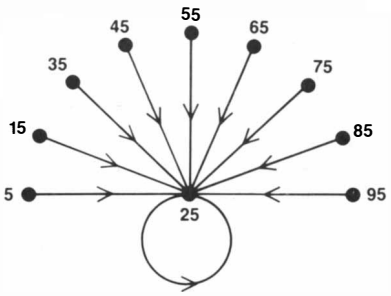
Each of the six parts of the iteration diagram for the integers 0 through 99 includes one attractor. Geometrically the attractor can be represented as a polygon, and the sets of numbers that

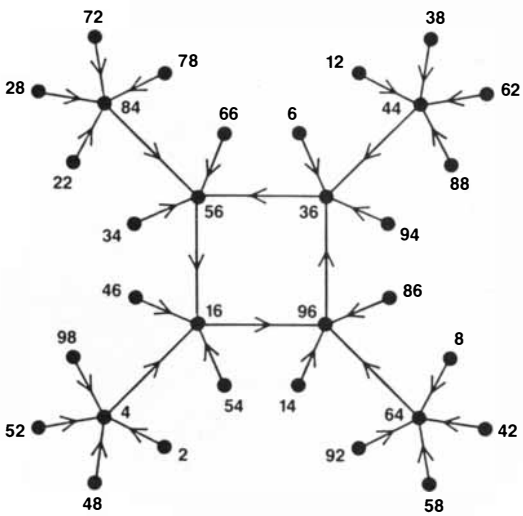
lead into it can be represented as trees.

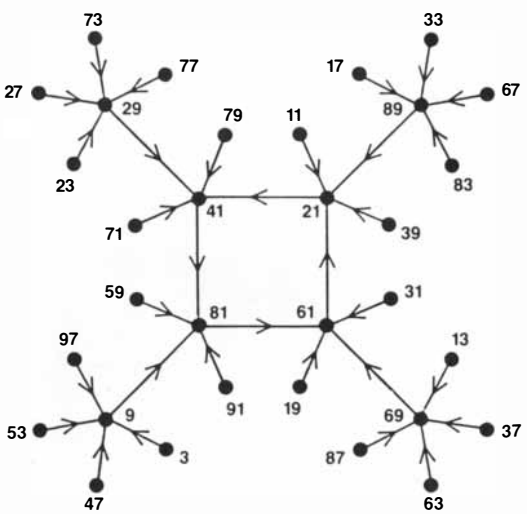
One way to find an attractor by computer is to store each newly generated number in a specially designated array. Compare the new number with all the numbers previously stored in the array. If a match is found, print all the numbers in the array from the matching number to the number just created. The method is straightforward and easy to program. Nevertheless, it can take a long time if the array is large. An attractor cycle within an array that includes n numbers would take on the order of n^2 comparisons to discover: each new number must be compared with up to n numbers in the array.

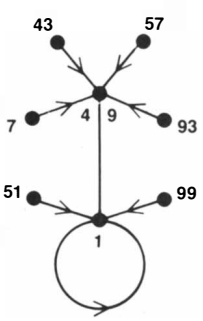
There is a clever little program that will find an attractor much faster. The program requires not n words of memory but only two, and it can be encoded on the simplest of programmable pocket calculators. The program is found in a remarkable book titled *Mathematical Recreations for the Programmable Calculator,* by Dean Hoffman of Auburn University and Lee Mohler of the University of Alabama. Needless to say, many of the topics that are covered in the book can be

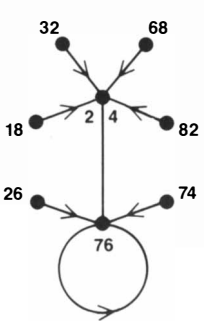












The six components of the iteration diagram for squaring the first 100 integers

AT&T has shattered the information barrier—with a beam of light.

Recently, AT&T Bell Laboratories set the world record for transmission capacity of a lightwave communications system—20 billion pulses of light per second. The equivalent of 300,000 conversations, sent 42 miles, on a hair-thin fiber of super-transparent glass. But that's really getting ahead of the story.

Actually, the 20-gigabit record is only one of a series of AT&T achievements in the technology of lightwave communications.

But what does that record mean?

The Light Solution To A Heavy Problem

All of us face a major problem in this Information Age: too much data and too little information. The 20-gigabit lightwave record means AT&T is helping to solve the problem.

For data to become useful information, it must first be quickly, accurately and securely moved to a data transformer—a computer, for instance. Getting there, however, hasn't always been half the fun.

Metallic pathways have a limited transmission speed, sensitivity to electrical interference and potential for interception—factors that reduce the effectiveness of today's powerful computers. Factors that are eliminated by lightwave communications technology.

Ten Goes Into One 20 Billion Times

Three primary components make up any lightwave communications system. On the transmitting end, a laser or light-emitting diode; on the receiving end, a highly sensitive photodetector; and in the middle, super-transparent glass fibers we call lightguides.

Installing these fibers is a major cost of a lightwave communications system. So, once installed they should stay put—increased capacity should come from fibers carrying more, rather than from more fibers.

Which brings us to the 20-billion bit-per-second story—about experimental technology that has the potential to upgrade installed fiber to meet any foreseeable capacity needs.

Using new, sophisticated lightwave system components, we multiplexed (combined) the outputs from 10 slightly different colored 2-billion bit-per-second laser beams into a single 20-billion bit-per-second data stream.

Playing Both Ends Against The Middle

But, let's start at the beginning—the 10 distributed feedback laser transmitters.

These powerful semiconductor lasers can be grown to produce light of different, but very precise, wavelengths. The lasers we used transmitted in the 1.55 micron (infrared) range, with only minuscule fractions of a micron between their wavelengths. The purity and stability of the beams let us pack their ten colors into the most efficient transmitting region of our single-mode, silica-core fiber.

To make the original 10 beams into one, a fiber from each laser was fed into a new lightwave multiplexer—a prism-like grating that exactly aimed each beam into the single transmission fiber. Over 42 miles later, a second grating fanned the beam back into its original 10 colors for delivery to 10 exceptionally sensitive avalanche photodetectors—receivers that convert the light pulses back into electrical signals and amplify them many times.

20-gigabit multiplexer

A similar avalanche photodetector was the receiver when AT&T Bell Laboratories set the world record for unboosted lightwave transmission—125 miles at 420 million bits per second.

From Sea To Shining Sea

System capacity is important. But system reliability is vital. Especially when the system is going under 10 thousand miles of water—and is expected to last for 25 years.

AT&T is going to build the first lightwave communications system under the Atlantic Ocean. A similar system is planned for the Pacific. In 1988, laser beams traveling through two pairs of glass fibers will carry the equivalent of 37,800 simultaneous conversations overseas, underwater, from the U.S. to Europe and the Far East.

AT&T has manufactured and installed lightwave systems—as large as the 780-mile Northeast Corridor and as small as single-office local area networks—containing enough fiber to stretch to the moon and back. And the capacity of each network is tailored to meet the unique needs of its users.

Systems being installed in 1985 will be able to grow from 6,000 up to 24,000 simultaneous conversations on a single pair of fibers.

AT&T is meeting today's needs with lightwave systems that are growable, flexible and ultra-reliable. And anticipating tomorrow's needs with a whole spectrum of leading-edge lightwave communications technologies.

AT&T

The right choice.

readily adapted to computer programs.

The program is called RHOP because the sequence of numbers that eventually repeats itself resembles a piece of rope with a loop at one end. It also resembles the Greek letter rho (ρ). There are two variables in the program called *slow* and *fast*. Initially both variables are assigned the value of the starting number. The iterative cycle of the program includes just three instructions:

$$fast \leftarrow fast \times fast \pmod{100}$$
$$fast \leftarrow fast \times fast \pmod{100}$$
$$slow \leftarrow slow \times slow \pmod{100}$$

The operation mod 100 extracts the last two digits of the products. Note that the squaring is done twice on the number *fast* but only once on the number *slow*. *Fast* makes its way from the tail to the head of the rho twice as fast as *slow* does. Within the head *fast* catches up with *slow* by the time *slow* has gone partway around. The program exits from its iterative cycle when *fast* is equal to *slow*.

The attractor is identified by reiterating the squaring process for the number currently assigned to *slow*. When that number recurs, halt the program and print the intervening sequence of numbers.

I should be delighted to see readers' diagrams that explore the effects of iterative squaring on finite realms of varying size. The diagrams can be done on a computer or by hand. Discrete iteration is a newly developing mathematical field with applications in computer science, biomathematics, physics and sociology. Theorists might watch for a book on the subject by François Robert of the University of Grenoble.

The two-dimensional beings who inhabit the planet Arde are deeply grateful to the many readers who tried to improve the crossover circuit I described in May. That circuit is made up of 12 two-input *nand*-gates. I asked readers to find the minimum number of *nand*-gates—and *nand*-gates only—from which a crossover circuit can be built. Most of the circuits submitted have 10 gates, a mild improvement, but three readers found an eight-gate crossover [*see illustration on this page*].

In the eight-gate circuit there is one three-input *nand*-gate and two single-input *nand*-gates. The latter act as inverters, converting a 0 signal into a 1 signal and vice versa. The three readers who discovered the eight-gate solution are Eric D. Carlson of Cambridge, Mass., Dale C. Koepp of San Jose, Calif., and Steve Sullivan of Beaverton, Ore. I have passed their names along with the improved crossover circuit to my Ardean friends. Believe it or not, the same crossover circuit appears under U.S. Patent 3,248,573 (April 26, 1966). Robert L. Frank, who is a systems consultant in Birmingham, Mich., wrote that the patent was awarded to Lester M. Spandorfer of Cheltenham, Pa., Albert B. Tonik of Dresher, Pa., and Shimon Even of Cambridge, Mass.

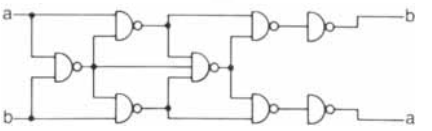
It seems natural to wonder whether the circuit actually appears in any present-day device. It is also natural to wonder whether there is an even smaller *nand* crossover. One supposes not.

C. Walter Johnson of Long Beach, Calif., wrote to me describing a wide variety of planar circuits that incorporate several types of gate. Apparently it is possible to build not only crossover circuits in two dimensions but also planar flip-flops. The flip-flops provide memory for a two-dimensional computer.

One-dimensional computers in the form of cellular automata have been investigated by Stephen Wolfram of the Institute for Advanced Study in Princeton. It is too early to say what contributions readers may have made to this field after reading Wolfram's "Glider Gun Guidelines," but I can pass along some initial reactions. A sin of commission sent a few readers off chasing gliders in the line automaton code-numbered 792. Wolfram and I meant to specify code 357. A sin of omission was my decision not to mention the line automata known to be capable of universal computation. I thought of describing such a line automaton, first constructed by Alvy Ray Smith in 1970. At the time Smith was a graduate student at Stanford University. I was afraid that the description of Smith's universal line automaton would unduly complicate the article: the automaton has 18 states ($k = 18$) and three-cell neighborhoods ($r = 1$).

Arthur L. Rubin of Los Angeles has made a sensible suggestion for defining the speed of light in an arbitrary line automaton. Rubin's suggestion corrects a defect in an earlier definition that sets the speed of light equal to one cell per unit of time. The old definition ignores the possibility that not all automata can attain such speeds. The revised speed of light is "the maximum speed of propagation of any impulse (say to the right)." The leading edge of the impulse is defined by the condition that only 0's can lie to its right. Rubin goes on to prove that the speed of light is 1/3 for the line automaton code-numbered 792.

In my May column I also asked whether the line automaton called



A crossover circuit with eight **nand***-gates*

Ripple has a one-way glider gun. Gliders fired from such a gun would spew out unendingly to the right but never to the left. William B. Lipp of Milford, Conn., has made a simple and charming argument against the existence of such a gun. "Consider a pattern," he writes, "that never has nonzero values to the left of some block labeled 0. Observe that the leftmost nonzero value in the pattern must always be a 1. If it were a 2, the 2 would ripple to the left forever, thus contradicting the assumption that no nonzero entries lie to the left of block 0. But the leftmost 1 must become 0 on the next cycle, moving the left boundary of the pattern at least one block to the right." Thus either a glider ripples to the left or its gun is eaten away by 0's.

Other readers sought to show that Ripple is not capable of universal computation. For some automata one can prove a sufficient condition, namely that the halting problem is decidable. Ripple halts when all its cells contain 0's, but the halting conditions for any universal computing machine it might contain could be quite different.

Several readers attempted constructions of line automata capable of universal computation, among them Frank Adams of East Hartford, Conn., Jonathan Amsterdam of Cambridge, Mass., Kiyoshi Igusa of Brandeis University and Carl Kadie of East Peoria, Ill. The constructions are all straightforward and believable, but Kadie, not content with his one-dimensional automaton, went on to suggest a zero-dimensional one. It would consist of a single cell, and it seems reasonable to call it a point automaton. Readers with a theoretical bent might enjoy pondering the universality of a point automaton. Is it possible?

Alvy Ray Smith published his proof for the existence of a computation-universal line automaton in 1971, in *Journal of the Association for Computing Machinery*. One might have thought a career with such an auspicious start would today be blossoming in some well-known academic institution. Instead it has blossomed in a quite different setting: Smith is director of computer-graphics research for Lucasfilm, Ltd., in San Rafael, Calif. In a future column I hope to report on some of the amazing cinematic effects produced at the Lucasfilm laboratory.

1,310 PROBLEM-SOLVING PAGES IN 4 VOLUMES. EVERYTHING FROM "ARISTOTLE'S WHEEL" TO ALGORITHMS. A $111.40 VALUE FOR ONLY $3.95. A $107.45 SAVINGS!

Yours when you join The Library of Science. You simply agree to buy 3 more books—at handsome discounts—within the next 12 months.

It's that easy, that inexpensive to solve your problem-solving problems. Our Selection Committee assembled four volumes into a set called Practical Tools for Problem-Solving because they all share a common concern: the systematic investigation of the most efficient procedures you can use to solve specific problems. By implementing methods expounded in these books, you can solve virtually any type of abstract problem in math, language and meaning, computer science, physics, etc.

And remember, you get all 4 volumes—a $111.40 value—for only $3.95. A savings of $107.45.

A. Wheels, Life, and Other Mathematical Amusements by Martin Gardner. Join the author in the cosmic fun by testing your wits against mathematical puzzles, paradoxes, and other mathematical recreations. The problems represent such diverse topics as: diophantine analysis, probability paradoxes, and combinatorial theory. Publisher's Price: $16.95.

B. Bypasses by Z. A. Melzak. Subtitled "A Simple Approach to Complexity," this volume offers a new and unique approach to problem-solving and demonstrates its use in a wide variety of scientific and engineering concerns. Publisher's Price: $40.50.

C. How To Solve It, second edition, by George Polya. Any problem can be solved scientifically. But how? The answer lies in the skill of heuristics—defined as the study of how problems are solved. In HOW TO SOLVE IT, mathematical expert G. Polya spells out the methods you can use to solve any problem. Regardless of the subject at hand, Polya's methods can help you quickly arrive at a solution. Publisher's Price: $25.00.

D. Algorithms by Robert Sedgewick. Surveys the most important algorithms and data structures in use on computers today. Covers seven major areas: mathematical algorithms, sorting, searching, string processing, geometric algorithms, graph algorithms, and advanced topics. Publisher's Price: $28.95.

The Library of Science is the nation's oldest, most respected science book club.

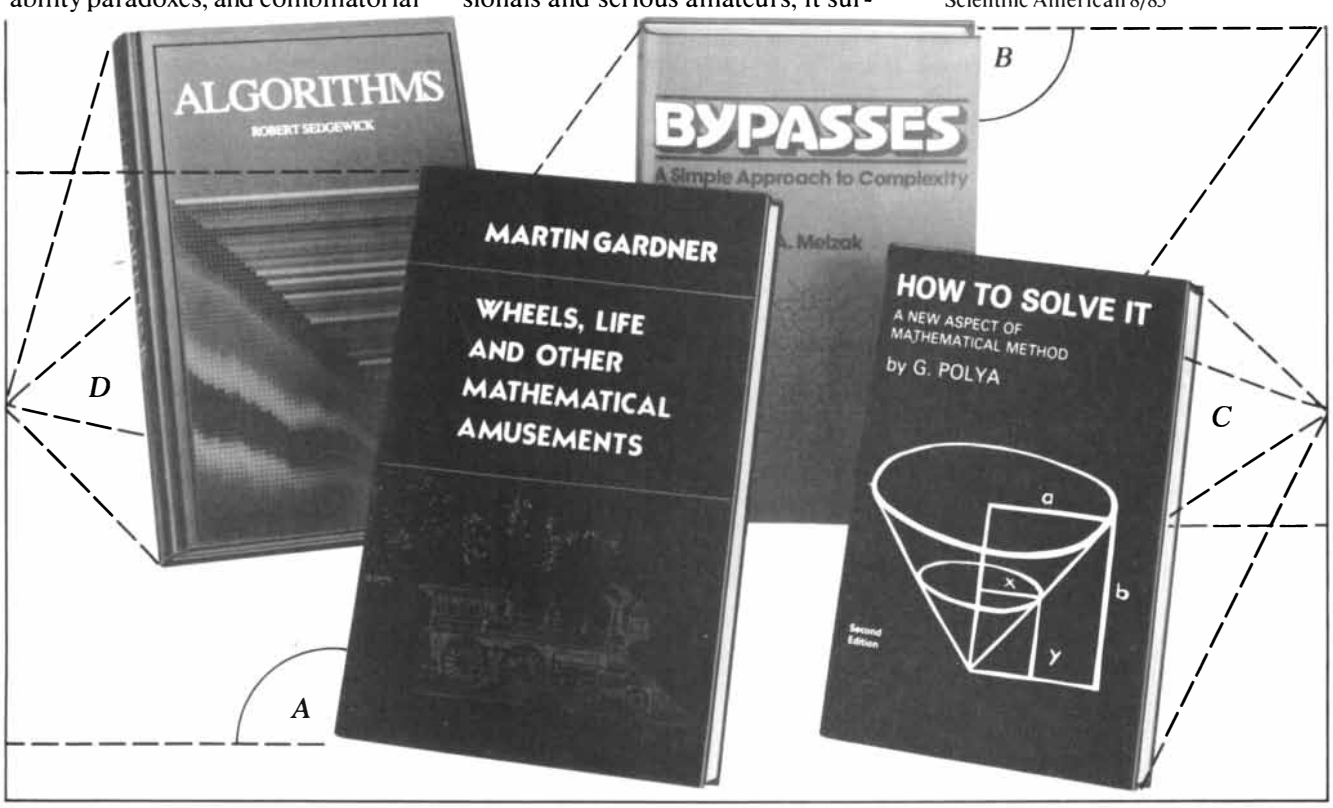
Dedicated to serving both professionals and serious amateurs, it surveys the entire field of published sources on scientific subjects to make the very finest works available to club members at substantial savings.

Membership Benefits:

In addition to getting Practical Tools for Problem-Solving for only $3.95 when you join, you keep saving substantially on the books you buy. Also, you will immediately become eligible to participate in our Bonus Book Plan, with savings of up to 70% off publishers' prices. At 3–4 week intervals (16 times per year), you will receive The Library of Science 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 to: The Library of Science, Dept. 2BF9-00676, Riverside, N.J. 08075 to obtain membership information and an application.

Scientific American 8/85



BOOKS

Downthrust Everest, cod and oil, variable man, women of the world, leaves of ferns

by Philip Morrison

GENERAL BATHYMETRIC CHART OF THE OCEANS: GEBCO (5TH EDITION), by the International Hydrographic Organization, the Intergovernmental Oceanographic Commission and the Canadian Hydrographic Service, Department of Fisheries and Oceans, Ottawa (18 map sheets, boxed with world map sheet and supporting manual, $105; any single sheet, $6). OIL AND WATER: THE STRUGGLE FOR GEORGES BANK, by William H. MacLeish. The Atlantic Monthly Press ($19.95). At the turn of the century His Serene Highness Prince Albert I of Monaco offered to organize and finance for his fellow oceanographers the construction of a series of maps depicting the depths of the world ocean, features properly named, terms defined, the entire work buttressed by text in French and English. His generous offer was accepted and work soon began.

Here is the great-great-grandchild of that pioneer effort, carefully and handsomely executed by the Canadian cartographers with plenty of help from their friends everywhere, area specialists around the world. The floor of the world ocean is fully mapped; all agreed place-names are entered, depth contours are drawn in detail and depth zones are tinted in an attractive gamut of blues. The entire globe is taken in 16 sheets (with overlap strips) projected à la Mercator at a scale of one to 10 million. One big sheet shows the North Atlantic from Halifax to Lagos. Two other meter-square maps present the polar caps. The lands are tinted in a range of browns to show zones of altitude; the important rivers and lakes are noted, but no contours or boundaries and few place-names are to be found ashore. This is not a map for landlubbers, nor are its companion sheets.

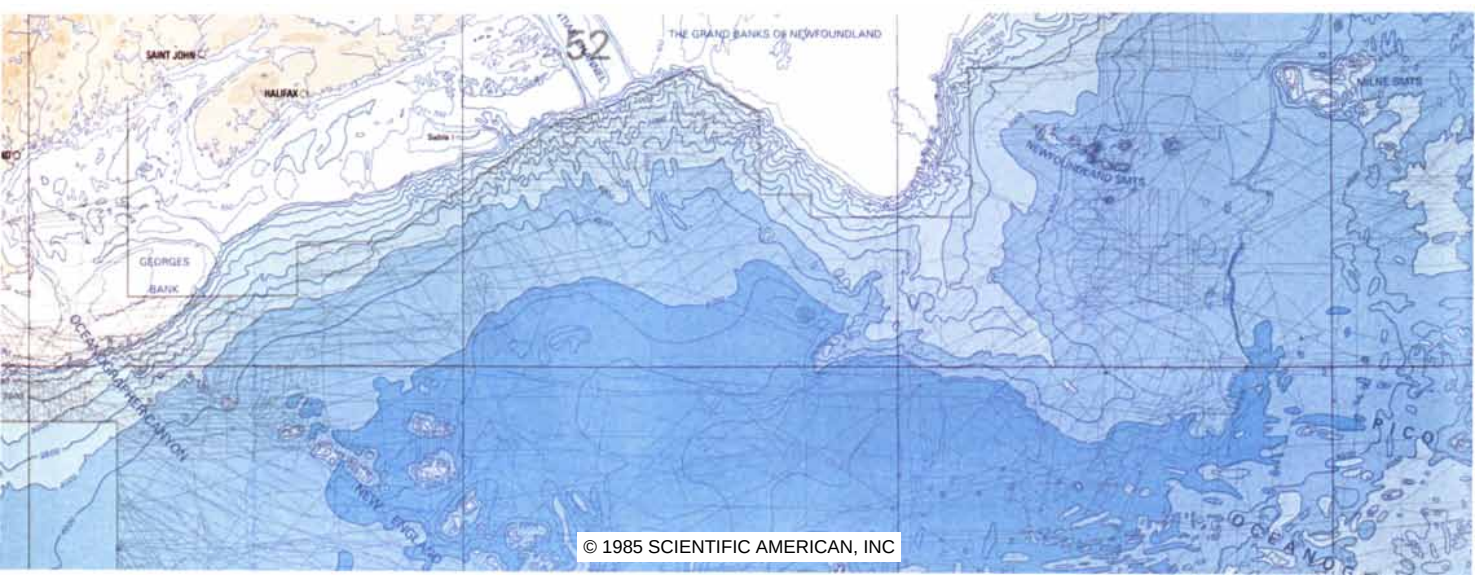
One map has been reduced in scale to show the entire world (it is identified as GEBCO 5.00). It is a striking depiction. The geographic patterns are familiar to few of us. But there they all are, world-girdling midocean ridges marked by the serried offsets of transform faults and fracture zones. Here are chains of seamounts and the lonely islands they sometimes bear, long dotted weld marks burned into the drifting plates by upwelling plumes; the submarine delta cones of the Congo and the Amazon sprawl onto the near ocean floor; the continents themselves assume unfamiliar shapes, their shores here broadened by wide, flat shelves, there cut short by deep trenches. The world sheet will be the map most sought by and most rewarding to general users.

Yet all the detailed sheets repay a browser. On one, near Guam, the ship tracks cross and recross the Challenger Deep, our downthrust Everest, plunging 11 kilometers toward the planet's core. Not so far off is the Yellow Sea, its shoal bottom less than 100 meters underwater, as big and flat as the Dakotas. These earth patterns should one day become as familiar as the Italian boot, a form not much more than a transient accident of present sea level.

The entire set bears a heavy load of documentation; all the random ship tracks used are drawn in lightly, except where marked areas have been mapped in detail. Then documentation is cited for the close-grained soundings. The meticulous scientific standard for the maps is supplemented with a bilingual text and an official register of consultants, so that the map merits its formal authority as the joint publication of the International Hydrographic Organization and the Intergovernmental Oceanographic Commission. All this apparatus is something of a mental and typographic burden, but it transmits a quiet air of real and growing consensus.

Modern navigation fixes ship positions by radio timing systems such as Omega and by signaling clocks borne aloft in orbit. Automatic crystal-timed echo recorders measure depths precisely everywhere the vessel steams. These modern aids have at last given tough substance to Prince Albert's old aspiration, an accurate world bathymetric chart. To be sure, data for the wide deep seas that lie far off the major lanes are sparse to this day. But overall soundings by the million are entered on those 655 internationally maintained master charts that support these more generalized public maps.

The intellectual heirs of Prince Albert have neatly set the stage for William H. MacLeish. On their maps 5.04 and 5.08 the continental shelf of Atlan-



tic North America trends southwesterly, its shoreline stretching 1,100 miles from its most easterly outpost, a circular swell called Flemish Cap, past the Grand Banks of Newfoundland, to a thumb-shaped shoal called Georges Bank, nearly the size of Massachusetts, whose center is 100 miles seaward from the elbow of Cape Cod. There for a couple of centuries seafaring men have eagerly hunted the teeming groundfish. The place was held "one of the greatest sources of wealth then known to the world." Its shallow waters are well mixed, sunlit, nutrient-rich, partly confined by a broad gyre of ocean currents. Not without reason did the Great and General Court of Massachusetts in the year 1784 hang a representation of a codfish high in the room where they sit today.

The fishery is the outcome of the choices of a few hundred captains, each assessing what his boat, his crew, the sea and next week's market might do. Fishermen are all venturers; their lives and fortunes ebb and flow with fish and weather. In 1981 a quite new group of venturers appeared on the Bank. The wealth they sought was rich food still, not for men but for engines. It was oil, held deep below the ocean floor in porous reservoir rocks the geologists knew were buried there in the reef of earlier epochs.

But deep test wells sunk farther to the north revealed no source shales to furnish the hydrocarbons. The prognosis was daunting. Under Georges Bank, however, there was one chance in 20 of finding 100 billion dollars' worth of oil and gas. Within a few decades that gain would match a few centuries of the Bank's yield. The Federal power leased the right to drill under Georges Bank to those corporations prepared to gamble mightily for offshore oil at a heaving green table.

In the summer of 1981 the high towers of huge steel drilling rigs appeared on the Bank. Not without a certain impiety did drill stems stir the bottom silt. Would chronic oil leaks come to threaten fish spawn? Could acute spills of oil one day place at risk the lucrative and venerable precincts of the Sacred Cod? The lawyers gathered, the Woods Hole experts prepared their reports, the Commonwealth of Massachusetts entered its suit against the United States, Greenpeacers in rubber rafts dogged the rigs and worried their Cajun crews.

As the seasons sped by, the husky tool pushers on the drill platforms in all their frustrated skill had to report eight dry holes in a row from all the rigs on the Bank. During the third auction of leases in the fall of 1984, not one oil company bid for the 1,000 submarine tracts along the shelf offered by the Department of the Interior. There was one unwelcome bidder: it was derisive, feisty Greenpeace. Economics and the geology of the Bank have for now ended that venture for petroleum. The rigs and their teams have gone by air and sea; the shallow productive waters remain to draw hunters for scallop and cod as they did when the Bristolmen sailed an ocean to fish there before John Cabot's day.

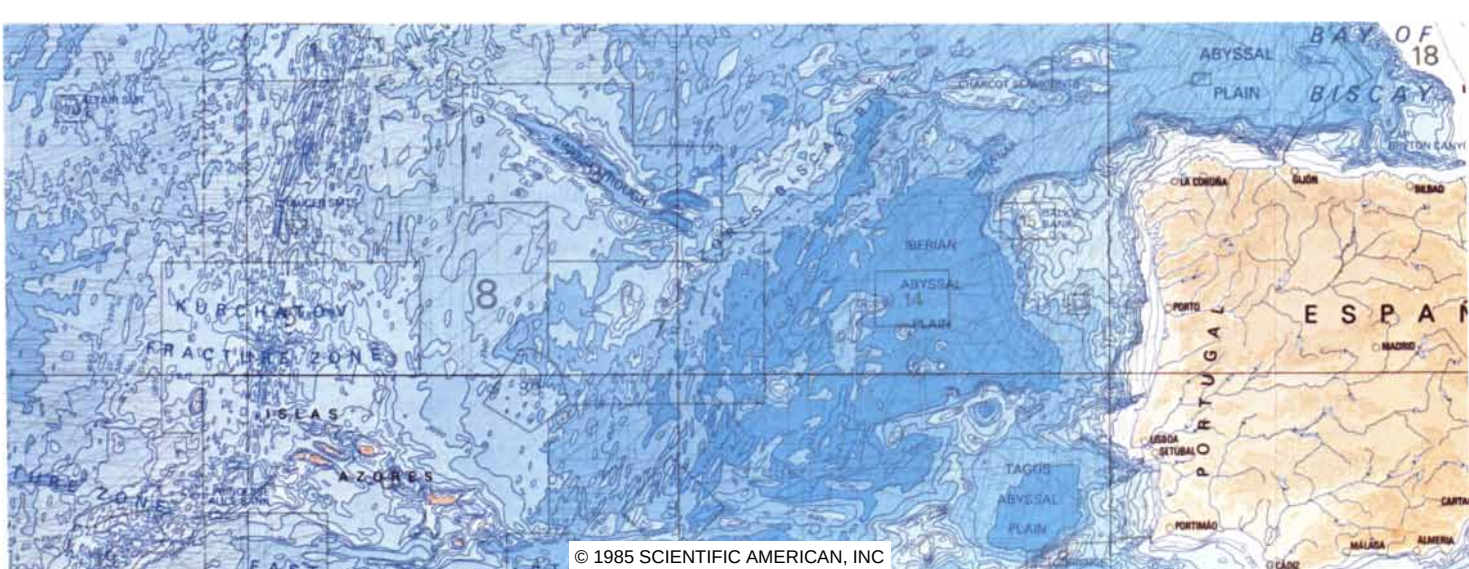
The real events here told seem to adhere almost perfectly to the canons of dramatic art: the conflict arose, blossomed and died well within a decade. Mr. MacLeish has told the story with concern for meaning; the background in oceanography, history, geology and law is clearly and nontechnically put. Yet what is strongest about this intricately textured and colorful story is MacLeish's direct experience, even companionship, with so many of the actors there on the scene, with the bidders in the first Providence lease auction, or the eloquent witnesses from the Gloucester Fishermen's Wives Association. "I will remember longest those I met at sea, on the boats, in the planes and helicopters, on the drilling rig. Theirs is a dignity of dangerous work done well."

The story opens with a patrol to Georges Bank in a slab-sided ancient Grumman amphibian out of Otis Air National Guard Base. A Coast Guard observer remarks on the stern trawler below, after name and number are seen, that the boat is Norwegian-run, out of Fairhaven, across from New Bedford. Not long after, we sail on such a boat, family-owned, its men best-educated among fishermen. They believe in equipment and spend more days at sea than any other group, and they are the most likely to fish for cod and haddock on Georges Bank.

"The cod end... breaches.... The seabirds are like bats from a cave... two tons of fish... in the bag. They are sorted and cut and lie in the bloody water of the washing basket, eyes bulging." By the fifth day the haul is big but the net is torn. "The needles and knives are flashing. An art, I say to Lars." "*Sixth day.* Bitter cold... covered with rime. The radars are icing up..."

Meet the surveyors just up from the Gulf, on their way somewhere else, three men in the wheelhouse of the supply boat, readying to anchor the first rig on the Bank in its legal tract. They talk in meters, "so many hundred thousand meters east of a central meridian and so many million meters north of the equator." Those places are fixed by shore-based timing and by satellite. The crewmen are all employees of one contractor, "the most precise of positioners." "Like so many in the oil patch, they are migrants with a single skill, the world's most sophisticated tinkers," their tools a bank of computers topped by a small screen and lashed down with duct tape.

The scientists from the several labs at Woods Hole are here in variety too, one insightful anthropologist among



the fishermen, as well as geochemists, ecologists and even a voice out of the past: Henry Bryant Bigelow. He was the first head of the Woods Hole Oceanographic Institution, a towering figure in science and in life.

Direct witness marks this entire account. MacLeish reports a surprisingly expressive phrase common now among the usually guarded Yankees, frequent on the VHF radio that is the squawking party line of shoal waters. It signifies approval, whatever is first rate. The book deserves no weaker term: "Finest kind."

WOMEN . . . A WORLD SURVEY, by Ruth Leger Sivard. World Priorities, Box 25140, Washington, D.C. 20007 (paperbound, $5). Lines wiggle across the years, bars stand shoulder to shoulder to be compared, sectored pies are arrayed on the page as on a shelf. All of them appear in attractive colors in telling quantitative support of what this clear and compact brochure concludes: that the real changes in women's status since World War II "have been extremely uneven and, on the whole, modest. Whether in the economy, education, health or government, there is no major field of activity and no country in which women have attained equality with men."

Women today hold up rather more than half the sky. They are entering the world labor force in large numbers, at a rate that about doubles the increase in male workers since 1950. By ILO projections they are now more than a third of the total number of paid workers. Their unpaid contribution to the household economy worldwide adds a largely uncounted additional third to the global G.N.P. Yet they suffer in comparison with men with regard to low pay, part-time work, lack of skills and training as well as increased unemployment. More women than men are poor now, and their numbers are growing. Increasingly women are left with children to support, a trend long familiar in urban centers and lately intensifying in rural regions as opportunity draws men to the shantytowns around mushrooming cities.

It is in the field of education that women can count their greatest recent gains. Equal access to school is widely guaranteed. The UN reports that in 1980 law imposed evenhandedness in 161 of the 194 countries maintaining "autonomous school systems." The rise in the enrollment of girls has been swift: fourfold since 1950 worldwide, perhaps ninefold in the developing countries. The enrollment curves for both girls and boys at primary and secondary levels in some regions, including Latin America, rise together. Africa, the Far East, the Middle East and particularly South Asia still show marked disparity. A world map scores Third World nations on the gender gap in education on a five-step scale; overall, women's adult literacy and school enrollment runs about three-fourths that of men; their share in higher education falls to just above half.

Measures of health follow. Women are genetically programmed to live longer than men; the seven-year gap between the life expectancy of men and that of women recorded in the developed countries for 1985 has widened from five years in 1950. Women there have been slower than men to adopt the "unhealthy habits" of affluence—smoking, alcohol, fast driving and high stress—although perhaps that difference is now eroding. In the Third World life expectancy for women is rising sharply, by 18 years on the average between 1950 and 1985, but it varies markedly from region to region. Mortality during pregnancy and delivery may account for a fourth of the deaths among women of childbearing age in the developing lands; in the U.S. that rate falls to 1 percent. World fertility rates are going down, but in developing lands (China excluded) one woman in six between the ages of 15 and 49 is now pregnant, compared with one among 17 in the developed world. Nutritional anemia is a serious problem for Third World women; where protein is scarce it goes first to the men. More than half of the pregnant women in the Third World bear risks from that form of malnutrition, eightfold the rate reckoned for the developed countries.

The last category treated is that of government and law. Since 1893, when New Zealand first granted women the right to vote, women's suffrage has become all but universal. The map shows a handful of states where the majority of citizens still cannot vote, but only Kuwait is a holdout against suffrage for women in particular. There have been and there are now women first ministers, as we know, but although a few national legislatures in the U.S.S.R. and its neighbors and in Scandinavia reported a fourth or more of their members to be women in 1984, a third of all national decision-making councils have no women members and most of the others have but one woman. Only in the Nordic countries, as well as in Trinidad and Tobago and in Panama, do women make up as much as 20 percent of the cabinet. Power flows only sluggishly under the pressure of participation and the vote. The Lord Mayor of London, Dorothy May Donaldson, the 656th incumbent of that prestigious office, is the first woman to hold it. In the U.S. only 7 percent of the college and university presidents are women, yet 51 percent of the students are female.

The experienced analyst Ruth Leger Sivard, with some help from her friends, has prepared this interesting and timely reference summary. The presentation is fundamentally statistical, but a number of significant highlights together with the colorful graphics add vividness. Half a dozen close-set pages of scrupulously documented data provide ample objective support. This volume is published at the close of the first International Decade of Women, in part to inform a big UN-sponsored conference this summer in Nairobi. Development cannot succeed without new attention to the special role of women. Nor can complacency be justified in our richer lands; there "most women remain highly segregated in low-paid jobs." We need at last to heed the post-Enlightenment insight, to "hold these truths to be self-evident, that all men and women are created equal." So spoke ringingly one of our Founding Mothers, Elizabeth Cady Stanton, in the Seventy-second Year of the Independence of the United States.

VARIATIONS IN HUMAN PHYSIOLOGY, edited by R. M. Case. Manchester University Press, 51 Washington Street, Dover, N.H. 03820 ($21; paperbound, $9). The usual course in human physiology considers the systems of the body one by one, attention focused on one nominal example, a healthy young man who weighs 70 kilograms and lives at sea level. There is little time or room for integration of the organism and still less for the wide variations of real human bodies that must be born and in general grow, that often give birth, that function by design, displaying fascinating differences under a variety of changed circumstances. A great deal is known of such matters, but it is usually communicated by bulky, specialized, intimidating volumes, as hard to come by as they are to enjoy.

This brief and lively text is the work of six authors, all in or around the department of physiology at the University of Manchester. They presume a basic knowledge of the body, perhaps more detailed and technical than most scientific readers will hold, yet rarely too difficult to synthesize adequately from background and context. On that base they have set 10 chapters, each one reviewing a distinct variant from the standard subject. We learn of the adjustments to pregnancy and to fetal and neonatal life as well as to inexorable aging; we follow the body to high ambient pressure and low, to

heat and to cold, to altered diurnal rhythm, to changing gravity, to demanding exercise and even into the shock of trauma. By their very nature these chapters integrate, for it is the organism as a whole that confronts these several conditions, not the standard lung or the nervous system alone. The outcome is an unusually fresh overview for the serious nonspecialist, a kind of widely applied physiology.

Pregnancy and birth nicely illustrate the approach. Engineers rarely design a mechanism capable of reproducing itself by internal fabrication. Yet that is a fair physiological description of the pregnant mother. Dramatic changes are requisite. There is a weight gain of 12.5 kilograms on the average (the range is wide); about eight kilograms of this are on the mother's part, maternal investments in tissue, fat and plasma volume. The products of conception make up the other five kilograms, of which about two-thirds consists of the fetus itself; the amniotic fluid and that novel nurturing organ, the intricate glovelike placenta, account for the rest.

Naturally maternal oxygen uptake increases. There the most important factor appears to be the large increase in tidal volume; the diaphragm is pushed up, the lungs are left more collapsed at expiration. The result is high oxygen levels in the arterial blood. The placenta, acting on behalf of the fetus at once as lung, kidney and gut, is less permeable to oxygen than it is to carbon dioxide. The fetus is constrained to obtain its oxygen through a high gradient. Within the fetus the arterial oxygen tension falls as low as that in a climber high on Everest. But fetal hemoglobin, a somewhat different protein from the one we all hold, has a higher oxygen affinity. These engineering changes run deep.

Adaptive design is subtler still. We have no direct measurements for human infants at high altitude, but the arterial oxygen in the lamb fetus is no lower than it is at sea level. This fits the observation that the red cells in the newborn mountain infant do not differ from those of sea-level babies. Fetal oxygen levels have not dropped, although the high altitude has certainly lowered the mother's oxygen flow; the placenta at high altitude must diffuse oxygen more easily than the placenta at sea level does. Again we have a clue to design: the average birth weight at altitude is less than it is at sea level, yet the placenta is of similar size. The system, to cope with thin air, has altered its pattern of investment.

In adjustment to altitude the newborn highlander, the adult highlander and the visitor have acclimatized to

low pressure; they have not adapted in the genetic sense. The oxygen cascade through the four main steps from inspired air to capillary diffusion proceeds at a lower gradient by quantitative modifications of all the barriers: more red cells, increased ventilation and so on. The same old mechanism is being operated differently within its range of settings. In the llama, in contrast, the oxygen cascade is simply more efficient, closer to the inbuilt fetal scheme than it is to the mountaineer's. We have all endured the oxygen deprivation experienced by those who climb Everest unaided; it was a legacy that saw us through.

Just as it does not experience respiration, so too the fetus does not know heat loss; the mother regulates the temperature of the bathing amniotic fluid until delivery into the cold, cruel world exposes the infant to a sudden drop in temperature that ranges from eight to 10 degrees Celsius. At that moment heat loss and gain become personal, lifelong responsibilities, cued by core temperature. For major heat loss there is a supply-side remedy, the generation of metabolic heat by oxidizing body fuels. For major heat gain there is a remedy more on the demand side: sweating.

Humans can produce more sweat per day per unit area than any other animal. Evaporation sheds body heat at rates up to 20 times the power of basal metabolism, providing that ambient humidity allows the process. Each sweat gland feeds through a remarkable tiny coil on its way to the final duct that goes straight to the surface. Within that tortuous channel the dissolved salt is reabsorbed as the fluid moves outward. When the rate of secretion is high, there is no time for full reabsorption; consequently salt loss can rise fifteenfold. Most people can tolerate a water deficit of 3 percent or so before fatigue begins, but salt loss is also exhausting. The extracellular volume contracts under the loss of osmotic tension, and cramps may follow.

Shivering, an involuntary oscillation in muscle-fiber tension, burns ATP during that muscular motion, and so generates heat. It can increase internal heat supply as much as fivefold over the short term. A newborn lacks the mechanism, perhaps because the reaction might lead to heat loss through the relative movement of air over the infant's large surface area. It is also possible that shivering would compromise temporarily the fine muscle control needed for suckling. Instead infants have under the skin a thin shawl of brown fatty tissue wrapped around shoulders and elsewhere. That shawl may amount to several percent of total body weight. The brown tissue stores fuel; it is yellow when its small vacuoles are fat-filled but turns brown as they empty because it contains so many mitochondria, chock-full of pigmented cytochrome enzymes. Brown fat enjoys a rich supply of blood. In contrast ordinary fatty tissue has only a meager blood supply; its large droplets of cellular fat cannot be mobilized rapidly. Hormonal control initiates the enzymatic generation of this provident heat; the biochemistry is complicated. (Quite generally physiological emphasis is at a scale above the molecular.)

Given a flow of oxygen, ordinary fat is the chief requisite for ATP supply, particularly in muscles relied on for long-term activity, such as those of the heart. The marathoner draws on the fat depots, setting fatty acids free to circulate in the blood, there to be oxidized to supply energy for ATP synthesis. Independent of the complicated paths of oxygen flow, supplementary energy is also derived from liver glycogen, released as blood glucose. That second source of ATP is a limited ready reserve intended for major peak effort; it is "grossly uneconomic" of fuel intake.

Such lore is all but street knowledge in our exercising times; it is clear that it is steady long-distance jogging, not

sudden sprints, that offers a way to burn fat. Sprints, like all bursts of muscular energy, use up mainly carbohydrate, the short-term reserve. Chemical energy is notoriously compact, however; for those who would slim, prevention of intake by the pound is much swifter and easier than cure by the oxidized ounce.

Only a single theme or two running through a few of the chapters of this interesting volume have been traced. The text sets out many more; for readers who would explore beyond lucid summary, medical monographs and topical research reviews are referenced at the end of each chapter. There are no footnotes; those attractive if taxing references are generally not annotated. Just the same, plenty is here, compact, timely and diverse, at a bargain price in muscular current dollars.

THE FERN HERBAL, INCLUDING THE FERNS, THE HORSETAILS AND THE CLUB MOSSES, written and illustrated by Elfriede Abbe. Cornell University Press ($35). The mysterious ferns, their delicate green fronds thrifty amidst woodland shadows, "bear neither flower nor seed." The hidden, lichenlike tiny heart-shaped prothallium grows from a spore to bear sperm and egg that, given an essential drop of water, fuse to send up in time another spore-bearing frond. But the real process of the alternation of generations was unknown before modern science.

Ferns took on the air of magic. No flowers to be found, the fronds of the coarsest and largest species were called male ferns; more delicate species were called female, lady ferns. The thickest and tallest ferns of limestone areas bore seed that, being invisible, itself could confer invisibility: when hunted prey went to ground in such a place, it must have been that fern seed fell on the vanished animal. The four-inch moonwort, its leaflets formed like half-moons, was dear to the alchemists: moonwort, "if it be put into the keyhole," has the "power to break lockes" and even to "draw off the shoes of a horse that passeth over it." Sir Thomas Browne was skeptical: the "magicall conceit seemes unto me," he wrote, "to have no deeper root in reason" than the shape. The origin of that belief is not yet clear, even considering magical theory.

Ferns share real mysteries enough. Their ancient lineage, much older than the Atlantic Ocean, has meant that the selfsame plant can often be found in both the Old World and the New. The luxuriant bracken fern is rich in potassium; its ash long supplied lye for soap and bleach; it also entered general use in glass manufacture, where it produced a more fluid melt than does the soda ash from the seashore plants first used by the glassmakers of antiquity. "And yet is glass not like ashes of fern," says *The Canterbury Tales.* The scouring rush, a fern relative, its rough, ribbed, unbranched stalk four feet high, is so rich in silica as to have become the abrasive of choice over a wide range of old crafts, for smoothing arrow shafts and bone combs, for polishing pewter and for scrubbing kitchen floors. Scouring rush grew profusely along freshwater margins in Holland and was there bundled and traded widely abroad. Club-moss spores make an extremely fine flammable and nonwetting powder. They have long been used in fireworks and for coating pills. The material is called vegetable sulfur in *The Merck Index.*

These are minor if unexpected attributes of living forms that are above all beautiful. Thoreau said, "Nature made ferns for pure leaves, to see what she could do in that line." The artist-author has represented a couple of dozen important fern species in full-page wood engravings, in three or four colors as required. That medium exactly suits these plants, evoked all the more beautifully when a single black stem, say, cuts across the flattened lace forms and subtly varied greens of true maidenhair.

BREAKTHROUGH:
COMPUTER GRAPHICS THAT CREATE MODEL PATIENTS FOR SURGEONS.

We're using computers to give reconstructive surgeons a startling new perspective on their work—three-dimensional images that depict the patient's face and skull from any angle, inside or out.

These 3D images show key relationships between bone and soft tissue. They help surgeons predetermine precisely how bone and skin grafts should be placed to achieve the desired results. They can be used to create physical models that let a surgeon perform a trial procedure or envision the results of his work before surgery begins.

The McDonnell Douglas Information Systems breakthrough in computer modeling* helps surgeons correct birth defects and undo the damage of disfiguring accidents.

We're creating breakthroughs not only in health care and information handling but also in communications and equipment leasing.

We're McDonnell Douglas.

**In collaboration with the Mallinckrodt Institute of Radiology and the Division of Plastic Surgery at Washington University Medical Center, St. Louis.*

MCDONNELL DOUGLAS

Established 1845 SCIENTIFIC AMERICAN August 1985 Volume 253 Number 2

A Critical Time for Nuclear Nonproliferation

This month the parties to the nuclear nonproliferation treaty meet again to debate how well the agreement is working. The continuing arms buildup by the nuclear powers puts the treaty's future in doubt

by William Epstein

Every five years since the Treaty on the Nonproliferation of Nuclear Weapons came into force in 1970 the parties have convened to discuss how well the agreement is achieving its purpose. This month the signatories will gather in Geneva for the third time, at what will probably prove to be the most crucial meeting in the series. It is already clear that an overwhelming majority of the nonnuclear nations believe the nuclear powers, and in particular the U.S. and the U.S.S.R., have not made serious efforts to achieve arms-control agreements or reduce the size of their strategic forces and have thus failed to hold up their end of the bargain. Indeed, this conference is almost certain to be the first one confronted with a total absence of progress on nuclear arms control since the preceding review. The continued credibility—even the viability—of the treaty is therefore put in question.

The treaty (usually called the NPT) was designed to prevent both the horizontal and the vertical proliferation of nuclear weapons. Horizontal proliferation is the spread of such weapons to nonnuclear states; vertical proliferation is the further development, production and deployment of nuclear weapons by the nuclear powers.

The NPT is the cornerstone of a "nonproliferation regime" that the signatory nations have built up over some 30 years. In addition to the treaty the regime has four other main components: the International Atomic Energy Agency (IAEA), the partial-test-ban treaty of 1963, the Treaty of Tlatelolco of 1967 (creating a Latin-American zone free of nuclear weapons) and the Nuclear Suppliers Group.

The IAEA was established in 1957 to promote the peaceful use of atomic energy. The nuclear nations agreed to assist nonnuclear nations in the development of the peaceful uses of nuclear energy on the condition that the nuclear material or equipment provided not be used for the manufacture of nuclear weapons. Hence the statute of the agency constituted a sort of nonproliferation bargain between the nuclear and the nonnuclear countries. At the time the U.S., the U.S.S.R. and the U.K. were the only nuclear powers. France set off its first nuclear explosion in 1960, China its first in 1964. In 1974 India exploded a "peaceful nuclear device," which incorporates the same technology as a nuclear bomb. All six countries are members of the IAEA, together with 106 others that have not exploded nuclear devices.

The partial-test-ban treaty prohibits tests of nuclear weapons in the atmosphere, under water and in outer space. It permits underground tests that do not create any radioactive debris outside the boundaries of the testing state. The treaty also declares that the parties seek to end "all test explosions of nuclear weapons for all time" and will continue negotiations to that end. The treaty was concluded by the three original nuclear powers and now has 112 signatories. China and France are not among them.

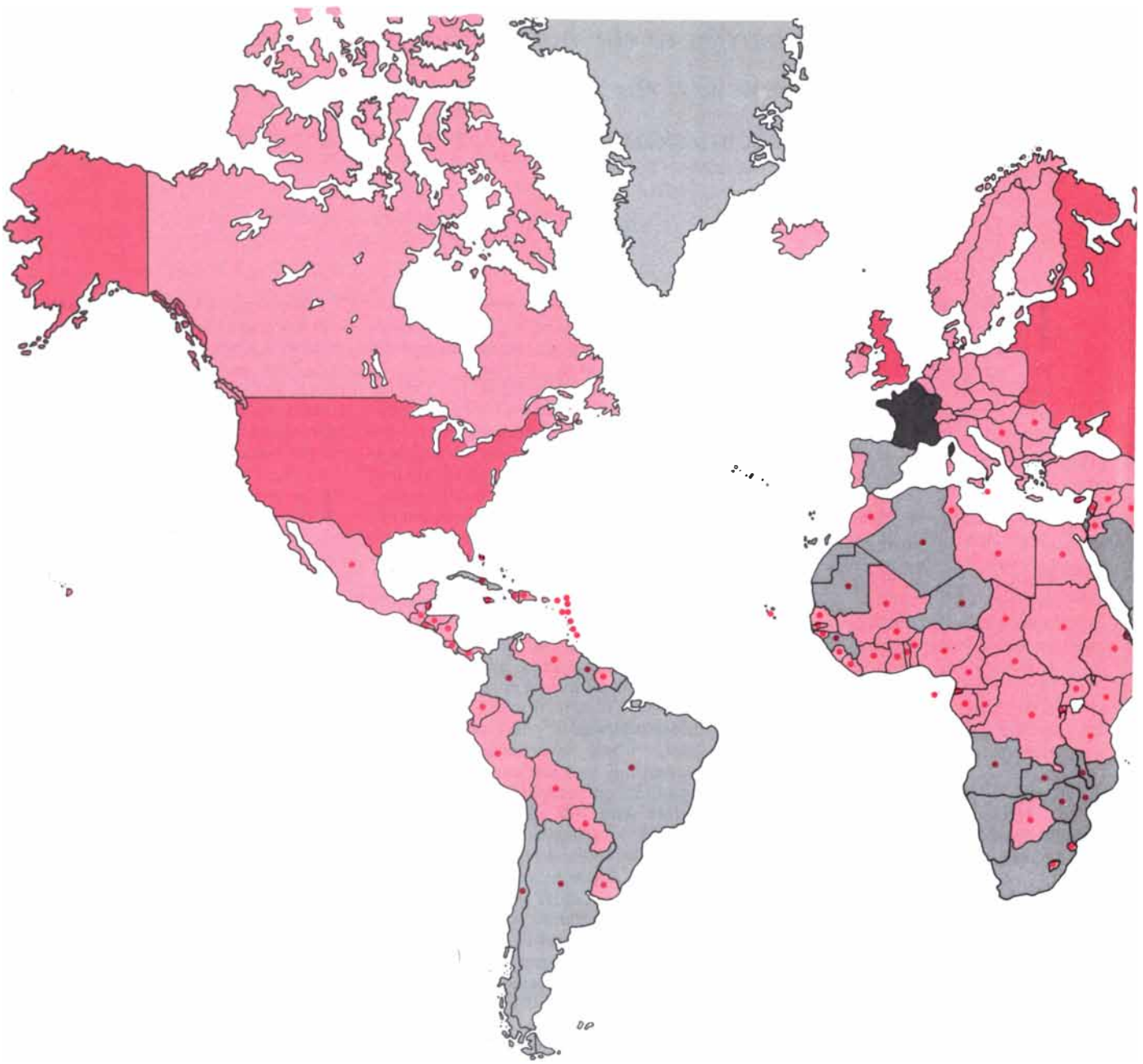
It is noteworthy that the U.S. and the U.S.S.R. have conducted underground tests at a higher rate since the treaty than they conducted atmospheric tests before it. Consequently the treaty has turned out to be more of an environmental and health measure than a curb on the nuclear arms race.

The Treaty of Tlatelolco prohibits the signatories from making, testing, deploying or using nuclear weapons in Latin America. Cuba and Guyana are the only countries in the area that have not signed. (The treaty has 23 full parties but is not yet in force for Argentina, Brazil and Chile.) All five of the nuclear powers have signed a protocol wherein they say they will respect the treaty provisions and will not use or threaten to use nuclear weapons against the parties. This is the only treaty on nuclear weapons to which all the nuclear powers are parties, the only one that puts a legal limit on their right to make, test, deploy or use nuclear weapons and the only one forbidding such weapons in an inhabited area.

The Nuclear Suppliers Group consists of 15 nations (Western, Eastern and neutral) that seek to prevent the diversion of nuclear materials, equipment and technology to weapons. They set up restrictive export guidelines in 1974 and 1975. In 1977 a group of 66 suppliers and importers, with the assistance of the IAEA, expanded discussion of the arrangements in a meeting called the International Nuclear

Fuel Cycle Evaluation. The aim was to minimize the danger of the proliferation of nuclear weapons without jeopardizing the development of nuclear energy for peaceful purposes. In 1980 the expanded group concluded that no technical measures could by themselves ensure the separation of peaceful and military applications in a way that would prevent the proliferation of nuclear weapons to additional countries. The group did envision some precautionary or ameliorative measures: the adoption of more effective and credible international safeguards and the establishment of regional organizations that might exert a restraining or deterring effect on nations edging toward nuclear weaponry. It seems clear, however, that such measures would not stop a determined nation from acquiring a capability to make nuclear weapons.

What, then, of the cornerstone of this edifice: the nonproliferation treaty? During the negotiations it was clear that both the nuclear and the nonnuclear countries wanted to prevent further proliferation. The nonnuclear countries insisted on a straightforward bargain: if they undertook to prevent horizontal proliferation, the nuclear powers must undertake to stop their vertical proliferation. Accordingly the preamble to the treaty recalled the pledge of the three nuclear powers in the test-ban treaty to negotiate an end to all testing, and Article VI obligated them "to pursue negotiations in good faith on effective measures relating to cessation of the nuclear arms



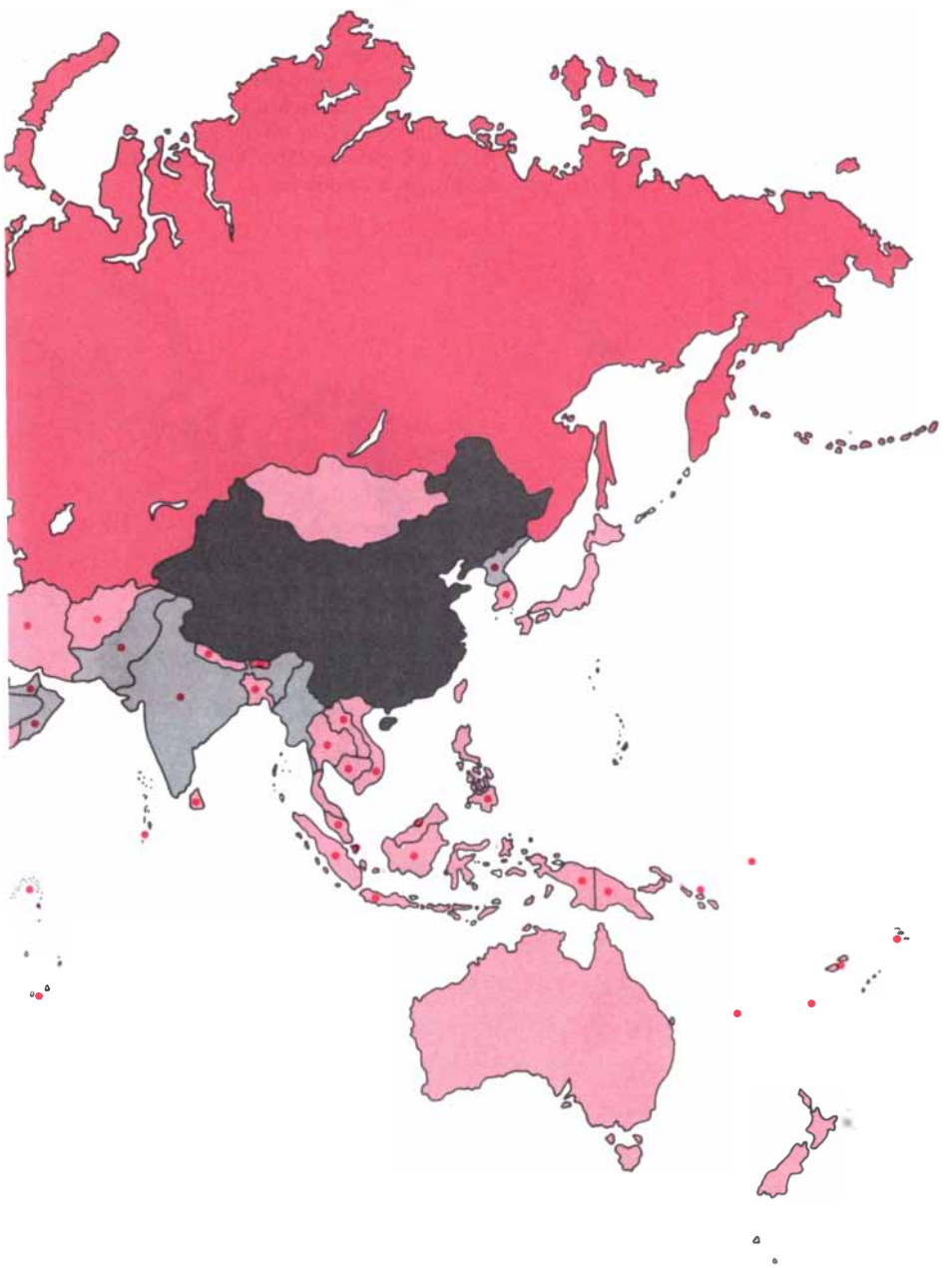
STATUS OF NONPROLIFERATION TREATY is shown on this map. The nuclear nations that are parties to the treaty appear in dark color, the nuclear nations not party to the treaty in dark gray. Nonnuclear nations party to the treaty appear in light color; nonnuclear nations that are not members are shown in light gray. The colored circles identify members of the Group of 77. They can be

race at an early date and to nuclear disarmament."

The treaty came into force in 1970 and now has 128 parties—more than any other arms-control agreement. China and France have not signed the treaty; neither have some 40 other nations, including several that could make nuclear weapons now or will be able to do so soon. They include Argentina, Brazil, India, Israel, Pakistan, South Africa and Spain. Several of them are presumably avoiding the treaty because they have actual or potential national-security problems; others contend that the NPT is discriminatory, favoring nuclear haves over nuclear have-nots, and that in any case the provisions imposing obligations on the nuclear powers are not being carried out.

The continued refusal of some 40 countries, several with advanced nuclear programs, to become parties to the NPT tends to diminish the effectiveness of the treaty. Yet even the parties can withdraw on three months' notice. Therefore adherence to the treaty and acceptance of IAEA safeguards do not provide any absolute guarantee that a country will not at some stage decide to "go nuclear." (Under the terms of the treaty a conference must be held in 1995 to decide on the future of the treaty.)

Moreover, the treaty does not prevent nonnuclear countries from developing a capability or option of having nuclear weapons; it merely prohibits them from acquiring or making the actual weapons or explosive devices. The only conclusive proof that a country has done so is a nuclear test. A country could therefore prepare or acquire a small stockpile of plutonium or weapon-grade uranium without testing an explosive device and then begin testing after giving the required notice of withdrawal.



expected to oppose the policies of the nuclear powers at this month's conference on the treaty in Geneva. This group argues that the nuclear powers have not met their treaty obligation to make serious efforts to halt the nuclear arms race and to reduce their nuclear forces.

In the light of such possibilities the neutral and nonaligned nonnuclear countries regard a cessation of the nuclear arms race (the first obligation under Article VI of the NPT) as requiring a halt to the production, testing and deployment of nuclear weapons—in effect a freeze on nuclear weapons. Their view is that Article VI gives this objective priority over other measures of nuclear disarmament and that it is logical to stop the development of new nuclear weapons before undertaking to either reduce the number of or eliminate such weapons.

Hence these nations regard a comprehensive test ban not only as the most important first step toward ending the nuclear arms race but also as the most feasible and most easily attainable measure to halt both the horizontal and the vertical proliferation of nuclear weapons. The reasoning is that with testing forbidden the nonnuclear countries would be unlikely to commit the resources required for a weapons program and the nuclear powers would not develop new weapons. Then over a longer term the deterioration of existing weapons would gradually make reliance on them less likely and their use less certain. Finally, the nonnuclear nations regard a willingness to enter into a complete test ban as a definitive test of the intentions of the two superpowers and of their will to live up to their obligations under the NPT.

From 1977 to 1980 the U.S., the U.S.S.R. and the U.K. did discuss a comprehensive test ban, but after making considerable progress they recessed the talks and have not resumed them. The reason is that the U.S. changed its position after the election of 1980, abandoning its long-time support of a ban and declaring that, although a test ban "remains an element in . . . [our] arms control objectives, we

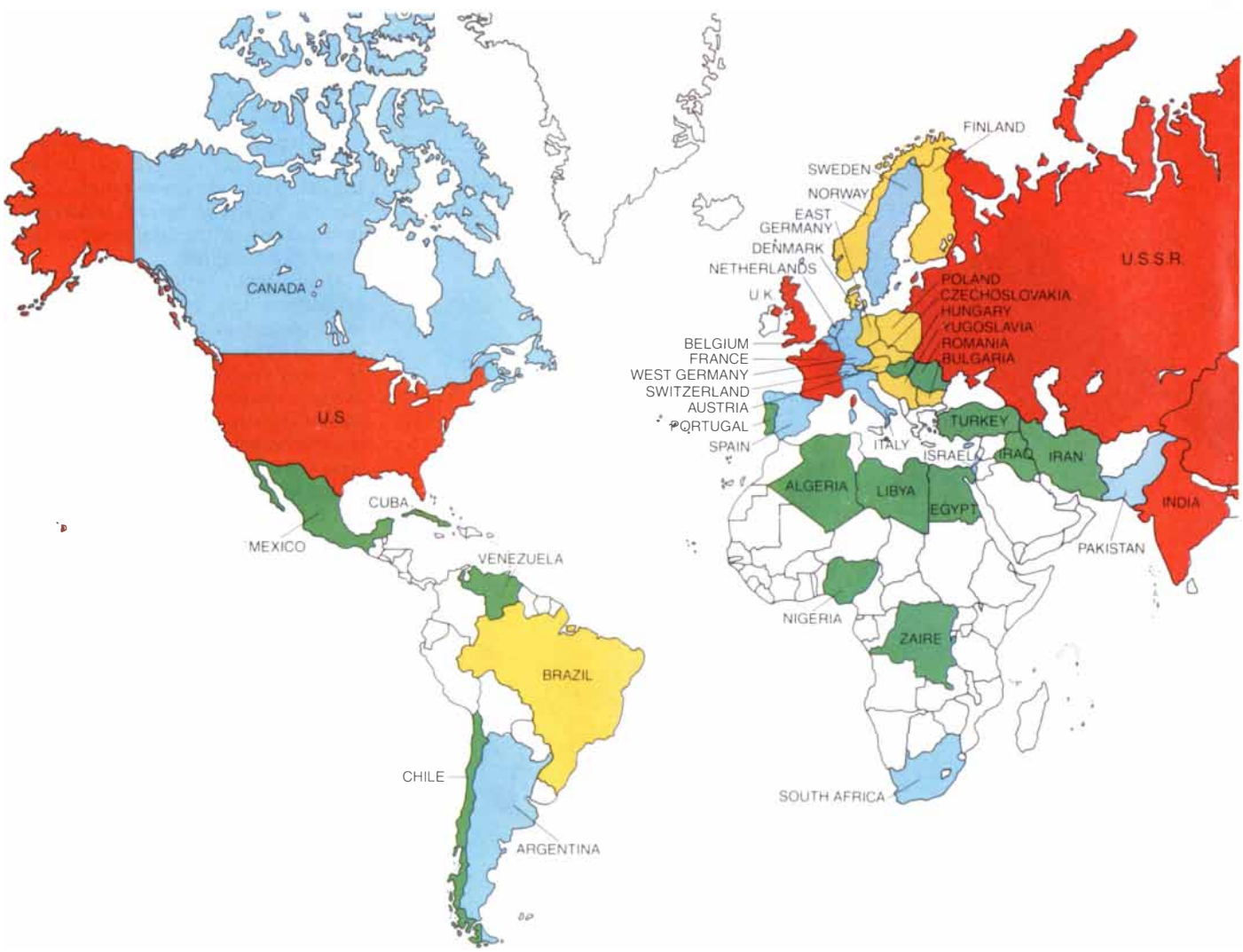
do not believe that, under present circumstances, a comprehensive test ban could help to reduce the threat of nuclear weapons or to maintain the stability of the nuclear balance."

This position did not go over well with the nonnuclear states. Their disappointment and frustration increased when the U.S. decided in 1982 not to resume the trilateral negotiations for a test ban, not only because of doubts about the verifiability of a ban but also because of a perceived need to keep testing new nuclear weapons. As the U.S. Arms Control and Disarmament Agency explained in 1983, "nuclear tests are specifically required for the development, modernization and certification of warheads, the maintenance of stockpile reliability and the evaluation of nuclear weapon effects." The UN secretary general seemed to lend support to the view of the neutral and nonaligned countries when he said earlier this year: "It is of direct importance to the future of humanity to end all nuclear explosions. No other means would be as effective in limiting the further development of nuclear weapons."

Another component of the background of the NPT review conference is the record of the two earlier conferences (in 1975 and 1980). Both of them featured an unusual alignment. Sharing common interests, the U.S., the U.S.S.R. and the U.K. cooperated in resisting the demands of the nonnuclear countries, particularly the neutral and nonaligned nations (mainly of the Third World) known as the Group of 77. (The group now has more than 100 members, but the original name persists.)

At the first conference in 1975 the participants from the Group of 77 asserted that they had fully lived up to their commitments under the NPT whereas the nuclear powers had not done so. They emphasized the failure of the nuclear powers to implement the treaty's provisions for stopping the nuclear arms race and cooperating in the peaceful uses of nuclear energy.

The group made several demands, among them an end to underground testing, a substantial reduction in nuclear arsenals, a pledge not to use or threaten to use nuclear weapons against nonnuclear parties to the treaty and substantial aid to the developing countries in the peaceful uses of nuclear energy. Several of the nonnuclear nations allied with the U.S. and the U.S.S.R. were sympathetic to these demands. In the end the 1975 conference achieved a fragile consensus, expressed in a declaration in which the nuclear powers in effect promised to try harder to meet the demands of the nonnuclear countries.



PRESENT AND POTENTIAL NUCLEAR NATIONS now number 50. This map groups them according to whether they have already conducted explosions of nuclear weapons or devices (*red*) or will be able to acquire the technical capability to explode such a device within a year or two (*blue*), five or six years (*yellow*) or by the year 2000 (*green*). India is classified here as a nuclear power

In 1980, when the second conference was held, the nuclear powers had not met any of the demands. The conference failed to reach agreement on any aspect of stopping the nuclear arms race. The U.S. led the nuclear powers in refusing to make concessions on measures for the control of nuclear arms. As a result, even though the participants had achieved a consensus on plans to advance international cooperation in the peaceful uses of nuclear energy, the conference ended without a final declaration or even any formal reaffirmation of support for the NPT.

The outlook for this year's conference is as bleak or bleaker. The frustration of the neutral and nonaligned countries appears to be turning into resentment and anger because they believe the nuclear powers have misled them. These nations will no doubt renew their long-standing demands for a comprehensive test ban, nuclear disarmament, assurances that nuclear weapons will not be used or held out as threats against nonnuclear countries and greater assistance in the peaceful use of nuclear energy. It seems clear that they will concentrate more than they did in 1980 on pressing their view that the nuclear powers must stop vertical proliferation. Indeed, many members of the Group of 77 believe that an end to vertical proliferation is the only way to stop horizontal proliferation and that both must be halted or neither will be.

In the view of that group the situation has worsened considerably since 1980. The members cite the rapid pace of the nuclear arms race; the abandonment of negotiations for a comprehensive test ban and of talks on intermediate-range nuclear weapons and on the reduction of strategic arms; the spread of the arms race to earth orbit and near space; the unlikely prospect of early progress in the resumed talks between the U.S. and the U.S.S.R. on nuclear and space-based weapons, and the escalation of global military expenditures to some trillion dollars per year. Members of the group also cite the absence of progress over the past five years on any substantive issue of nuclear arms control.

At the same time the fear of a possible "nuclear winter" has helped to create the growing demand for greater efforts to reduce and prevent the risk of nuclear war and to end the nuclear arms race. Many governments, organizations and individuals are calling for a freeze on nuclear weapons, to be followed immediately by substantial reductions in nuclear forces, and for a declaration by each of the nuclear powers that it will not be the first to use nuclear weapons.

The nonaligned countries are therefore considering new approaches to their goals of first halting and then reversing the nuclear arms race. They seem to be searching for ways to awaken the nuclear powers to the seriousness with which they regard the current situation and its impact on the proliferation problem.

One strategy under consideration relates to voting at the conference. The previous conferences operated on the basis of consensus, not purely as a matter of preference but because neither the nuclear powers and their allies nor the Group of 77 could muster the two-thirds vote needed to make a decision. In 1975 only 57 of the 96 countries then party to the treaty attended and in 1980 only 75 of 115. Most of the absentees were small nations belonging to the Group of 77. The group is therefore pressing for a larger turnout this year in the hope of assembling a two-thirds majority.

If the conference is able to muster a decisive vote, the decisions would not be legally binding on the nuclear powers or other parties to the treaty. Nevertheless, they would carry a good deal of political, moral and psychological weight. Their impact could affect the credibility of the policies of the three nuclear powers and also the viability of the NPT if the three powers failed to bring their policies into line with their treaty obligations. Some members of the Group of 77 believe their views should be put to a vote even if only a simple majority favors them.

Another scenario is envisioned by some members of the Group of 77. They would concentrate their efforts at the conference on ending the nuclear arms race and controlling nuclear arms. Discussion of such other issues as peaceful uses and international safeguards on nuclear programs and facilities and on traffic in nuclear materials and technology would be deferred or downgraded.

In a third scenario some members of the Group of 77 would consider walking out of the conference if it becomes clear that the nuclear powers are not ready to proceed with speed and diligence toward fashioning a comprehensive test ban and halting the nuclear arms race. A related idea is for some parties to threaten to give notice of withdrawal from the treaty unless the nuclear powers move promptly to implement their treaty obligations.

Also under consideration is the notion of not ending the conference if the discussions are fruitless. Instead the Group of 77 would call for a recess of a year or so as a means of exerting pressure on the nuclear powers to come up with positive proposals.

Going further, some members of the group talk of taking matters into their own hands by calling a conference to amend the test-ban treaty of 1963 in order to prevent all nuclear tests. It would only be necessary to delete the provisions that permit underground tests. The treaty stipulates that the depository governments (the U.S., the U.K. and the U.S.S.R.) "shall convene" such a conference if one-third (38 or more) of the 112 parties call for it. An amendment would have to gain the approval of a majority (57) of all the parties, including the depository governments.

One cannot suppose the nuclear powers, particularly the U.S. and the U.K., would approve an amendment prohibiting underground tests. Hence the amending conference would not create new legal obligations for the sig-

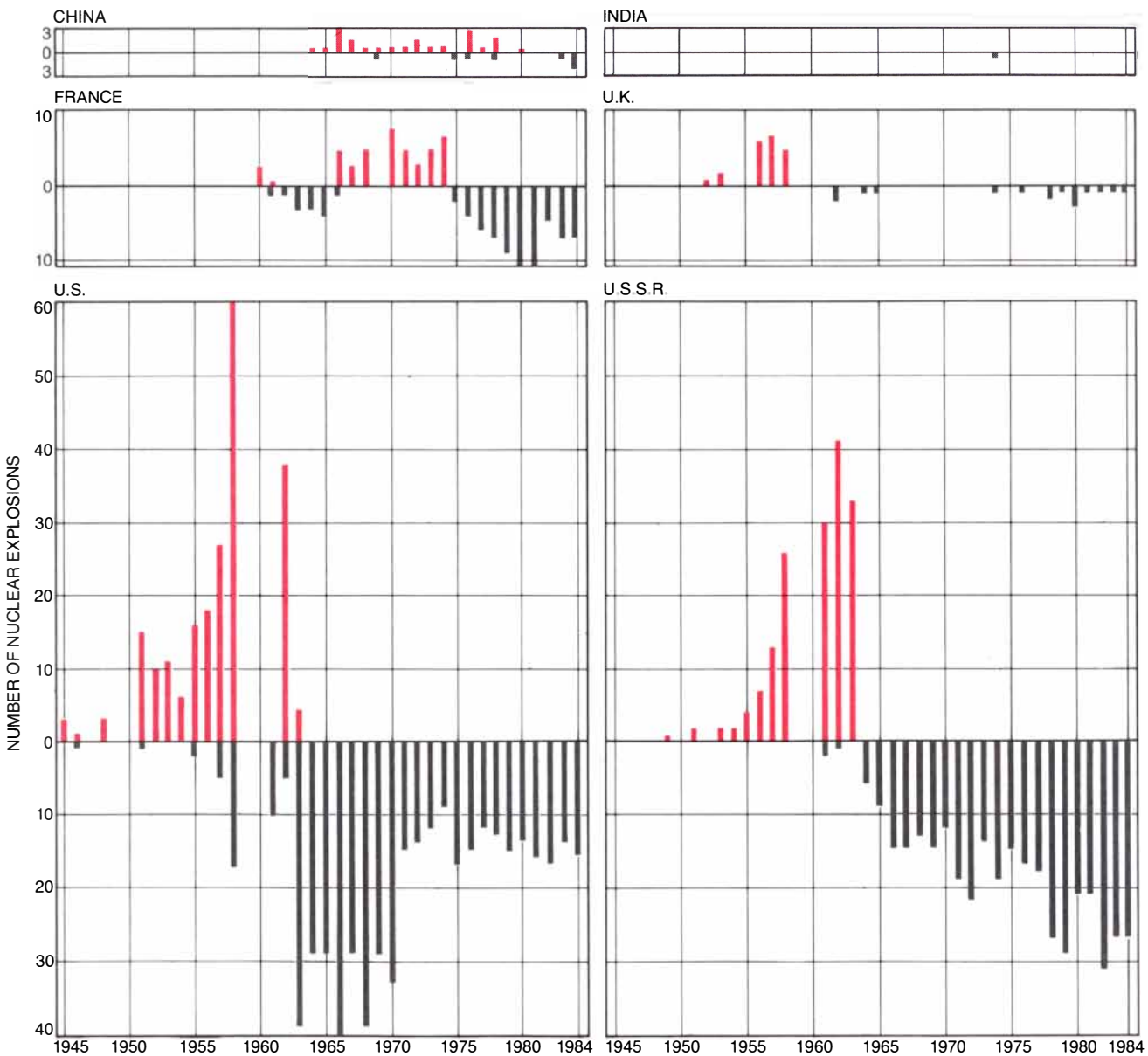
because in 1974 it exploded a "peaceful nuclear device," which in fact incorporates the same basic technology as a nuclear bomb.

natories. Nevertheless, majority approval of such an amendment would put considerable pressure on the three nuclear parties. A failure on their part to respond could have profound political repercussions.

How might the nuclear powers deal with the demands they will encounter at the conference? The U.S. and its allies (the Western group) appear to have no new strategies. They will probably content themselves with the policies they pursued at the previous review conferences. Doubtless they will again emphasize the importance of the treaty and stress the fact that it is in the interest of all states—nuclear and nonnuclear, large and small—to strengthen the treaty and enlarge the membership.

In addition they will probably emphasize again their willingness to expand their cooperation in the peaceful uses of nuclear energy and to provide assurances of supply to nonnuclear parties to the NPT. They will also again propose various plans for international arrangements to manage the nuclear fuel cycle better, to strengthen and improve the IAEA's system of safeguards and to make the entire nonproliferation regime more effective.

Doubtless too the U.S. and its allies will try to put a good face on the lack of progress toward halting the nuclear arms race or negotiating a comprehensive test ban. At the previous conferences they could claim some degree of movement in those directions. They cannot do so this time, but they can point to the recent resumption in Geneva of arms negotiations between the U.S. and the U.S.S.R. and to their stated desire for drastic reductions in the size of the nuclear arsenals. They may also maintain that President Reagan's strategic-defense initiative (the "Star Wars" program) points the way to the elimination of nuclear weapons by rendering them obsolete.

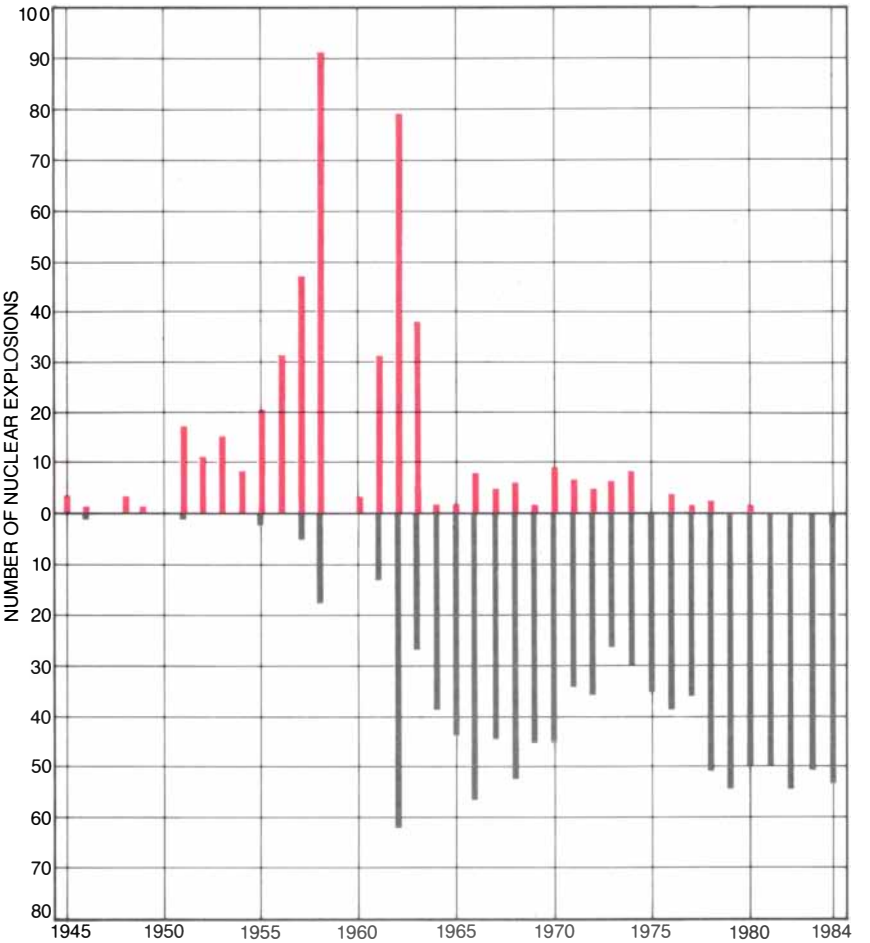


NUCLEAR-EXPLOSIVE TESTS by the six nuclear powers are charted from 1945, when the U.S. exploded the first fission bombs, through 1984. The colored bars represent tests aboveground; the gray bars represent underground tests. The partial-test-ban treaty of 1963 prohibits tests of nuclear weapons in the atmosphere, under water and in outer space, but it allows underground tests. The 33 aboveground tests attributed to the U.S.S.R. in that year were actually carried out before the signing of the treaty in 1963.

The Western nuclear powers are also hoping to achieve a consensus on a short and simple final declaration that would reaffirm the importance of the NPT in enhancing the security of all nations. The declaration would also reaffirm support for the treaty as well as the need to strengthen it with strict adherence to its obligations and goals by moving seriously toward nuclear disarmament and by providing greater assistance in the peaceful uses of nuclear energy.

The Western nations are fully aware of the demands of the Group of 77. They recognize also that the nonaligned nations are most unlikely to be satisfied with Western positions and proposals. Accordingly it seems likely that the Western nations will continue their previous reliance on a "damage limitation" strategy, wherein they will try to prevent defections from the NPT. Perhaps the best they can hope for is that the debates of the conference will proceed without undue animosity and confrontation and that the conference will end, as in 1980, without adopting any final declaration.

In aid of this strategy the U.S. initiated talks with the U.S.S.R. last November to coordinate nonproliferation efforts. The two nations have agreed to hold such talks every six months. The U.S. also arranged to meet with 12 other members of the Nuclear Suppliers Group to reaffirm their policies on nonproliferation and their restrictions on the export of nuclear material, equipment and technology.



TOTAL NUMBER OF NUCLEAR TESTS through 1984 was 1,522, of which 461 were aboveground (*color*) and 1,061 underground (*gray*). The data were assembled by the Swedish Defense Research Institute and published by the Conference on Disarmament. Tests by the nuclear powers are China 29, France 127, India 1, U.K. 37, U.S. 772, U.S.S.R. 556.

The U.S.S.R. and its allies (the Eastern group) will have an easier time at the review conference. They support a comprehensive test ban, a freeze on nuclear weapons, no first use of nuclear weapons (which is tantamount to no use against nonnuclear states) and the prevention of an arms race in outer space. In addition the U.S.S.R. will no doubt emphasize its recent agreement with the IAEA to accept safeguards over certain peaceful nuclear reactors in the U.S.S.R. (The U.S. and the U.K. had already made such an agreement.)

Some members of the Group of 77 have expressed the hope that the Eastern group might even decide to side with them, thereby ensuring a two-thirds vote for a final declaration supporting their common positions. A declaration with that level of support would give greater legitimacy to the positions and put heavier pressure on the Western group. Such an alliance seems unlikely, however, because the U.S.S.R. believes in maintaining solidarity with the U.S. on policies toward nuclear nonproliferation and strengthening the nonproliferation regime.

It is noteworthy that notwithstanding the disagreements between the U.S. and the U.S.S.R. on other aspects of the arms race and arms control, both countries want to prevent the spread of nuclear weapons to more nations. Nonproliferation is the only area where the two countries have maintained a common front, whatever the current tensions between them or the current state of their relations in other fields. Still, the U.S.S.R. has the option of joining forces with the nonaligned states at the review conference in order to increase the pressure on the U.S. and the Western powers. The U.S.S.R. also has the option of abstaining on any vote, ensuring that the proposal would be adopted but would not have the formal support of the U.S.S.R.

It may also come about that the Group of 77 will round up enough supporters (without the backing of the Eastern group) to muster a two-thirds vote for a final declaration supporting the Group of 77 position. In that case the Western group would find itself in an uncomfortable and embarrassing position. Even if no declaration can be formulated and the conference ends, as it did in 1980, without a consensus, such a second failure would have a negative effect on the future of the NPT. Hence there may be considerable support for a move to recess.

Given the present position of the parties it is difficult to see how the conference can avoid a confrontation between the Group of 77 and the nuclear powers, mainly the U.S. That situation would serve not to strengthen the NPT but to weaken it, creating doubt and fear for its future effectiveness. As long as the treaty retains good credibility it exercises a restraining influence even on nonsigners who may contemplate acquiring or making nuclear weapons. Erosion of the treaty's credibility would make it easier for such countries to go nuclear. Since the treaty is the main bulwark against the further proliferation of nuclear weapons, the explosion of a nuclear device by even one or two more countries would put the entire nonproliferation regime in jeopardy.

The Adrenal Chromaffin Cell

This cell synthesizes, stores and secretes a complex mixture containing adrenaline, proteins and peptides. Studies of these processes elucidate mechanisms relevant to other secretory cells, which include neurons

by Stephen W. Carmichael and Hans Winkler

Under conditions of fear or stress a surge of the hormone adrenaline mobilizes the body for peak physical response. Flooding the bloodstream at up to 300 times the normal concentration, the adrenaline interacts with receptors on cells in various organs, increasing the heart rate and blood pressure and prompting the release from the liver of extra sugar to fuel muscular work. Taken together these reactions constitute a "fight or flight" response that prepares one to combat an enemy or flee from danger. They are the end result of a secretory event in the adrenal medulla: the inner part of the two adrenal glands just above the kidneys. There specialized cells known as chromaffin cells manufacture, store and secrete a complex mixture of hormones, the most important of which is adrenaline.

Chromaffin cells are of interest not only as the root of the fight-or-flight reaction but also because they offer insights into the workings of other secretory cells, notably neurons, or nerve cells. The adrenal medulla is an endocrine gland: a ductless gland that affects other tissues and organs by discharging hormones into the bloodstream. It can also be thought of, however, as part of the sympathetic nervous system, which helps to regulate such involuntary functions as heart rate, intestinal movements and the dilation of the pupil. Like the neurons of the sympathetic system, the adrenal medulla is controlled by nerves originating in the spinal cord; its primary hormone, adrenaline, is closely related to noradrenaline, the characteristic neurotransmitter of the sympathetic nerves. (Adrenaline and noradrenaline are also known respectively as epinephrine and norepinephrine.) Moreover, the adrenal medulla itself secretes some noradrenaline, and it also releases the neurologically active substances known as neuropeptides.

The secretion of hormones from the chromaffin cells elicits a widespread response. In contrast, neurons, which have fine axons that extend into their target organs, exert localized control at their axon terminals. When chromaffin cells are grown in culture, however, they extend axonlike processes, which indicates their close kinship with neurons. Adrenal chromaffin cells are available in a relatively pure form, unlike sympathetic neurons, which are scattered throughout the body; the cells are therefore accessible to biochemical analysis. As a result they have served as a laboratory model of neurons, and much of what is known about the production and secretion of neurotransmitters was established through studies of the chromaffin cell.

Chromaffin cells owe their name to the color produced by a chemical reaction of adrenaline. In the 19th century it was found that an unidentified substance in the adrenal medulla reacts with ferric chloride to produce a greenish color and with chromium salts to produce a yellowish brown. At the beginning of this century Alfred Kohn, working in Prague, coined the term chromaffin for cells stained by chromium. At about the same time adrenaline, the cellular compound responsible for the reaction, was isolated and synthesized in the laboratory.

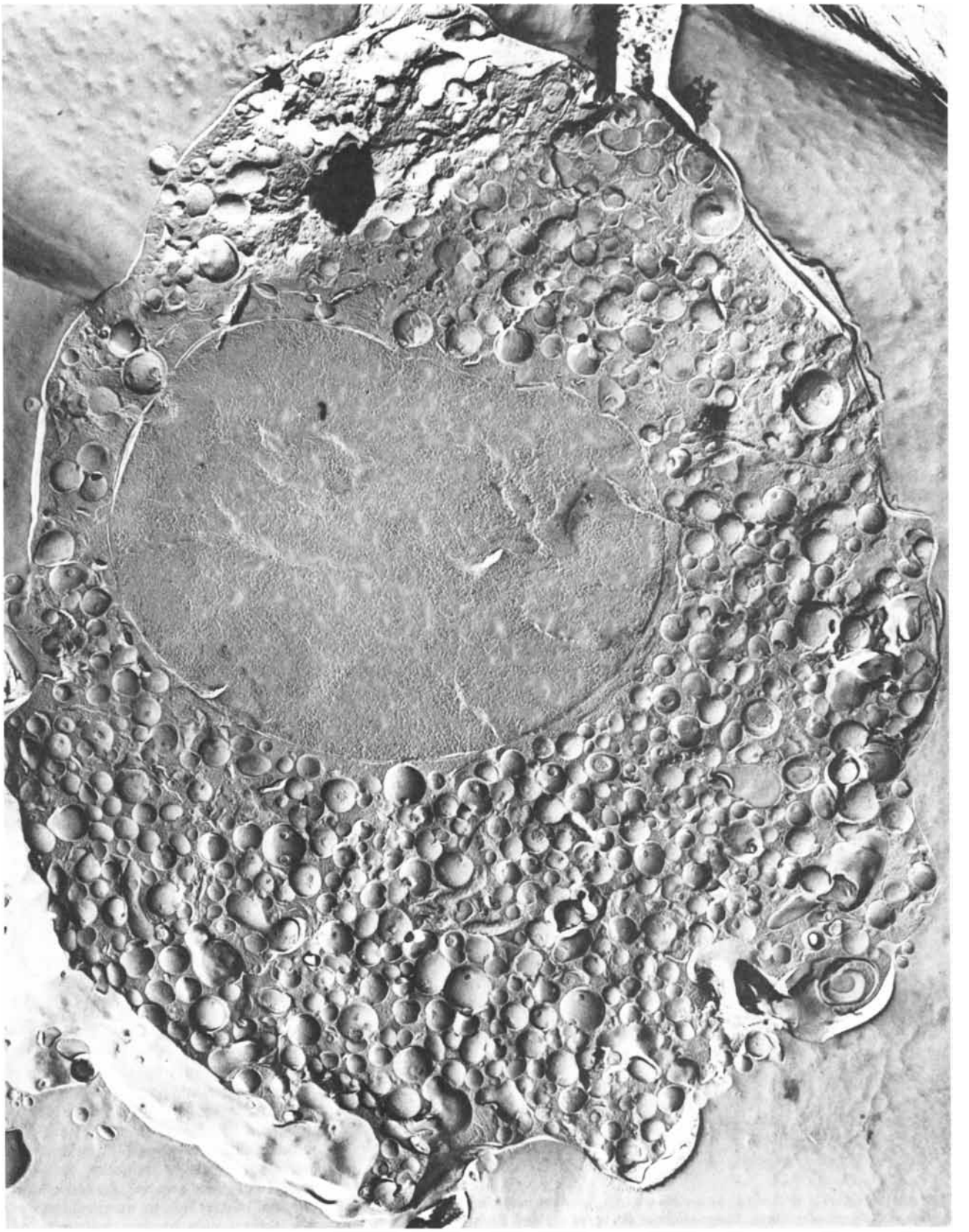
The first hint of how the chromaffin cell stores adrenaline came in 1953. In that year Hermann Blaschko and Arnold Welch of the University of Oxford and Nils-Åke Hillarp and his coworkers at the Karolinska Institute disrupted cells of the adrenal medulla using special homogenizers and searched for adrenaline and noradrenaline in the resulting mixture. They found the hormones were present not in the cytosol (the watery substance in which the intracellular elements are dispersed) but in a form that settled to the bottom of the tube when the mixture was spun in a centrifuge.

It was apparent that the hormones are stored not as free molecules in the cytosol but within subcellular organelles. Jeffrey D. Lever, now at University College Cardiff, and other workers made electron micrographs that revealed the organelles. They are small vesicles, or membrane-bounded sacs, about .3 micrometer in diameter. A single chromaffin cell, as calculated by John H. Phillips of the University of Edinburgh, contains about 30,000 so-called chromaffin granules, or vesicles. The discovery of chromaffin vesicles established for the first time in an endocrine gland a mode of storage now known to be common to almost all cells that secrete hormones or neurotransmitters.

How are the hormones that are contained in chromaffin vesicles synthesized? In 1939 Blaschko proposed that the body makes adrenaline from the amino acid tyrosine in a four-step process, each step being catalyzed by a different enzyme. Since then other workers have traced the biosynthesis in detail.

The first step, the conversion of tyrosine into dopa, is catalyzed by tyrosine hydroxylase, an enzyme found in the cytosol of chromaffin cells. A second enzyme, dopa decarboxylase, also present in the cytosol, rapidly converts the dopa into dopamine. The enzyme needed to transform dopamine into noradrenaline in the third step of the process, dopamine beta-hydroxylase, is present only within the chromaffin vesicles. Therefore dopamine must be transported into the vesicles for the synthesis to proceed.

Torgeir Flatmark of the University of Bergen and his colleagues have shown the vesicles to be rich in a second substance that is crucial to the production of noradrenaline: ascorbic acid, or vitamin C. Dopamine beta-hy-



CHROMAFFIN CELL encloses a multitude of chromaffin vesicles: intracellular sacs containing adrenaline or noradrenaline and a variety of proteins and peptides. The large oval body is the cell nucleus. The subcellular features were highlighted by the freeze-fracture technique. An isolated chromaffin cell was quick-frozen and then fractured. Platinum and carbon, sprayed on the fractured surface, formed a replica of the surface relief; the tissue was then dissolved and a transmission electron micrograph was made of the replica. The cell is enlarged 15,300 times in this micrograph, which was made by Wolfgang Schmidt of the University of Innsbruck.

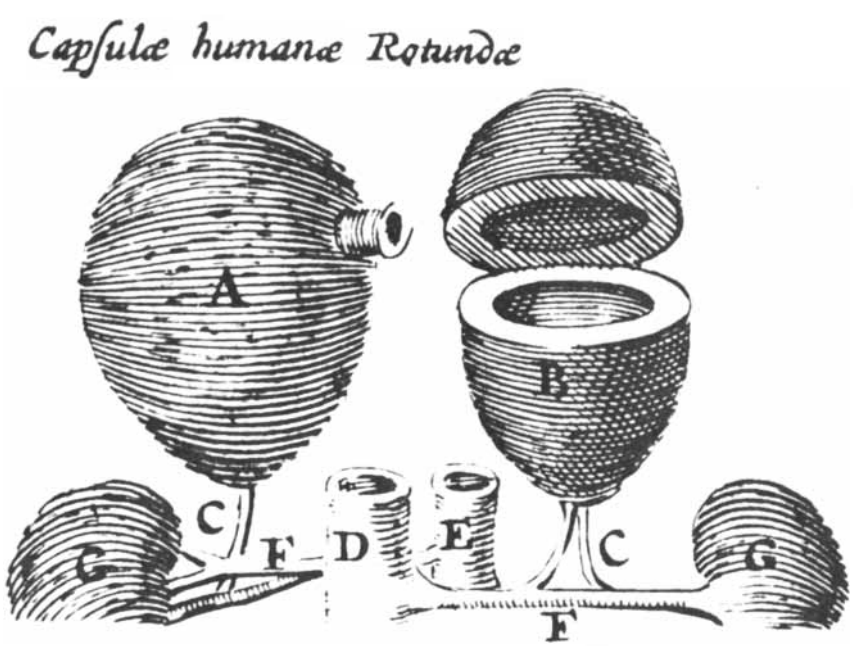
droxylase probably depends on ascorbic acid as a cofactor. The ascorbic acid most likely serves by donating electrons; in the process it is changed to semidehydroascorbate, as Emanuel J. Diliberto and his colleagues at the Wellcome Research Laboratories have demonstrated. The enzyme needed to reconstitute the ascorbic acid and keep the synthesis of noradrenaline under way is found only in the cytosol, however, necessitating a biochemical link between the vesicle interior and the cytosol. Cytochrome b_{561}, one of the most abundant proteins in the vesicle membrane, is thought to transfer electrons across the membrane to the semidehydroascorbate, thus renewing the supply of ascorbic acid.

In some cells of the adrenal medulla (10 percent of the total in human glands) the process of synthesis ends with noradrenaline, just as it does in the sympathetic nerves. In most adrenal cells, however, noradrenaline is converted into adrenaline by phenylethanolamine N-methyltransferase. Because this enzyme is found only in the cytosol, noradrenaline must leak out of the vesicles to be changed into adrenaline, which is subsequently returned to the vesicles for storage.

Regulatory mechanisms controlling the rate at which adrenaline is synthesized through this four-step process act mainly at the first stage, in which tyrosine is converted into dopa by tyrosine hydroxylase. Both the number of tyrosine hydroxylase molecules and their state of activation determine the rate of synthesis. When the organism is placed under sudden stress, the level of activity of the cell's existing complement of enzyme increases, leading to a spurt in adrenaline production. Norman Weiner and his collaborators at the University of Colorado at Denver School of Medicine showed that in an acutely stressed rat about 50 percent of the tyrosine hydroxylase is in an activated form, whereas only about 5 percent of the enzyme is activated in an unstressed rat.

When stress is sustained, resulting in prolonged stimulation of the adrenal medulla through the splanchnic nerve, a different process leads to a long-term increase in the rate of adrenaline synthesis. Working in the laboratory of Julius Axelrod at the National Institute of Mental Health, Hans Thoenen (now at the Max Planck Institute for Psychiatry in Munich) established that continued stimulation of the chromaffin cell results in the synthesis of extra tyrosine hydroxylase. Coupled with the short-term mechanism of enzyme activation, this long-term adaptive device means the adrenal medulla need not store an excess of hormone sufficient to meet any demand. Similar mechanisms, it is now known, control the synthesis of noradrenaline in sympathetic neurons.

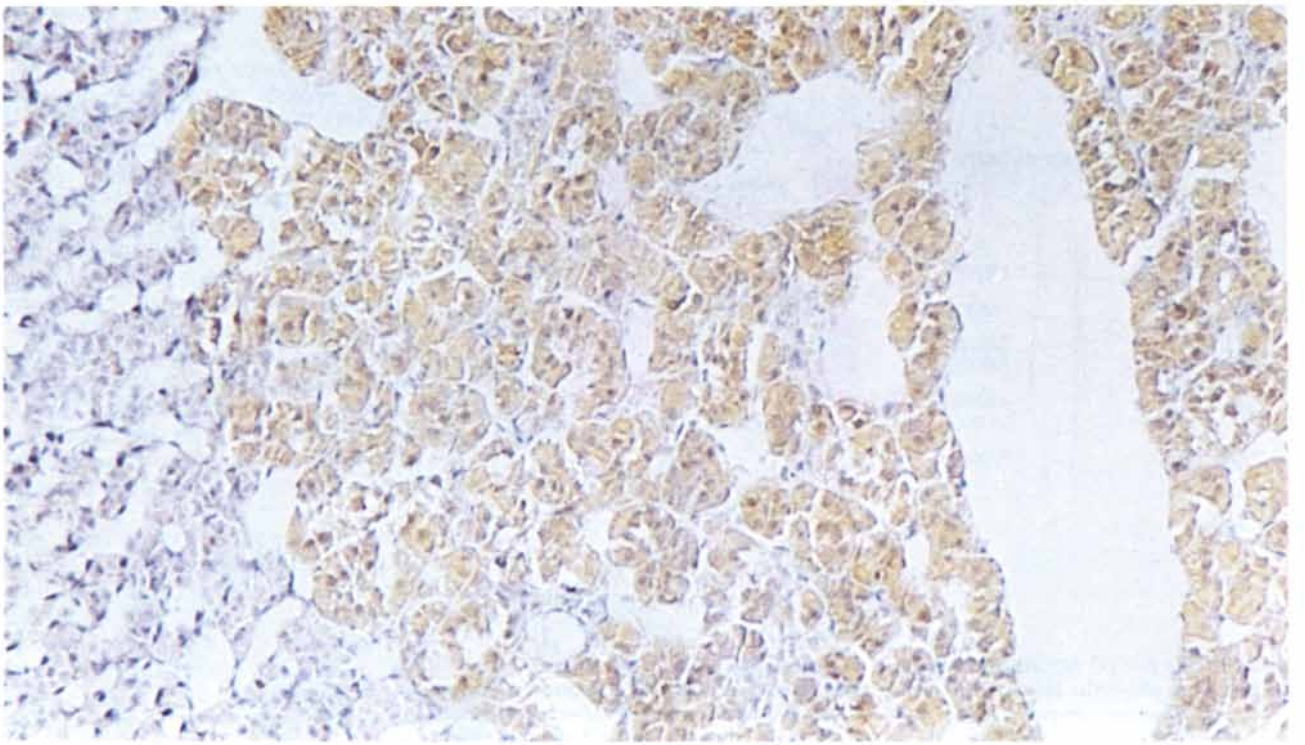
Adrenaline synthesis in the adrenal medulla is also subject to the biochemical influence of the adrenal cortex: the surrounding area of the adrenal gland. Because blood from the cortex passes through the medulla, most of the chromaffin cells are exposed to high levels of cortical steroids, the hormones released by the cortex. When the steroids are absent, Roland Ciaranello and Dona Wong of the Stanford University School of Medicine have found, the enzyme that catalyzes the conversion of noradrenaline into adrenaline is degraded at an unusually high rate; its level in the cell declines, slowing the synthesis of adrenaline. Thus the intimate anatomical relation of the cortex and the medulla extends to the molecular level as well. Stress is known to induce the secretion of hormones from the adrenal cortex as well as from the medulla. The cortical hormones apparently ensure that the synthesis of adrenaline is maintained.



THE ADRENAL GLANDS (*A, B*) are shown in a woodcut made in 1611 by the Danish anatomist Caspar Bartholinus. The medulla of the adrenal gland, the central part that contains the chromaffin cells, was not yet recognized as a distinct entity; Bartholinus described the glands as capsulelike organs filled with "black bile." Two centuries passed before the French anatomist Georges Cuvier established that the adrenal gland has a solid medulla. It was later determined that the dark liquid Bartholinus noted results from the rapid breakdown of the medulla following death. The structures labeled *G* in the woodcut are the kidneys, shown disproportionately small; the other features that are shown are blood vessels.

At two junctures in the process of adrenaline synthesis crucial substances synthesized in the cytosol of the chromaffin cell must be transported into the chromaffin vesicles. Dopamine must enter the vesicles for conversion into noradrenaline; adrenaline, the end product of the synthesis, must pass from the cytosol into the vesicles for storage. Within the vesicles the concentration of adrenaline is at least 25,000 times as great as it is in the cytosol, testifying to the efficacy of the transport mechanism. How are hormones pumped across the vesicle membrane?

In 1962 Arvid Carlsson and his colleagues at the University of Göteborg and Norman Kirshner of the Duke University School of Medicine reported an early clue to the nature of the transport mechanism. They showed that isolated chromaffin vesicles can accumulate and store adrenaline from a dilute solution. They also discovered that the accumulation took place only when magnesium ions and the compound adenosine triphosphate (ATP), the cellular energy carrier, were included in the solution. Noting that the vesicle membrane contains the enzyme ATPase, which degrades ATP when it is activated by magnesium, Peter Banks, now at the University of Sheffield, proposed that the uptake of adrenaline depends on the energy-yielding breakdown of ATP by ATPase. But the actual role of the ATPase in the process remained elusive



CHROMAFFIN REACTION yields a yellowish brown color; it takes place when chromium salts applied to adrenal tissue react with oxidized adrenaline in the chromaffin cells. In this micrograph of rat adrenal gland the color distinguishes the adrenal medulla, the site of the chromaffin cells, from the blue-stained adrenal cortex: the part of the adrenal gland that surrounds the medulla. The oval space corresponds to the central vein, which carries off secretions. Rex E. Coupland of the University of Nottingham made the image.

for 20 years after Banks's proposal.

The role of ATPase became clearer in the late 1970's, when George K. Radda, David Njus and their co-workers at Oxford and Robert J. Johnson and Antonio Scarpa of the University of Pennsylvania School of Medicine demonstrated in an elegant series of experiments that as the ATPase of the vesicle membrane degrades ATP it drives protons (hydrogen ions) from the cytosol into the interior of the vesicles. The accumulation of protons within the vesicle gives its interior a positive charge and makes it more acidic than the cytosol. The difference in charge and acidity across the vesicle membrane is referred to as the electrochemical proton gradient.

The gradient represents stored energy, which drives the uptake of catecholamines (a group that includes adrenaline, noradrenaline and dopamine) as protons flow back down the gradient and out into the cytoplasm. The precise mechanism by which the gradient powers the transport process remains to be determined.

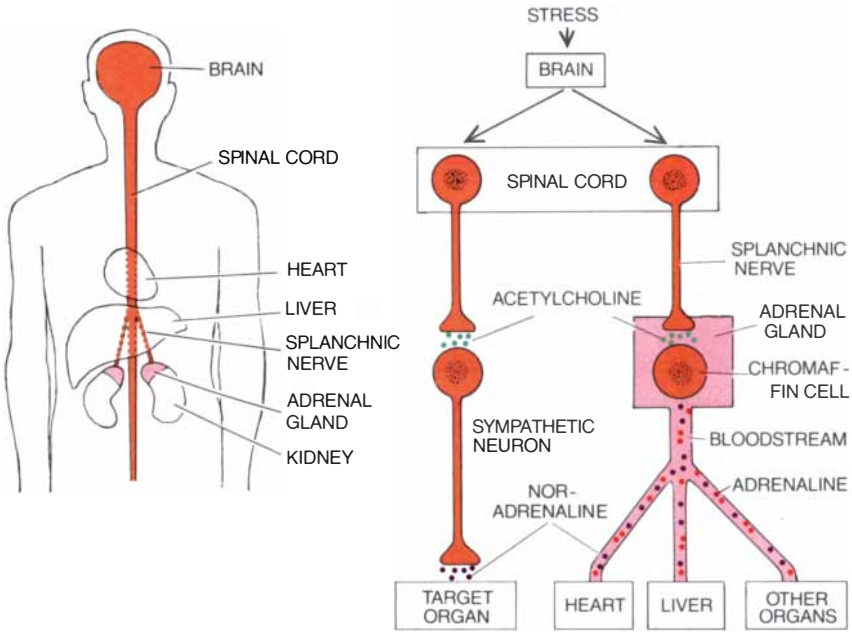
It is clear that to be conveyed into the vesicles the hormones must combine with a special transporter protein. Shimon Schuldiner and Ruth Gabizon of the Hebrew University of Jerusalem recently searched for the protein with a radioactive-labeled molecular probe that was functionally related to the catecholamines. The probe became linked with a vesicle-membrane protein not yet characterized in detail but presumed to be the transporter.

The transport of the catecholamines into the chromaffin vesicles was the first reported instance of a process driven by a proton gradient in an animal-cell organelle other than the mitochondrion. The process in mitochondria was described in 1961 by Peter Mitchell of the Glynn Research Institute in England, who called it the chemiosmotic mechanism. In mitochondria the oxidation of pyruvate (produced by the breakdown of glucose) and other substances generates a proton gradient, which in turn powers the synthesis of ATP. In chromaffin vesicles the relation of ATP and the proton gradient is normally inverted: instead of powering ATP synthesis the proton gradient is sustained by ATP breakdown. (Jean-Pierre Henry and Daniel Scherman of the Institute of Physico-Chemical Biology in Paris and Gabriele Taugner of the Max Planck Institute for Medical Research in Heidelberg have shown, however, that when a proton gradient is artificially imposed across the vesicle membrane, the outflow of protons can lead to ATP synthesis.)

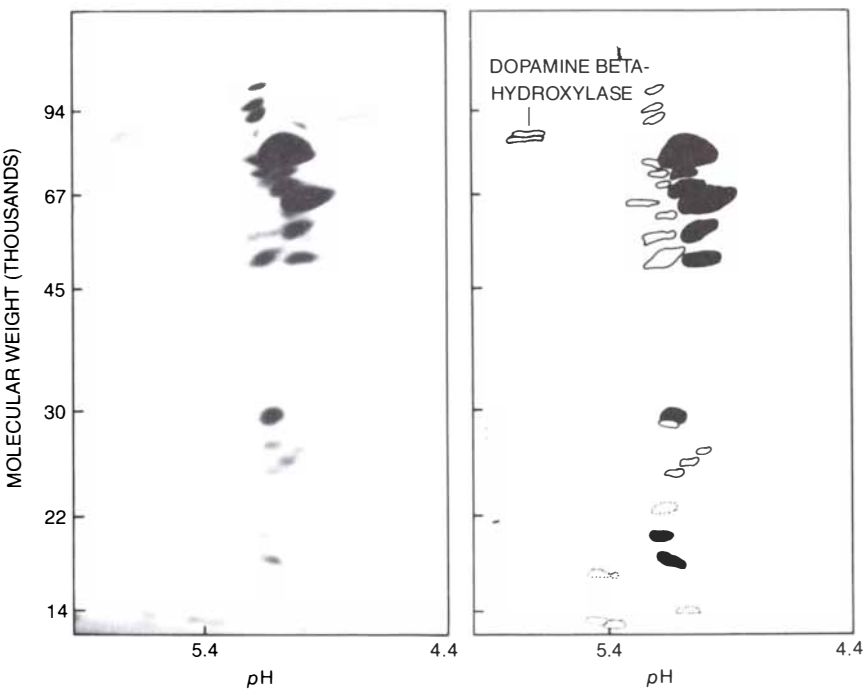
The ATPase enzymes crucial to the proton gradients of the chromaffin vesicle and of the mitochondrion are made up of distinct protein subunits, as shown by David K. Apps, working in Edinburgh, and Nathan Nelson and Shulamit Cidon of the Technion–Israel Institute of Technology. The enzymes have a similar two-part structure, however: a component on the membrane surface that takes part in ATP synthesis or breakdown and another component that serves to conduct protons through the membrane.

The uptake mechanism described for catecholamines seems to be common to a number of secretory cells. For example, a proton gradient concentrates serotonin, a substance that is released during blood clotting, in the storage vesicles of blood platelets. A number of investigators have recently established that the accumulation of the neurotransmitter acetylcholine in the synaptic vesicles of neurons also depends on a proton gradient.

Working with our colleagues at the University of Innsbruck, we have demonstrated a second role for the proton gradient in the chromaffin cell. The concentration of ATP is at least 30 times as high in the chromaffin vesicles



STRESS LEADS TO SECRETION both in the chromaffin cells of the adrenal glands, which are just above the kidneys (*left*), and in sympathetic neurons, which infiltrate a range of tissues and organs. Impulses arrive at both kinds of cells through nerves originating in the spinal cord (*right*). Chromaffin cells secrete adrenaline, noradrenaline and other substances into the bloodstream, thereby exerting widespread control over tissues and organs; sympathetic neurons discharge noradrenaline locally. Because of the similarities in the activation of chromaffin cells and sympathetic neurons, in their secretions and in their effects, and because they develop from the same stem cells (in the neural crest) during embryogenesis, the adrenal medulla is considered to be part of the sympathetic nervous system.



PROTEINS IN THE CHROMAFFIN VESICLE were separated by two-dimensional gel electrophoresis, in which proteins embedded in a gel are sorted by electric charge (here expressed as *p*H) and molecular size (expressed as molecular weight in thousands). The image at the left shows the gel after the application of a stain that revealed the positions of the proteins. Antibodies to specific proteins were then applied to resolve the vesicle contents into the groups of proteins shown schematically at the right: chromogranin *A* (*spots*), chromogranins *B* (*outlines*), enkephalin precursors (*broken outlines*) and the enzyme dopamine beta-hydroxylase. Reiner Fischer-Colbrie of the University of Innsbruck did the analysis.

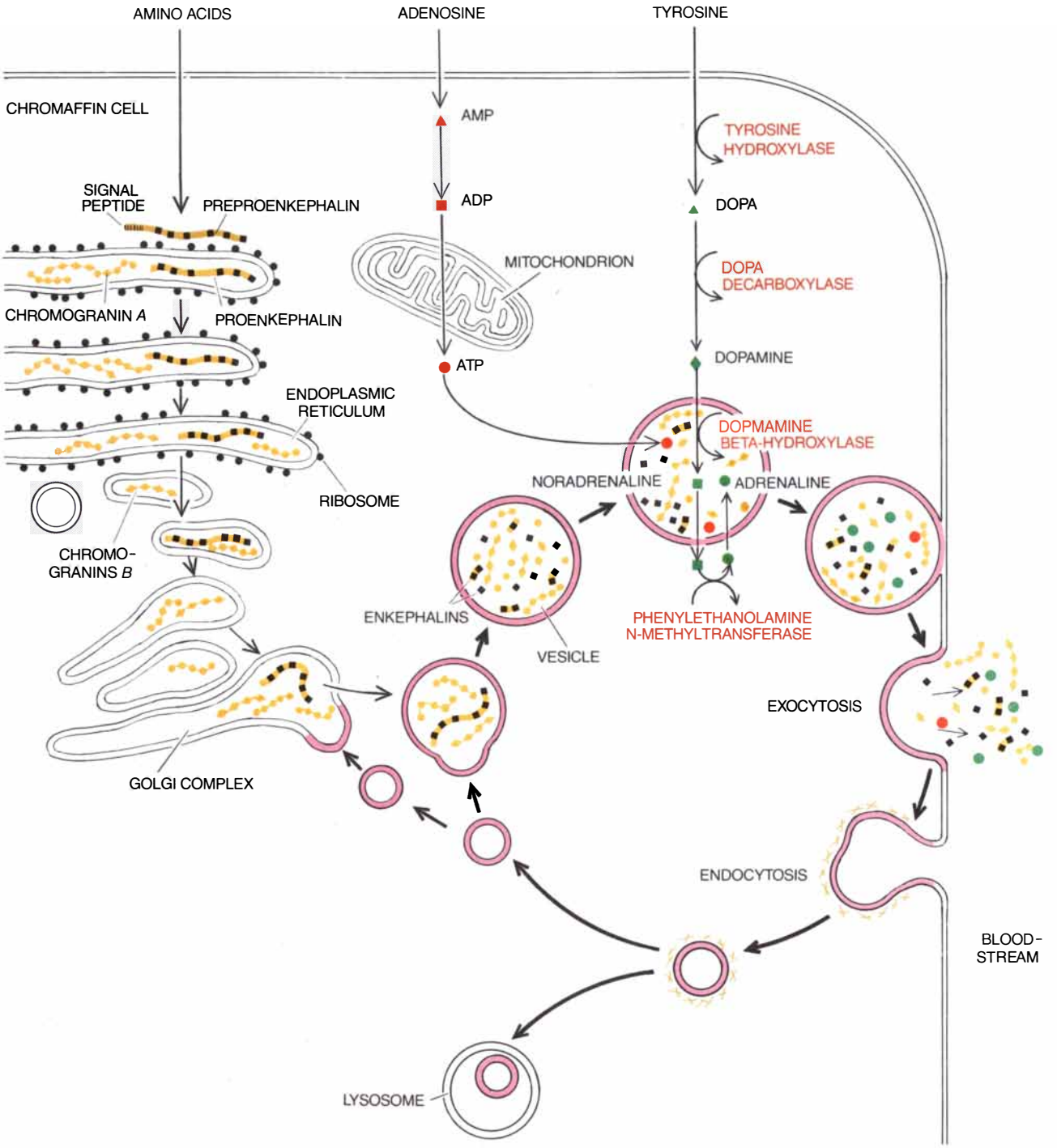
as it is in the cytosol. Our group has established that the electrochemical gradient, coupled with a distinct transporter protein, carries ATP into the vesicles. Work done by Edward W. Westhead, Jr., and his colleagues at the University of Massachusetts at Amherst suggests that within the vesicles ATP serves to maintain osmotic stability. Ordinarily a concentrated solution separated from a dilute one by a membrane takes on water by osmosis. Inside chromaffin vesicles the catecholamines form a very concentrated solution; unchecked, osmosis would rupture a chromaffin vesicle. Interaction of the ATP with the catecholamines lowers their effective concentration and thereby averts the process.

The chromaffin vesicles store not only catecholamines and ATP but also a diverse mixture of proteins that may have widespread effects on the nervous system and on other organs. The presence of proteins in the vesicles was recognized in the 1950's by Hillarp; since then the technique of two-dimensional gel electrophoresis, which makes it possible to separate proteins by both their size and their electric charge, has enabled workers to resolve the proteins in the vesicles into at least 30 components. The most abundant component is an acidic protein named chromogranin *A*.

Although chromogranin *A* has been isolated and characterized, its function remains obscure and its distribution is still being elucidated. It was originally thought to be confined to the adrenal medulla and the sympathetic nerves. A collaboration between groups led by David Cohn of the Veterans Administration Medical Center in Kansas City and by one of us (Winkler) revealed a similar if not identical protein in the parathyroid gland, however, and chromogranin *A* has since been found in other endocrine tissues and the brain.

Whatever the other puzzles are that chromogranin *A* presents, its production within the chromaffin cell is well understood. It is synthesized in the endoplasmic reticulum, the subcellular structure in which many proteins are made, as a proprotein: a precursor of other proteins. It then passes into the Golgi apparatus, the organelle in which intracellular containers such as the chromaffin vesicles are formed, and is packaged into vesicles. Within the chromaffin vesicles the proprotein is split by proteolytic enzymes, freeing smaller molecules.

The features of chromogranin *A*—its widespread distribution in endocrine and nervous tissues and its proteolytic processing—are common to a group of related but less abundant proteins re-



PATHWAYS IN THE CHROMAFFIN CELL include the biochemical routes by which the cell synthesizes the contents of the chromaffin vesicles and the physical cycle through which each vesicle is emptied and its membrane retrieved for subsequent reuse. The primary secretory product of most chromaffin cells, adrenaline, is synthesized from tyrosine in four steps, each one catalyzed by a different enzyme. Tyrosine is transported into the cell, changed into dopa and then into dopamine. The dopamine is carried into the chromaffin vesicles and there is transformed into noradrenaline, which subsequently leaks out of the vesicles into the cytosol and is converted into adrenaline; the adrenaline is pumped back into the vesicles for storage. A second substance stored in the vesicles, adenosine triphosphate (ATP), is synthesized from adenosine, which is converted first into adenosine monophosphate (AMP) and then into adenosine diphosphate (ADP). The ADP enters the mitochondrion and is changed into ATP. Amino acids serve as raw material for the manufacture of enkephalin precursors and chromogranins *A* and *B*. Preproenkephalin, a long polypeptide chain containing enkephalin sequences, is synthesized on ribosomes along the endoplasmic reticulum. As it passes into the reticulum the preproenkephalin loses a short signal peptide; the shortened molecule (now known as proenkephalin) then moves into the Golgi complex and is packaged in newly formed chromaffin vesicles. There it is broken down into smaller pieces, including some free enkephalins. The chromogranins are also synthesized as large proproteins, most of which are then broken down. The vesicles themselves go through a cycle of exocytosis and endocytosis (*heavy black arrows*). In exocytosis they move to the cell surface, fuse with the plasma membrane and discharge their contents. The vesicle membrane then takes on a "coated," or fuzzy, appearance, a sign of imminent endocytosis. It pinches off into an endocytotic vesicle and returns to the Golgi region, where the membrane is recycled to make new chromaffin vesicles. Some of the retrieved vesicle membrane does not complete the cycle but is broken down in digestive organelles known as lysosomes.

cently characterized by Reiner Fischer-Colbrie in Innsbruck: chromogranins *B*. The same features also distinguish the neuropeptides. Neuropeptides are short amino acid chains, some of which act both as neurotransmitters when released by neurons and as hormones when secreted by endocrine glands; they are currently the focus of considerable interest because of their extraordinarily elaborate and varied roles, which may include the regulation of blood pressure, the suppression of pain and the control of behavior [see "Neuropeptides," by Floyd E. Bloom; SCIENTIFIC AMERICAN, October, 1981].

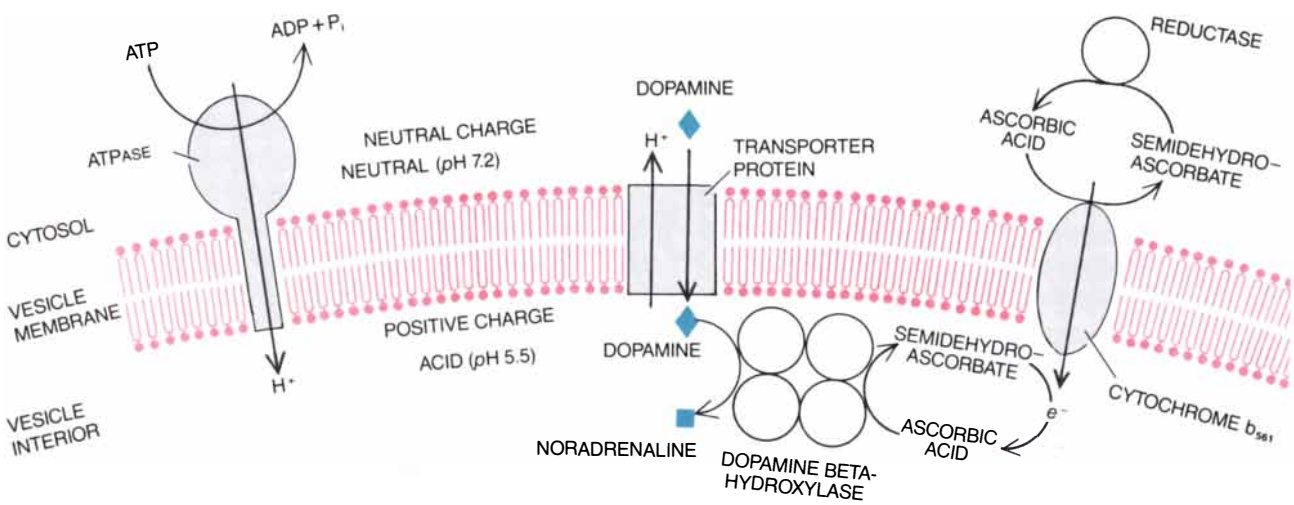
The exciting story of neuropeptides in the adrenal medulla began in 1978, when they were found there by Marianne Schultzberg and Tomas G. M. Hökfelt and their colleagues at the Karolinska Institute. The Karolinska workers exploited immunohistochemistry, in which an antibody to a particular molecule is applied in order to determine the distribution of the molecule in tissues. They showed that the adrenal medulla contains material resembling enkephalins: a group of neuropeptide molecules five amino acids long previously identified in the brain, where they have analgesic properties. Soon afterward O. Humberto Viveros and his colleagues at the Wellcome Research Laboratories and other groups established that in the adrenal medulla enkephalins are stored in chromaffin vesicles. While characterizing these enkephalins Sidney Udenfriend of the Roche Institute of Molecular Biology and Jean Rossier of the CNRS Laboratory for Nervous Physiology at Gif-sur-Yvette and their colleagues also found longer peptides that seem to be the precursors of enkephalins.

Subsequent research showed how enkephalins are produced in the chromaffin cell. Using the techniques of molecular cloning, Michael Comb of the University of Oregon, Ueli Gubler of the Roche Institute and Masaharu Noda and Shosaku Numa of Kyoto University found that in the adrenal medulla the initial protein, called preproenkephalin, consists of 267 amino acids and contains seven enkephalin sequences. The protein enters the endoplasmic reticulum as it is synthesized; there it is transformed into a proenkephalin through the removal of an initial signal peptide (which enabled the protein to enter the reticulum). The proenkephalin then moves into the Golgi region and is enclosed in chromaffin vesicles, where proteolytic enzymes process it to yield free enkephalins.

Like the pace of catecholamine synthesis, the rate at which chromaffin cells produce enkephalins has been found to vary. When chromaffin cells isolated from the adrenal gland of a cow are artificially stimulated, the measured level of enkephalin activity increases along with that of the catecholamines. Lee E. Eiden and Erminio Costa and their colleagues, working separately at the National Institute of Mental Health, have shown that stimulation increases the amount of messenger RNA coding for preproenkephalin. In the rat stimulation of the chromaffin cell has the opposite effect: Ira Black and his colleagues at Cornell University have demonstrated that it dramatically decreases the enkephalin content even as the synthesis of catecholamines rises.

Although much is known about the synthesis of neuropeptides in the adrenal gland and, by extension, in other tissues, we are far from understanding how these substances function after they are released. James Lewis and John C. Liebeskind and their colleagues at the University of California at Los Angeles have shown that the secreted products of the adrenal medulla may provide some analgesia during stress—an effect for which the enkephalins are probably responsible. On the whole, however, the presence of neuropeptides of various sizes in the "cocktail" of substances stored in the adrenal medulla remains mysterious. Insight into why the adrenal medulla mixes such an elaborate cocktail will eventually help to explain the secretion of similar mixtures by other endocrine glands and by nerves.

The chromaffin cell not only stores but also releases its characteristic substances as a mixture. The finding in the 1960's that the adrenal medulla secretes chromogranins and other large



PROTEIN BRIDGES carry catecholamines, protons and electrons across the membrane of the chromaffin vesicle. The enzyme ATPase breaks down ATP in the cytosol into ADP and inorganic phosphate (P_i), releasing energy that pumps protons (H^+) across the membrane into the vesicle. A proton gradient is thereby created across the vesicle membrane: the *p*H of the vesicle interior is lowered, and the vesicle acquires a positive charge. The energy stored in the gradient drives the transport of catecholamines such as dopamine into the vesicle. As protons flow back out into the cytosol a transporter protein carries dopamine into the vesicle. One step in the synthesis of adrenaline, the transformation of dopamine into noradrenaline, takes place inside the vesicle. It is catalyzed by a four-part enzyme, dopamine beta-hydroxylase, with ascorbic acid as a cofactor. In the process the ascorbic acid loses an electron and becomes semidehydroascorbate. A third protein in the vesicle membrane, cytochrome b_{561}, transfers electrons (e^-) into the vesicles from a complementary process in the cytosol, thereby renewing the ascorbic acid and enabling the synthesis of noradrenaline to proceed.

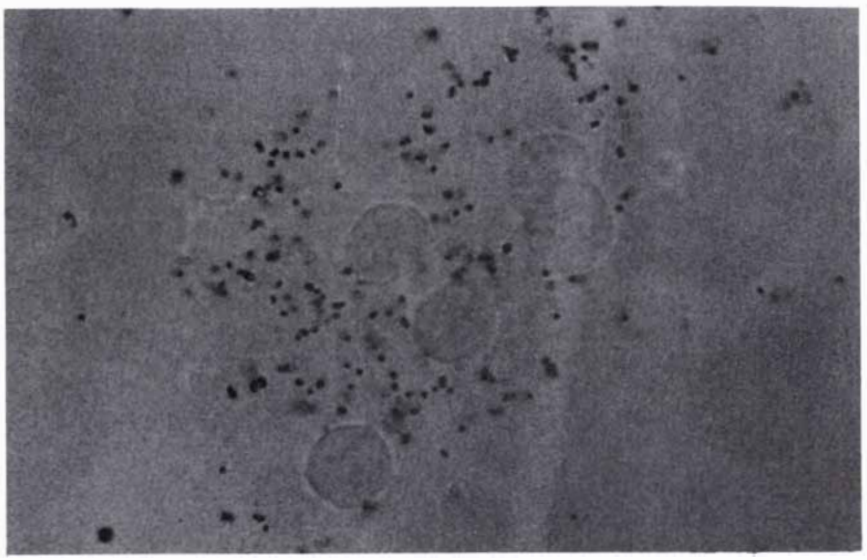
molecules at the same time as it releases adrenaline suggested the secretion occurs by exocytosis. In exocytosis a vesicle moves toward the surface of the cell, the vesicle membrane fuses with the plasma membrane (the cell's outer membrane) and the vesicle is opened to the exterior of the cell. The entire content of the vesicle is thereby emptied into the extracellular space.

The first convincing morphological evidence that exocytosis is responsible for secretion from chromaffin cells came in 1967 in the form of electron micrographs of the hamster adrenal gland made by Odile Grynszpan-Winograd (née Diner) of the University of Paris. The micrographs showed chromaffin vesicles opening directly into the extracellular space after fusing with the plasma membrane. Since then exocytosis has been established as the primary release mechanism for all neurotransmitters and hormones, except for the steroids secreted by the adrenal cortex and the gonads.

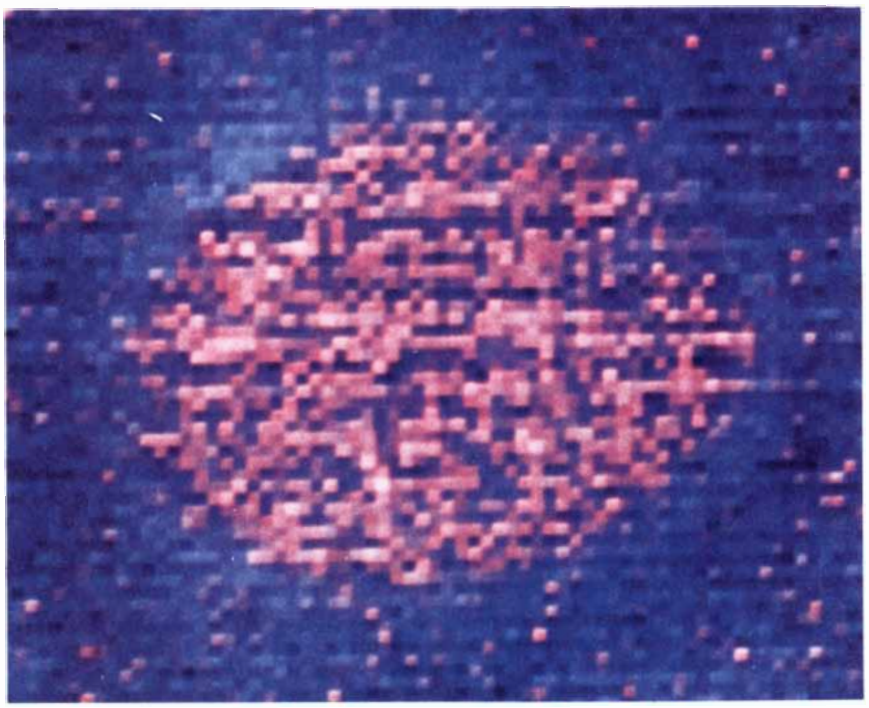
Exocytosis from vesicles is an elegant and efficient mechanism of release. It delivers compounds to the exterior of the cell at a much higher concentration than would be possible if they were secreted from the cytosol. It renders unnecessary a specialized mechanism for transporting large molecules across the plasma membrane, which would be needed if secretion took place through an intact membrane. And it enables a cell to release a package of diverse substances as readily as if its membrane were ruptured, but without a loss of cytosol. One of us (Winkler) recently calculated that when a vesicle fuses with the plasma membrane in the chromaffin cell of the cow, it releases about three million molecules of catecholamines, 800,000 molecules of nucleotide compounds such as ATP, 5,000 molecules of chromogranin *A*, 80 molecules of chromogranins *B* and several thousand molecules of enkephalin precursors and free enkephalins.

How is exocytosis triggered in the chromaffin cell? When a nerve impulse causes the neurons innervating the adrenal medulla to fire, they release the neurotransmitter acetylcholine. The acetylcholine interacts with receptors on the chromaffin cells, setting in motion a process in which calcium channels (molecular pores in the cell membrane) open and allow the concentration of calcium ions within the cells to increase. Secretion ensues.

Just how an increase in intracellular calcium leads to secretion by exocytosis is still not known. One step toward establishing the role of calcium was to separate the events associated with the



PREPROENKEPHALIN MESSAGE in cultured bovine chromaffin cells that were stimulated artificially was made visible by Ruth E. Siegel of the National Institute of Mental Health by means of in situ hybridization histochemistry. A sequence of DNA complementary to the messenger RNA (mRNA) coding for preproenkephalin was prepared and labeled with a radioactive isotope. When the complementary DNA (cDNA) was introduced into chromaffin cells, it hybridized with (bound to) the mRNA. The cells were then coated with a photographic emulsion to make an autoradiograph, in which black dots reveal the location of radioactive cDNA that is bound to the mRNA. (Some dots are visible outside cells because the radioactive particles traveled beyond the cell before interacting with the emulsion.) The image shows about three times as many dots as images of unstimulated cells show. Such results suggest that the increase in the synthesis of preproenkephalin that occurs in stimulated chromaffin cells stems from an increase in mRNA encoding the protein.



ELECTRON ENERGY-LOSS SPECTROSCOPY reveals the relative abundances of elements in a chromaffin vesicle and the surrounding cytosol. The technique uses variations in the energy loss of the electrons passing through a specimen under an electron microscope to map the distribution of elements in the tissue. In this computer-generated image nitrogen is represented as pink and carbon as blue. The suffusion of pink in the vesicle reflects its high concentration of nitrogen-containing substances such as adrenaline, noradrenaline, dopamine and ATP. The blue halo around the vesicle probably results from the relatively high carbon content of its membrane. Richard Ornberg and Richard D. Leapman of the National Institutes of Health made the image, in which the enlargement is 300,000 diameters.

influx of calcium ions from the effects of calcium within chromaffin cells. The step became possible when Bruce G. Livett, now at the University of Melbourne, and his colleagues pioneered techniques for isolating chromaffin cells. Piercing the membrane of the cell makes it possible to change the intracellular calcium level without opening membrane channels. Peter F. Baker and Derek E. Knight of King's College in London caused a selective breakdown of the membrane with brief, intense electric fields; Jack Brooks of Marquette University, Ronald W. Holz of the University of Michigan Medical School and others found that somewhat larger holes can be made using detergents. Because the pierced membrane is permeable to ions, the addition of calcium to the fluid bathing the cells leads directly to an increase in intracellular calcium.

These techniques enabled investigators to determine that not only calcium but also ATP and magnesium had to be present in the bathing solution for exocytosis to take place. It seems likely that a magnesium-activated ATPase breaks down ATP to provide energy for exocytosis. It is clear as well that calcium's effect is exerted within the cell. What does it interact with there to precipitate exocytosis?

Like most cells, chromaffin cells have a cytoskeleton: an internal network of microtubules and microfilaments made up of structural and contractile proteins. For a while it was thought that in the presence of calcium and possibly ATP the mesh of proteins pulls the chromaffin vesicles toward the edge of the cell, enabling them to fuse with the plasma membrane. This scenario now appears unlikely. Dominique Aunis and his colleagues at the National Institute for Health and Medical Research in Strasbourg have suggested recently that in the presence of calcium the cytoskeletal mesh simply releases the vesicles, allowing them to move toward the plasma membrane. In a related finding Velia M. Fowler and Harvey B. Pollard of the National Institute of Arthritis, Diabetes, and Digestive and Kidney Diseases showed that the calcium decreases the viscosity of solutions containing vesicles and contractile proteins such as those of the cytoskeleton. The necessary concentration of calcium was comparable to the level measured in the chromaffin cell during secretion.

Once the vesicles reach the surface of the cell, other proteins presumably mediate the fusion of the vesicles to the plasma membrane. Pollard and Carl E. Creutz and their colleagues noted that chromaffin vesicles isolated from the cell fuse when they are exposed to calcium and proteins from the cytoplasm. They isolated one of the proteins, which was particularly effective at promoting fusion, and called it synexin, from a Greek word for "meeting." Other synexins have since been identified, but the role of these proteins in exocytosis is uncertain: the fusion of chromaffin vesicles in vitro may not be a valid model for the fusion of vesicles to the plasma membrane in vivo. Some other proteins, which bind to chromaffin vesicles if calcium is present, have been identified and called chromobindins. Their significance for exocytosis is similarly unclear.

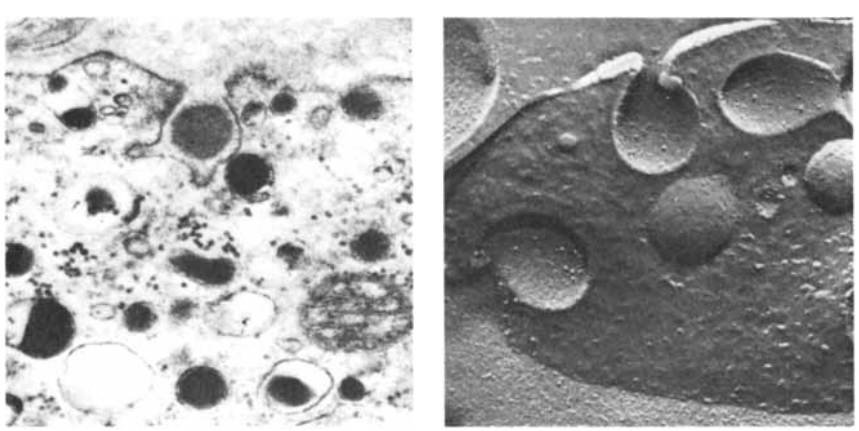
Immunologic techniques may establish the role of these or other proteins in the process of exocytosis. One introduces an antibody into a chromaffin cell, where it binds to, and thus neutralizes, a specific protein. By then determining the point at which exocytosis is interrupted it should be possible to define the precise function of the neutralized protein. Labeling the antibody with a fluorescent dye or small gold particles, which are visible under the electron microscope, should make it possible to trace the protein's distribution within the chromaffin cell.

Introducing such an immunologic probe into the chromaffin cell is not simple. One solution was devised by Jose M. Trifaró and his colleagues at McGill University; their results emphasize the importance of calcium in exocytosis. They ruptured red blood cells and allowed them to re-form in a medium containing antibody to calmodulin, a protein that is crucial to the action of calcium ions within cells. The red cells were in effect transformed into cellular containers of antibody. The workers then introduced the anti-calmodulin into chromaffin cells by fusing them with the red cells. The result was an inhibition of secretion from the chromaffin cells, indicating a direct role for calmodulin, and therefore for calcium, in exocytosis.

The release of a hormone or a neurotransmitter by exocytosis would seem to present a mechanical problem. When a vesicle discharges its contents, its membrane is incorporated into the cell-surface membrane. As secretion proceeds and vesicle membranes are added to the cell surface, the cell might be expected to enlarge to an untenable size. Yet chromaffin cells stay much the same size during secretion.

It was long suspected that some of the surplus surface membrane is returned to the interior of the cell. Eric Holtzman and his colleagues at Columbia University and William W. Douglas and his colleagues at the Yale University School of Medicine demonstrated that the retrieval takes place. They introduced tracer molecules into the medium surrounding chromaffin cells and later detected the tracers in the cell interior. The workers concluded that chromaffin cells remove surplus membrane by endocytosis. Endocytosis reverses the sequence of exocytosis: part of the surface membrane forms a pocket, pinches off into an endocytotic vesicle and passes into the cell interior, carrying with it some of the extracellular medium.

Since then Erwin Neher and Alain Marty at the Max Planck Institute for Biophysical Chemistry in Göttingen and their colleagues have adapted the electrophysiological technique of patch clamping to detect individual events of endocytosis as well as exocytosis. A micropipette filled with an electrically conducting fluid is at-



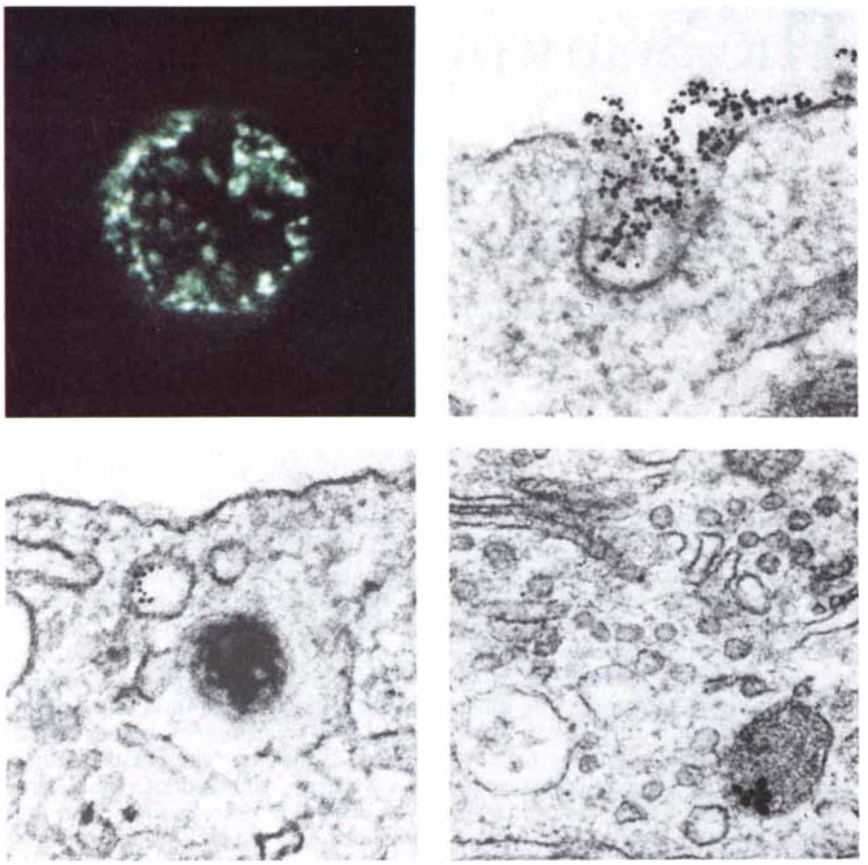
EXOCYTOSIS takes place as the membrane of a chromaffin vesicle fuses with the cell's surface membrane, thus opening the vesicle to the exterior of the cell and releasing its cargo of hormones and other substances. The exocytosis of a single vesicle is shown at a magnification of 40,000 diameters in both a conventional electron micrograph (*left*), made by Odile Grynszpan-Winograd, and a freeze-fracture micrograph (*right*), by Wolfgang Schmidt.

tached to a cell, forming a tight electrical seal. By this means electrical events occurring in the patch of cell membrane encircled by the pipette can be measured; the patch of membrane can also be ruptured, allowing electrical events in the rest of the cell membrane to be recorded. The technique reveals minute changes (on the order of 10^{-15} farad) in the electrical capacitance of the plasma membrane. Under certain conditions such electrical events may be associated with the addition or removal of a small piece of membrane, such as that of a chromaffin vesicle.

Two questions remained: Does endocytosis single out what had been a vesicle membrane for retrieval instead of simply absorbing an equivalent patch of plasma membrane? If it does, how long does the cycle of exocytosis and subsequent endocytosis take? To answer these questions several groups, including ours in Innsbruck, exposed chromaffin cells to antibodies against proteins known to be unique to the inner surface of the vesicle membrane. The antibodies were labeled with a fluorescent dye, making it possible to identify the membrane following exocytosis. When isolated chromaffin cells were stimulated, the labeled vesicle membranes became visible as discrete spots on the plasma membrane, indicating that exocytosis had taken place. After the stimulation was stopped the fluorescent patches corresponding to vesicle membranes disappeared from the cell surface, having been internalized by endocytosis; the retrieval was complete within 30 minutes.

To determine the fate of the vesicle membranes after retrieval we needed an immunologic probe that could be resolved at the level of individual organelles, under the electron microscope. In place of fluorescent antibodies we used antibodies tagged with electron-dense gold particles to mark vesicle membranes during their exposure on the cell surface. We then followed the retrieval of the membranes in a series of electron micrographs. The labeled patches first took on a fuzzy aspect characteristic of "coated pits": patches of cell membrane about to undergo endocytosis. The gold-labeled coated pits then formed pockets and pinched off to become endocytotic vesicles within the cell. In time the vesicle membranes lost their coating. Some of them could later be seen in the Golgi apparatus, where new vesicles are formed. In the end we found gold-labeled vesicle membrane in what appeared to be newly formed vesicles.

Earlier studies of several endocrine glands by Marilyn Farquhar of Yale, Volker Herzog of the University of Munich and Berton C. Pressman and Robert W. Rubin of the University of Miami School of Medicine used nonspecific markers to demonstrate that membrane is recycled between the Golgi region and the cell surface. The use of specific antibodies has now enabled us to establish that in the chromaffin cell the membrane of the secretory vesicle is recycled through vesicle formation, exocytosis and endocytosis. Is it not an elegant and economical strategy to use the vesicle membrane as a recyclable container?



MEMBRANE RECYCLING is documented in micrographs made by Alexandra Patzak of the University of Innsbruck. The photomicrograph (*top left*) shows the surface membrane, enlarged 5,150 diameters, of a cell that was stimulated to secrete. The cell was treated with a fluorescent antibody to a protein specific to the inner surface of chromaffin vesicles. Patches of the membrane give off a green glow, showing that they are bits of vesicle membrane that have been incorporated into the surface membrane through exocytosis. Labeling the antibody with electron-dense colloidal gold enables one to follow the recycling of such patches of membrane in electron micrographs. First the vesicle membrane, decorated with gold-labeled antibodies (*visible as black dots*), takes on a coated appearance and forms a pocket (*top right*), which will later become an endocytotic vesicle. Within such a vesicle labeled membrane is visible (*bottom left*). A micrograph made in a different experiment records the completion of the cycle: former vesicle membrane is evident in an organelle that appears to be a new chromaffin vesicle (*bottom right*). The magnification of the first two electron micrographs is 104,000 diameters and that of the final image is 58,900 diameters.

The adrenal chromaffin cell has demonstrated its worth as a laboratory model for neurons and other secretory cells. The chromaffin cell may soon prove itself as a substitute for neurons in a clinical setting as well. In 1982 Lars Olson and his colleagues at the Karolinska Institute did the first neural transplants performed in human beings: they transferred chromaffin cells from the adrenal medulla to the brain in patients with severe Parkinson's disease. In Parkinsonism the dopamine-containing neurons in a certain area of the brain stem degenerate. From a wealth of basic research the workers reasoned that chromaffin cells, which produce dopamine as a precursor to noradrenaline, could compensate for the dopamine deficiency.

Of the two patients treated with the experimental surgery, one showed significant improvement and the other did not get worse. The possibility that the adrenal chromaffin cell will one day figure in the treatment of Parkinson's disease typifies the way basic biological research, done for the mere joy of discovery, can yield unexpected benefits for humanity.

The Scanning Tunneling Microscope

A new kind of microscope reveals the structures of surfaces atom by atom. The instrument's versatility may extend to investigators in the fields of physics, chemistry and biology

by Gerd Binnig and Heinrich Rohrer

"The surface was invented by the devil," said the illustrious physicist Wolfgang Pauli. Pauli's frustration was based on the simple fact that the surface of a solid serves as the boundary between it and the outer world. Whereas an atom within a solid is surrounded by other atoms, an atom at a surface can interact only with other atoms on the surface, with atoms beyond the surface or with those immediately under it. The properties of the surface of a solid therefore differ radically from those of the interior. For instance, to minimize energy, surface atoms often arrange themselves differently from the other atoms in a solid. The resulting complexities of surface structures have long thwarted efforts to derive precise experimental and theoretical descriptions of them.

At the IBM Zurich Research Laboratory we have developed a device that makes it possible to characterize in a quantitative way such surface complexities: the scanning tunneling microscope. Our microscope enables one to "see" surfaces atom by atom. It can even resolve features that are only about a hundredth the size of an atom. Such a tool has important implications, for example, in the microelectronic industry. As the silicon chip, which is the key element in computer architecture, decreases in size its surface area increases sharply in relation to its volume. Therefore the surface becomes increasingly important in the chip's operation and in its interactions with other logic elements. The scanning tunneling microscope will probably contribute to the understanding of other physical, chemical and biological phenomena as well.

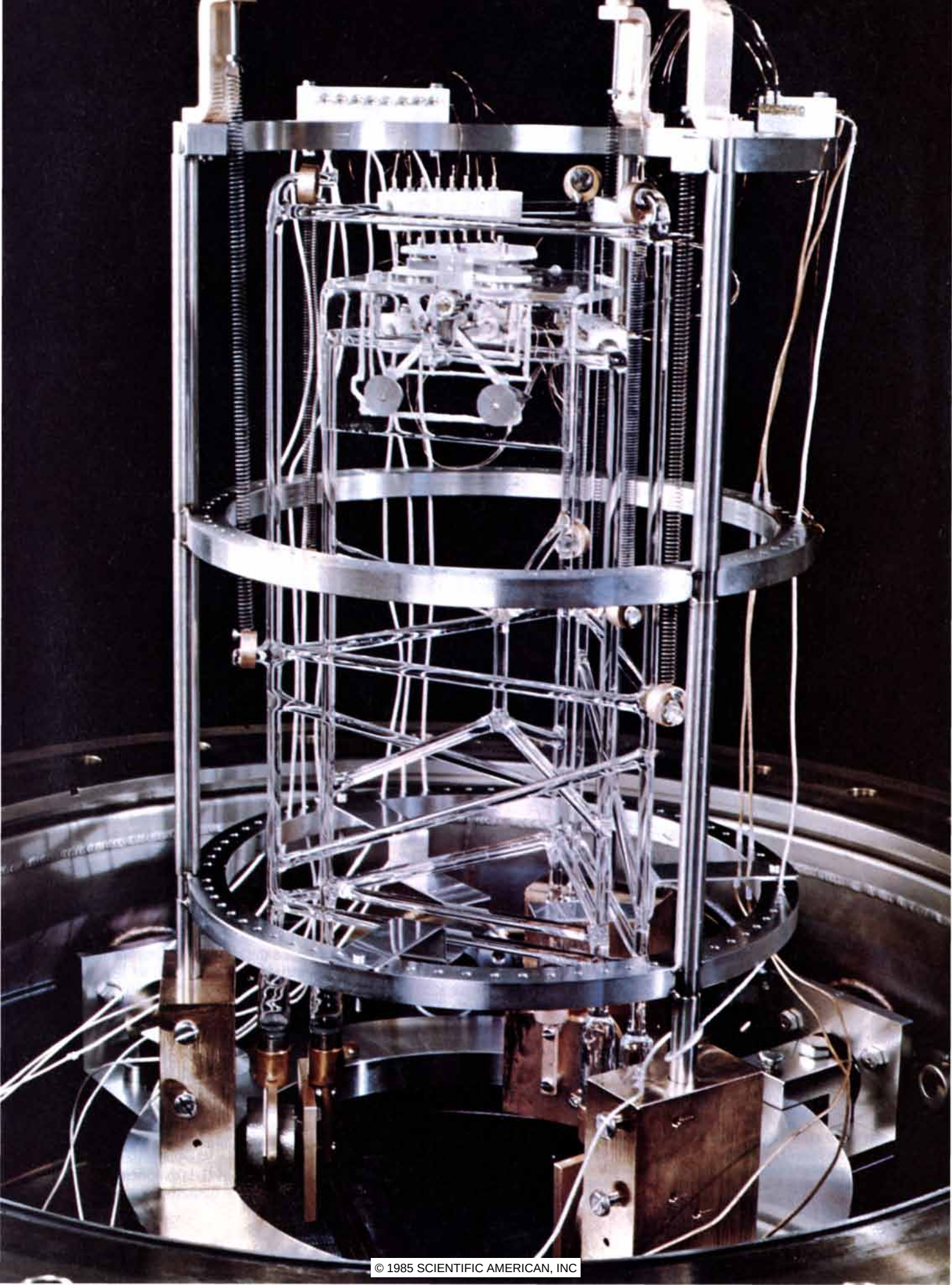
Scanning tunneling microscopy is the product of considerable evolution. Microscopy appears to have begun in the 15th century when simple magnifying glasses were made with which to observe insects. In the late 17th century Antony van Leeuwenhoek developed the optical microscope, which revealed the existence of single cells, pathogenic agents and bacteria. Although optical microscopy has developed into a sophisticated, versatile technique, a physical limit hampers it: the optical microscope cannot resolve atomic structures. The reason is that the average wavelength of visible light is about 2,000 times greater than the diameter of a typical atom, which is about three angstrom units. (One angstrom unit is one ten-billionth of a meter.) In other words, trying to probe atomic structures with visible light is like trying to find hairline cracks on a tennis court by bouncing tennis balls off its surface.

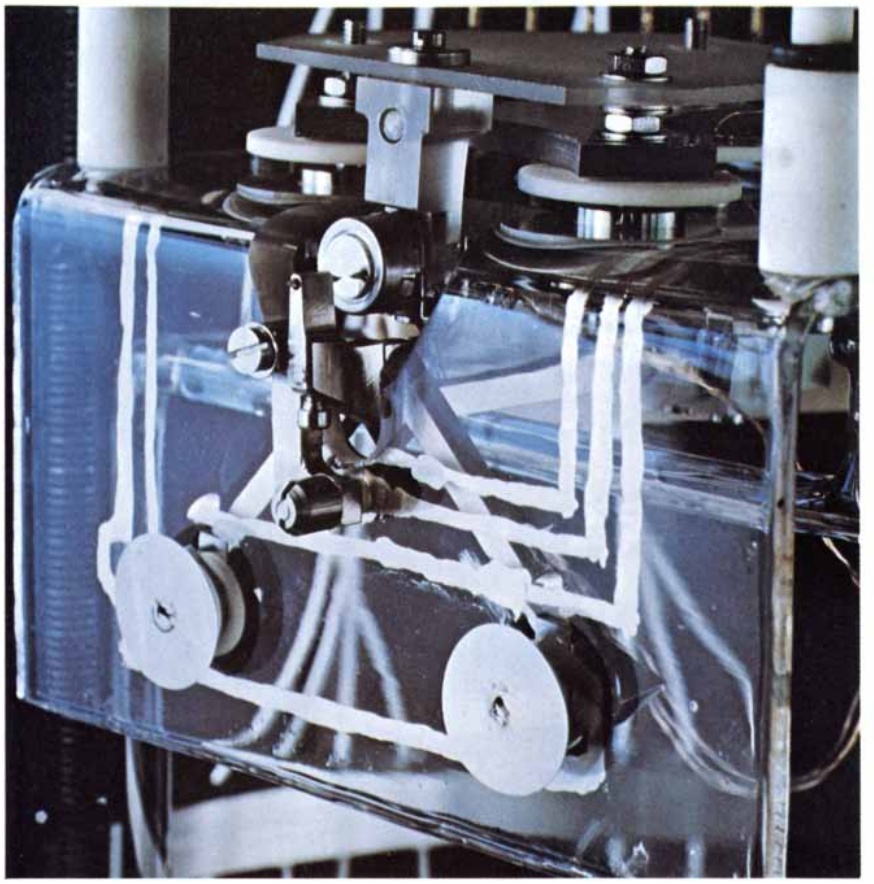
The first successful exploration of atomic structures grew out of a basic discovery of quantum mechanics. It is that light and other kinds of energy exhibit characteristics of both particles and waves. In 1927 Clinton J. Davisson and Lester H. Germer of the Bell Telephone Laboratories confirmed the wave nature of the electron. They also found that a high-energy electron has a shorter wavelength than a low-energy electron. An electron of sufficient energy exhibits a wavelength comparable to the diameter of an atom. This fact led to the invention of the electron microscope. With electron microscopy projections of atomic rows and even atomic orbitals in thin crystalline films have been observed.

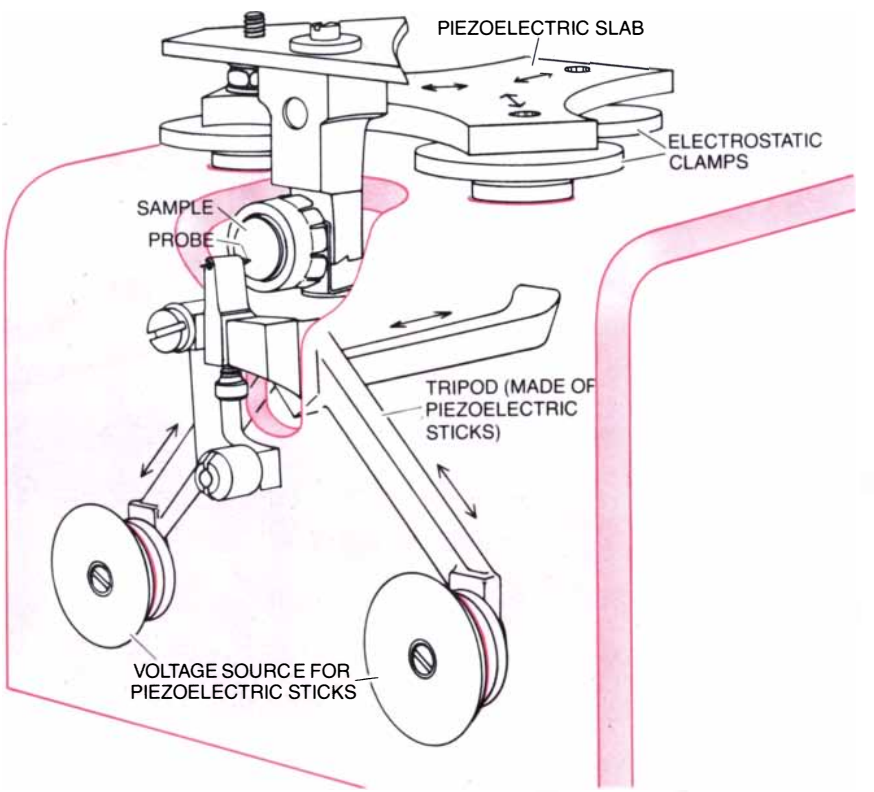
Since the electron microscope has such a high resolving power, why was it necessary to develop a new kind of microscope? Although electron microscopy has proved to be extremely successful in observing the bulk features of crystalline materials, it cannot resolve surface structures except under very special circumstances. A high-speed electron penetrates deep into matter and so reveals little of the surface structure. A slow-moving electron is easily deflected by the charges and the electric and magnetic fields of the sample. In the 1950's Edwin W. Müller made some progress when he invented the field-ion microscope, an instrument that is highly sensitive to surfaces. Unfortunately its range of applicability is narrow: a sample must sit on a fine needle tip that is only a few angstroms wide and the sample must be stable against the high electric fields characteristic of the technique.

The principle of operation of the scanning tunneling microscope makes it possible to avoid these difficulties. The main difference between the scanning tunneling microscope and all other microscopes is that it uses no free particles; consequently there is no need for lenses and special light or electron sources. Instead the bound electrons already existing in the sample under

SCANNING TUNNELING MICROSCOPE has two stages, suspended from springs, that nestle within a cylindrical stainless-steel frame. The innermost stage contains the microscope mechanism. To achieve high-resolution images of surface structures the microscope must be shielded from even such small vibrations as those caused by footsteps and sound. The copper plates (attached to the bottom of the stainless-steel frame) and the magnets (attached to the bottom of the inner and outer stages) damp vibrations. Any disturbance causes the copper plates to move up and down in the field generated by the magnets. The movement induces eddy currents in the plates. The interaction of the eddy currents with the magnetic field retards the motion of the plates and hence the motion of the stages. For work required in a vacuum a steel cover is placed over the outer frame of the microscope.







MICROSCOPE DEVICE contains a sample and a scanning needle. Piezoelectric materials, which expand or contract when voltage is applied to them, enable the device to resolve features that are only about a hundredth the size of an atom. A piezoelectric drive positions the sample on a horizontal metal plate. A piezoelectric tripod then sweeps the scanning needle over the surface of the sample, simultaneously achieving high stability and precision.

investigation serve as the exclusive source of radiation.

To understand this principle imagine that the electrons bound to the surface of the sample are analogous to the water of a lake locked in by the shore. Just as some of the lake water seeps into the surrounding land to form groundwater, so some of the electrons on the sample's surface leak out and form an electron cloud around the sample. According to classical physics, no electron cloud exists because reflection at the sharp boundaries of surfaces confines the particles. In quantum mechanics, however, each electron behaves like a wave: its position is "smeared out." This accounts for the existence of electrons beyond the surface of matter. The probability of finding an electron beyond the surface of a conductor falls rapidly, in fact exponentially, with distance from the surface. Since the electrons appear to be digging tunnels beyond the surface boundary, the effect is traditionally known as tunneling.

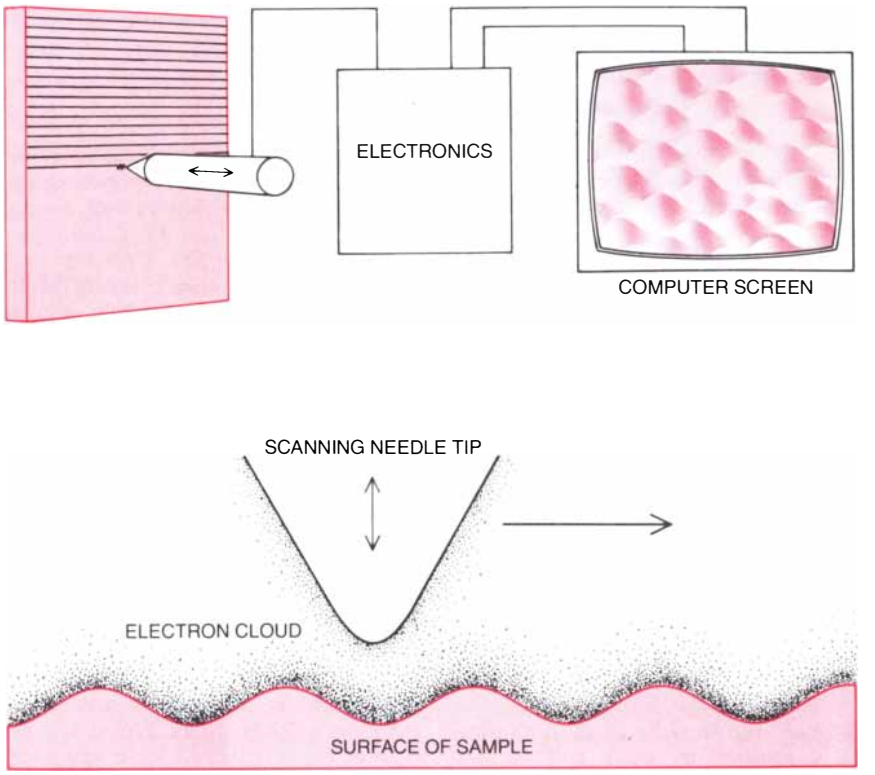
The first experimental verification of tunneling was made about a quarter of a century ago by Ivar Giaever of the General Electric Company. A thin, rigid insulating layer was used to separate two metal plates called electrodes. The gap between the electrodes was small enough to allow the electron clouds associated with the electrodes to overlap slightly. A potential difference between the electrodes, induced by applying voltage to them, causes electrons to flow from one electrode to the other through the overlapping clouds. The flow is analogous to the flow of groundwater between two adjacent lakes when one lake is higher than the other.

We built our scanning tunneling microscope by making a few basic changes in the standard tunneling configuration. First we replaced one of the electrodes with the sample we wanted to investigate. Then we replaced the other electrode with a sharp, needlelike probe. Finally we replaced the rigid insulating layer with a nonrigid insulator such as liquid, gas or vacuum so that we could scan the needle tip along the contours of the sample's surface.

To scan the surface we push the tip toward the sample until the electron clouds of each gently touch. The application of a voltage between the tip and the sample causes electrons to flow through a narrow channel in the electron clouds. This flow is called the tunneling current. Since the density of an electron cloud falls exponentially with distance, the tunneling current is ex-

tremely sensitive to the distance between the tip and the surface. A change in the distance by an amount equal to the diameter of a single atom causes the tunneling current to change by a factor of as much as 1,000.

We exploit the sensitivity of the tunneling current to produce exquisitely precise measurements of the vertical positions of the atoms on the sample's surface. As the tip is swept across the surface a feedback mechanism senses the tunneling current and maintains the tip at a constant height above the surface atoms. In this way the tip follows the contours of the surface. The motion of the tip is read and processed by a computer and displayed on a screen or a plotter. By sweeping the tip through a pattern of parallel lines a three-dimensional image of the surface is obtained. A distance of 10 centimeters on the image represents a distance of 10 angstroms on the surface: a magnification of 100 million.



ELECTRON TUNNELING is the phenomenon that underlies the operation of the microscope. An electron cloud occupies the space between the surface of the sample and the needle tip (*bottom*). The cloud is a consequence of the indeterminacy of the electron's location (a result of its wavelike properties); because the electron is "smeared out," there is a probability that it can lie beyond the surface boundary of a conductor. The density of the electron cloud decreases exponentially with distance. A voltage-induced flow of electrons through the cloud is therefore extremely sensitive to the distance between the surface and the tip. As the tip is swept across the surface a feedback mechanism senses the flow (called the tunneling current) and holds constant the height of the tip above the surface atoms (*top*). In this manner the tip follows the contours of the surface. The motion of the tip is read and processed by a computer and displayed on a screen or a plotter. Sweeping the tip through a pattern of parallel lines yields a high-resolution, three-dimensional image of the surface.

How is it possible to move the needle over a sample while maintaining a gap between the tip and the surface that is less than 10 angstroms and achieve a stability and precision that is better than .1 angstrom? First, the microscope must be shielded from vibrations such as those caused by sound in the air and by people walking around in a building. Second, the drives of the needle must be highly precise. Finally, the tip must be as sharp as the limits of rigidity and stability allow.

Two stages, or sections, suspended from springs, nestle within the stainless-steel cylindrical frame of the microscope and protect the tunneling gap from vibration. Both stages, triangular in cross section, are made of glass rods. The second stage slips into the first stage, from which it is suspended by three springs. The first stage in turn is suspended from the outer frame, also by three springs. The second stage carries the heart of the microscope: it contains both the sample and the scanning needle.

When the entire microscope sits in a vacuum, air resistance is minimal and the first and second stages could, if they were disturbed, bounce up and down almost indefinitely. To stop this motion we rely on the phenomenon of eddy-current damping. We let copper plates attached to the bottom of the first and second stages slide between magnets attached to the outer frame. As each plate slides up and down, the magnetic field causes the conducting electrons of the copper to move around, inducing a so-called eddy current. The reaction between the eddy current and the magnetic field retards the motion of the plate and thereby protects the microscope from even the smallest vibrations.

Once the gross vibrations have been stopped the sample can be positioned. This is done with a specially developed drive that carries the sample across a horizontal metal plate on the second stage. The body of the drive consists of a slab of piezoelectric material that expands or contracts when voltage is applied. The drive has three metallic feet, arranged in triangular fashion, that are coated with a thin layer of insulating material. They can be clamped to the metal plate by establishing a voltage between them and the metal plate.

We move the drive in the following manner. Suppose, for instance, we clamp only one foot and apply a voltage to the piezoelectric body so that it contracts. The other two feet will move slightly. We then clamp those two feet, release the third foot and remove the applied voltage so that the body expands back to its original size. The drive has just moved one step. The step width can be varied between 100 and 1,000 angstroms. Since the drive can rotate about each of its feet, it can walk along the plate in any desired direction.

When the drive has carried the sample to the wanted tunneling position, we begin scanning the surface of the sample. We use a rigid tripod made of three piezoelectric sticks to move the tip of the scanning needle. When we apply a voltage to expand or contract one of the sticks, the other two bend slightly. Consequently the tip moves in a straight line over distances as great as 10,000 angstroms. Furthermore, this motion is quite sensitive to the magnitude of the applied voltage: a voltage on the order of .1 volt results in a motion of 1.0 angstrom. The precision of the tripod's drive is so good that at present only vibration limits the vertical resolution of the sample's surface. This resolution at present is in the

range of approximately a few hundredths of an angstrom.

The lateral resolution of the surface is limited by the sharpness of the tip. In this instance nature has been kind to the vacuum tunneler. It is relatively easy to make a sharp tip that yields a lateral resolution of about six to 12 angstroms: one simply grinds the end of a needle, which is usually made of tungsten.

To achieve a lateral resolution of two angstroms, however, the needle must have a single atom sitting securely on top of its tip. Such an atom usually comes from the sample itself. It is dislodged by high electric fields that are caused by applying a voltage difference of from two to 10 volts between the sample and the tip. Since luck plays a large role in the final stage, we are trying to sharpen the tip by bombarding it with a high-energy beam of ions. This causes the atoms on the surface to sputter away in a highly controlled manner.

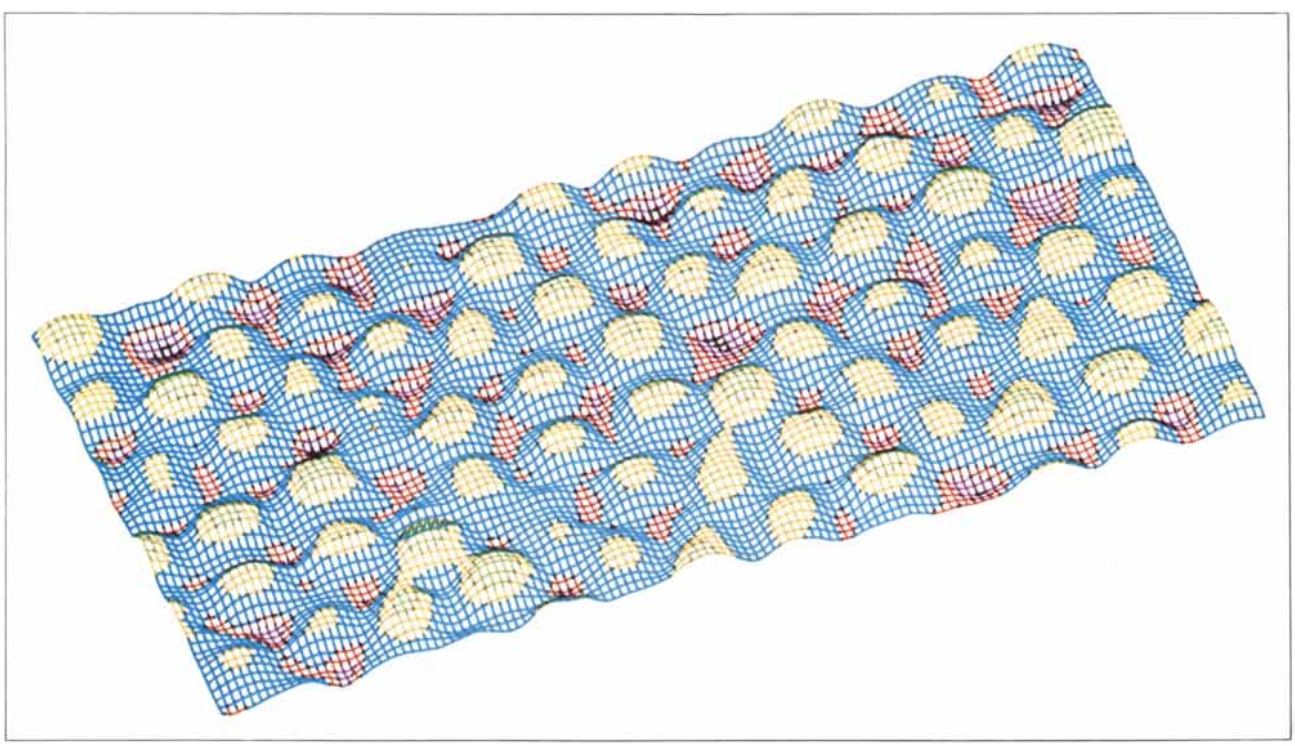
In addition to delineating the atomic topography of a surface, the scanning tunneling microscope reveals atomic composition. The tunneling current depends both on the tunnel distance and the electronic structure of the surface and on the fact that each atomic element has an electronic structure uniquely its own.

The ability of the microscope to resolve both topography and electronic structure will make it useful to investigators in physics, chemistry and biology. We first pursued the simplest case: the topographic structures of single crystals characterized by a homogeneous surface structure. Crystals consist of identical atomic layers built one on top of another. While results from scattering experiments indicate that the top layer is different from and more complex than the others, the precise structure of this layer was hard to determine.

The best-known surface structure is the diamond-shaped unit cell of silicon. Since each of the four edges of the cell measures seven atomic spacings, the cell is referred to as the 7-by-7. Each 7-by-7 contains 12 bumps that have not been visualized before. Each bump apparently corresponds to a single atom. The arrangement of the surface atoms is, although aesthetically pleasing, quite complex. This is in contrast to the relatively simple structure of any bulk layer found in silicon. Its unit cell, 49 times smaller in area than the 7-by-7, contains only two atoms. Another great difference between the two kinds of layers is that the surface layer is much rougher than any bulk layer. Although the surface pattern is now known and a vast amount of information about it has been gained from other experiments, the reason this and not a different structure forms is not yet understood.

Another crystal whose surface structure is now better understood is the gold crystal. We found that when we cut the crystal in a direction parallel to its atomic layers, the resulting face is smooth. A cut in a direction diagonal to the atomic layers results in a rougher face. Just as one learns from studying the earth's crust how it was formed millions of years ago, so we have learned from studying these surfaces how they took shape. Current theories reveal that the diagonally cut surface assumes its jagged nature because such a configuration has a lower energy and is consequently stabler than a smooth configuration.

A more exotic branch of physics, the study of superconductivity, has also benefited from the application of scanning tunneling microscopy. A superconducting material is characterized by its complete lack of electrical resistance. The use of superconductors to make cables that are free from power losses could save enormous amounts of energy. The colliding-beam acceler-



SURFACE OF SILICON as disclosed by the scanning tunneling microscope consists of a pattern of diamond-shaped unit cells. Each cell measures 27 angstrom units (one angstrom unit is one ten-billionth of a meter) on a side. The cell is called the 7-by-7 because each side measures seven atomic units. Each 7-by-7 contains 12 bumps that are arranged in two groups of six. The bumps, which have never before been resolved, apparently correspond to the surfaces of individual atoms. They stand as much as 1.3 angstroms above the rest of the surface. The image was formed by applying a voltage so that electrons flowed from the needle tip to the surface.

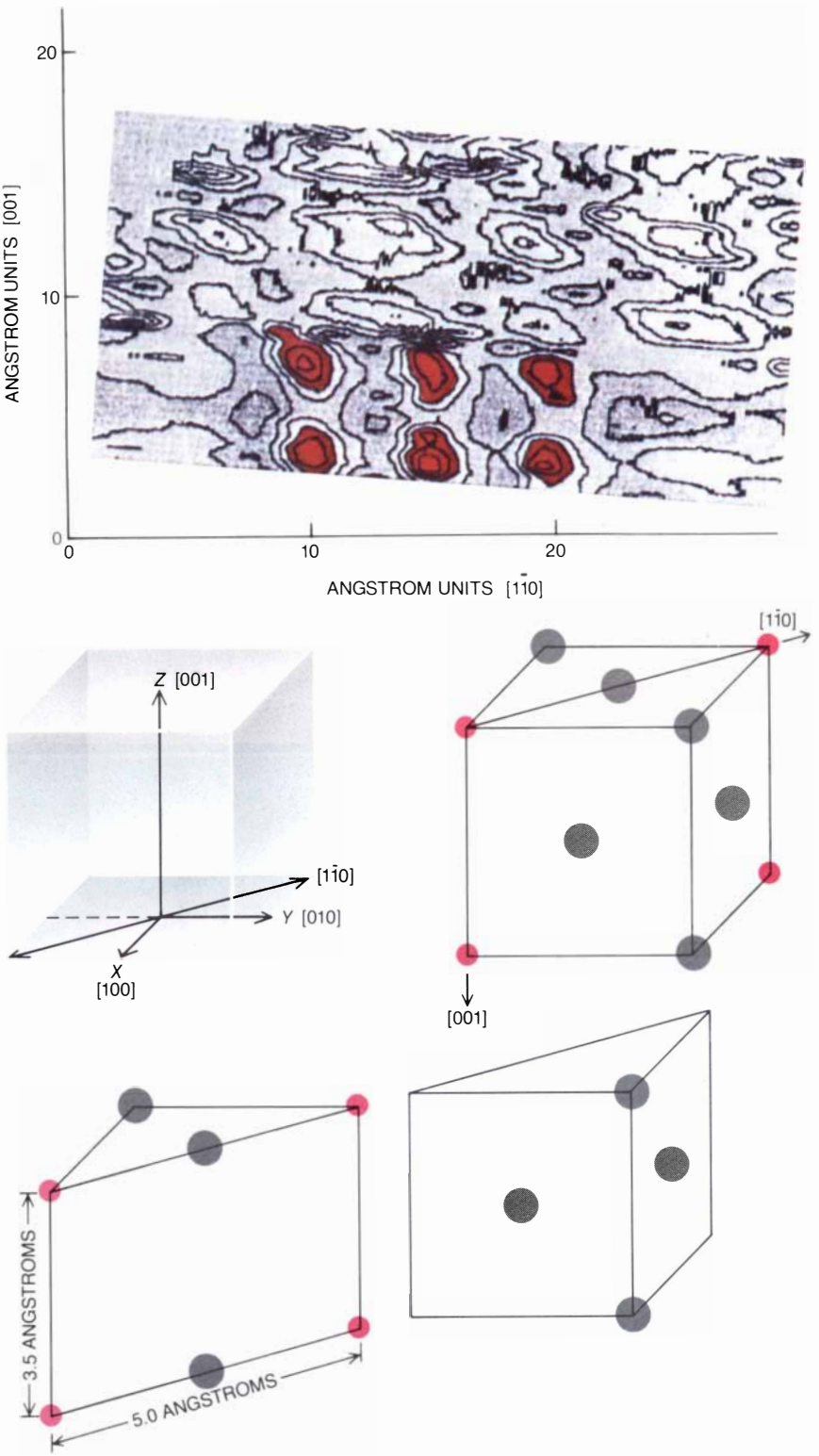
ator at Fermilab uses superconducting magnets to achieve high magnetic fields while saving energy. There is a catch, however. Superconductivity is only known to occur in some conductors that have been chilled below a critical temperature, typically a few degrees above absolute zero (−273 degrees Celsius).

A group of Stanford University investigators led by Calvin F. Quate has developed a scanning tunneling microscope that operates effectively at low temperatures. The workers first used their microscope to map the electronic structure of the surfaces of several conductors at room temperature. Then they chilled the conductors below the critical temperature of each and recorded the changes in electronic structure. The group can now document the growth of regions of superconductivity on surfaces.

The scanning tunneling microscope has also led to new understanding of certain chemical interactions. Our group has now observed on an atomic scale the adsorption of oxygen by nickel [*see illustration at right*]. Our finding confirms the results from earlier scattering experiments: The spacing of the oxygen atoms bound to the nickel surface varies according to direction. In particular, oxygen atoms that lie in one specific direction, designated [001], are separated by one lattice spacing, or the distance between two adjacent nickel atoms in that direction. Oxygen atoms that lie in another direction, designated [1$\bar{1}$0], are separated by two or five lattice spacings but never by one, three or four. We suspect that some kind of screening effect between the electric charges of the nickel and oxygen atoms is responsible for the anomaly, but more investigation is needed to determine the actual details of the physical interaction.

All the applications we have so far discussed have hinged on the ability of the microscope to detect structures whose dimensions are measured in mere fractions of an angstrom. Such high resolution is not always necessary. Even where the resolution of the scanning tunneling microscope is no better than some tens of angstroms, we expect on the basis of previous results that it will yield novel information and stimulate significant progress. In particular the possibility of operating the scanning tunneling microscope in air at ordinary pressure will in many applications more than compensate for any loss in resolution.

One such application is found in the study of friction. In order to minimize friction energy losses, investigators are



OXYGEN ADSORBED ON NICKEL (*top*) **is observed on an atomic scale. The oxygen atoms** (*color*) **are 3.5 angstroms apart in one direction, designated [001], and 5.0 angstroms apart in the other, [1$\bar{1}$0]. The face-centered cubic model of the nickel crystal** (*bottom*) **suggests the reason. The geometry of the model dictates that if the lattice spacing between two adjacent nickel atoms in the [001] direction is 3.5 angstroms, the spacing in the [1$\bar{1}$0] direction must be 2.5 angstroms. The electronic repulsion between two oxygen atoms is too great, however, to allow them to rest stably at lattice points separated by a mere 2.5 angstroms. Therefore oxygen atoms lying along the [1$\bar{1}$0] direction must be separated by at least two lattice spacings, which corresponds to the observed 5.0-angstrom distance. Separations of five lattice spacings are sometimes seen, but never separations of three or four. Additional investigation may yet disclose the underlying reasons for the anomaly.**

interested in learning more about the structure and causes of surface roughness of industrial materials. Recent studies suggest that the scanning tunneling microscope is ideally suited to the required work.

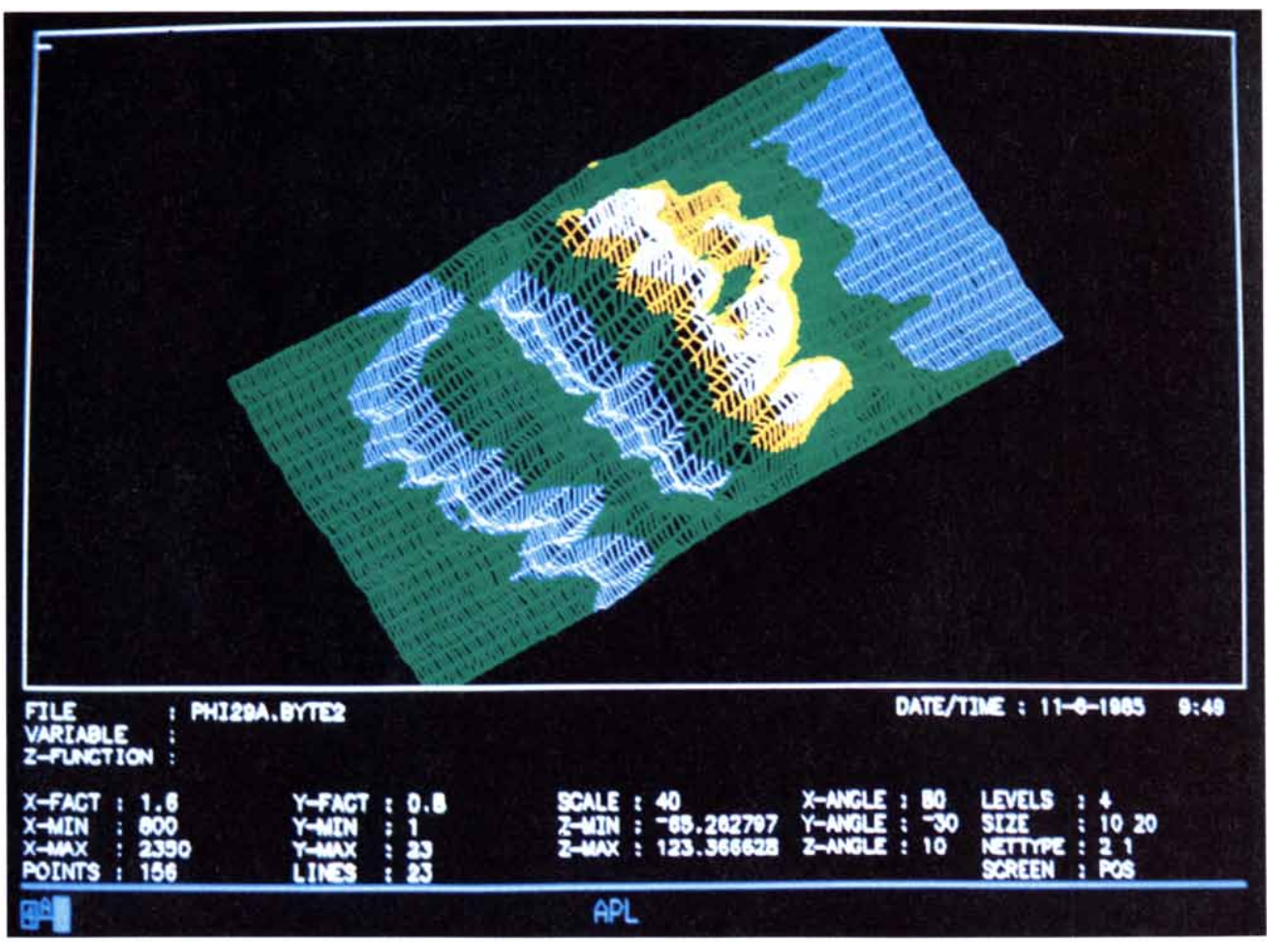
Our microscope has also demonstrated its usefulness in biology, even though at present it can achieve lateral resolutions of only 10 angstroms. In this case the relatively poor resolving power of the microscope is more than compensated for by its ability to provide a direct and nondestructive method of viewing biological samples.

Other microscopes in some sense partially destroy the samples on which they have been focused. In standard electron microscopy, for instance, samples must be coated with a thin layer of metal and, because they must be studied in a vacuum, they dry out. Since water molecules are an important part of biological substances, this might change the samples in an undesirable—and uncontrollable—way. In the scanning tunneling microscope water can even be used as the insulating layer between the sample and the probing needle. (Water is a relatively poor conductor unless it contains ions such as those formed when sodium chloride dissolves in it.) Exploiting the sensitivity of the scanning tunneling microscope, we have, with the help of E. Courtens of the IBM Zurich Research Laboratory and H. Gross and J. Sogo of the Swiss Federal Institute of Technology, scanned the surface of the nucleic acid DNA. We observed a series of zigzags corresponding to its helical structure.

In a collaborative effort with Arturo Baró, Nicolas Garcia and Rodolfo Miranda of the Autonomous University of Madrid and José L. Carascosa of IBM Spain, we found that the head of the virus known as phi 29 measures $400 \times 300 \times 200$ angstroms. The structure of the connection between the head and the tail of the virus, called the collar, which appears to play a significant role in the process of infection, has been unraveled; the results agree with those obtained by means of image-processed electron micrographs.

Apart from imaging, the probe tip will also be useful for testing electronic circuits. As components continue to shrink, the probes that test them must also be continuously miniaturized. The tip then serves as both a local voltage probe and a current source.

In all the foregoing applications it is vital that the imaging process not destroy or even alter the object. But the scanning tunneling microscope also offers promise as a tool for spurring specific chemical processes. One of the unique features of the microscope is its highly focused low-energy electron beam, or tunnel current. The energies of the beam lie within the range of those of most chemical processes. Therefore by tuning the beam to specific energies workers can cause particular reactions to occur. This mode of operation and the other capabilities of the instrument appear to open an entire new gamut of investigative possibilities.



COLLAR OF VIRUS PHI 29 connects the head of the virus to its tail. Raw, unprocessed electron micrographs such as this one have aided in unraveling the structure of the collar, an understanding of which is critical in controlling the spread of some viral infections.

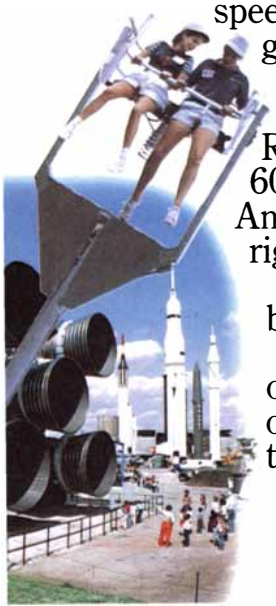
COME FACE TO FACE WITH THE SPACE PROGRAM.

Climb inside a real space capsule. Pilot a spacecraft by computer. Leap 15 feet high with the zero-gravity machines. Fire a rocket engine or lasers at the speed of light. Walk among giant rockets. And even feel the sensation of liftoff.

At The Space and Rocket Center, more than 60 hands-on exhibits put America's space program right at your fingertips.

And that's just the beginning.

From there, you're off to enjoy the thrills of our spectacular Spacedome theater.



Engulfed by a wraparound screen and 48 speaker sound system, you'll travel aboard the space shuttle in our breathtaking new film, "The Dream is Alive."

Then comes the real stuff — A guided tour of the adjoining Marshall Space Flight Center, where you'll see the ongoing development of the first NASA Space Station.

Plus, you'll also witness several training and testing operations, often manned by real astronauts.

It's a full day of exciting, enlightening adventures.

And best of all, it's just hours away in Huntsville.

So, plan to Come Do NASA. And do it soon.

NASA

THE SPACE AND ROCKET CENTER.

Huntsville, Alabama

For more information about NASA or Huntsville, call: **800-633-7280**. Ask for Operator 8.

SCIENCE/SCOPE

A new technique may expand the use of lasers in commercial and military applications. The approach, called optical phase conjugation, is considered a major advance in optics because it offers a solution to distortion problems that have limited the use of lasers. When a laser beam passes through a turbulent atmosphere or a severely strained optical component, the beam is distorted and the information it carries is degraded. Hughes Aircraft Company's technique, however, forces the laser to retrace its path through the distorting medium so the beam emerges free of distortion. The method eliminates the need for complex electro-optical and mechanical components to correct the distortions.

A future generation of infrared "eyes" for space surveillance systems will be far more capable as the result of technology advances at Hughes. These systems will be able to see distant targets in space, in the air, or on the ground—and relay data instantly to ground stations. Advances are being made in focal plane design, signal processing architecture, and in the design of a unique sensor with very steady telescoping optics. By building modularity and programmability into the new technologies, researchers are making it possible for systems to use tailored combinations from a single family of hardware and software. For its advances, Hughes received an Award for Technical Achievement from the Strategic Technology Office of the Defense Advanced Research Projects Agency. This effort was sponsored by DARPA and monitored by the U.S. Air Force Space Technology Center.

A laser that won't cause blindness or other eye injuries will be used in a rangefinder now under development by Hughes for the U.S. Army. The lightweight device, designated the AN/PVS-6 Mini Eyesafe Laser Infrared Observation Set (MELIOS), resembles a binocular case. Its neodymium yttrium aluminum garnet laser beam is sent through a chamber, or cell, filled with high-pressure methane gas. There the 1.06-micron wavelength is transformed into a wavelength of 1.54 microns. The new signal is safe because it never reaches the retina, but instead is absorbed in the vitreous humor, the white area of the eye between the retina and the lens. MELIOS is being developed under a competitive contract from the U.S. Army Night Vision and Electro-Optics Laboratory.

Large ceramic circuit boards will be built into the Amraam missile to help keep the missile reliable, lightweight, and low in cost. These boards, measuring 5x7 inches, are used instead of standard printed wiring boards where many components must be crammed into a small space and where a lot of heat must be dissipated. The cost of these circuit boards has been lowered significantly by replacing gold circuits with copper. The boards are manufactured by a thick film process in which layers of copper and glass dielectric are alternately applied to provide a multilayer circuit board. Hughes designed and developed the advanced medium-range air-to-air missile for the U.S. Air Force and Navy. The manufacturing facility is located in Tucson, Arizona.

Excellence in communications systems engineering has placed Hughes in a leading position in many of the major U.S. Air Force, Army, Navy, and Marine Corps communications programs, including PLRS, PJH, MILSTAR, and JTIDS. Our Communications Systems Division is committed to meeting the strategic and tactical communications requirements of the 1990s and beyond. We have a continuing need for qualified engineers in all communications systems disciplines at all levels. If your career goals include design of advanced antijam communications systems or HF through millimeter-wave radios, please send your resume to Hughes Ground Systems Group, Professional Employment, P.O. Box 4275, Dept. S2, Fullerton, CA 92634. Equal opportunity employer. U.S. citizenship required.

For more information write to: P.O. Box 45068, Dept. 71-10, Los Angeles, CA 90045-0068

HUGHES
AIRCRAFT COMPANY

SCIENCE AND THE CITIZEN

More, or Less?

Lumpectomy (a surgical procedure that saves the breast) followed by radiation therapy is just as effective in treating breast cancer as mastectomy, according to the first large-scale randomized study comparing the two forms of therapy. In lumpectomy only the tumor and a small margin of surrounding tissue are removed, whereas in mastectomy the entire breast is removed. About 85 percent of women with operable breast cancer currently undergo mastectomies and about 15 percent undergo lumpectomies. The results of the recent study, however, suggest that the less disfiguring procedure is appropriate for about half of all women with operable breast tumors.

The study was carried out by the National Surgical Adjuvant Breast Project, a cooperative group that includes physicians from many institutions. From 1976 through 1984 about 2,100 women were enrolled in the study from medical centers throughout the U.S. and Canada. The women were randomly assigned to receive mastectomy, lumpectomy with radiation or lumpectomy without radiation. After surgery the groups were followed and actuarial methods were employed to determine the probability of survival and of the reappearance of cancer. Results for the first five years after surgery were published in a recent issue of the *New England Journal of Medicine.*

Only women with Stage I or Stage II breast cancer and tumors less than four centimeters in diameter were eligible for the study. The clinical stages of the disease are determined by tumor size and other factors, including the number of lymph nodes in the armpit that are found to contain cancer cells. About 50 percent of women with operable breast cancer meet the criteria of the project. Among such women, the study shows, five years after lumpectomy and radiation therapy the probability of survival is slightly more than 80 percent; five years after mastectomy the probability is slightly more than 70 percent.

After five years about 70 percent of the women who had received lumpectomies and radiation were free of cancer in the lymph nodes, skin and chest muscles around the breast that had been operated on; the figure was about 65 percent for those who had received mastectomies. Slightly more than 70 percent of those who had lumpectomies and radiation were free of cancer outside the region of the breast that had been operated on; for those who underwent mastectomy the figure was slightly less than 70 percent. Lumpectomy with radiation yielded better results than lumpectomy without radiation therapy, particularly in preventing a recurrence of the tumor in the same breast.

The results of the cooperative study could contribute to a fundamental change that is already under way in the understanding and treatment of breast cancer. It was once thought that breast cancer is disseminated by the passage of cancer cells through the lymphatic system (along with tumor growth); if the diseased lymph nodes were removed along with enough breast and muscle tissue, the spread could be checked. It is now thought the cancer spreads in a more complex fashion that probably involves the circulatory system, according to Lawrence Wickerham, a member of the project staff.

Wickerham said that as long as the entire tumor is removed, more extensive operations are no more effective than less extensive ones in stopping the spread of the disease in women who meet the study's criteria. The emphasis is currently shifting to systemic forms of treatment such as chemotherapy, which can reach cancer cells that have been disseminated through the circulatory system, he added. Wickerham noted that lumpectomy entails "very minimal" disfigurement and said the women in the study had found the procedure "most satisfactory" cosmetically. The cooperative study is being continued in order to provide results for up to 10 years after the surgical procedure.

Illegal Ecstasy

Street chemists have a new product called Ecstasy. The drug is an amphetamine analogue, 3,4-methylenedioxymethamphetamine, or MDMA, and it has recently become popular in several parts of the U.S. Its name describes the state of mind its users say it produces: a feeling of euphoria, a lack of concern about social barriers and a strong sense of rapport with other people. At the University of Chicago Medical Center, Charles Schuster and Lewis Seiden, who are respectively head of the drug-abuse research center and professor of pharmacological and physiological sciences, believe the drug may have very dangerous side effects. Chemically similar drugs they have studied, Schuster and Seiden say, can damage brain cells containing the neurotransmitters serotonin and dopamine. Serotonin is involved in the control of such functions as sleep, sexual behavior, mood, aggression and sensitivity to pain. Dopamine has a number of biological effects, including a role in initiating movement.

Schuster and Seiden had been working with two drugs only slightly different in chemical structure from MDMA: 3,4-methylenedioxyamphetamine, or MDA, and methamphetamine. They found that even a single dose of MDA seriously depletes levels of serotonin in the brain and that the deficit persists for at least two weeks. They also found that methamphetamine can cause the degeneration of nerve cells containing dopamine.

Until recently the use of Ecstasy was legal. Since it is made by street chemists and not by a pharmaceutical manufacturer, the agent was not controlled by the laws governing prescription drugs. The U.S. Drug Enforcement Administration has now closed that loophole by declaring Ecstasy illegal. Meanwhile Schuster and Seiden have begun studies to determine whether MDMA does indeed have neurotoxic effects similar to those of its analogues.

Transplacental Mediation

The most devastating effects of the rubella (German measles) virus on an unborn child are recognizable at birth or soon thereafter. What about the long-term effects of in utero exposure to this viral infection and others?

To answer the question a team of investigators at the London School of Hygiene and Tropical Medicine analyzed data from a long-term study of 3,076 individuals, up to the age of 40, who were known to have been exposed to a viral infection in utero. P. E. M. Fine, A. M. Adelstein, J. Snowman, J. A. Clarkson and S. M. Evans report in the *British Medical Journal* that people exposed prenatally to chickenpox virus (varicella zoster) or cytomegalovirus had a slightly increased risk of cancer. Both viruses belong to the herpes group, which has long been suspected of having major carcinogenic potential.

Among individuals whose mother had mumps during the first trimester of pregnancy the investigators found a small cluster of diabetes cases. The finding is consistent with a known association of childhood mumps and the subsequent onset of acute diabetes. Exposure to the rubella virus was not

Where engine

The more you rely on sophisticated graphics, the more you need a Smart Desk equipped with an IBM 3270 Personal Computer/GX.

The 3270 PC/GX gives you easy access to host-based graphics and data. And, with IBM's new Graphics Editor and Picture Plotting software, you can work offline locally to draw your own charts, diagrams and free-form sketches—all with excellent resolution.

You can actually create or edit directly on the screen, using either a mouse or a tablet. What's more, you can select from a wide range of colors, lines and type styles to arrive at a design or presentation that works best for you.

If you want to make foils for a meeting or reproduce your design sketches, just zoom and scroll to position the graphics you want to reproduce. Then use IBM's Picture Plotting to plot the results. You also get a handy graphics index, so you can file and retrieve pictures at your convenience.

But there's more to the 3270 PC/GX than just interesting pictures. It can display up to four host graphics sessions, one PC session and two notepads concurrently. And the keyboard has a variety of ergonomic features including special function keys and a numeric keypad.

There's also a smaller model of the 3270 PC with advanced color graphics—the 3270 PC/G. Whether you choose the G or the GX, you're sure to find it a graphic example of the state of the art.

IBM 8-85
DRM, Dept. KW/428
400 Parson's Pond Drive
Franklin Lakes, NJ 07417

☐ Please send me literature on the IBM 3270 PC/G and GX.
☐ Please have an IBM marketing representative contact me.

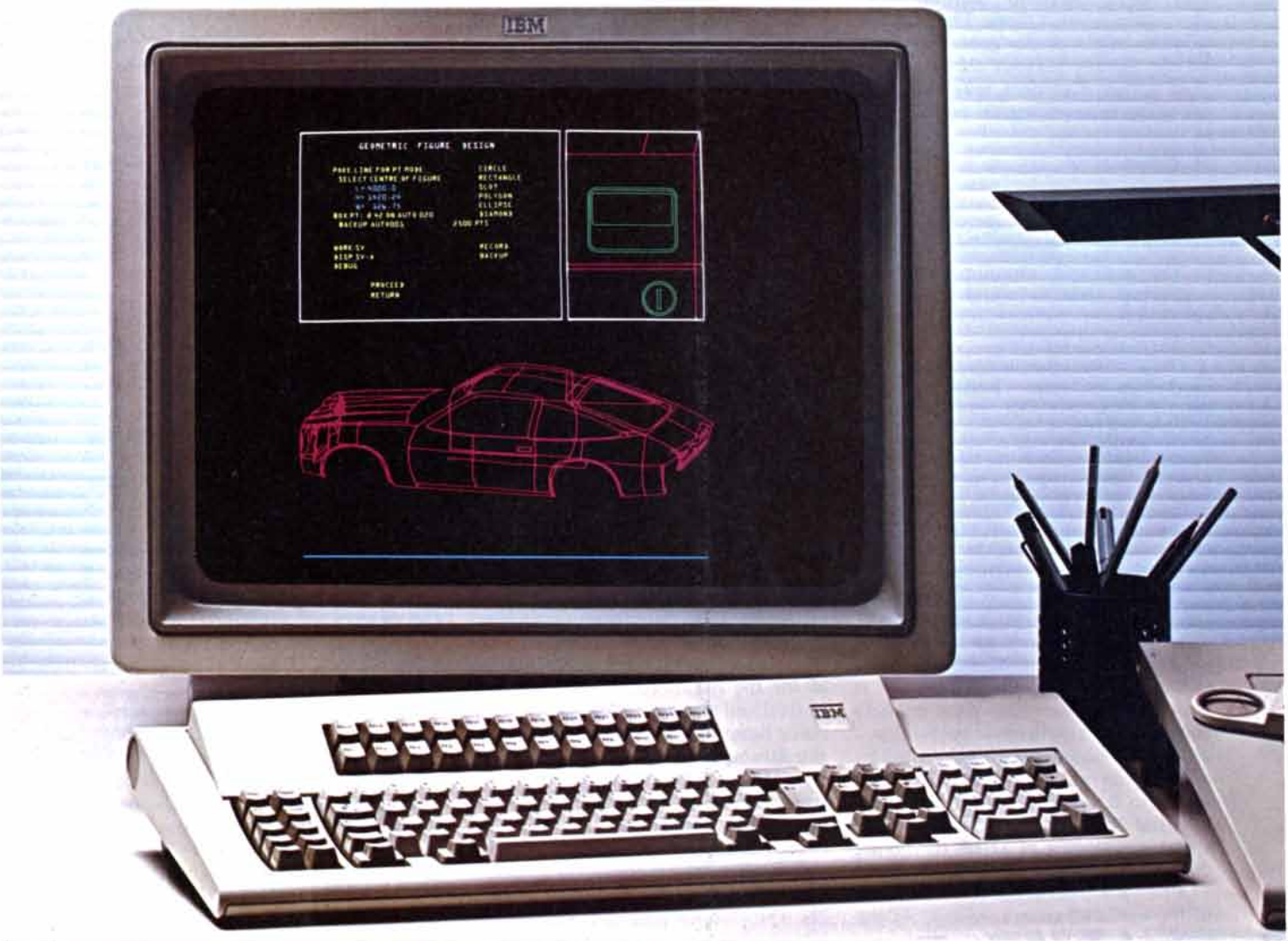
Name_____
Title_____
Company_____
Address_____
City_____ State_____ Zip_____
Phone_____

To receive literature, call 1 800 IBM-2468, Ext. 428, contact your IBM marketing representative or send in this coupon.

IBM

The Smart Desk from IBM

ers practice the art of decision making.



found to be associated with increased risk of cancer or diabetes.

It may be that not all the effects of viral exposure in utero are harmful. Individuals whose mother had chickenpox during pregnancy appeared to experience appreciably fewer diseases of the skin, subcutaneous tissue and nervous system later in life than unexposed controls did. Fine's group speculates that some autoantibodies made by an infected mother against her own virus-infected skin and neural tissue may possibly cross the placenta; the fetus might develop "anti-idiotype" antibodies against them. The fetal antibodies might confer protection later in life against various skin and nervous-system disorders, many of which are of an immunological nature.

Retrovaccine

The introduction into commercial use earlier this year of a vaccine against a feline disease with parallels to the acquired immune deficiency syndrome (AIDS) could ultimately contribute to the prevention of the deadly human ailment.

The vaccine immunizes cats against one form of feline leukemia virus (FeLV). Like HTLV-III, the infectious agent of AIDS, FeLV is a retrovirus. Retroviruses, which insert their own genetic material into the genome of the host, are known to cause a variety of tumors in animals and one form of cancer in human beings. The FeLV vaccine, which is manufactured by Norden Laboratories, is the first successful vaccine against a retrovirus.

FeLV infects from 1 to 3 percent of all domestic cats; in households with more than one cat the proportion is about 30 percent. The virus particle consists of a protein envelope wrapped around a dense core that contains a molecule of the genetic material RNA. In the host cell a viral enzyme uses the RNA as a template to construct a molecule of DNA (reversing the usual process of transcription). The DNA is inserted into one or more of the host-cell chromosomes. At the same time the virus usurps the host's protein-making machinery to manufacture new virus particles.

In spite of the virus's name, few infected cats die of leukemia. The virus can, however, cause death two other ways. The viral DNA can combine with other genes to cause tumors. In addition a protein called *p15E* from the viral envelope suppresses the cat's immune response. Like AIDS victims, cats infected with FeLV frequently die of opportunistic diseases that would be resisted by a healthy immune system.

The FeLV vaccine is based on work done by Richard G. Olsen of Ohio State University. Olsen worked with a cell line designated FL-74, which is derived from a tumor caused by FeLV. Ordinarily FL-74 cells contain intact and infectious virus particles. Olsen found, however, that when the tumor cells are transferred to a poor medium after being grown in a rich one, they shed viral antigens (proteins) and tumor antigens but no live virus particles. Hence the vaccine protects against tumor formation and against viral infection without causing infection or promoting tumor growth.

Furthermore, the FL-74 cells shed no *p15E* in active form and therefore the vaccine does not suppress the immune response. In that respect it does what an AIDS vaccine would have to do. Indeed, Jeffrey Laurence of the Cornell University Medical Center noted that the overall process of FeLV infection provides "a pretty good model" for AIDS. Laurence cautioned, however, that knowledge gained in formulating the feline vaccine cannot be applied directly to the search for an AIDS vaccine, which is under way in several laboratories. HTLV-III suppresses immune function by a mechanism much more complex than that of FeLV. In addition the fact that the AIDS pathogen has a much higher mutation rate than the feline virus could enable it to alter its antigens and avoid the effect of a vaccine.

Grace under Pressure

Silicon is among the most protean of solids. In its pure form it is an insulator. The addition of impurities turns it into a semiconductor: a basic material for the manufacture of solid-state optical and electronic devices. At pressures hundreds of thousands of times the atmospheric pressure at the surface of the earth it undergoes a further transformation, one that makes it a conductor, specifically a metal. Indeed, it undergoes a sequence of transformations, at least one of them unparallelled in previous studies of materials. The metallic phases of silicon may all turn out to be superconductors. Two of them already have.

The metallic sequence begins at a pressure of 110 kilobars, or roughly 1.6 million pounds per square inch. At that point silicon abruptly assumes a structure that makes it similar to tin—the beta form of tin. (The alpha form is only a semimetal.) The transition is striking. The silicon sample is compressed in the jaws of a diamond anvil cell. At 110 kilobars the sample becomes opaque, shiny and metallic. At a temperature of six degrees Kelvin (six degrees Celsius above absolute zero) the metal becomes superconducting: it offers no resistance to the passage of electrons.

At a pressure of about 130 kilobars the beta-tin form of silicon transforms itself into a phase discovered only last year. Designated the primitive hexagonal phase, it consists of hexagons of silicon atoms stacked on top of each other. No one had predicted that such a phase might exist in the crystal of a chemical element. In the wake of the discovery Marvin L. Cohen, KeeJoo Chang and Michel M. Dacorogna of the University of California at Berkeley undertook a theoretical calculation of the properties of the phase. They found that the bonds linking the atoms in each of the planes defined by the hexagons should be weaker than the bonds linking atoms in adjacent planes. Hence the electronic charge distribution should be inhomogeneous along one dimension. Such inhomogeneity is a hallmark of good superconductors; in effect it provides corridors through which electrons can move.

Intrigued, the Berkeley group calculated the properties of phonons: the quantized vibrations of the atoms in the crystal lattice of the phase. In quantum mechanics a phonon can be treated as a particle, one that interacts with electrons. The stronger the interaction, the better the chances for superconductivity. Too strong a coupling, however, and the integrity of the lattice collapses: a structural phase transition occurs. The calculations were so lengthy that they required one of the world's fastest computers, the CRAY/X-MP computer at the National Magnetic Fusion Energy Computer Center of the Lawrence Livermore National Laboratory.

The calculations indeed supported a prediction of superconductivity. Seeking experimental verification, Cohen consulted with Gérard Martinez of the French National Center for Scientific Research (CNRS). Martinez, in collaboration with J. M. Mignot and G. Chouteau, polished a crystal of *p*-type silicon (a kind employed in transistors) to a thickness of 25 micrometers and cut it to a length of .8 millimeter and a width of .08 millimeter. Placed in a diamond-anvil cell, it was subjected to pressures as great as 250 kilobars. The attainment of the pressure required for the transition from the beta-tin phase to the primitive hexagonal phase proved to be accompanied by a sudden upward jump of about one degree in the temperature required for the onset of superconductivity. The maximum superconducting temperature, about eight degrees Kelvin, came at a pressure of about 150 kilobars.

At a pressure of some 400 kilobars,

F R E S H

A I R

F U N D

GIVE THEM LOTS OF SUN AND WATER AND WATCH THEM GROW.

For over 108 years, the Fresh Air Fund has been bringing city children to the country each summer for a fresh perspective on life.

In the country they can learn how to swim, what it's like to milk a cow, and why fireflies light up. And in the process they'll discover new ideas and feelings about themselves and the world.

Two weeks of country life cost only $114. So help give more city children a breath of fresh air this summer. Send your tax deductible contribution to the Fresh Air Fund, 70 West 40th Street, New York, NY 10018. And you'll be giving them a beautiful chance to grow.
Fresh Air Fun. It lasts a lifetime.

the primitive hexagonal phase of silicon verifies a further prediction of the Berkeley group by transforming itself into a form called the hexagonal-close-packed phase. The hexagons shift so that the edge of one hexagon lies under the interior of another. No one has yet tried to test for superconductivity, but then the experiment will be exceedingly difficult: it will demand that probes be kept in place when a tiny specimen of silicon is subjected to pressures of six million pounds per square inch and cooled to temperatures near absolute zero. The Berkeley group predicts that the superconducting temperature will rise to a value greater than 10 degrees as the pressure nears the value required for the transition from the simple hexagonal phase to the hexagonal-close-packed phase. That would make silicon a more willing superconductor than any other chemical element.

Macrowhisker

Workers at the Los Alamos National Laboratory have developed a technique for growing silicon carbide whiskers, or fibers, that are stronger and stiffer than commercially available polycrystalline fibers. The superior mechanical properties of the whiskers derive from the fact that each one is actually a single crystal. Their lengths range from half an inch to four inches; their diameter is about five micrometers (millionths of a meter). According to George Hurley, manager of the whiskers project, the fibers could reinforce ceramics: brittle materials that are more resistant to corrosion, heat and wear than metal is. The modified ceramics could offer advantages as a material from which valves, turbine blades and heat exchangers could be made.

To grow the whiskers, workers first coat plates of graphite with tiny granules of stainless steel whose average diameter is about 10 microns. Then they put the treated plates in a container with pieces of porous insulation brick saturated with silicon and carbon atoms. Finally, they heat the container and its contents to about 1,400 degrees Celsius for 10 hours. As the container heats up, a gaseous vapor of silicon and carbon forms, from which liquid droplets condense. Whiskers sprout from the droplets, lodge on the graphite plate and grow like blades of grass. The whiskers are then harvested, sorted by length and cleaned.

The Los Alamos workers are now readying the whisker-growing process for commercialization by the ARCO Chemical Company in Greer, S.C. They are also developing methods of growing single-crystal whiskers from other materials.

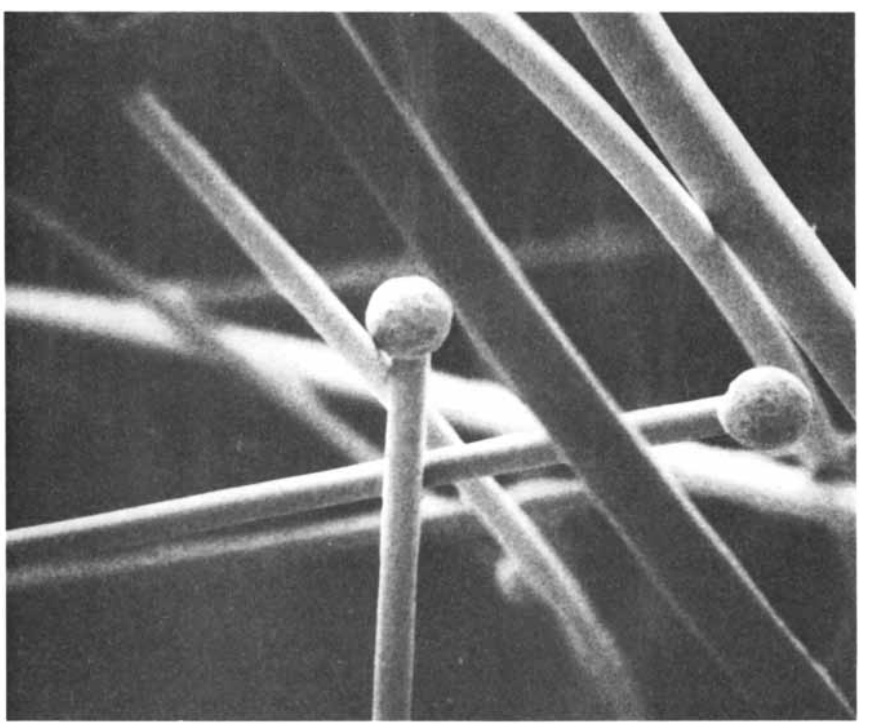
Prospecting

Thomas Gold of Cornell University, an astrophysicist of repute, is also known for his forays into other disciplines. Perhaps the most controversial of those ventures has been Gold's long-standing effort to show that the primary source of petroleum and natural gas is not organic sediments buried at the earth's surface but primordial, abiotic methane rising from deep in the mantle. Most geologists have been and continue to be extremely skeptical of the "deep-earth gas" hypothesis. This fall it will undergo a test in Sweden, a country without major sedimentary formations and lacking in "fossil fuels." The Swedish State Power Board plans to drill 5,000 meters into granite bedrock in the hope of finding deep-earth gas.

Gold believes the granite shield under Sweden blocks the upward flow of methane from a huge reservoir in the mantle; in his view methane flowing up around the edges of the shield is responsible for the oil and gas fields in the North Sea and the southern Baltic. The putative reservoir itself is beyond the reach of any foreseeable drilling technology. Where the granite is fractured and porous, however, the methane might have risen into the crust and become trapped at accessible depths.

The granite at the planned drilling site, in the 44-kilometer-wide Siljan meteorite crater in south-central Sweden, was presumably shattered by the impact of the meteorite 360 million years ago. Although the sediment layer in the region is only 300 meters thick, oil and gas seeps are observed at the surface, and test wells drilled to a depth of 700 meters in the crater have detected small amounts of methane. According to Gold, the tests have also detected an isotopically distinct type of carbonate known to be derived from the oxidation of methane and often associated with oil and gas fields. Furthermore, seismic studies have revealed strongly reflecting layers at various depths. The layers must have formed since the meteorite impact, says Gold; they might consist of shattered granite cemented together by methane-derived carbonate.

Gold's many critics advance at least three basic arguments in predicting the Siljan hole will be dry. First, according to John S. Lewis of the University of Arizona, no evidence supports the notion that a large amount of methane was incorporated in the earth when it formed; indeed, the available geochemical evidence suggests that the early atmosphere, which was produced by outgassing of the planetary interior, could not have been hydrogen-rich. Second, if the earth did once contain primordial methane or other hydrocarbons, most of this volatile material would long since have escaped through volcanism or diffusion. Third, analysis of volcanic basalts sug-



Silicon carbide whiskers after harvesting

gests the rock in the upper mantle is highly oxidizing. In such an environment one would expect methane to be oxidized to carbon dioxide.

Gold, undaunted, is confident that a large natural-gas strike at Siljan will force his critics to rethink the standard view of how oil and gas are formed. Such a find, he has written, "will make clear that gas can be found in many hitherto unsuspected locations in many countries," thereby ending the "present precarious and dangerous pattern of energy dependence."

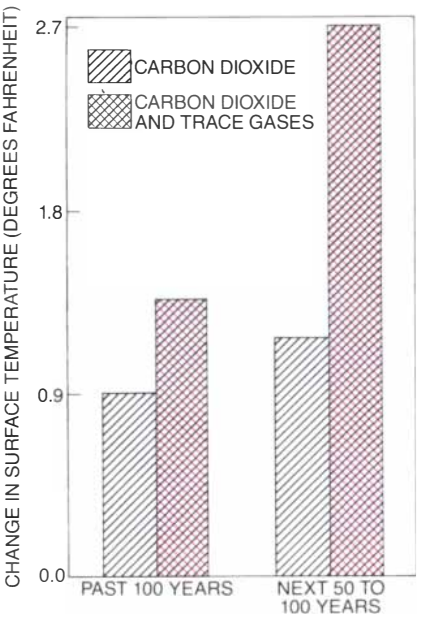
Heat Beat

For many years climatologists have warned that the increasing level of carbon dioxide in the earth's atmosphere could escalate the greenhouse effect, a warming of the earth's climate that takes place because atmospheric gases trap heat radiated from the surface of the earth. It has been estimated that the effect could raise the average global temperature by as much as nine degrees Fahrenheit in the next 30 years, melting the polar ice caps and causing extremes in drought and rainfall.

New evidence indicates that trace gases other than carbon dioxide may contribute significantly to the warming of the climate, writes Harold Johnston of the Lawrence Berkeley Laboratory in *Annual Review of Physical Chemistry.* Ralph J. Cicerone, director of the National Center for Atmospheric Research (NCAR), and Veerabhadran Ramanathan, also of NCAR, corroborate Johnston's account in a recent issue of the *Journal of Geophysical Research.* Cicerone estimates that the "trace gases can double or triple the carbon dioxide effect in the next 40 to 50 years" [*see illustration*].

The workers base their conclusions on an investigation of nearly 40 trace gases. The levels of these gases, most notably chlorofluorocarbon-11 (CFC-11), CFC-12, ozone, methane and nitrous oxide, have increased dramatically over the past 20 years. Methane and nitrous oxide are products of natural life cycles. CFC-11 and CFC-12 have been employed in aerosol propellants and refrigerator fluids and in the manufacture of industrial foam.

The NCAR workers looked for effects caused both by individual gases and by interactions among gases. They report that one molecule of CFC-11 can absorb as much infrared radiation as 10,000 molecules of carbon dioxide. They also note that an increase in the concentration of carbon monoxide may augment the amount of methane and ozone in the troposphere (the air layer closest to the earth). This effect



Potential warming of the earth's surface

would be the result of carbon monoxide's propensity to attack hydroxyl radicals, agents that break down the other two gases.

All told, the increase in trace gases and carbon dioxide should have resulted by now in a perceptible global warming. Two things may account for the fact that temperature records do not reflect such a trend. First, clouds may modulate the greenhouse effect by reflecting solar radiation. Second, the oceans may act as a sink for the added heat. Even if both mechanisms were to fail in the future, Johnston cautions against taking an "alarmist view" of the situation. He comments that whereas the climate of the earth has been unusally warm, "a cooler period is due and it's possible that the greenhouse effect can partially compensate for this."

Soil Conditioner

In many arid regions there is not enough rain to wash out of the soil the sodium salts released from minerals by weathering. The sodium makes the soil dense and impermeable to air and water, rendering it uncultivable.

A group of workers from the U.S. Agricultural Research Service has discovered a remedy for the problem. It is Sordan grass, a fairly recent hybrid of sorghum and Sudan grass. The roots of Sordan grass produce unusually large quantities of carbon dioxide. In soil that is at least moderately moist the carbon dioxide forms carbonic acid, which in turn dissolves calcium carbonates. Calcium in solution replaces sodium attached to the clay in the soil, and the released sodium can be washed out by irrigation.

The leader of the group, Charles W. Robbins, and his colleagues report their discovery in *Agricultural Research.* In collecting data on the ability of the roots of certain plants to produce carbon dioxide they noted that Sordan grass put out at least twice the amount yielded by cotton, barley, alfalfa and tall wheat grass.

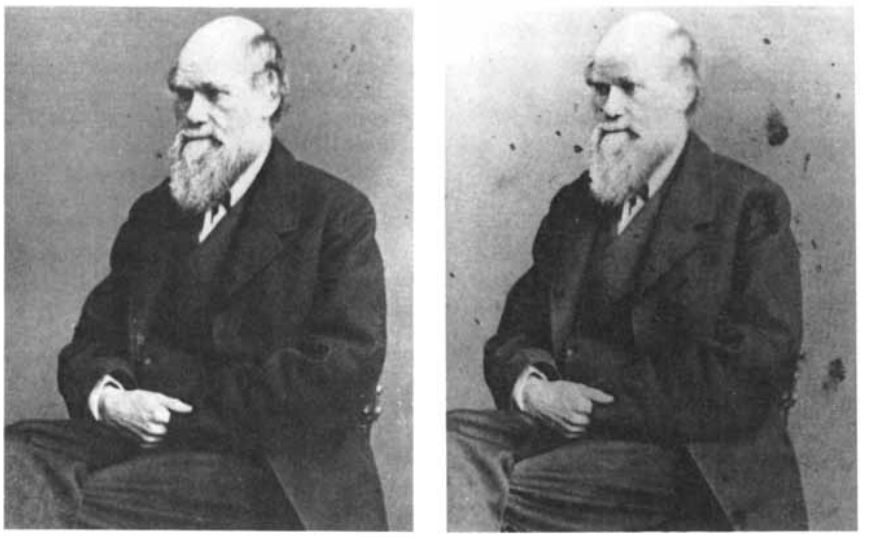
The potential of Sordan for reconditioning salty soil was immediately apparent. "We planted Sordan grass in a small salt-affected field," Robbins says, "with barely enough moisture permeability to keep a crop growing. The rancher who owned the field was unable to raise silage corn the previous year. Last fall, after one planting of Sordan, the field produced an average of about 20 tons of Sordan grass per acre, which the rancher cropped and put in his corn silage pit."

At present the standard procedure for removing salt and reclaiming marginal soil is to apply gypsum (a form of calcium sulfate) at a rate of from five to 30 tons per acre. The gypsum alone costs about $65 per ton, and the costs of bringing it in and applying it can make the procedure prohibitively expensive. Planting Sordan, Robbins points out, will both eliminate those costs and provide food for livestock.

Darwin in 3-D

In 1864, five years after the publication of *The Origin of Species,* Charles Darwin posed in the studios of the London Stereoscopic Company while three stereoscopic photographs of him were made. Such images were popular in the 19th century, and often hundreds of copies were printed; many can still be found in the collections of antique dealers. The three 1864 Darwin stereoscopic images were exhibited by the British Museum on the 100th anniversary of Darwin's birth in 1909; then, unfortunately, they seem to have been lost from view. Detective work by Gene Kritsky of the College of Mount St. Joseph in Cincinnati has now brought one of the pairs together again for the first time since the exhibition, and they are published here for the first time.

According to Kritsky, the right half of the stereo pair came to light in 1977, when James W. Valentine, a collector of Darwiniana at the University of California at Santa Barbara, bought a copy of Darwin's *Descent of Man* at a secondhand bookstall in Shropshire. Inside the front cover of the book Valentine discovered a postcard bearing Darwin's likeness. No one appreciated the significance of Valentine's good



Charles Darwin: a reconstituted stereo pair

fortune, however, until late last year. Earlier in 1984 Kritsky had identified an image held by the Hunt Institute for Botanical Documentation at Carnegie-Mellon University as the left half of a stereo pair. Valentine sent his photograph to Kritsky, and Kritsky mounted it next to a copy of the image from the Hunt Institute. "Darwin's image just jumped out of the page," Kritsky said.

To view the stereo images, bring the page close to your nose and stare at each half of the pair with one eye for about a minute. Then, as you keep your gaze fixed, move the page slowly to a position about 25 centimeters from your face. Adjust the page by small amounts until the two images seem to fuse. For those who have difficulty fusing the images a piece of cardboard can be helpful: it should be held vertically against the nose to block half of the field of vision for each eye. Inexpensive stereoscopic viewing glasses are also available.

Kritsky is convinced many more stereo pairs, as well as other uncatalogued images of Darwin, remain to be found. An exhibit of Darwin photographs, including the newly reunited pair, will begin a national tour with an opening next month at the Cincinnati Museum of Natural History.

Heart of Darkness

What lies at the center of our galaxy? Clouds of dust hide the region from view, and radio observations, although they yield a broad picture of radio-emitting structures, give little information about their dynamics. A group of physicists and astronomers from the University of California at Berkeley, led by Charles H. Townes and Reinhard Genzel, has now made detailed and extensive spectroscopic observations at infrared and far-infrared wavelengths. The data show that the gas and dust at the heart of the galaxy are gravitationally bound by a massive, pointlike central object—most likely a black hole.

The Berkeley workers, who describe their findings in *Nature,* used newly designed spectrometers to examine a region of hot, rarefied ionized gas extending out from the galactic center for about 5.5 light-years and a disk of cool, neutral gas and dust stretching for tens of light-years beyond the ionized cavity. Their objective was to determine the velocity pattern of the gas and dust; to do so they measured Doppler shifts in the emission spectrum of neon (for the ionized gas) and in the spectra of atomic oxygen, carbon and carbon monoxide (for the neutral disk). Such spectral shifts occur when a radiation source is moving toward or away from the observer.

The measured velocities show that the gas and dust are held in orbit around the center of the galaxy by the gravitational pull of a mass about four million times greater than the mass of the sun. The orbital velocity of the interstellar matter decreases with distance from the galactic center; the falloff is characteristic of orbital velocities in a system in which the central mass is concentrated at a point.

The finding of great mass in a highly compact form strongly suggests a black hole. The central mass could also consist of a cluster of stars, in which the observed gas and dust would be dispersed; in ordinary clusters, however, the velocity of matter orbiting the center does not diminish with distance. An atypical cluster in which most of the stars were densely packed at the core might give the observed gradient of velocities. But estimates of the concentration of stars at the galactic center fall well short of the needed density.

A black hole would explain not only the measured velocities but also the evidence of recent upheaval at the center of the galaxy: the low density of the ionized region and the turbulence in the gas surrounding it. The authors calculate that a quantity of gas and dust averaging between a ten-thousandth and a thousandth of the mass of the sun falls into the black hole each year, releasing considerable energy as it is accelerated and heated. The calculated release of energy is far greater than the amount that currently emanates from the region, which suggests the energy-generating infall is episodic. The cavity and the turbulence may be relics of a flare-up that occurred within the past million years as a large clump of material plunged into the black hole.

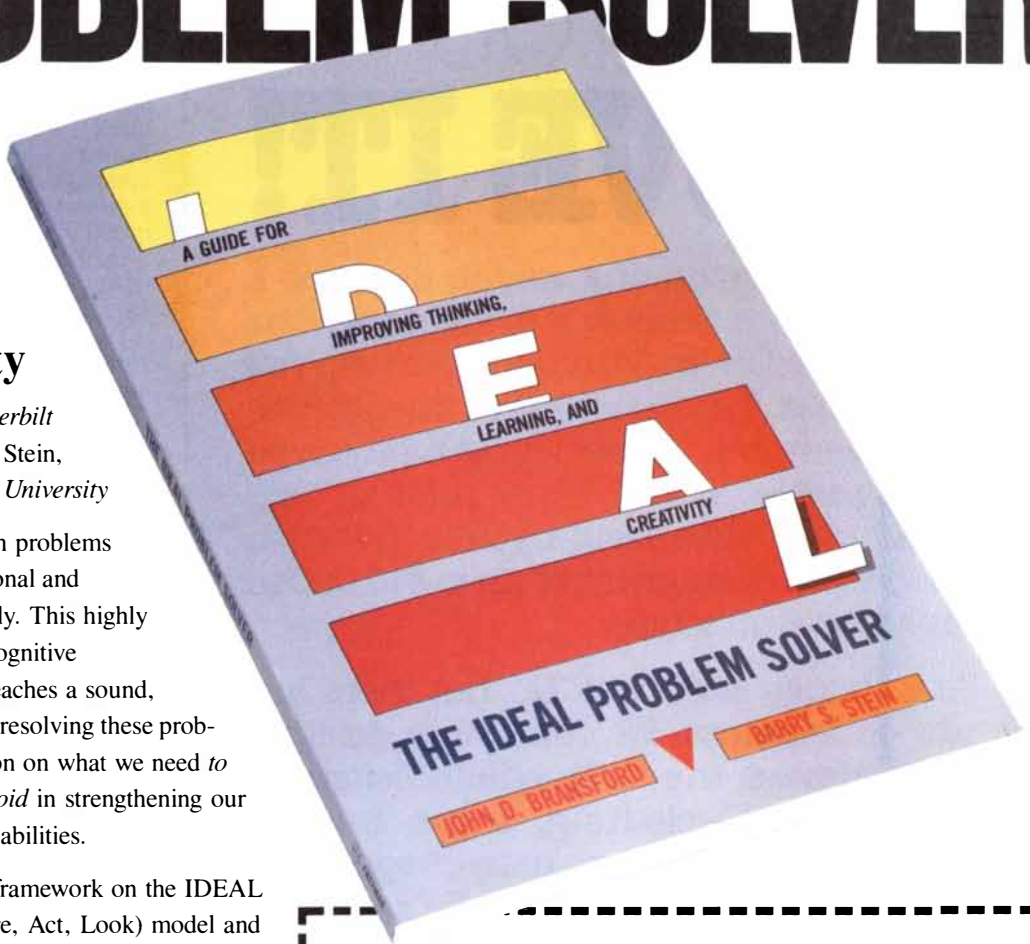
A recent radio study of the galactic center also points to a black hole there—probably but not necessarily the same one that seems to govern the motions of the gas and dust. Kwok-Yung Lo of the California Institute of Technology, Donald C. Backer of Berkeley and their colleagues used very-long-baseline interferometry and extremely sensitive radio receivers to set an upper limit on the diameter of a powerful radio source that lies at the center of the ionized cavity: about 20 astronomical units, or 20 times the distance from the earth to the sun. The source therefore approaches stellar size, but known stellar sources do not match its characteristics. The workers write in *Nature* that the evidence suggests "a single massive collapsed object at the galactic center."

Two Birds

Two outstanding observational lacunae mar contemporary astronomy: the deficiency in the observed flux of high-energy neutrinos from the sun and the lack of enough visible mass to account for the rotational dynamics of galaxies. Scores of solutions have been proposed for each problem, but many of them have had to postulate the existence of some exotic but so far undiscovered elementary particle. The longstanding failure to detect such a particle has caused some astrophysicists to lose their enthusiasm for the exercise, but now the stakes in the game have been raised. Two groups

A positive, practical "I can do it" approach to solving problems

THE IDEAL PROBLEM SOLVER



A GUIDE TO
Improving Thinking, Learning, and Creativity

John D. Bransford, *Vanderbilt University,* and Barry S. Stein, *Tennessee Technological University*

We all face and deal with problems and predicaments—personal and professional—almost daily. This highly readable book, by two cognitive psychologists, actually teaches a sound, methodical approach for resolving these problems by focusing attention on what we need *to do* as well as what *to avoid* in strengthening our natural problem-solving abilities.

The authors build their framework on the IDEAL (Identify, Define, Explore, Act, Look) model and show how potential problems both fit and are solved within this framework. They suggest new strategies for improving memory, for criticizing ideas and generating alternatives, for overcoming blocks to creativity, and for communicating more effectively with a wider range of people. Provocative, challenging, and fun, *The IDEAL Problem Solver* is liberally sprinkled with everyday examples, brain-teasing drawings, and amusing anecdotes. It is the ideal remedy for the myriad problems that confront and confound us daily.

Available at fine bookstores. 150 pages, 39 illustrations. Paper.

Please send me _____ copy (copies) of *The IDEAL Problem Solver* (ISBN 0-7167-1669-0) at $8.95 each.
☐ 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 add appropriate sales tax).
Charge my ☐ VISA ☐ MasterCard Acct. #_____
Exp. Date _____/_____ 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
0161

have independently considered the effects of a hypothetical particle whose properties correspond roughly to those of particles that are predicted on independent grounds by some physical theories. If such a particle were found, it would solve both problems at once.

The solar-neutrino problem arose in the late 1960's, when Raymond Davis, Jr., of the Brookhaven National Laboratory began to measure the flux of high-energy neutrinos emitted by the sun. Davis has consistently found that the measured flux is smaller, by a factor of three, than the flux predicted by the standard theory of nuclear burning in the sun.

The missing-mass problem arises because the orbital radii of visible matter in many galaxies are too small to be explained by the mass of the visible matter. If the galaxy is in a state of dynamic equilibrium, there must be approximately as much dark matter in the galaxy as there is matter bound up in stars.

In 1977 John Faulkner of the University of California at Santa Cruz and one of his students, Ronald Gilliland (who is now at the National Center for Atmospheric Research), considered the effects of a weakly interacting massive particle (WIMP) on the flux of neutrinos by the sun. Their argument is straightforward: The flux of high-energy neutrinos depends sensitively on the temperature at the core of the sun. Faulkner and Gilliland adjust the mass of the WIMP to between five and 60 times the mass of the proton and its so-called scattering cross section to about 4×10^{-36} square centimeter. If such particles reside in the core, they would interact just often enough with other particles there to carry off some of the thermal energy from the core to a region outside it. The consequent reduction of temperature in the core is just enough to lower the neutrino flux to the levels observed. Faulkner and Gilliland did not publish their result until recently because it seemed unlikely that such a particle could exist.

Last year, unaware of the earlier work, David N. Spergel and William H. Press of the Center for Astrophysics of the Harvard College Observatory and the Smithsonian Astrophysical Observatory proposed virtually the same solution to the solar-neutrino problem. They noted that WIMP's—or cosmions, as they call the particles—would be similar in mass and cross section to the photino, which is a massive counterpart of the photon in a theory of elementary-particle interactions known as supersymmetry. Spergel and Press then took an additional step: they proposed that if

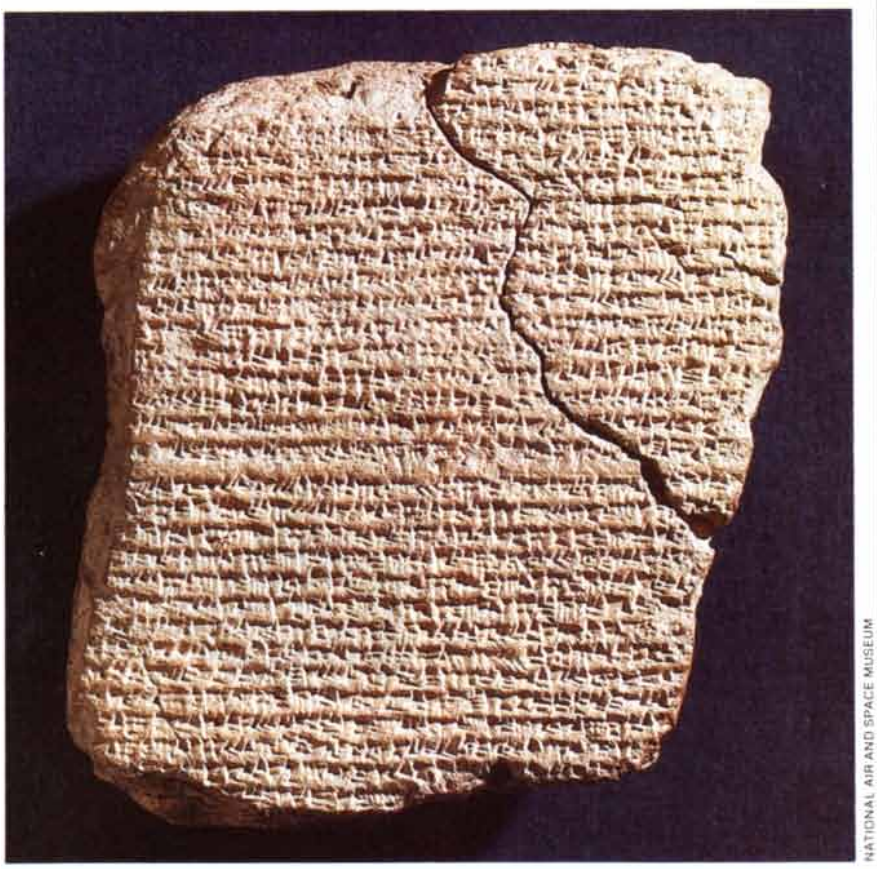
WIMP's (or cosmions) exist in the cosmic abundance required to explain solar neutrinos, there would be just enough of them to explain the missing mass of galaxies as well.

Old Faithful

Every 75 to 80 years Halley's comet sweeps through the inner solar system and past the earth, following an elongated elliptical orbit around the sun. As astronomers prepare for the well-publicized arrival of the comet this winter a retrospective observation has been reported: the discovery of three clay tablets recording sightings of the comet by Babylonian astronomers in 164 B.C. and 87 B.C. Although there is a Chinese record of 240 B.C. that may refer to an even earlier sighting, the Babylonian tablets appear to provide the most accurate records predating 12 B.C.

F. R. Stephenson and K. K. C. Yau of the University of Durham and H. Hunger of the University of Vienna write in *Nature* that they discovered the records on clay tablets in the British Museum. The late Abraham J. Sachs of Brown University deciphered the cuneiform text. The records are part of a collection of some 1,200 surviving fragments of astronomical diaries that, C. B. F. Walker of the British Museum notes, were kept by the Babylonians from the time of King Nabonassar (747–734 B.C.). Although the Babylonian astronomers were primarily interested in the movement of the moon and the planets, comets are mentioned in several of the diaries. A line of text pertaining to the 164 B.C. sighting of Halley's comet, for example, appears on tablet BM 41462: "The comet which previously had appeared in the east in the path of Anu in the area of Pleiades and Taurus, to the west [...] and passed along in the path of Ea."

The records, although badly damaged, have enabled the investigators to place fairly narrow limits on the dates of perihelion (the point of closest approach to the sun). The accuracy of the dates is important because, even though the comet exhibits periodic motion, planetary gravitational fields cause the period to vary by as much as two and a half years on each side of the mean. Consequently the very early history of the comet (before 12 B.C.) has been obscure. The Babylonian sightings may help historians of astronomy to calculate the motion of Halley's comet further back in time and so perhaps lead to the discovery of still earlier records.



NATIONAL AIR AND SPACE MUSEUM

Babylonian clay tablet recording a sighting of Halley's comet in 164 B.C.

The Classical Vacuum

It is not empty. Even when all matter and heat radiation have been removed from a region of space, the vacuum of classical physics remains filled with a distinctive pattern of electromagnetic fields

by Timothy H. Boyer

Aristotle and his followers believed no region of space could be totally empty. This notion that "nature abhors a vacuum" was rejected in the scientific revolution of the 17th century; ironically, though, modern physics has come to hold a similar view. Today there is no doubt that a region of space can be emptied of ordinary matter, at least in principle. In the modern view, however, a region of vacuum is far from being empty or featureless. It has a complex structure, which cannot be eliminated by any conceivable means.

This use of words may seem puzzling. If the vacuum is not empty, how can it be called a vacuum? Physicists today define the vacuum as whatever is left in a region of space when it has been emptied of everything that can possibly be removed from it by experimental means. The vacuum is the experimentally attainable void. Obviously a first step in creating a region of vacuum is to eliminate all visible matter, such as solids and liquids. Gases must also be removed. When all matter has been excluded, however, space is not empty; it remains filled with electromagnetic radiation. A part of the radiation is thermal, and it can be removed by cooling, but another component of the radiation has a subtler origin. Even if the temperature of a vacuum could be reduced to absolute zero, a pattern of fluctuating electromagnetic waves would persist. This residual radiation, which has been analyzed only in recent years, is an inherent feature of the vacuum, and it cannot be suppressed.

A full account of the contemporary theory of the vacuum would have to include the ideas of quantum mechanics, which are curious indeed. For example, it would be necessary to describe the spontaneous creation of matter and antimatter from the vacuum. Nevertheless, some of the remarkable properties of the vacuum do not depend at all on the peculiar logic of the quantum theory, and they can best be understood in a classical description (one that ignores quantum effects). Accordingly I shall discuss the vacuum entirely in terms of classical ideas. Even in the comparatively simple world of classical physics the vacuum is amply strange.

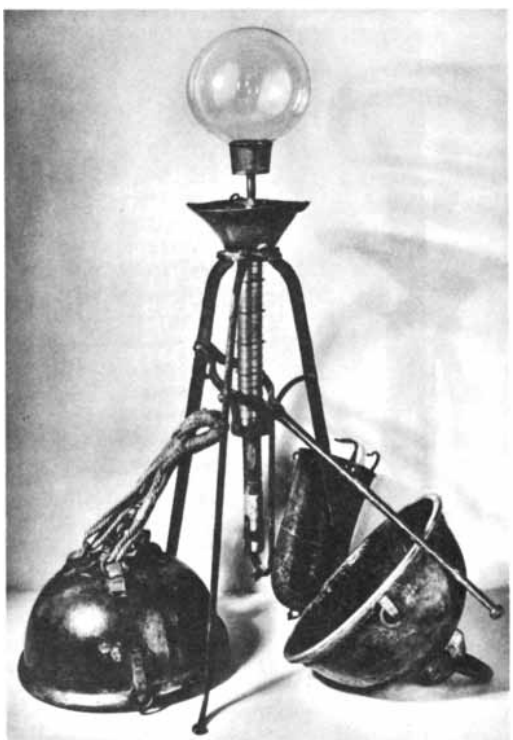
The Discovery of the Vacuum

Aristotle's doctrine that a vacuum is physically impossible was overthrown in the 17th century. The crucial development was the invention of the barometer in 1644 by Evangelista Torricelli, who had been secretary to Galileo. Torricelli poured mercury into a glass tube closed at one end and then inverted the tube, with the open end in a vessel filled with mercury. The column of liquid fell to a height of about 30 inches above the level of the mercury in the vessel, leaving a space at the top of the tube. The space was clearly empty of any visible matter; Torricelli proposed that it was also free of gas and so was a region of vacuum. A lively controversy ensued between supporters of the Aristotelian view and those who believed Torricelli had indeed created a vacuum. A few years later Blaise Pascal supervised a series of ingenious experiments, all tending to confirm Torricelli's hypothesis.

In the following decades experiments with the vacuum had a great vogue. The best-remembered of these demonstrations is one conducted by Otto von Guericke, the burgomaster of Magdeburg, who made a globe from two copper hemispheres and evacuated the space within. Two teams of eight draft horses were unable to separate the hemispheres. Other experiments of the era were less spectacular but perhaps more informative. For example, they led to the discovery that a vacuum transmits light but not sound.

The understanding of the vacuum changed again in the 19th century. The nature of the change can be illustrated by a thought experiment to be done with imaginary ideal apparatus.

Suppose one had a piston and cylinder machined so perfectly that the piston could move freely and yet nothing could leak past it. Initially the piston is at the closed end of the cylinder and there is no vacant space at all. When a steady force is applied to withdraw the



MAGDEBURG HEMISPHERES made in 1654 by Otto von Guericke demonstrated the existence of the vacuum. When the hemispheres were put together and the air was pumped out, two teams of eight draft horses could not separate them. The apparatus is now in the Deutsches Museum in Munich.

piston against the pressure of the air outside, the space developed between the piston and the end of the cylinder is a region of vacuum. If the piston is immediately released, it moves back into the cylinder, eliminating the vacuum space. If the piston is withdrawn and held for some time at room temperature, however, the result is quite different. External air pressure pushes on the piston, tending to restore the original configuration. Nevertheless, the piston does not go all the way back into the cylinder, even if additional force is applied. Evidently something is inside the cylinder. What appeared to be an empty space is not empty after the wait.

The physicists of the 19th century were able to explain this curious result. During the period when the piston was withdrawn the walls of the cylinder were emitting heat radiation into the vacuum region. When the piston was forced back in, the radiation was compressed. Thermal radiation responds to compression much as a gas does: both the pressure and the temperature rise. Thus the compressed radiation exerts a force opposing the reinsertion of the piston. The piston and cylinder could be closed again only if one waited long enough for the higher-temperature radiation to be reabsorbed by the walls of the cylinder.

The form of thermal radiation is intimately connected with the structure of the vacuum in classical physics. Nothing in my discussion so far has indicated that this should be so, and indeed the physicists of the 19th century were unaware of the connection.

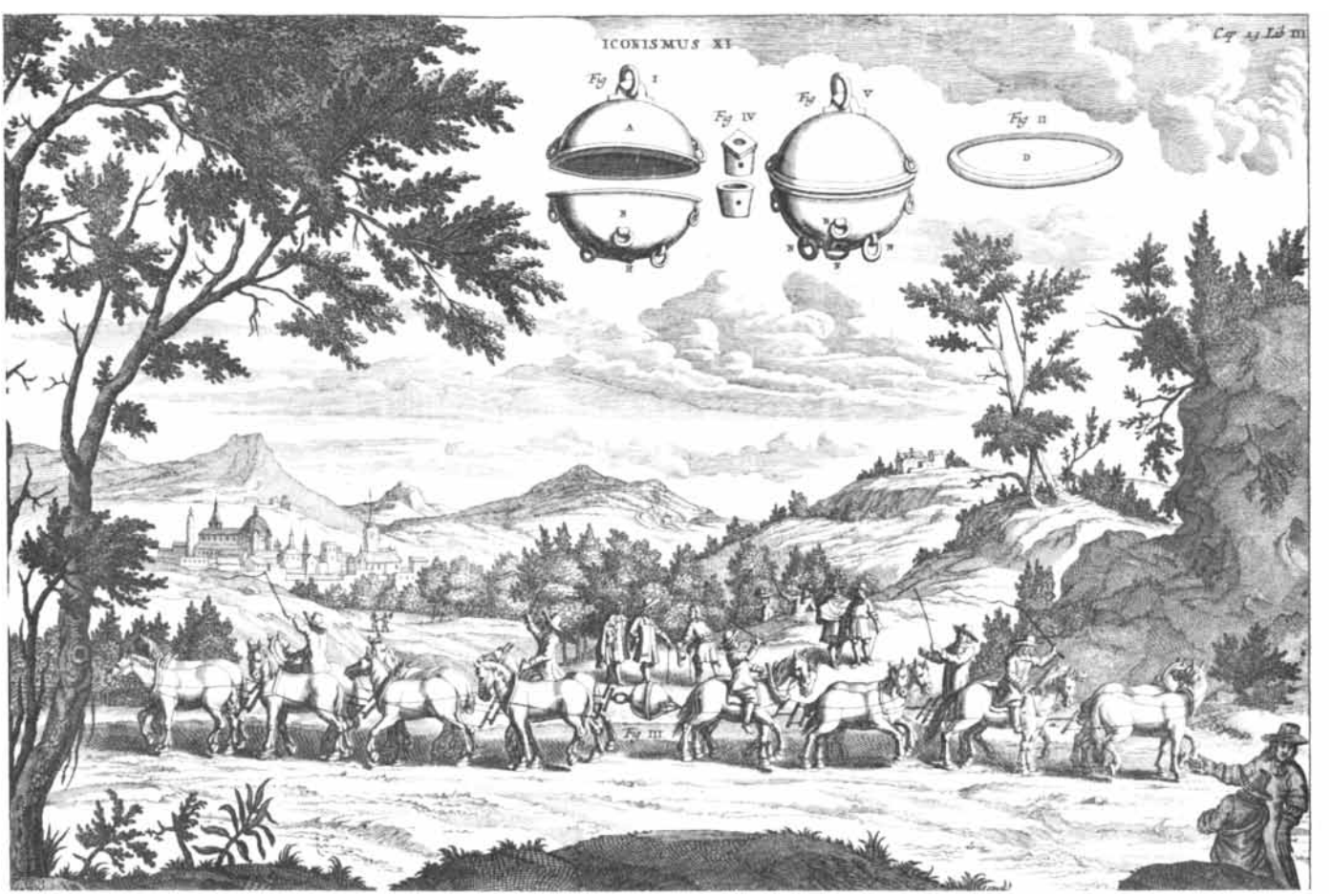
The Thermal Spectrum

Thermal radiation consists of electromagnetic fields that fluctuate in the most random way possible. Paradoxically this maximum randomness gives the radiation great statistical regularity. Under conditions of thermal equilibrium, in which the temperature is uniform everywhere, the radiation is both homogeneous and isotropic: its properties are the same at every point in space and in every direction. An instrument capable of measuring any property of the radiation would give the same reading no matter where it was placed and what direction it was pointed in.

The one physical quantity that determines the character of thermal radiation is its temperature. In 1879 the Austrian physicist Josef Stefan investigated the total energy density (or energy per unit volume) of thermal radiation and on the basis of some preliminary experiments suggested that the energy density varies as the fourth power of the absolute temperature. Five years later Stefan's student Ludwig Boltzmann found the same relation through a theoretical analysis.

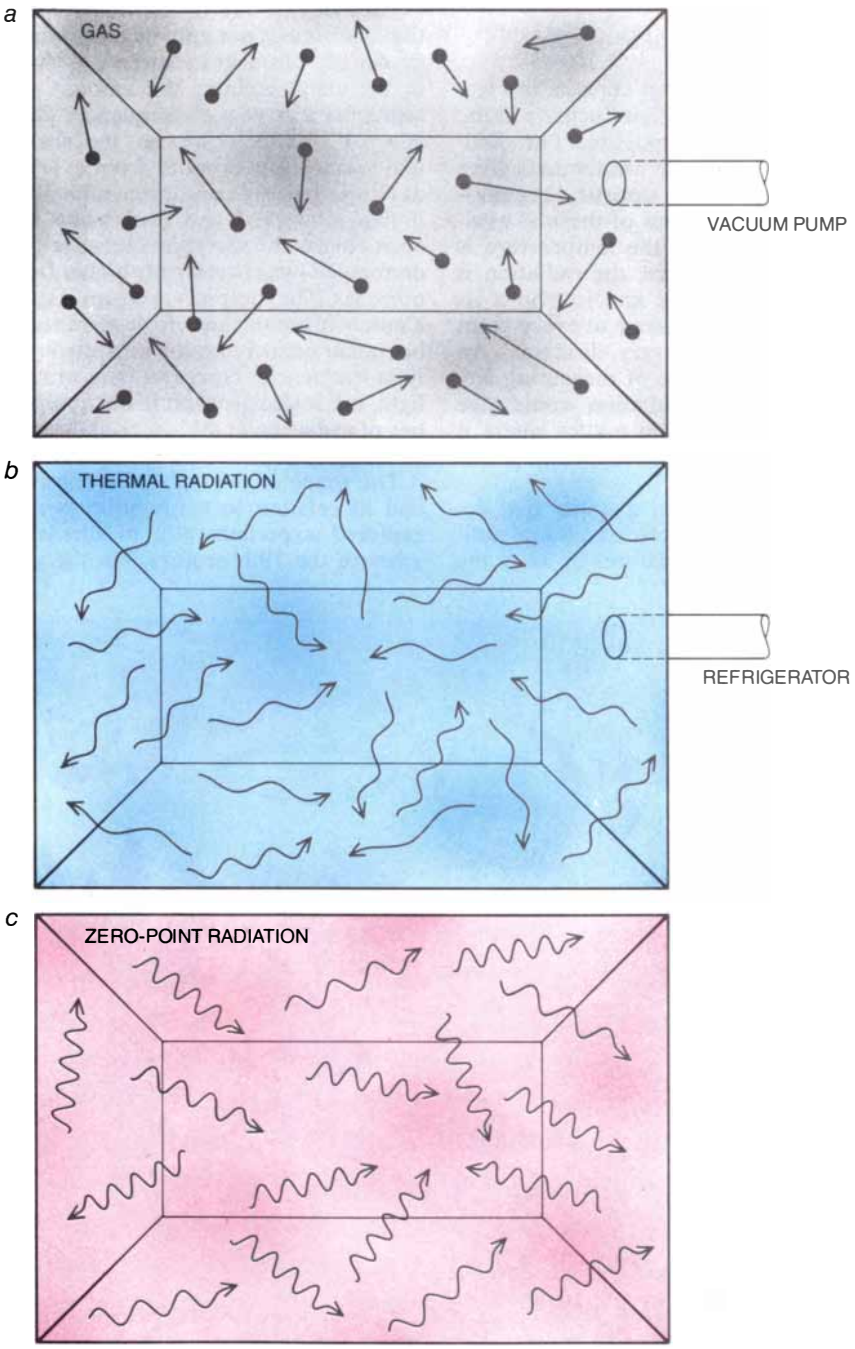
The temperature of thermal radiation determines not only its total energy density but also its spectrum, that is, the curve defining the amount of radiant energy at each frequency. The effect of temperature on the thermal spectrum is familiar from everyday experience; as an object is heated it first glows red and then white or even blue as the spectrum comes to be dominated by progressively higher frequencies. The thermal spectrum is not a monochromatic one, however; a red-hot poker emits radiation most strongly at frequencies corresponding to red light, but it also gives off lesser quantities of radiation at all higher and lower frequencies.

The shape of the thermal spectrum and its relation to temperature were explored experimentally in the last years of the 19th century, but the at-



tempt to formulate a consistent theoretical explanation met with only limited success. The aim was to find a mathematical expression that would give the intensity of the radiation as a function of the frequency and the temperature. In other words, given some specified temperature, the expression had to predict the intensity of radiation that would be measured at any chosen frequency.

A sophisticated classical analysis of the thermal spectrum was given by the German physicist Wilhelm Wien in 1893. Wien based his analysis on a thought experiment much like the one described above, but with the added provision that the cylinder be perfectly insulated so that no heat could be gained or lost. Wien calculated the change in the spectrum that would be brought about by an infinitesimal change in the internal volume of the cylinder. From this calculation he was able to deduce that the mathematical expression describing the spectrum must have two factors, which are multiplied to yield the intensity at a given frequency and temperature. One factor is the cube of the frequency. The second factor is a function of the absolute temperature divided by the frequency, but Wien was not able to determine the correct form of the function. (He made a proposal, but it was soon shown to be wrong.)



CREATION OF A VACUUM proceeds in stages that parallel the historical development of ideas about the vacuum. In the 17th century (*a*) it was thought a totally empty volume of space could be created by removing all matter, and in particular all gases. Late in the 19th century (*b*) it became apparent that such a region still contains thermal radiation, but it seemed the radiation might be eliminated by cooling. Since then both theory and experiment have shown there is nonthermal radiation in the vacuum (*c*), and it would persist even if the temperature could be lowered to absolute zero. It is called zero-point radiation.

Classical Electron Theory

The mathematical function needed to describe the thermal spectrum was suggested by Max Planck in 1900. Planck emphasized that an understanding of thermal radiation required the introduction of a new fundamental constant, now called Planck's constant, with a value of 6.626×10^{-27} erg-second. In the course of his struggle to explain his function for the thermal spectrum Planck launched the quantum theory. The start of quantum physics, however, did not mark the end of the story of classical physics.

Stefan's and Boltzmann's proposal that the total energy density of the thermal radiation is proportional to the fourth power of the temperature implies that the energy density falls to zero at a temperature of absolute zero. The thermal radiation simply disappears at zero temperature. The possibility of eliminating all thermal radiation led to a conception of the classical vacuum that was an extension of the 17th-century view. A perfect vacuum was still a totally empty region of space, but to attain this state one had to remove not only all visible matter and all gas but also all electromagnetic radiation. The last requirement could be met in principle by cooling the region to absolute zero.

This conception of the vacuum within classical physics was embodied in the fundamental physical theory of the time, which has since come to be known as classical electron theory. It views electrons as pointlike particles whose only properties are mass and electric charge. They can be set in motion by electric and magnetic fields, and their motion in turn gives rise to such fields. (An electron in steady oscillation, for example, radiates electromagnetic waves at the frequency of

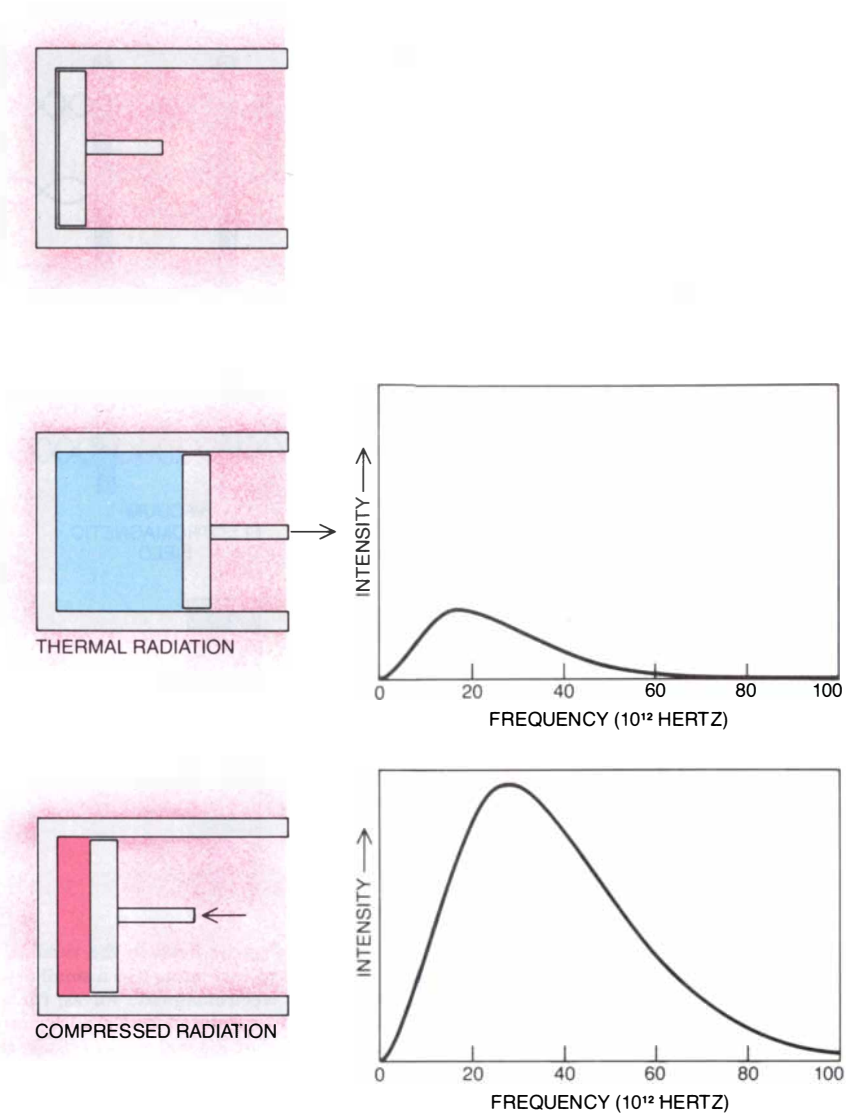
oscillation.) The interactions between particles and fields are accounted for by Newton's laws of motion and by James Clerk Maxwell's equations of electromagnetism. In addition certain boundary conditions must be specified if the theory is to make definite predictions. Maxwell's equations describe how an electromagnetic field changes from place to place and from moment to moment, but to calculate the actual value of the field one must know the initial, or boundary, values of the field, which provide a baseline for all subsequent changes.

It is through the choice of initial conditions that the nature of the vacuum enters classical electron theory. Since in the 19th-century view the vacuum was empty of all radiation, the initial conditions set on Maxwell's equations were the absence of electric and magnetic radiation. Roughly speaking, the 19th-century version of classical electron theory assumed that at some time in the distant past the universe contained matter (electrons) but no radiation. All electromagnetic radiation evolved from the acceleration of electric charges.

The Casimir Effect

Classical electron theory remains a viable field of investigation today, but it has taken a new form in the 20th century. The need for a revision is easily seen from an experiment proposed in 1948 by Hendrik B. G. Casimir of the Philips Research Laboratories in the Netherlands. Casimir analyzed the forces that would act on two electrically conducting, parallel plates mounted a small distance apart in a vacuum. If the plates carry an electric charge, the laws of elementary electrostatics predict a force between them, but Casimir considered the case in which the plates are uncharged. Even then a force can arise from electromagnetic radiation surrounding the plates. The origin of this force is not immediately obvious, but a mechanical analogy serves to make it clear.

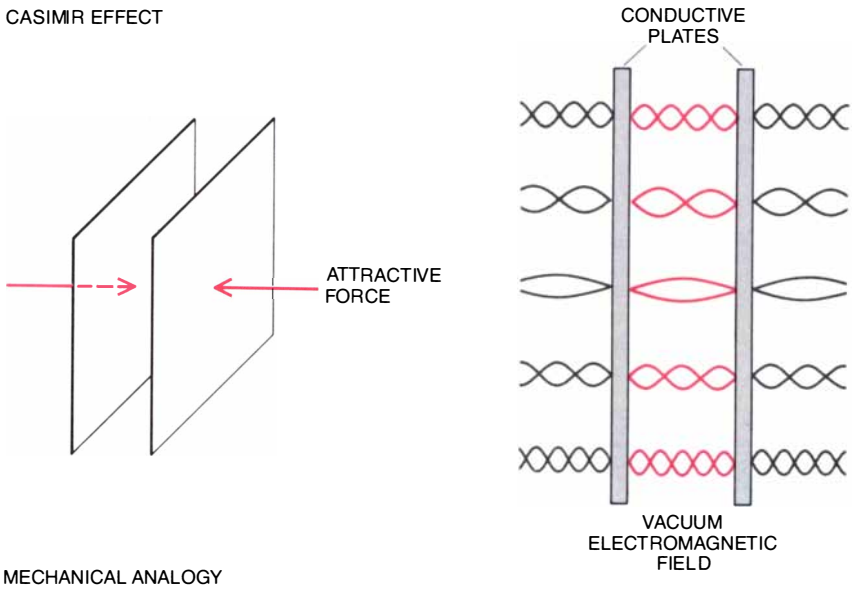
Suppose a smooth cord is threaded snuggly through holes in two wood blocks, as in the upper illustration on the next page. The cord is not tied to the blocks, and so at rest it neither pushes them apart nor pulls them together. Nevertheless, if the part of the cord between the blocks is made to vibrate transversely, a force acts on the blocks and they tend to slide along the cord away from each other. The force arises because transverse motion of the cord is not possible where it passes through a block, and so waves in the cord are reflected there. When a wave is reflected, some of its momentum is transferred to the reflector.



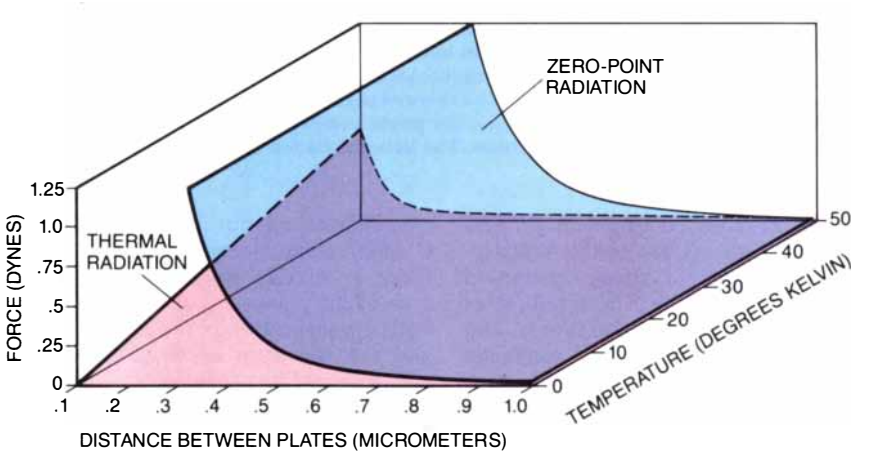
IDEAL PISTON AND CYLINDER provide the apparatus for a thought experiment revealing the presence of thermal radiation. The piston is initially at the closed end of the cylinder, leaving no free space; then it is withdrawn partway and held in this position for some time at room temperature. The space enclosed would seem to be a vacuum, and yet when the piston is released, it does not return to its initial position; indeed, it cannot be pushed all the way back into the cylinder even with additional force. While the piston was held in the open position the walls of the cavity emitted thermal radiation with a spectrum determined by the temperature. An attempt to reinsert the piston compresses the radiation, raising its temperature and thus altering its spectrum. The hotter radiation opposes the compression.

The situation in Casimir's proposed experiment is similar. The metal plates are analogous to the wood blocks, and the fluctuating electric and magnetic radiation fields represent the vibrating cord. The analogue of the hole in the wood block is the conducting quality of the metal plates; just as waves on the cord are reflected by the block, so electromagnetic waves are reflected by a conductor. In this case there is radiation on both sides of each plate, and thus the forces tend to cancel. The cancellation is not exact, however; a small residual force remains. The force is directly proportional to the area of the plates and also depends on both the separation between the plates and the spectrum of the fluctuating electromagnetic radiation.

So far this analysis is wholly consistent with the 19th-century view of the vacuum. The force acting on the plates is attributed to fluctuating thermal radiation. When the temperature is reduced to absolute zero, both the thermal radiation and the force between the plates should disappear. Experi-



CASIMIR EFFECT demonstrates the existence of electromagnetic fields in the vacuum. Two metal plates in a vacuum chamber are mounted parallel to each other and a small distance apart. Because the plates are conducting, they reflect electromagnetic waves; for a wave to be reflected there must be a node of the electric field—a point of zero electric amplitude—at the surface of the plate. The resulting arrangement of the waves gives rise to a force of attraction. The origin of the force can be understood in part through a mechanical analogy. If a cord threaded through holes in two wood blocks is made to vibrate, waves in the cord are reflected at the holes and generate forces on the blocks. The forces on a single block act in opposite directions, but a small net force remains. Its magnitude and direction depend on the separation between the blocks and the spectrum of waves along the cord.



FORCE OBSERVED IN THE CASIMIR EXPERIMENT has two components. At high temperature thermal radiation gives rise to a force directly proportional to the temperature and inversely proportional to the cube of the distance between the plates. This force disappears at absolute zero, as the thermal radiation itself does. The force associated with the zero-point radiation is independent of temperature and inversely proportional to the fourth power of the distance between the plates. The forces shown are for plates with an area of one square centimeter; the thermal force is an approximation valid at high temperature.

ment contradicts this prediction. In 1958 the Dutch physicist M. J. Sparnaay carried out a series of experiments based on Casimir's proposal and found that the force did not approach zero when the thermal radiation was reduced to low intensity. Instead there was a residual attractive force that would persist even at absolute zero.

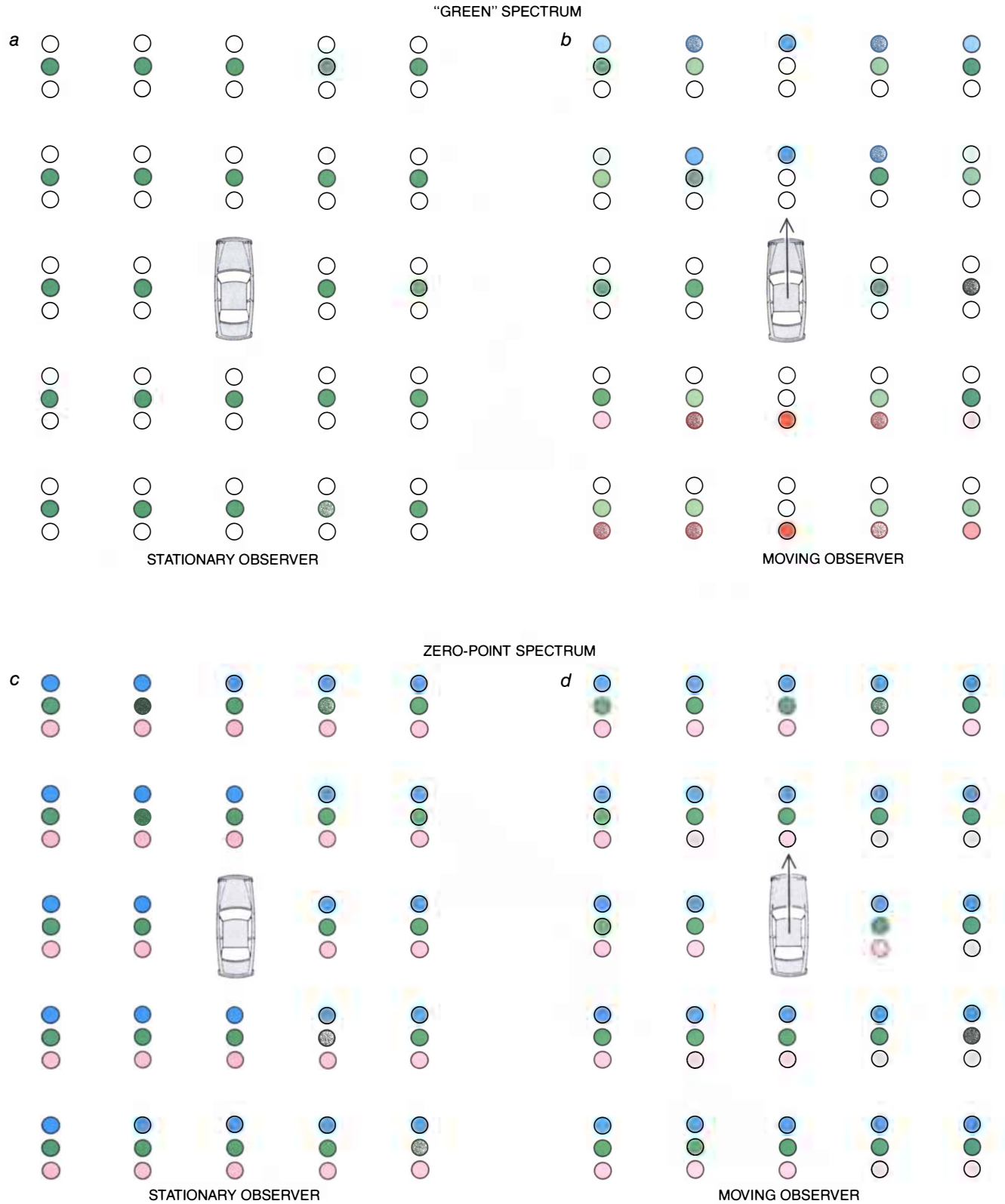
The residual force is directly proportional to the area of the plates and inversely proportional to the fourth power of their separation; the constant of proportionality is 1.3×10^{-18} erg-centimeter. Although such a force is small, it is measurable if the plates are sufficiently close together. For plates with an area of one square centimeter separated by .5 micrometer the Casimir force is equivalent to the weight of .2 milligram.

Whatever the magnitude of the Casimir effect, its very existence indicates that there is something fundamentally wrong with the 19th-century idea of the classical vacuum. If one is to fit classical theory with experiment, then even at zero temperature the classical vacuum cannot be completely empty; it must be filled with the classical electromagnetic fields responsible for the attractive force Sparnaay measured. Those vacuum fields are now referred to as classical electromagnetic zero-point radiation.

The Zero-Point Spectrum

What are the characteristics of the zero-point radiation in the classical vacuum? Much can be deduced from the fact that it exists in a vacuum: it must conform to accepted basic ideas about the nature of the vacuum. For example, it seems essential that the vacuum define no special places or directions, no landmarks in space or time; it should look the same at all positions and in all directions. Hence the zero-point radiation, like thermal radiation, must be homogeneous and isotropic. Furthermore, the vacuum should not define any special velocity through space; it should look the same to any two observers no matter what their velocity is with respect to each other, provided the velocity is constant. This last requirement is expressed by saying the zero-point radiation must be invariant with respect to Lorentz transformation. (The Lorentz transformation, named for the Dutch physicist H. A. Lorentz, is a conversion from one constant-velocity frame of reference to another, taking into account that the speed of light is the same in all frames of reference.)

The requirement of Lorentz invariance is a serious constraint. A railroad



LORENTZ INVARIANCE of the zero-point radiation ensures that the vacuum looks the same to observers moving through it at different velocities, provided each observer's velocity is constant. The Lorentz transformation relates frames of reference that differ in velocity; for radiation to be Lorentz-invariant its spectrum must be unchanged by the transformation. The effect of motion on the spectrum is illustrated by an observer surrounded by peculiar traffic signals, which always indicate the intensity of the zero-point radiation at three frequencies, namely those of red, green and blue light. Suppose an observer at rest with respect to the array of signals finds they all show green (*a*), meaning that all the zero-point radiation is concentrated in the green part of the electromagnetic spectrum. If the observer then begins to move (*b*), the pattern is altered by the Doppler effect: the signals ahead appear blue and those behind red. The Lorentz transformation also makes the approaching signals brighter and the receding ones dimmer. It turns out that only one spectral form has the property of Lorentz invariance: the intensity must be proportional to the cube of the frequency. When the traffic signals are illuminated according to this rule, an observer at rest (*c*) and an observer in motion (*d*) see the same pattern.

passenger may be momentarily unsure whether his own train or the one on the next track is moving relative to the earth, but the ambiguity can be resolved simply by looking at some landmark known to be fixed. Lorentz invariance implies that there are no such landmarks in the vacuum and that no experiment could ever reveal an observer's velocity with respect to the background of zero-point radiation. To meet this condition the spectrum of the radiation must have quite specific properties.

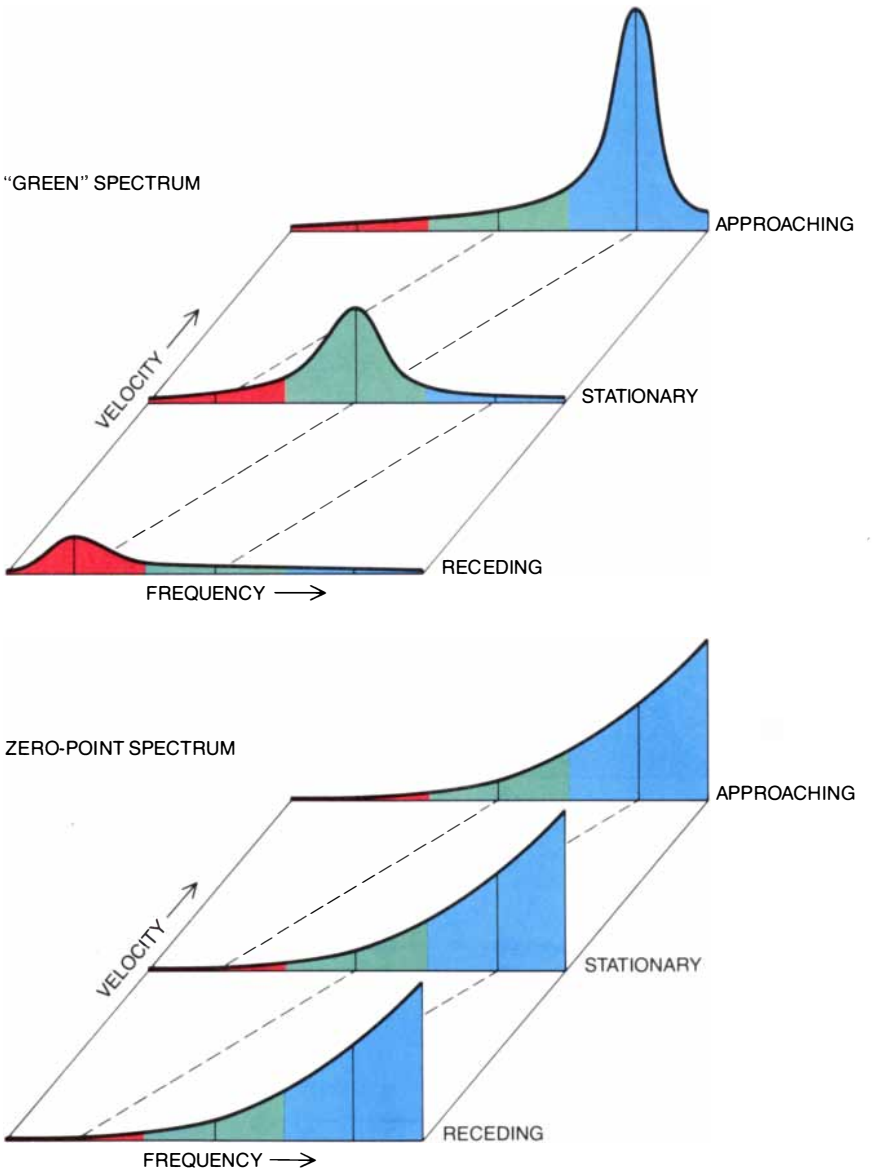
Suppose for the moment that the zero-point radiation, as perceived by some observer, were all in the region of the electromagnetic spectrum corresponding to green light. No matter where the observer stood and no matter in what direction he looked, the vacuum would appear to be filled with uniform green radiation. Such a spectrum satisfies the requirements of homogeneity and isotropy for this one observer, but now suppose there is another observer moving toward the first one at a constant speed. Because of the Doppler effect, the moving observer would see the radiation in front of him shifted toward the blue end of the spectrum and the radiation behind him shifted toward the red end. The Lorentz transformation also alters the intensity of the radiation: it would be brighter in front and dimmer behind. Thus the radiation does not look the same to both observers; it is isotropic to one but not to the other.

It turns out that the zero-point spectrum can have only one possible shape if the radiation is to be Lorentz-invariant. The intensity of the radiation at any frequency must be proportional to the cube of that frequency. A spectrum defined by such a cubic curve is the same for all unaccelerated observers, no matter what their velocity; moreover, it is the only spectrum that has this property.

One immediate objection might be made to the cubic form of the zero-point spectrum: because the intensity of the radiation increases steadily at higher frequencies, the spectrum predicts an infinite energy density for the vacuum. In the 19th century such a prediction might well have been considered a fatal flaw, but since the 1940's infinities have turned up in several areas of physics, and methods have been developed for dealing with them. In this case the infinite energy is confronted directly only in the realm of gravitational forces. All other calculations are based on changes or differences in energy, which are invariably finite.

If the universe is permeated by classical zero-point radiation, one might suppose it would make its presence known in phenomena less subtle than the Casimir effect. For example, one might think it would alter the outcome of the piston-and-cylinder experiment by resisting the insertion of the piston even after all thermal radiation had been eliminated.

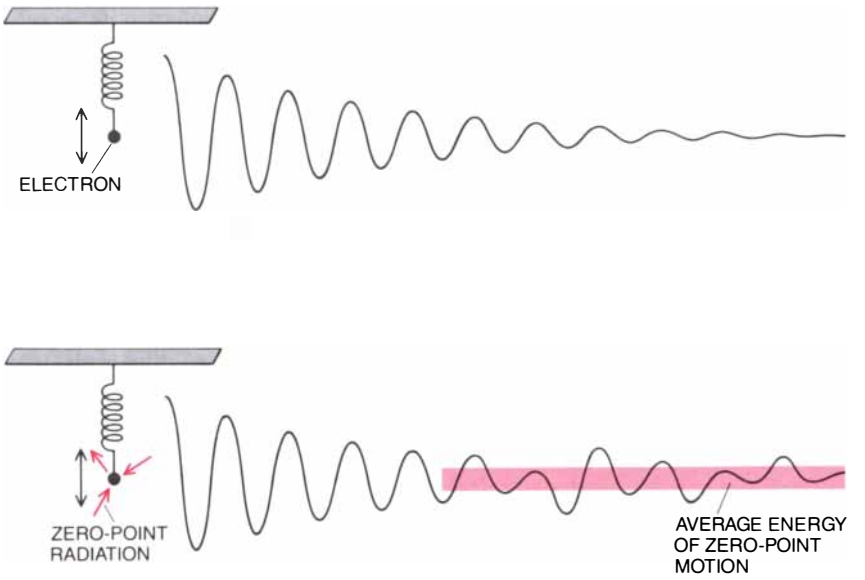
Analysis indicates otherwise. Under equilibrium conditions, when no external force is applied to the piston, there is radiation both inside and outside the cylinder, and the radiation pressures acting on the piston are balanced. This balance holds for both thermal and zero-point radiation. When the piston is pushed into the cylinder, the radiation is compressed. Wien's calculation of the change in the spectrum as a result of a change in volume indicates that the thermal radiation resists such compression; it increases in temperature and exerts a greater pressure



ZERO-POINT SPECTRUM is independent of the observer's velocity because of compensating changes in frequency and intensity. When an observer is approaching a source of radiation, all frequencies are shifted to higher values and all intensities are increased; moving away from the source has the opposite effect. Thus a spectrum that has a peak in the green region for a stationary observer has a larger blue peak for an approaching observer and a smaller red peak for a receding observer. The cubic curve that defines the zero-point spectrum balances the shifts in frequency and intensity. Light that appears green in the stationary frame of reference becomes blue to an approaching observer, but its intensity matches that of the blue light seen by an observer at rest. By the same token, green light is shifted to red frequencies for a receding observer, but its intensity is diminished correspondingly.

against the piston. When the same analysis is made for the zero-point radiation, however, the result is different: the zero-point spectrum does not change at all in response to compression. Indeed, a spectrum described by a cubic curve is the only one that has this remarkable property.

The other experiment in which the cubic zero-point spectrum should be checked is the Casimir effect itself. A theoretical calculation based on the spectrum predicts a force between the plates directly proportional to their area and inversely proportional to the fourth power of their separation, in agreement with Sparnaay's results. Again it can be shown that the spectrum is unique in supporting this prediction; no other spectral curve yields an inverse-fourth-power dependence on distance.



HARMONIC OSCILLATOR reveals the effects of zero-point radiation on matter. The oscillator consists of an electron attached to an ideal, frictionless spring. When the electron is set in motion, it oscillates about its point of equilibrium, emitting electromagnetic radiation at the frequency of oscillation. The radiation dissipates energy, and so in the absence of zero-point radiation and at a temperature of absolute zero the electron eventually comes to rest. Actually zero-point radiation continually imparts random impulses to the electron, so that it never comes to a complete stop. Zero-point radiation gives the oscillator an average energy equal to the frequency of oscillation multiplied by one-half of Planck's constant.

The New Classical Electron Theory

The statement that a spectrum described by a cubic curve is unique refers only to the shape of the curve; actually there are infinitely many curves with the same shape but different scales. In all the curves the intensity of the radiation is proportional to the cube of the frequency, but the magnitude of the intensity in each spectrum depends on a constant, which sets the scale of the curve.

The value of the constant cannot be calculated theoretically, but Sparnaay's measurement of the force in the Casimir effect allows the value to be determined from experiment. After some preliminary algebraic manipulation it is found that the constant is equal to 3.3×10^{-27} erg-second, a magnitude corresponding to one-half of Planck's constant. Thus Planck's constant, the hallmark of all quantum physics, appears in a purely classical context.

The introduction of classical zero-point radiation in the vacuum mandates an important change in classical electron theory. The revised version of the theory is still based on Newton's laws of motion for the electrons and Maxwell's equations for the electromagnetic field, but the boundary conditions imposed on Maxwell's equations must be altered. No longer is the vacuum empty of all electromagnetic fields; it is now filled with randomly fluctuating fields having the zero-point spectrum. The modified theory is called classical electron theory with classical electromagnetic zero-point radiation, a name often shortened to stochastic electrodynamics.

The altered boundary conditions change the predictions of the theory. The changes can be understood by considering one of the favorite models of modern physics: a harmonic oscillator made up of an electron attached to a perfectly elastic and frictionless spring. This imaginary mechanical system is to be set up in the classical vacuum. If the spring is stretched and then released, the electron oscillates about its equilibrium position and gives off electromagnetic radiation at the frequency of oscillation.

The harmonic oscillator is a convenient model because the motion of the electron is readily calculated. Under the older version of classical electron theory just two forces act on the electron: the restoring force from the spring and a reaction force arising from the emission of radiation. Because the reaction force is directed opposite to the electron's motion, the theory predicts that the oscillations will be steadily damped and the electron will eventually come to rest. In the new version of classical electron theory, however, the zero-point radiation provides an additional force on the electron. The charged particle is continually buffeted by the randomly fluctuating fields of the zero-point radiation, so that it never comes to rest. It turns out the harmonic oscillator retains an average energy related to the zero-point spectrum, namely one-half of Planck's constant multiplied by the frequency of oscillation.

Up to now the classical vacuum has been described from the point of view of an observer at rest or moving with constant velocity. The consequences of zero-point radiation are even more remarkable for an accelerated observer, that is, one whose velocity is changing in magnitude or direction.

Effects of Acceleration

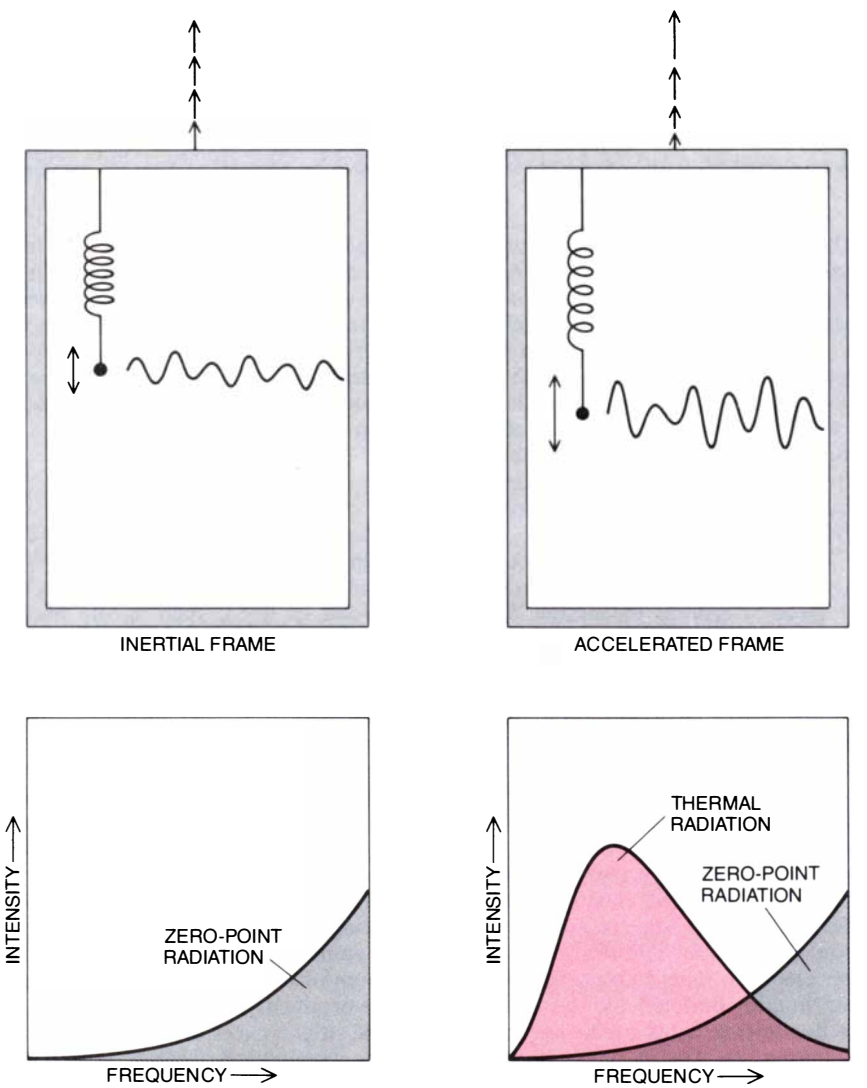
Consider an observer in a rocket continuously accelerating with respect to some frame of reference that can be regarded as fixed, such as the background of distant stars. What does the classical vacuum look like to the rocket-borne observer? To find out, one must perform a mathematical transformation from the fixed frame of reference to the accelerated one. The Lorentz transformation mediates between frames that differ in velocity, but the situation is more complex here because the velocity of the accelerated observer is continuously changing. By carrying out Lorentz transformations over some time interval, however, the vacuum observed from the rocket can be determined.

One might guess that the spectrum for an accelerated observer would no longer be isotropic, and in particular that some difference would be detected between the forward and the backward directions. The spectrum might also be predicted to change as the accelera-

tion continued. In fact the spectrum remains homogeneous and isotropic, and no change is observed as long as the rate of acceleration itself does not change. Nevertheless, the spectrum is not the one seen by an unaccelerated observer. At any given frequency the intensity of the radiation is greater in the accelerated frame than it is in the frame at rest.

The form of the classical electromagnetic spectrum seen by an accelerated observer is not one immediately familiar to physicists, but it can be interpreted by analyzing the motion of a harmonic oscillator carried along in the rocket. The equation of motion for the accelerated oscillator is much like the one valid in a fixed frame of reference. There are two differences: the radiation-reaction force has a new term proportional to the square of the acceleration, and the oscillator is exposed to a new spectrum of random radiation associated with the acceleration. The effect of these changes is to increase the average energy above the energy associated with the zero-point motion. In other words, when an oscillator is accelerated, it jiggles more vigorously than it would if it were at rest in the vacuum.

One way of understanding the effect of acceleration on the harmonic oscillator is to ask what additional electromagnetic spectrum could be added to the zero-point radiation to cause the extra motion. To answer this question one can turn to the equivalence principle on which Einstein founded his theory of gravitation. The principle states that an observer in a small laboratory supported in a gravitational field makes exactly the same measurements as an observer in a small accelerating rocket. The laws of thermodynamics are found to hold in a gravitational field. From the equivalence principle one therefore expects the laws of thermodynamics to hold in an accelerating rocket. There is then only one possible equilibrium spectrum that can be added to the zero-point radiation: the additional radiation must have a thermal spectrum. With any other spectrum the oscillator would not be in thermal equilibrium with its surroundings, and so it could serve as the basis of a perpetual-motion machine. By this route one is led to a remarkable conclusion: a physical system accelerated through the vacuum has the same equilibrium properties as an unaccelerated system immersed in thermal radiation at a temperature above absolute zero.



EFFECT OF ACCELERATION through the vacuum is to change the spectrum of observed radiation. At a temperature of absolute zero a harmonic oscillator in a frame of reference at rest or moving with constant velocity is subject only to zero-point oscillations. In an accelerated frame the oscillator responds as if it were at a temperature greater than zero.

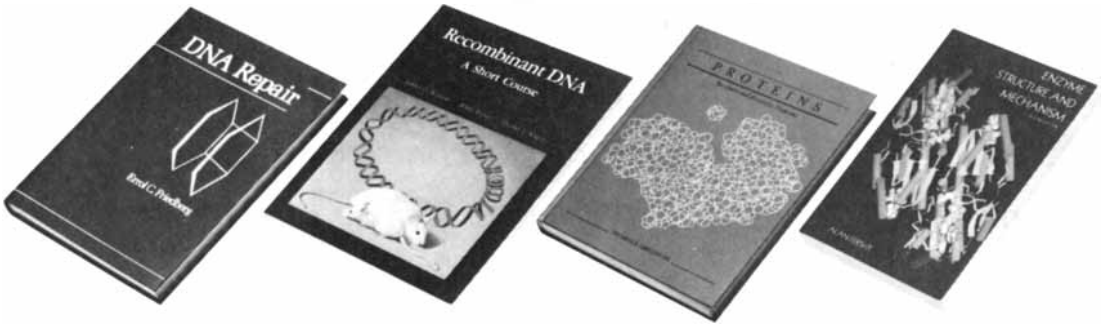
The mathematical relation connecting acceleration and temperature was found in about 1976 by William G. Unruh of the University of British Columbia and P. C. W. Davies of the University of Newcastle upon Tyne. The effective spectrum seen by an observer accelerated through the vacuum is the sum of two parts. One part is the zero-point radiation; the other is the spectrum of thermal radiation deduced by Planck in 1900. Planck was able to explain the form of that curve only by introducing quantum-mechanical ideas, which he did with some reluctance; it now turns out the curve can be derived from an entirely classical analysis of radiation in the vacuum.

At least one more intriguing result arises from this line of inquiry. If one again invokes the equivalence principle relating an observer in a gravitational field with an accelerating observer, one concludes that there is a minimum attainable temperature in a gravitational field. This limit is an absolute one, quite apart from any practical difficulties of reaching low temperatures. At the surface of the earth the limit is 4×10^{-20} degree Kelvin, far beyond the capabilities of real refrigerators but nonetheless greater than zero.

The discovery of a connection between thermal radiation and the structure of the classical vacuum reveals an unexpected unity in the laws of physics, but it also complicates our view of what was once considered mere empty space. Even with its pattern of electric and magnetic fields in continual fluctuation, the vacuum remains the simplest state of nature. But perhaps this statement reflects more on the subtlety of nature than it does on the simplicity of the vacuum.

Essential new works in molecular biology

FROM W. H. FREEMAN AND COMPANY



DNA REPAIR

Errol C. Friedberg, *Stanford University*

The repair of DNA is now regarded along with replication and recombination as an essential transaction of the genetic material in all life forms. Much of the excitement in the field comes as a result of the relevance of DNA repair to human health. How dangerous is a chest X-ray? Are some combinations of carcinogens more potent than the health threat posed by individual chemicals?

DNA Repair was written specifically to provide answers to these and other questions for those who possess a solid understanding of the fundamental molecular biology of nucleic acids and proteins. Beginning with a discussion of spontaneous DNA damage and damage from interactions with environmental physical and chemical agents, the book then treats the reversal of base damage and excision of damaged or inappropriate bases from DNA.

"Friedberg's account is likely to become a classic treatise for postgraduates and research workers." **—NATURE**

1984, 640 pages, 292 illustrations, $39.95

ENZYME STRUCTURE AND MECHANISM

Second Edition

Alan Fersht, *Imperial College, London*

This state-of-the-art guide takes as its theme the way in which the interaction of an enzyme with its substrate leads to catalysis and specificity. New chapters have been added to cover such issues as the application of recombinant DNA technology to enzymology, the basic enzymology behind the methods of genetic manipulation, protein modification, and new advances in stereochemistry.

". . . an exceedingly important book that should be owned and read by all students and researchers in the biochemical sciences." **—Roger Kornberg,** Stanford University, on the First Edition

1984, 496 pages, 125 illustrations, $24.95

RECOMBINANT DNA: A SHORT COURSE

James D. Watson, Director, *Cold Spring Harbor Laboratory*
John Tooze, *European Molecular Biology Organization*
David T. Kurtz, *Cold Spring Harbor Laboratory*

Nobel Laureate James Watson and his colleagues summarize the exciting, often stormy history of the recombinant DNA revolution, survey the recent techniques in recombinant molecular genetics, and explain both the experimental methods now in use and their dramatic results.

"In a swift, often aphoristic way these three molecular biologists and their talented illustrator George Kelvin have set out the state of their marvelous art . . . It is as successful a summary of a discipline in flower as one can find, a worthy follower to the senior author's earlier comprehensive text on the gene." **—Philip Morrison,** *Scientific American*

1983, 260 pages, 170 illustrations, $17.95

PROTEINS: Structures and Molecular Properties

Thomas E. Creighton, *Medical Research Laboratory of Molecular Biology, Cambridge*

A coherent and comprehensive description of the molecular, chemical, and physical properties of proteins. Organized on a combined basis of chronology and structural and functional hierarchy, the book treats all aspects of proteins—biosynthesis, evolution, dynamics, ligand binding, catalysis, and energy transduction—not only their structures. The treatment is often empirical, presented in a manner meant to convey principles without obscuring them in mathematical derivations.

"It will certainly become a standard text for the undergraduate in biochemistry, biophysics, and biology and also a most useful handbook for the researcher." **—Times Higher Education Supplement**

1984, 515 pages, 170 illustrations, $36.95

DNA REPAIR
☐ ISBN 0-7167-1674-7 $39.95 (cloth)

ENZYME STRUCTURE AND MECHANISM, Second Edition
☐ ISBN 0-7167-1614-3 $24.95 (cloth)

RECOMBINANT DNA: A SHORT COURSE
☐ ISBN 0-7167-1484-1 $17.95 (paper)

PROTEINS
☐ ISBN 0-7167-1566-X $36.95 (cloth)

☐ 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 add appropriate sales tax.)

☐ VISA ☐ MasterCard Acct. No. _____

Exp. Date _____ Signature _____

(All credit card orders must be signed.)

Name _____

Address _____

City/State/Zip _____

0216

W. H. FREEMAN AND COMPANY
4419 West 1980 South, Salt Lake City, UT 84104

The Mesozoic Vertebrates of Thailand

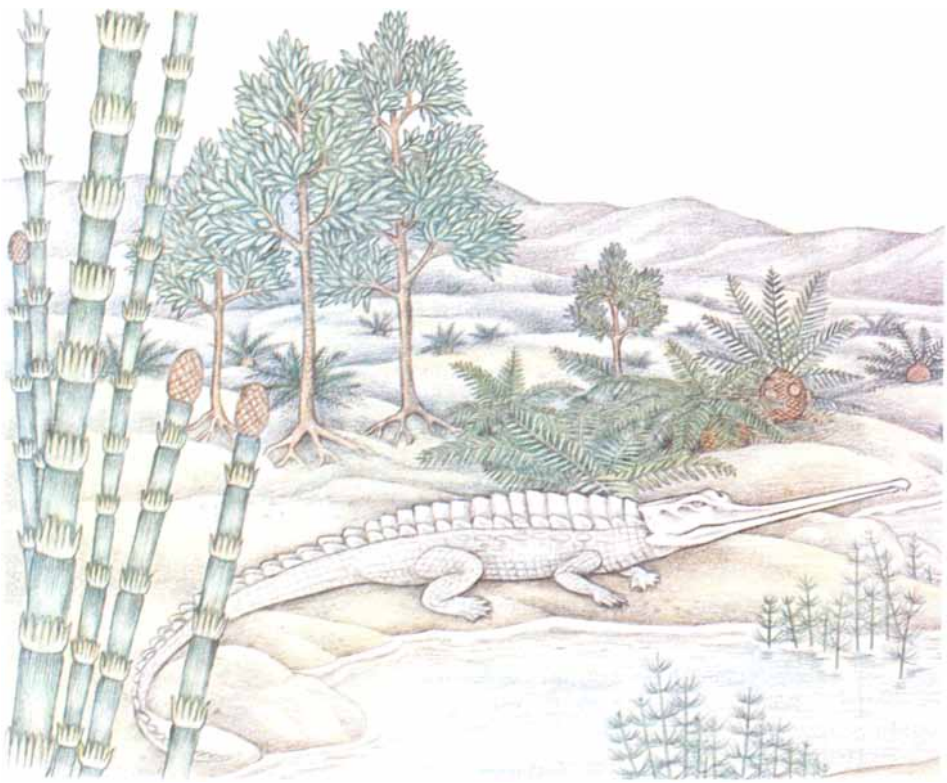
The northeastern part of the country holds a fossil record spanning more than 100 million years of vertebrate evolution. The record in turn helps to clarify the geologic history of southeastern Asia

by Eric Buffetaut and Rucha Ingavat

If one were looking for a place where it would be easy to hunt for fossils, one's first choice would probably not be southeastern Asia. The latitudes are tropical, monsoons dominate the climate and lush vegetation covers most of the land. Yet early investigators who visited southeastern Asia did make some discoveries. In Laos at the end of the 19th century the French geologist Henri Counillon found the skull of a reptile resembling a mammal. Then in the 1930's the French geologist Josué-Heilmann Hoffet collected the bones of dinosaurs in another part of Laos. Somewhat later Thailand yielded a few reptile teeth.

Today the situation is totally different: a series of Thai-French expeditions to northeastern Thailand has collected remains of fishes, amphibians and reptiles from the Mesozoic era. This was the age of the dominance of the reptiles. It began some 240 million years ago and ended 65 million years ago, with the ascendancy of the mammals. On the basis of these fossils it has now become possible to reconstruct the changing faunal communities in fresh water and on land in this part of southeastern Asia over a time span of more than 100 million years. The knowledge can be related to what is known about the Mesozoic fauna that inhabited what is now Europe and North America. Indeed, it turns out that northeastern Thailand is among the few regions of the world where a fossil record of the development of vertebrate life from the late Triassic through the Jurassic to the Cretaceous (spanning, in other words, the three geologic periods composing the Mesozoic era) has been preserved within a relatively limited area in a reasonably continuous sequence of continental rocks.

The record is also of great help to geology. It aids in establishing the age of the sedimentary rocks in which the fossils were found. More broadly, it illuminates the chronology of continental drift and collision in southeastern Asia. The chronology is eventful: Thailand consists of pieces of two distinct blocks, or microcontinents, whose geologic histories were independent until they collided. According to one hypothesis, the blocks were part of the northern margin of a southern supercontinent, called Gondwanaland. (They may have been part of the northern coast of what would eventually become Australia.) Sometime during the Paleozoic era, which preceded the Mesozoic, they separated from Gondwanaland and drifted northward, as several other such fragments did. Eventually they came in contact with the northern supercontinent, Laurasia. Specifically they met the southern margin of China. The process re-



LATE TRIASSIC LANDSCAPE depicts northeastern Thailand as it may have looked some 220 million years ago; the landscape is reconstructed on the basis of fossil fishes, amphibians and reptiles found in the shales and limestones of the Huai Hin Lat formation

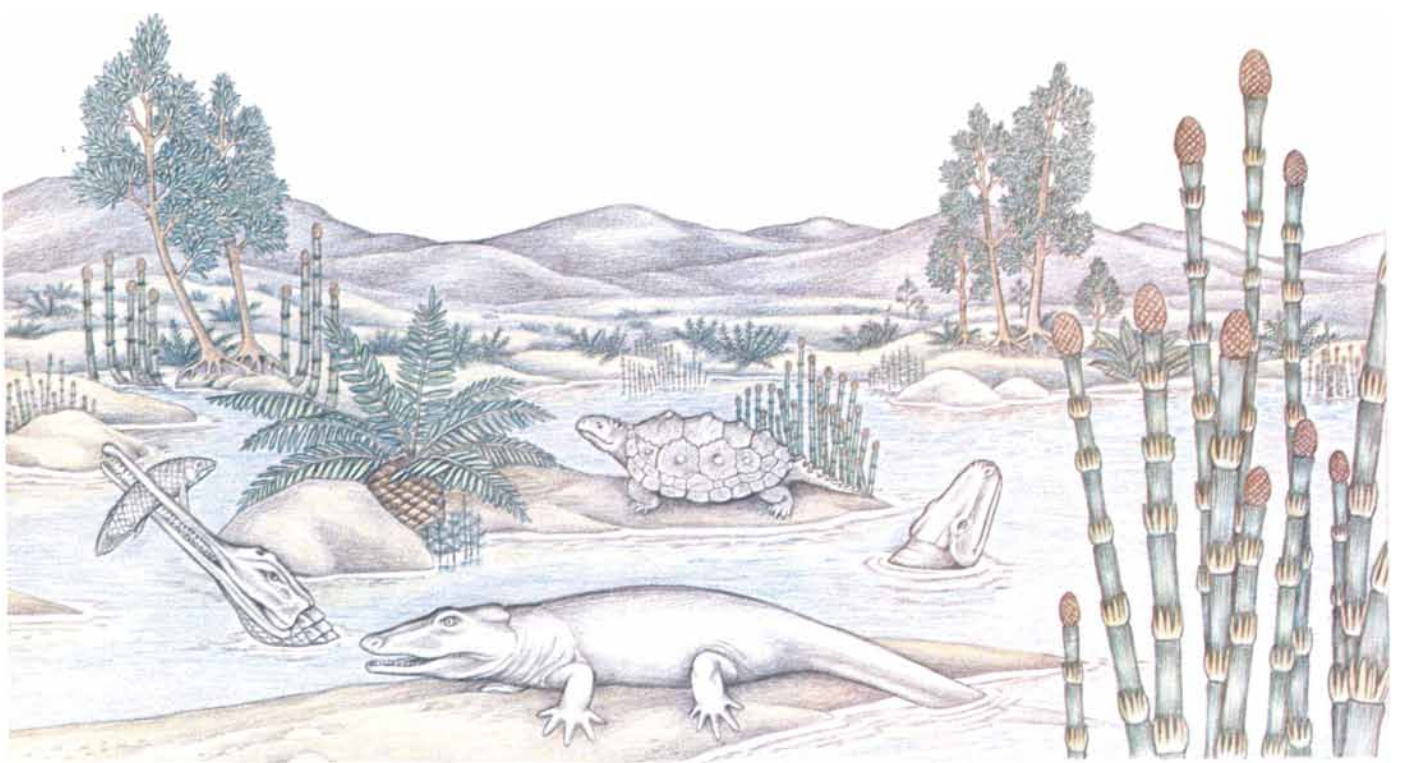
sembles what happened more than 100 million years later when India separated from Gondwanaland and drifted northward, colliding with what had become Asia.

One of the microcontinents whose collision created Thailand is called the Indochina block; it includes the Khorat plateau of northeastern Thailand as well as Kampuchea and most of Laos and Vietnam. The other microcontinent is known as the Shan-Thai block; Sangad Bunopas of the Thai Department of Mineral Resources in Bangkok has given it that name. The first part of the name refers to the Shan states of eastern Burma. In addition the block includes western Thailand and most of the Malay Peninsula. In recent years many geologists have attempted to reconstruct the detailed history of the blocks, and particularly to establish at what time in the geologic past the blocks became fused to each other and collided with mainland Asia. The fossil vertebrates of Thailand yield important evidence for that determination.

The record of Mesozoic life borne by the geology of Thailand could not have been assembled without the work of our colleagues in the Thai Department of Mineral Resources. Their nation exemplifies the difficulties of fossil hunting in southeastern Asia. Even atop the Khorat plateau of northeastern Thailand, in the driest part of the country, natural outcrops of rock are few and far between. The known sites of fossils tend to be small and unimpressive; many of the sites are just road cuts or quarries. The field geologists of the Department of Mineral Resources are adept at discovering sites. In the wake of a discovery our own team collects fossils in a systematic way. If need be, we start a regular excavation. We also prospect nearby outcrops, a strategy that often yields further discoveries. After preparation and study, which takes place partly in Paris, the specimens are returned to Thailand, where they augment the collections of the Department of Mineral Resources. In Bangkok the most interesting specimens are displayed to the public in the museum of geology.

Most of the fossils found in Thailand come from the Khorat plateau. There, during much of the Mesozoic era, great quantities of sediments eroded from newly raised mountains. The sediments were deposited on riverbeds, floodplains and at the bottom of lakes. The result is an accumulation of sandstones, clays and some lacustrine limestones (limestones deposited in lakes) known to geologists as the Khorat group. In places it reaches a thickness of several thousand meters. Some of the Khorat group's formations, or assemblages of layers, have yielded vertebrate bones, teeth or footprints.

The oldest fossils come from the Huai Hin Lat formation, consisting largely of dark gray or black bituminous shales and limestones. According to Chongpan Chonglakmani and Nares Sattayarak of the Department of Mineral Resources, the shales and limestones were laid down in great freshwater lakes after a mountain-building episode had forced the sea to retreat from the region in the later part of the Triassic period, some 225 million years ago. Our 1980 expedition discovered a good locality near the powerhouse of Chulabhorn Dam, toward the western edge of the Khorat plateau. Not unexpectedly, the most abundant vertebrate fossils were the remains of fishes, in particular the isolated scales of archaic bony fishes. In addition a few toothplates of a lungfish (an air-breathing fish) were found. Toothplates, which consist of fused individual teeth, help an animal to crush its food. We also uncovered the remains of the predators that fed on such fish. Some fragments of the amphibian *Cyclotosaurus,* including the well-



of northeastern Thailand by the authors and their colleagues. In the freshwater stream at the center *Mystriosuchus,* a phytosaur, or large, heavily armored reptile, holds a lungfish in its jaws. Another phytosaur is on the shoreline to the left. In the foreground to the right the amphibian *Cyclotosaurus* is crawling out of the water. A primitive turtle, *Proganochelys,* occupies the far shore to the right.

preserved back part of a skull, were found at Chulabhorn Dam. *Cyclotosaurus* may have attained a length as great as 1.5 meters. Its weak limbs, long, broad, flat head and eyes on top of the skull (so that it could see while almost completely submerged) suggest that it must have spent most of its time in the water.

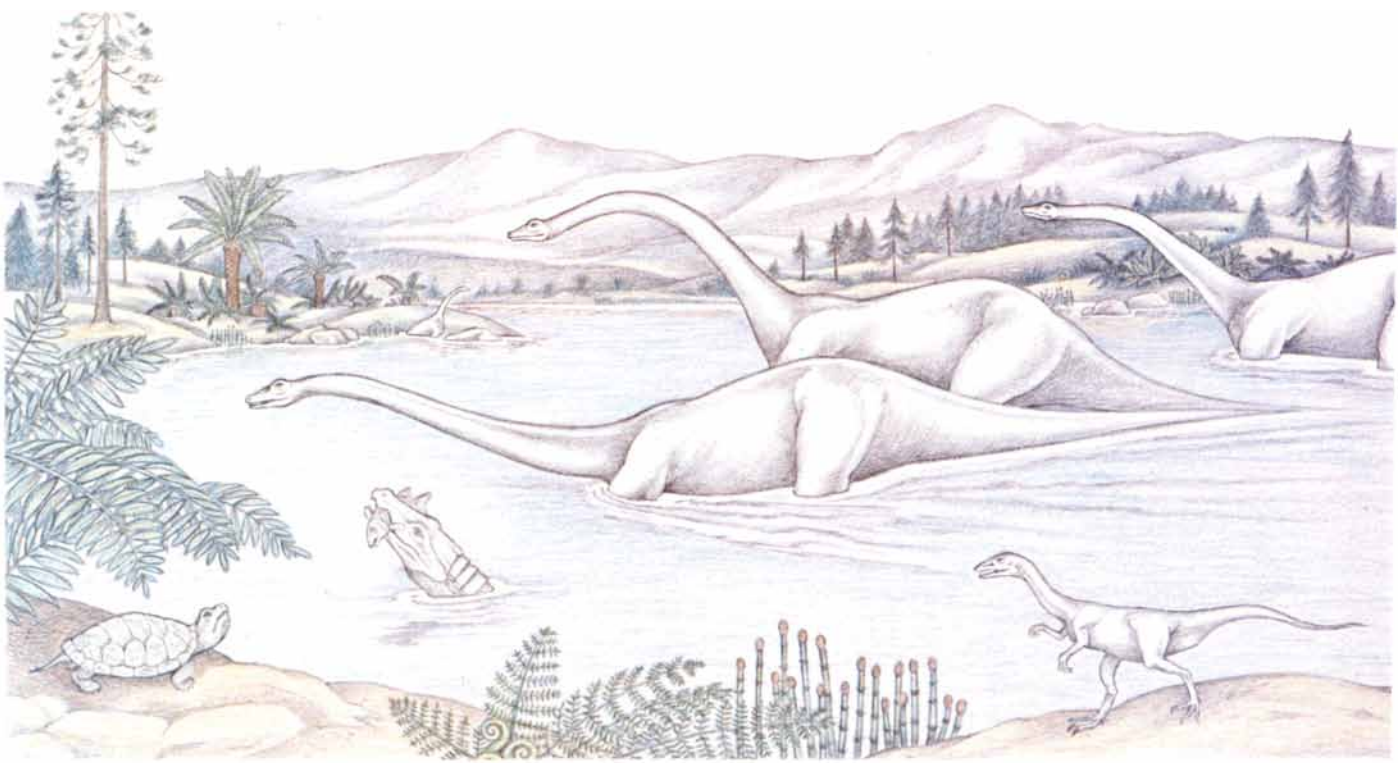
Bones and teeth of other predators, the phytosaurs, were also unearthed at Chulabhorn Dam. Phytosaurs were large, heavily armored, sharp-toothed reptiles that thrived in the late Triassic but apparently did not survive the end of the period. Although their elongated snout, short limbs and long tail gave them a resemblance to crocodiles, they were not closely related: the phytosaurs evolved from the thecodonts, the order of Triassic reptiles that also gave rise, independently and at a later time, to the crocodiles and the dinosaurs. Phytosaurs and crocodiles are thus a classic example of convergent evolution: the emergence of similar evolutionary adaptations to a particular environment from different groups of organisms. Unlike crocodiles, whose nostrils are at the tip of the snout, the phytosaurs had their nostrils far back on the top of the head.

Finally, the remains of turtles, a further group of reptiles, are found in the Huai Hin Lat formation. Although scarce at Chulabhorn Dam, they are extremely abundant elsewhere; indeed, fragments of turtle shells occasionally form distinct layers in the bituminous limestone. The shells were well developed; in other respects, however, the turtles of the Triassic were quite primitive. For instance, they had teeth on their palate, whereas more modern turtles are completely toothless. In general, finds of Triassic turtles are uncommon: outside Thailand the only well-known finds are some specimens from southern and central Germany.

The fossil bones and teeth from the Huai Hin Lat formation give a picture of the community of vertebrates that inhabited the late Triassic lakes and shorelines of northeastern Thailand. They also yield important stratigraphic and paleogeographic information. That is, they help to place the formation in the sequence of geologic strata based on evidence collected around the world. At the same time they help to delineate the geographic distribution of life in the late Triassic and so to determine the arrangement of the earth's landmasses then.

The arrangements can be deduced because the Huai Hin Lat fossils closely resemble fossils from other parts of the world. Phytosaurs provide an impressive example. They appear to have existed only during the late Triassic. On the other hand, their evolution was rapid. Hence the occurrence of the phytosaur genera *Belodon* and *Mystriosuchus* in the Huai Hin Lat formation shows that it must be roughly contemporaneous with formations in Germany that contain the same genera; the latter formations are assigned to a particular part of the late Triassic, the part called the Norian stage. The assignment is confirmed by the amphibians and turtles in the Huai Hin Lat formation, which also have counterparts in the German late Triassic. According to Philippe Janvier of the French National Center for Scientific Research (CNRS), the large amphibian *Cyclotosaurus* from Chulabhorn Dam is closely related to a species from the Norian of Germany. As for the turtles, France de Broin, also of the CNRS, thinks it is related to the genus *Proganochelys,* first described in the German Norian.

From a paleogeographic point of view the Huai Hin Lat vertebrate fauna is interesting because of its resemblance to Laurasian faunas. In the late



LATE JURASSIC LANDSCAPE depicts northeastern Thailand as it may have looked some 160 million years ago; the landscape, a floodplain crossed by large streams, is reconstructed on the basis of fossils found in the siltstones and sandstones of the Sao Khua formation. The dominant forms of life are the dinosaurs. A group of sauropods (quadrupedal, plant-eating dinosaurs) is entering the water; they are thought to have resembled the sauropod *Diplodocus*. The hindmost of them is pursued by a carnosaur (a bipedal, preda-

Triassic all the earth's continents were still largely united; the main "supercontinents," Laurasia in the north and Gondwanaland in the south, were still joined at their western end. Yet the animal life on Laurasia was becoming distinct from that on Gondwanaland. The phytosaurs, amphibians and turtles from Thailand are definitely more reminiscent of the forms from Laurasia. So are some of the fishes: according to Michel Martin working at the University of Paris, the lungfish toothplates from the Huai Hin Lat formation resemble toothplates from China.

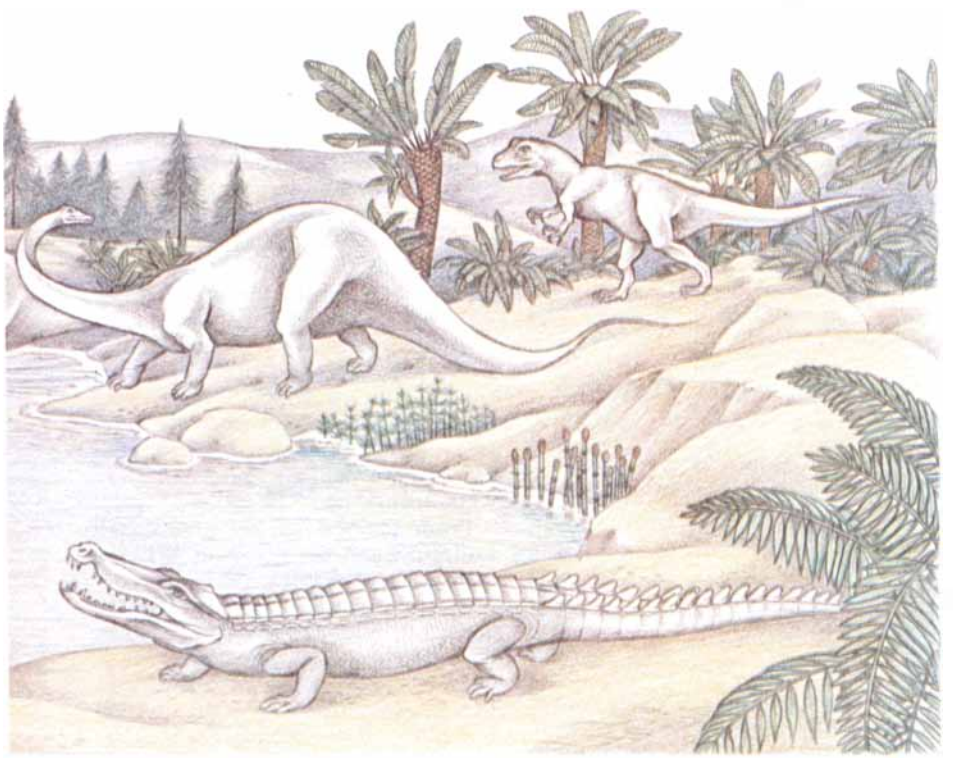
In the context of continental drift these resemblances are significant: they suggest that by late Triassic times there was a land connection between the Indochina microcontinent and Laurasia, which enabled Laurasian vertebrates to colonize Indochina; otherwise the Indochina block would have had a distinctive fauna of its own or a fauna similar to that of Gondwanaland. In this way the resemblances establish the late Triassic as the last possible date for the collision of the Indochina microcontinent with China. The date is in agreement with current theories, which require a Triassic collision, or an even older, Paleozoic one.

Above the Huai Hin Lat formation is the Nam Phong formation, which consists mainly of sandstones and has so far yielded no fossils. It in turn is overlain by the Phu Kradung formation, consisting of siltstones and sandstones. There again vertebrate fossils are found from the Mesozoic. One of the most spectacular finds was made in 1979 by Nares Sattayarak in a road cut near the town of Nong Bua Lam Phu in the northern part of the Khorat plateau; it is the lower jaw of a very large freshwater crocodile. When our 1980 expedition finished the excavation, the jaw proved to be nearly complete. It is 1.14 meters long, which suggests the animal was nearly eight meters long. Its snout was long but robust and held sturdy teeth; the animal must have fed on large fishes. We have assigned the animal to *Sunosuchus,* a genus known from the Jurassic of Gansu Province in China. The Chinese specimen is smaller and differs in certain respects; accordingly we have hypothesized a hitherto unknown species and coined for it the name *Sunosuchus thailandicus.* Nevertheless, the occurrence of closely related species of the same genus of freshwater crocodilian in China and Thailand shows that the two were probably part of the same zoogeographic province in the Jurassic.

The Phu Kradung formation has yielded further vertebrate remains at a conglomerate and sandstone quarry near Ban Wang Dingso, a village west of the Khorat plateau. Many of them were collected by Paladej Srisukh, a veterinarian who is also a keen amateur paleontologist. In addition to crocodilian teeth and bones that may belong to *Sunosuchus,* the remains include the teeth of freshwater sharks and the button-shaped teeth of bony fishes that must have fed on hard-shelled invertebrates. They also include turtle plates. The fossils are being studied; they should provide a better understanding of the Mesozoic fauna of northeastern Thailand. They may also help to establish the date of the Phu Kradung formation. Already the presence of *Sunosuchus* has shown that the formation cannot be Triassic. Its exact position within the Jurassic remains uncertain.

The next two formations present a contrast. The Phra Wihan formation, which lies just above the Phu Kradung formation, has yielded no vertebrate remains, but the Sao Khua formation (the next one up) has proved to be remarkably rich in bones and teeth. The most productive locality in the Sao Khua formation is at Phu Wiang, a great crescent-shaped mountain in the western part of the Khorat plateau. There, in 1976, Roon Song Siri, a Thai geologist prospecting for uranium, found the first dinosaur bone to be reported in Thailand; it turned out to be a fragment of the thighbone of a large sauropod, a long-necked dinosaur. Since the first find at Phu Wiang our team has collected many bones and teeth of a variety of vertebrates. Some remains could simply be picked up from a slope, where weathering had freed them from a greenish, fine-grained sandstone; others required excavation from a matrix of red clay. Apparently these abundant vertebrate remains accumulated in a floodplain environment; rivers must have acted as conduits, producing concentrations of small bones and teeth. The fossils reveal a mixture of communities: the remains of vertebrates that lived in the rivers are mingled with those of terrestrial forms.

The most spectacular fossils from the Sao Khua formation are the sauropod ones: the massive bones of sauropods have now been found in several outcrops of the Sao Khua formation, and a partial skeleton has been excavated at Phu Wiang. (It is being prepared in Bangkok.) The sauropods, with their long necks and tails, were perhaps the most impressive of the di-

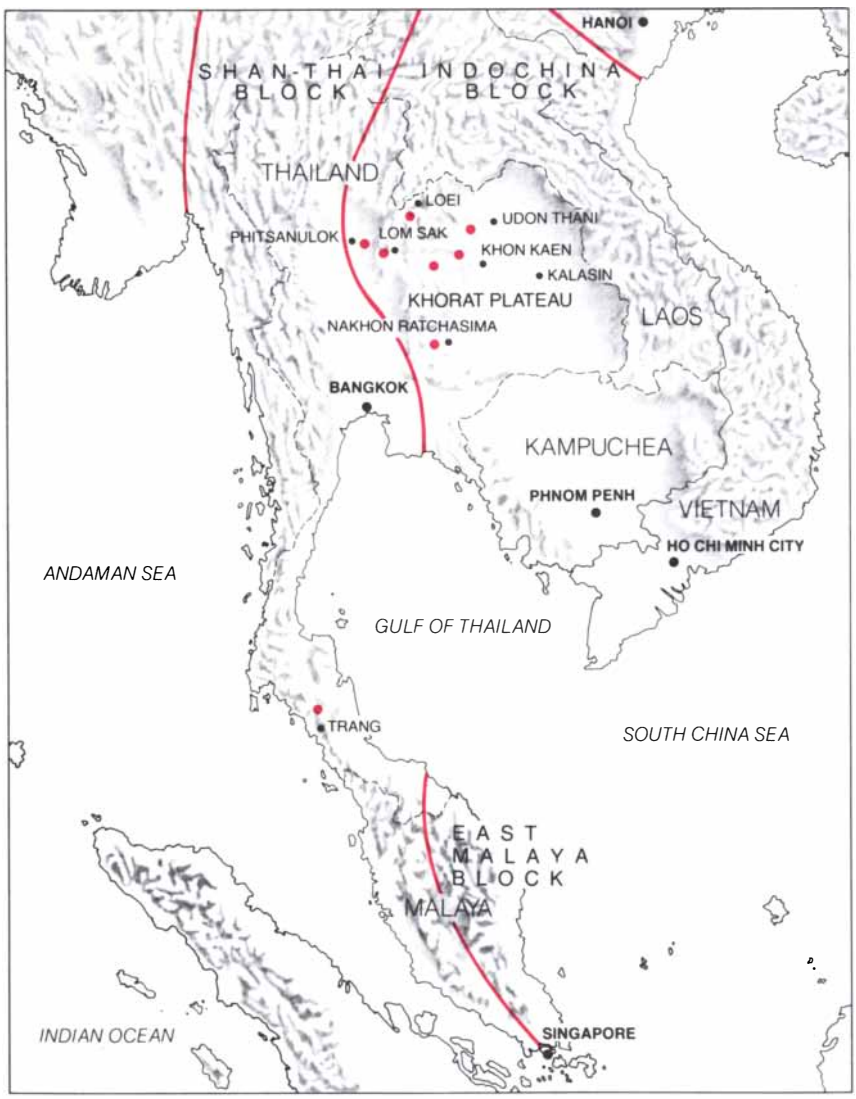


ceous dinosaur) that resembled the carnosaur *Allosaurus,* which is known from North American fossils. In the foreground *Goniopholis,* which resembled living crocodiles but was not ancestral to them, pursues *Compsognathus,* one of the smallest of dinosaurs. (It was no larger than a chicken.) In the water another *Goniopholis* holds the fish *Lepidotes* in its jaws.

nosaurs: the great reptiles of the Jurassic and the Cretaceous. They are certainly the most famous: they included such forms as *Diplodocus* and *Brontosaurus.* By sauropod standards the Thai specimens are not particularly large. Nevertheless, the fossils indicate animals some 15 meters long.

The sauropods were plant eaters. They shared their environment with large dinosaurian predators; bladelike, serrated teeth, which are fairly common at Phu Wiang, reveal the presence of carnosaurs, the largest of the bipedal predators. The carnosaurs at Phu Wiang were probably similar to the late Jurassic North American genus *Allosaurus,* a large animal as much as 10 meters long, and to related forms from China such as *Szechuanosaurus.* Finally, the group of small, slender dinosaurs known as the coelurosaurs is represented at Phu Wiang by a few very small, hollow bones, notably a fragmentary tibia. The bones suggest a slightly built bipedal animal no larger than a chicken, probably closely related to *Compsognathus,* one of the smallest dinosaurs, hitherto known only from two skeletons, one from Bavaria and the other from southern France.

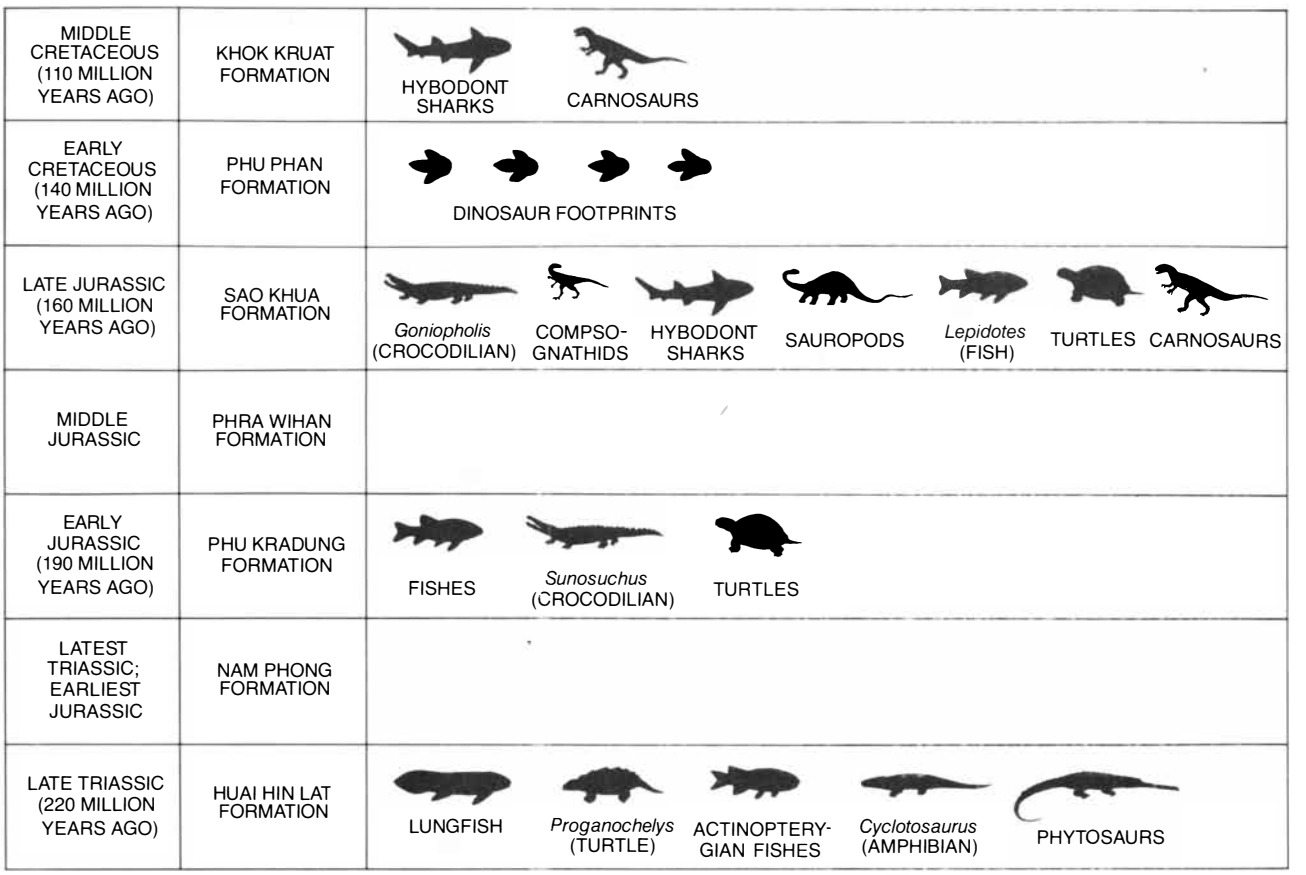
The sauropods also shared their environment with nondinosaurian reptiles. Turtle remains are abundant but fragmentary. Crocodilians are also present in the Sao Khua formation: they include the genus *Goniopholis,* a common crocodilian in the late Jurassic and early Cretaceous deposits of Europe and North America. In external appearance *Goniopholis* was fairly similar to the crocodiles living today; still, it was not ancestral to living crocodiles. Fishes are represented by the spines (that is, the fin reinforcements) and the flat, platelike crushing teeth of freshwater, mollusk-eating sharks and by the buttonlike teeth of *Lepidotes,* a genus that was widespread in the Jurassic. All in all, the fossils collected at Phu Wiang suggest a vertebrate community rather like the ones described for other late Jurassic fossiliferous formations, such as the Morrison formation of the western U.S. or the Shangshaximiao formation of Sichuan Province in China.



JOINING OF TWO MICROCONTINENTS to the southern margin of China created southeastern Asia in the Triassic period (the early part of the Mesozoic era), or probably even earlier, in the Paleozoic era. The microcontinents are the Shan-Thai block, which forms the western part of southeastern Asia, and the Indochina block, which forms the eastern part. The vertebrate fossils found in Thailand so far come from localities (*colored dots*) that lie for the most part on the Indochina block, and in particular the Khorat plateau of northeastern Thailand. Last year, however, a vertebrate-fossil locality was discovered on the Shan-Thai block, close to the city of Trang on the southern peninsula of Thailand.

Above the Sao Khua formation come the rather massive sandstone layers of the Phu Phan formation, which is probably of early Cretaceous age. It has yielded very few fossils. Quite recently, however, some dinosaur tracks have been found in the Phu Phan formation at the Phu Luang Wildlife Sanctuary, a scenic flat-topped mountain in the northwestern part of the Khorat plateau, near the city of Loei. We have studied them with Nares Sattayarak and Varavudh Suteetorn of the Department of Mineral Resources. Some 15 footprints have been mapped; they occupy a sandstone surface a few square meters in extent. Although some of them are badly weathered, others are well preserved. Each of the latter show three well-separated pointed toes and distinct claw marks. The length of each print is about 35 centimeters, the width about 30. They are large, rather birdlike prints, reminiscent of footprints found in early Cretaceous rocks in British Columbia and Texas. We conclude they were made by bipedal carnivorous dinosaurs, probably similar to the carnosaurs whose teeth are found in the Sao Khua formation.

From the size of the Phu Luang footprints and the distances between them it is possible (by means of equations developed by R. McNeill Alexander of the University of Leeds and Richard Thulborn of the University of Queensland) to estimate the size of the Phu Luang dinosaurs and the speed at which they were moving. The calculations indicate animals nearly 1.8 meters tall (from hip to foot), walking at about eight kilometers per hour. Such low speed (hardly twice the speed of a human walker) is in good agreement



GEOLOGIC FORMATIONS that come to the surface of the Khorat plateau make up the Khorat group. Five of the seven formations have now yielded vertebrate fossils, which range from fishes, amphibians (*Cyclotosaurus*) and reptiles (turtles and phytosaurs) of the late Triassic period to fishes (sharks and *Lepidotes*), crocodilians (*Goniopholis*) and reptiles (dinosaurs) of the late Jurassic and the Cretaceous. Together the formations span the Triassic, the Jurassic and the Cretaceous: the three periods composing the Mesozoic era.

with earlier estimates based on other carnosaur tracks—estimates that led Thulborn to conclude that carnosaurs, as well as other bipedal dinosaurs, preferred a walking gait and a pace that would quite likely have minimized the consumption of energy by locomotion. The Phu Wiang footprints show that several dinosaurs of roughly the same size were walking in the same direction, probably together. (The well-preserved footprints are closely grouped and all point in the same direction.) Perhaps carnosaurs exhibited some kind of group behavior; conceivably they hunted in packs for large prey such as sauropods.

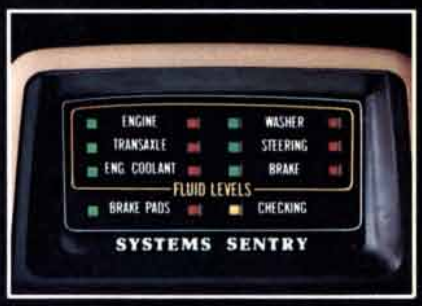
The remains of the last known Mesozoic vertebrates of Thailand come from the Khok Kruat formation, which overlies the Phu Phan formation and is thought to represent the middle part of the Cretaceous. As it happens, the red sandstones and conglomerates of the formation are quarried near the city of Nakhon Ratchasima, also known as Khorat (the Khorat plateau is named after the town), and there a few vertebrate fossils have been found. Land animals are represented by dinosaur fragments. Fish are more abundant. A few years ago an almost complete fish similar to *Lepidotes* was uncovered. The teeth of sharks (presumably freshwater sharks) are of special interest: their peculiar shapes suggest a hitherto unknown form, which may also be present in Cretaceous rocks in Tibet, where some teeth were collected during a recent French-Chinese geologic project. The Tibetan finds may help us to establish the age of the Khok Kruat formation more accurately.

The deposition of the layers of the Khok Kruat formation seems to have been followed by a relatively short-lived marine invasion of northeastern Thailand, followed in turn by the evaporation of the seawater and the deposition of rock salt and potash, found today at the center of the Khorat plateau. No vertebrate fossils have yet been unearthed from these deposits, which may be late Cretaceous or even early Tertiary in age. (The Tertiary period marks the beginning of the Cenozoic era, which followed the Mesozoic.) For now the known fossil record resumes with the mammals from the Miocene epoch of the Tertiary found in some of the coal-bearing basins of northern Thailand.

In spite of the gaps and uncertainties, the picture of the development of vertebrate faunas in Thailand between the late Triassic and the Cretaceous is emerging. Future work on the Khorat plateau will aim to fill the remaining gaps and extend the record both "downward" and "upward." Efforts will also be made to find Mesozoic vertebrates elsewhere in Thailand, in part so that the evidence of the past relations of the Indochina and Shan-Thai blocks can be augmented. The discovery of Jurassic vertebrates, including the crocodilian *Sunosuchus,* at Ban Wang Dingso, west of the Khorat plateau, shows that good opportunities exist. Another example is offered by the discovery of vertebrate fossils in the Lom Sak formation, which occu-

RENAULT ALLIANCE FEATURISTIC

If your car's features are relics of the past, enter the world of Renault. Where European technology is affordable. Look at the array of technological benefits only Renault Alliance offers:



SYSTEMS SENTRY computer continuously monitors all vital fluid levels and front disc brake linings.*

CHILDPROOF LOCKS on rear doors provide extra safety and peace of mind when traveling with children.

KEYLESS ENTRY system uses an infrared beam to unlock doors and get you inside without delay.*

☐ Sedan comfort for five. ☐ Spacious luggage area. ☐ Oil level gauge allows you to check the oil electronically, without opening the hood. ☐ Pedestal-mounted front seats allow rear seat passengers extra room to stretch their legs. ☐ AM/FM six-speaker cassette stereo system with electronic tuning.* ☐ Power door locks and windows.* ☐ Cruise control.* ☐ Twin coaxial rear torsion bars for a quiet, smooth ride. ☐ 40 MPG HWY. 34 MPG CITY.** Renault Alliance. *Built in America.* Featuring European technology and a lot more to spend less on.

Alliance 2-door **$5995**†

Buy or lease. From American Motors. Safety Belts Save Lives.

*Optional Equipment **Base Model 1.4-litre engine. Use EPA est. for comparison. Your results may differ. †Base model. List price. Tax, license, destination charges, optional and regional equipment extra. All options/features not available on all models.

RENAULT
THE ONE TO WATCH

pies a folded mountain belt just west of the Khorat plateau. On the basis of poorly preserved plant remains the formation had been assigned to the late Cretaceous; other evidence, more geologic in nature, suggested a Triassic age. Newly discovered vertebrate remains from the Lom Sak formation, which we have studied with Michel Martin, include fish scales and phytosaur teeth; they definitely indicate a late Triassic age.

Farther from the Khorat plateau a discovery made last December suggests that more evidence is waiting to be unearthed in southern Thailand. Not far from the city of Trang on the southern peninsula of Thailand, remains of various Mesozoic vertebrates have been collected from a layer of grayish clays whose precise geologic age is still uncertain. The remains include turtle plates, fragments of crocodile bones and the thick scales of large bony fishes (possibly *Lepidotes*). The remains are reminiscent of what we have found in the Phu Kradung formation at Ban Wang Dingso, 1,000 kilometers to the north; they suggest, therefore, that the locality is of Jurassic age and that the vertebrate fauna of the Shan-Thai block was similar at the time to the fauna of the Indochina block. The similarity is in good agreement with current hypotheses about the geologic evolution of southeastern Asia, which hold that the collision between the Shan-Thai and the Indochina blocks occurred before the Jurassic, during the Triassic or probably even earlier. The end of the Paleozoic and the beginning of the Triassic are clearly a crucial period, one in which the southeastern Asia blocks were drifting across a body of water called the Tethys Sea and perhaps already coming in contact with southern China.

Last year the British geologist Anthony Hallam wrote that "the changing relationships of eastern Gondwana, Tethys and Asia within the last few hundred million years remain perhaps the outstanding challenge to geotectonic and paleobiogeographic research in the coming years." We hope that our search for fossil vertebrates in Thailand will help to unravel this complex history.

DINOSAUR FOOTPRINTS were discovered in the sandstone of the Phu Phan formation at the Phu Luang Wildlife Sanctuary near the city of Loei in northeastern Thailand. The character of the three-toed prints shows that the prints were made by carnosaurs. In all, 15 footprints were found; perhaps some carnosaurs were walking together at a moderate pace.

Young Supernova Remnants

The remnants of recent stellar explosions in our galaxy are intense X-ray sources. An orbiting telescope has revealed their structure. One has a pulsar; others are expanding shells of shock-heated gas

by Frederick D. Seward, Paul Gorenstein and Wallace H. Tucker

The role of chance in scientific discovery is nowhere more apparent than in the study of supernovas. Only five supernovas, or exploding stars, are known to have been observed in our galaxy in the past 1,000 years—the last one in 1604, before the invention of the telescope. The first precise measurements were those of Tycho Brahe, who studied the spectacular "new star" of 1572 for more than a year, gauging its brightness by comparing it with successively dimmer planets and stars until it finally faded from view. Tycho's observations were a crucial episode in the history of astronomy: they led him to break with the Aristotelian tradition, which held that the realm of the "fixed stars" was immutable. Had Tycho's supernova happened a century or so earlier, before the intellectual authority of Aristotelianism had begun to wane, the course of modern astronomy might have been very different. From medieval Chinese records it is known that other explosions were visible on the earth in 1006, 1054 and 1181—the first of these was almost as bright as the half-moon—yet European astronomers largely ignored them.

Modern understanding of supernovas (including the knowledge that they represent the death and not, as Tycho thought, the birth of stars) is based largely on outbursts in distant galaxies. But whereas no one since Johannes Kepler in 1604 has had the chance to study a nearby supernova, workers today can examine the remnants of the explosions observed by Kepler, Tycho and earlier astronomers. Still among the most luminous objects in our galaxy, the remnants consist of hot stellar debris that is hurtling outward from the center of explosion at a speed of roughly 10,000 kilometers per second. The remnants are interesting in their own right, and they may hold clues to the properties of other, more distant energetic objects such as quasars. At the same time they offer a check on theoretical models of the supernova mechanism. By carefully studying a remnant one can derive a rough estimate of the exploded star's mass, which is a critically important variable in all the models.

Such observations are best conducted in the X-ray range of the spectrum. Although supernova remnants are readily detectable with radio telescopes and some are also visible at optical wavelengths, the young ones are so hot that they radiate the bulk of their energy as X rays. The remnant of Tycho's supernova, for example, emits several hundred times more energy in the X-ray band than the sun does at all wavelengths. Seen through the largest optical telescopes, however, it is an unprepossessing collection of faint wisps. Even the visually striking Crab Nebula, a fossil of the supernova of 1054, radiates most of its energy in the form of X rays.

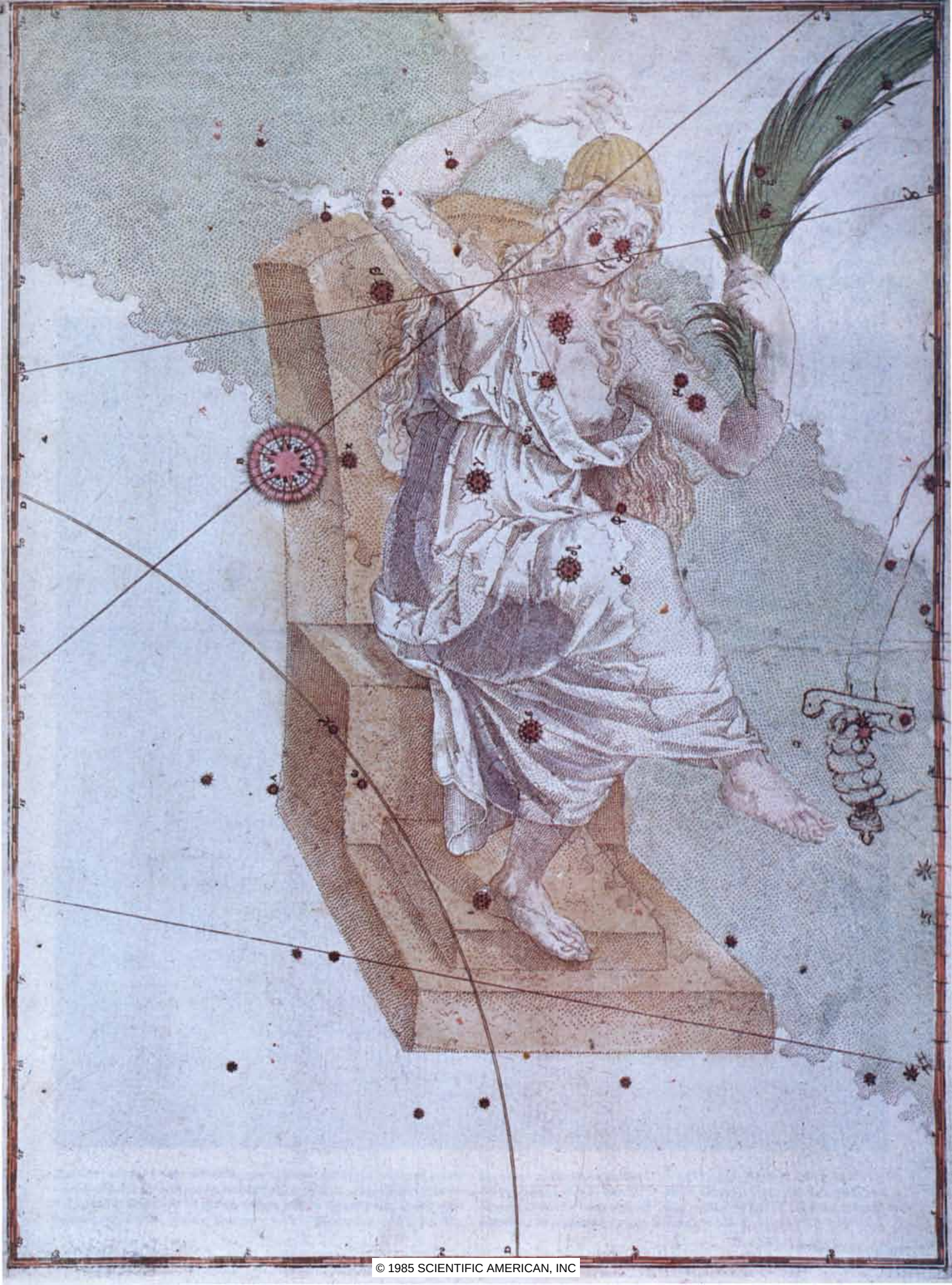
X rays are absorbed by the earth's atmosphere, and so X-ray astronomers depend on instruments carried aloft on rockets or satellites. The first large orbiting X-ray telescope was the Einstein Observatory, which operated from November, 1978, until April, 1981 [see "The Einstein X-Ray Observatory," by Riccardo Giacconi; SCIENTIFIC AMERICAN, February, 1980]. Both in angular resolution and in sensitivity to faint objects the Einstein telescope was a thousandfold improvement over previous instruments. Many of its primary targets were supernova remnants. For the past five years we have been engaged in interpreting the data gathered on these objects.

Long before the X-ray observation of supernova remnants became possible investigators had achieved a general understanding of supernovas. It is widely agreed that a supernova occurs when a star has depleted its nuclear fuel. Throughout its evolution the star maintains its shape by balancing internal pressure against the tendency to collapse under its own gravitation. The internal pressure results from heat generated by nuclear fusion reactions in the core. Initially hydrogen nuclei fuse to form helium; later helium is burned to form carbon. Eventually the fuel in the core is used up, the internal pressure drops and the star contracts.

Most stars do not then explode. A medium-size star (about the size of the sun) collapses until its radius is about equal to the radius of the earth. At that point it is stabilized by a quantum-mechanical effect called degenerate-electron pressure: the electrons resist being crowded too close together. The matter in a collapsed star of this size (called a white dwarf because the collapse has made it white-hot) is a million times denser than ordinary matter.

There is a limit, however, to the amount of gravitation that can be offset by electron degeneracy. In 1935 Subrahmanyan Chandrasekhar showed that a star whose mass is more than approximately 1.4 solar masses will continue to collapse. Ultimately the catastrophic collapse is converted into a catastrophic explosion. The manner of the conversion is thought to depend on

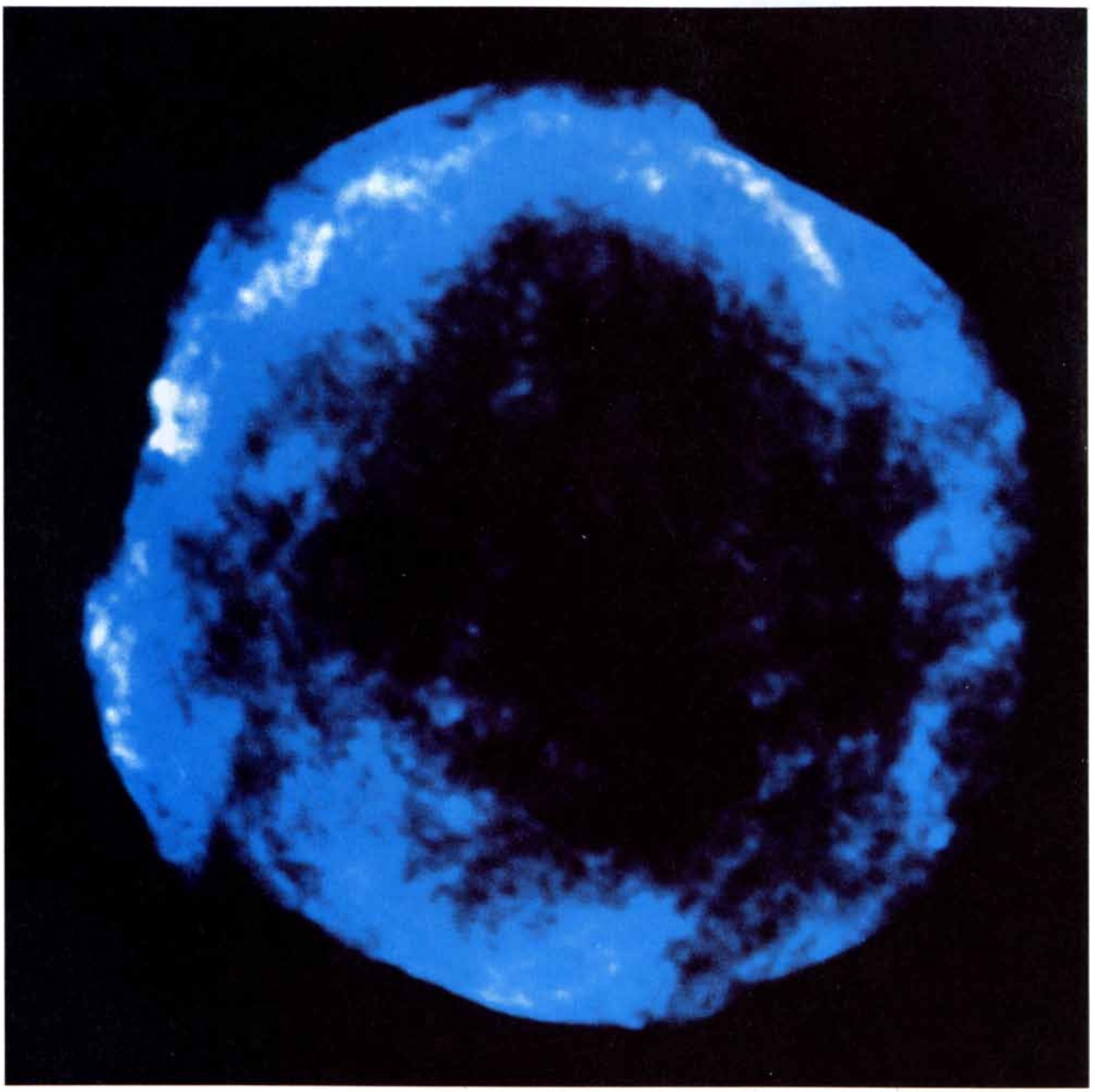
TYCHO'S "NEW STAR" appeared in November, 1572, in the constellation Cassiopeia. In this engraving made by the German astronomer Johann Bayer, originally published in 1603, the supernova is the large star under Cassiopeia's right arm. (In Greek mythology Cassiopeia was the wife of King Cepheus and the mother of Andromeda.) When the star appeared, it had a magnitude of −4 and was nearly as bright as Venus. Tycho observed it until it vanished in March, 1574. The remnant of the supernova was discovered 378 years later.



just how massive the star is; according to current theory, two fundamentally different mechanisms are involved. The mechanisms correspond to the observational classification of supernovas into two main types.

One of the principal bases for classifying a supernova is its light curve: the measure of how its brightness changes with time. Type I outbursts, the rarer of the two classes, all display similar light curves. The brightness of the star increases rapidly during the first few weeks, reaching a peak at which the star may be as luminous as several billion suns, and then falls off gradually over the next six months or so. The light curves of Type II supernovas are more varied, but typically they are about five times fainter at maximum brightness than Type I events, and they fade faster.

Type II supernovas are almost certainly produced by larger stars. Their spectra suggest the explosion takes place in a massive, hydrogen-rich envelope that absorbs much of the radiation. Furthermore, the distribution of Type II supernovas is closely correlated with the distribution of young, bright, massive stars: the outbursts always occur in the arms of spiral galaxies, near star-forming clouds of gas and dust, and they are rarely observed in elliptical galaxies, which consist primarily of old, dim stars. In contrast, Type I supernovas show no preference for spiral arms, and they occur in elliptical galaxies as well. This suggests they are produced by stars that are billions of years old. The mass of such stars cannot be more than a few times



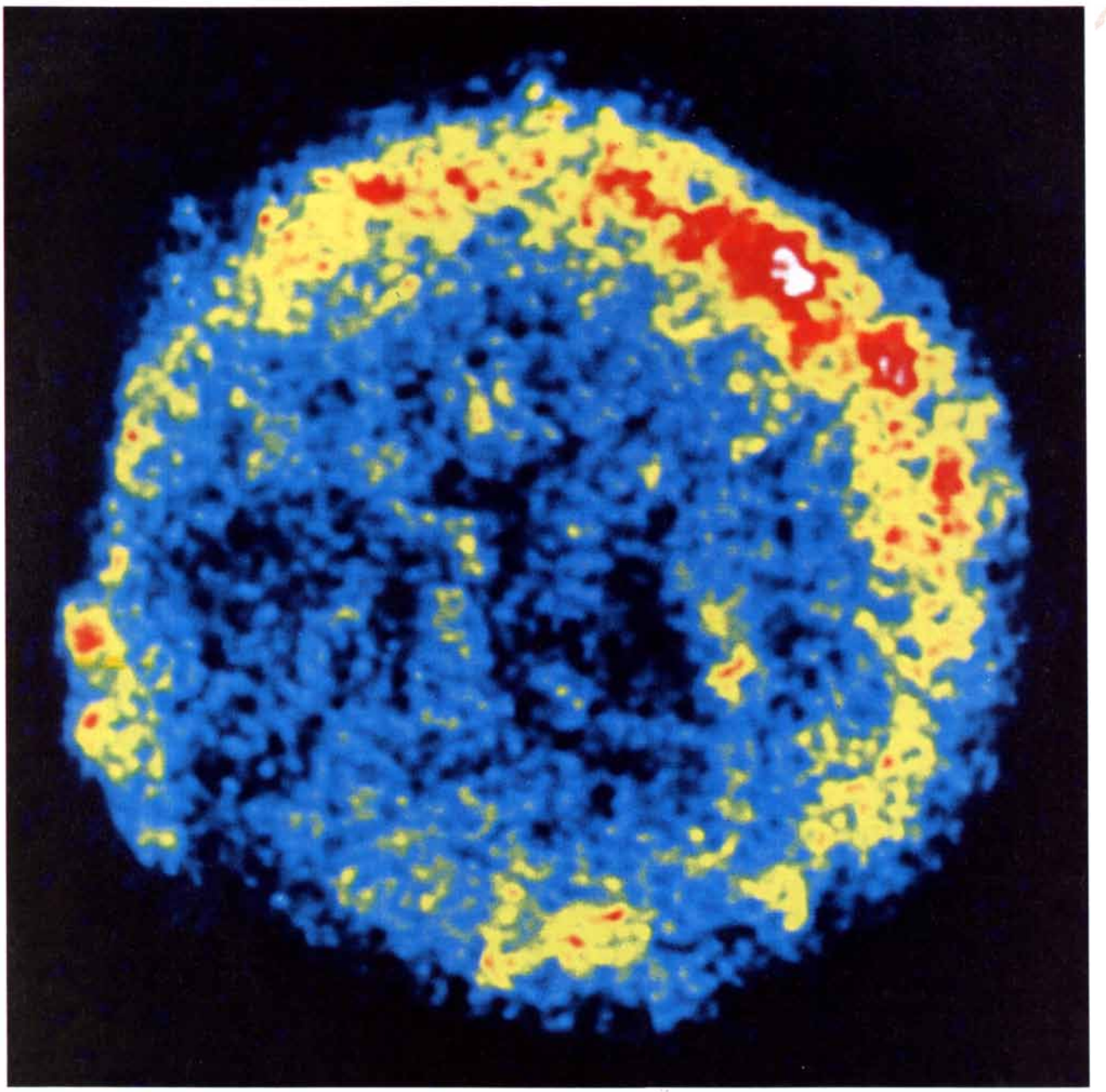
RADIO AND X-RAY IMAGES of Tycho's supernova remnant are strikingly similar. At both radio (*left*) and X-ray (*right*) wavelengths the bulk of the emissions come from a shell whose thickness is roughly 30 percent of the remnant's radius. The lack of evidence for a bright central neutron star supports the view that the remnant was formed by the complete disruption of a white dwarf. On the X-ray image the clumps of hot stellar ejecta in the shell account for the brightest emissions (*yellow, red and white*); individual clumps

that of the sun; otherwise they would have long since consumed their nuclear fuel. (The more massive a star is, the faster it evolves.)

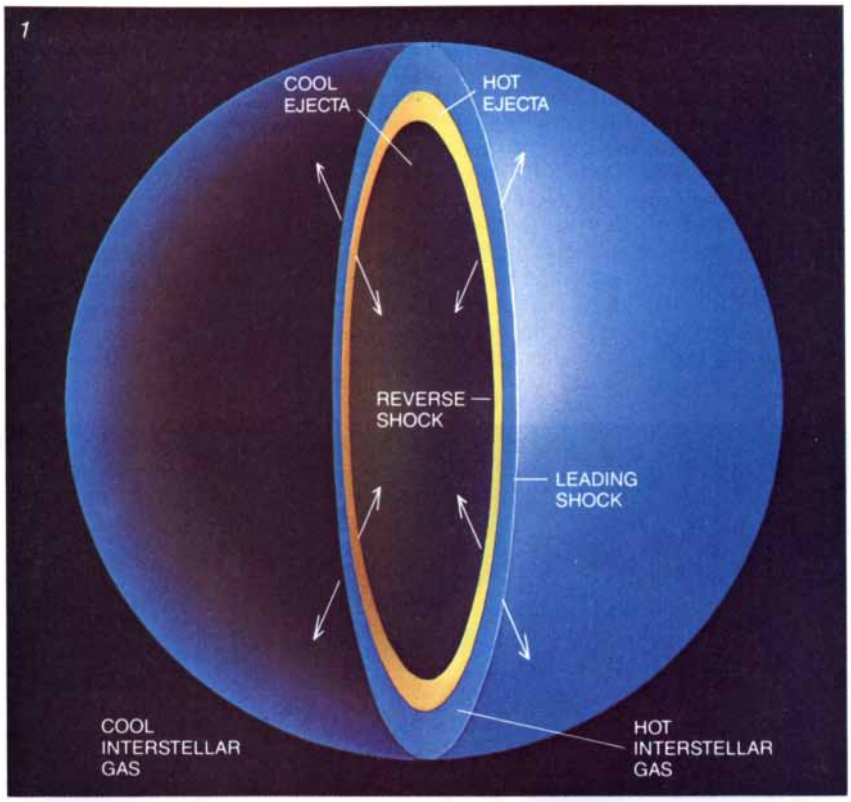
According to the most widely held hypothesis, a Type I supernova arises when a white dwarf has a nearby companion star. If the two are in a close binary orbit, the intense gravitation of the white dwarf can draw matter off the surface of its companion. Eventually the white dwarf, whose mass is initially less than 1.4 solar masses, accretes enough matter to drive it over the Chandrasekhar limit. At that point it begins to collapse again.

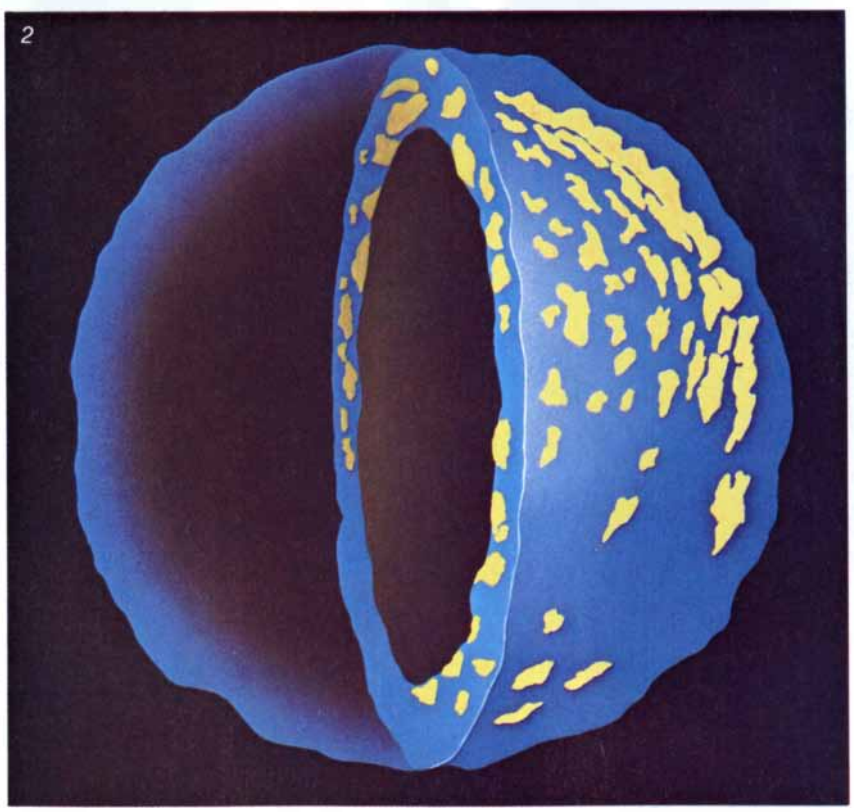
The resulting dramatic increase in temperature and density in the stellar core leads to a new sequence of thermonuclear reactions. Carbon and other light nuclei produced during the normal life of the star fuse to form heavier elements in conjunction with a substantial release of energy. The energy released is enough to disrupt the star completely and to eject the reaction products outward at high velocity. Radioactive nickel decaying to cobalt and iron supplies additional energy to the expanding debris, causing it to glow; many astronomers believe this additional energy explains why the brightness of Type I supernovas declines relatively slowly. The similarity of Type I light curves is attributable to the circumstance that the supernovas all undergo the same radioactive decay processes and are all produced by white dwarfs whose mass is just above the Chandrasekhar limit.

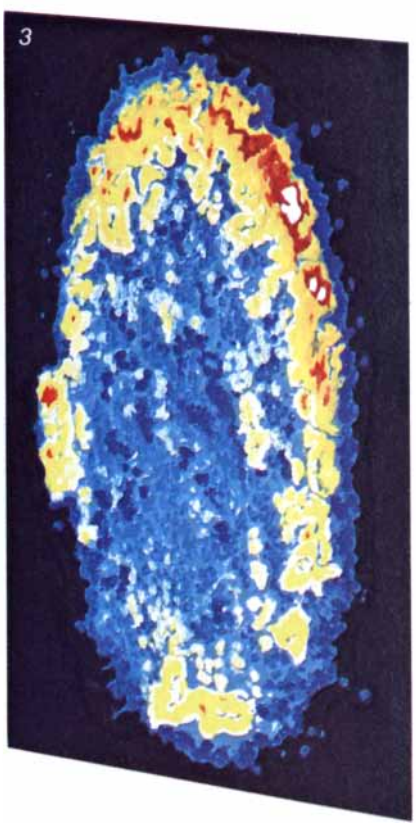
Type II supernovas, on the other



are resolved near the center of the image. Swept-up interstellar gas forms a layer of faint emission outside the shell. This layer is blue on the X-ray image and is most clearly visible at the upper-right edge. The radio map was made at a wavelength of 20 centimeters by David Green and Steven Gull with the five-kilometer telescope of the University of Cambridge. The X-ray image was prepared by the authors from data gathered with the high-resolution imager on the orbiting Einstein Observatory; 22 hours of satellite time was required.







hand, are thought to result from the death of stars at least eight times as massive as the sun. The internal temperature of large stars is much greater than that of small ones, and they evolve through a sequence of contractions and fusion reactions that form progressively heavier elements. At the end of the evolution the core consists primarily of iron, the nuclei of which do not fuse spontaneously.

When this inert central mass exceeds a critical limit, the core—but not the gaseous envelope that surrounds it—collapses suddenly. Electrons and protons in the core are compressed to the point where they combine to form neutrons and neutrinos. As the collapse proceeds the degenerate neutrons offer increasing resistance, analogous to the degenerate-electron pressure that stabilizes a white dwarf. When the radius of the core is about 10 kilometers and its density is equivalent to the density inside an atomic nucleus, the collapse comes to a sudden halt. The gravitational energy released by the collapse is carried outward by neutrinos and

SHOCK WAVES generated by a supernova explosion account for the X-ray and radio emissions from a shell-like remnant. As the expanding shell pushes into the interstellar medium a shock wave precedes it, heating the interstellar gas. The collision slows the expansion. A reverse shock front forms, moving inward with respect to the first shock and heating the ejecta. The hot ejecta, and to a lesser extent the interstellar gas, emit thermal X rays. Shock-accelerated electrons emit radio waves by the synchrotron process. In an idealized supernova explosion (*1*) the shock fronts are perfectly smooth and spherical. In reality (*2*) the shock wave in the interstellar gas is indeed relatively smooth (*see illustration on preceding two pages*), but the reverse shock wave is not because the unevenly distributed ejecta break into clumps. Seen in projection (*3*), as it is on an X-ray image, the shell of clumpy ejecta produces an incomplete ring.

by shock waves that blow off the envelope [see "How a Supernova Explodes," by Hans A. Bethe and Gerald Brown; SCIENTIFIC AMERICAN, May].

Although the Type II mechanism is different from the one underlying a Type I supernova, the amount of kinetic energy it imparts to the stellar ejecta is comparable. Whereas a Type I outburst completely disrupts a white dwarf, however, a Type II explosion leaves behind a tiny cinder a billion times denser than a white dwarf: a neutron star.

The theoretical account we have presented so far has gaps. For example, the fate of a star whose mass lies between 1.4 and eight solar masses is unclear. (It is conjectured that the intermediate stars lose enough mass during their lifetime, in the form of stellar wind, to enable them to evolve into white dwarfs.) More fundamental, the entire explanation of Type I and Type II supernovas in terms of the different masses of their stellar progenitors has to be considered unproved: no one has ever observed a star that subsequently exploded. It has been argued, for instance, that at least some Type I supernovas have occurred in regions where stars are being formed, suggesting they were the product of young massive stars rather than of old white dwarfs. In this connection the observation of supernova remnants can be of some help, because the remnants of large stars and of small stars are expected to be different. In particular, the former should include a dense stellar relict.

Such relicts are not always easy to detect. If the original star is very massive, its core will be unstable even as a neutron star and will collapse to a black hole. The gravitational field of a black hole is so strong that no radiation can escape from it; the object is detectable only by its gravitational pull on a companion star, if it has one. A remnant containing an undetected black hole might be mistakenly attributed to the complete disruption of a small star. Likewise a neutron star, which is only 20 kilometers in diameter, might be expected to escape observation from a distance of thousands of light-years.

Fortunately a neutron star, particularly a young one, is extremely energetic. Because a star's magnetic field is tied to the stellar material, the collapse intensifies the field enormously. The magnetic field of a neutron star is estimated to be about 10 trillion times stronger than that of the sun. Furthermore, the collapse increases the star's velocity of rotation by a factor of as much as 100 million because the angular momentum of the star is conserved. The rapidly rotating magnetic field induces a powerful electric field that pulls charged particles off the surface of the neutron star and accelerates them to high speeds. As the particles spiral around the magnetic field lines they emit photons. This effect, known as synchrotron radiation, is strongest near the magnetic poles, which give off intense beams of radiation. Because the star is rotating, the beams are observed on the earth as a regular sequence of pulses, and the neutron star that produces them is called a pulsar.

Most pulsars emit only radio waves. To radiate at shorter wavelengths the star must rotate rapidly and have an uncommonly powerful magnetic field. Only three pulsars have been found to radiate in the X-ray band. The most luminous one in our galaxy lies in the Crab Nebula, the remnant of the supernova of 1054.

The Einstein X-ray image of the Crab Nebula is dominated by a small region of synchrotron emission—a nebula within the Nebula—around the central pulsar. Ninety-six percent of the X-ray emission comes from electrons moving at velocities close to the speed of light in the synchrotron nebula. The other 4 percent comes from the pulsar itself, which is seen as a bright point source. The spinning star pulses rapidly (30 times per second), but a gradual increase in the pulse period indicates the rotation is slowing down. The star is losing rotational energy at about the same rate as energy is being radiated by the synchrotron nebula in the X-ray, optical and radio bands, which supports the notion that high-energy electrons streaming from the neutron star produce the radiation.

Unlike the Crab Nebula, the remnants of the supernovas of 1006, 1572 (Tycho) and 1604 (Kepler) give no evidence of a pulsar or of any kind of neutron star. If a neutron star were present in a remnant, one would expect to find it on high-resolution X-ray images. At the distance of Tycho's remnant, for example, the thermal radiation from a 400-year-old neutron star should be readily detectable; according to conventional models of the cooling process, such a star would have a surface temperature of several million degrees Celsius. Moreover, although a pulsar would not be visible if the beam did not point toward the earth, a rapidly rotating neutron star should still produce an observable synchrotron nebula. Einstein observations and radio surveys have indeed revealed Crab-like remnants that lack a pulsar, but the remnants of the 1006, 1572 and 1604 supernovas are not among them. It is conceivable that these objects contain unexpectedly cool, slowly spinning neutron stars or even black holes, but we consider it unlikely.

The X-ray images of all three remnants have a shell-like appearance, and all three have been attributed to Type I supernovas. In the case of the outbursts of 1572 and 1604 the evidence for the classification is particularly clear-cut. The observations by Tycho and Kepler were so precise that in the 1940's Walter Baade was able to reconstruct the light curves of the supernovas and show they were typical of Type I events. Chinese records of the outburst of 1006 do not allow a conclusive classification, but its extreme brightness suggests it too was a Type I explosion. On the other hand, some scholars have argued that the Crab supernova of 1054 was a Type II event, although again the recorded observations are inconclusive.

The X-ray observations thus support the idea that there are two distinct mechanisms for producing supernovas, each associated with a particular type of remnant: Type II explosions of massive stars leave Crab-like remnants with neutron stars, and Type I outbursts involving white dwarfs leave hollow spherical shells.

Thanks to its pulsar the Crab Nebula radiates about 70 to 200 times more energy in the X-ray range than the 1006, Tycho and Kepler remnants do. The X rays from the latter objects are not generated by the synchrotron mechanism. Rather they are thermal X rays resulting from collisions of atoms and charged particles in the expanding shell of hot gas. The Crab Nebula probably also has an X-radiating shell that is obscured on the Einstein images by the much brighter emissions from the synchrotron nebula. In the other remnants the large gaseous shell is apparent. It is a tenuous, transparent cloud, yet it may radiate as much energy as a thousand suns.

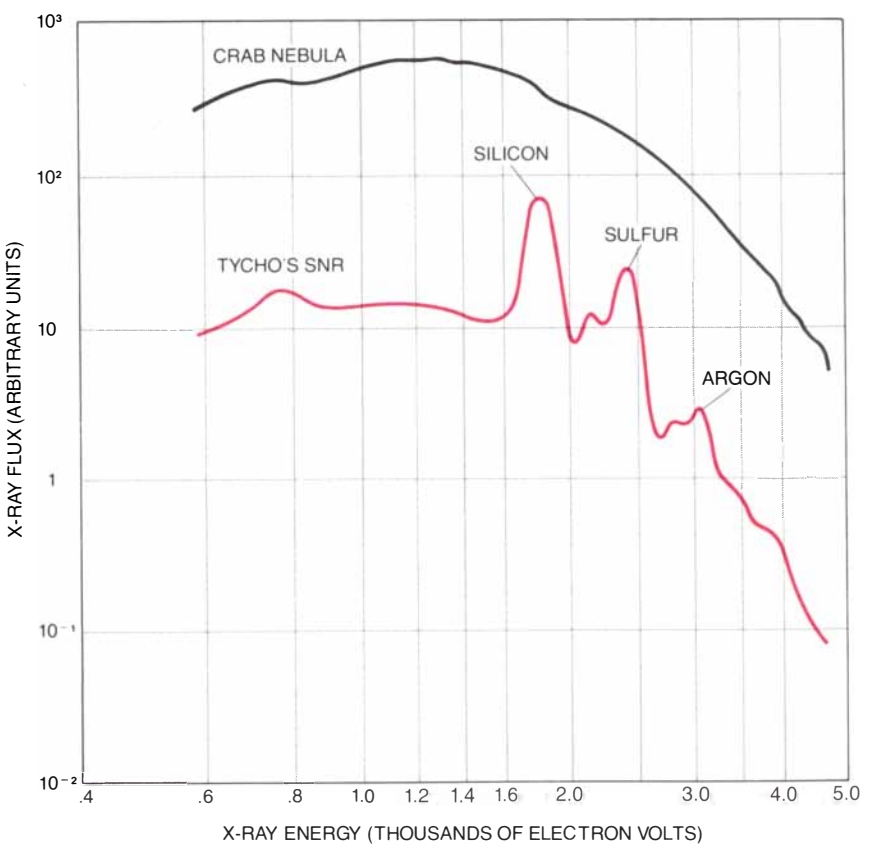
The shell has two components. One component is debris from the exploded star and the other is interstellar gas swept up by the debris. Initially the fastest stellar ejecta travel outward from the explosion with velocities of between 10,000 and 20,000 kilometers per second. Interstellar gas, however, forms a barrier that offers increasing resistance to the expansion. As the leading edge of the shell plows into the interstellar material, two shock waves form. One wave moves ahead of the ejecta and the other, called the reverse shock, moves (with respect to the first shock) back into the ejecta. Seen from the outside both waves move outward. The ejecta tend to become unstable and break into clumps as they push

into the interstellar gas, which is itself not distributed uniformly. As a result the shock fronts do not propagate with the same velocity in all directions, and so they are not perfectly spherical.

The hottest material lies between the two shock waves. Here the previously cool interstellar gas has been heated and compressed by the expanding ejecta, which have in turn been slowed and heated by the collision. Both the shocked ejecta and the shocked interstellar gas radiate in the X-ray range.

The spectra of the shell-like remnants prove that the X-radiation emanating from them is indeed of thermal origin. Most of their energy is radiated at discrete frequencies in the form of atomic emission lines. The lines are produced by atoms or ions in a hot gas that have been excited by collisions with charged particles. Such a collision raises one of the atom's electrons to a higher energy level; the electron then falls back to a lower energy level by emitting a photon. Each emission line is associated with a particular energy transition and hence reveals the presence in the gas of atoms of a particular substance. Synchrotron radiation, in contrast, is emitted across a continuous range of frequencies by free electrons in a magnetic field, and so it reveals nothing about the composition of the gas cloud.

From the strength of their emission lines it is apparent that the levels of silicon, sulfur, argon and calcium are significantly higher in the shell-like remnants than they are in the sun or in the interstellar gas. Tycho's remnant is enriched in these "medium weight" elements by about a factor of six. The enrichment can only be attributed to the supernova. Thus the X-ray spectra provide convincing evidence for a theory frequently stated but seldom verified: that the elements heavier than helium are produced inside stars and are scattered through space by supernova explosions. (As the shells expand over hundreds of thousands of years, the heavy elements become dispersed in the interstellar medium.) One puzzling feature of the spectra, however, is the lack of evidence for elements heavier than calcium—specifically, for the nickel, cobalt and iron expected to be manufactured in Type I explosions. Perhaps these iron-group elements are present at too low a temperature to radiate in the X-ray band.



X-RAY SPECTRA of the Crab Nebula and of Tycho's supernova remnant show that the two remnants produce X rays by different mechanisms. The Crab Nebula contains a pulsar (a spinning neutron star) that emits electrons moving at speeds close to the speed of light. As the electrons spiral around the magnetic field lines in the nebula, they emit synchrotron radiation in a smooth continuum across the entire X-ray band. In contrast the spectrum of Tycho's remnant shows strong thermal emission lines from ions of silicon, sulfur and argon formed during the supernova explosion. The ions are excited by collisions in the gas, whose temperature is roughly six million degrees Celsius. The spectra were prepared by workers at the Goddard Space Flight Center using the Einstein Observatory solid-state spectrometer.

In principle the X-ray observations enable one to calculate the mass as well as the composition of a thermally radiating remnant. A fairly simple formula relates the luminosity of a gas cloud to its temperature, volume and density. The temperature can be inferred from the spectrum of the gas, and so one can derive the density of a remnant from its observed luminosity. If the X-ray data make it possible to distinguish between stellar ejecta and swept-up interstellar gas in the shell, the mass of each can be computed separately. Because the mass of the ejecta should approximately equal the mass of the exploded star, the computation indicates whether the star could have been a white dwarf, as theory predicts for Type I supernovas.

We have done the calculation for Tycho's remnant. On its X-ray image the shocked interstellar gas and the ejecta are distinguishable [*see illustration on page 91*]. The interstellar gas is faintly visible at the edge of the image, just outside the main ring where the emissions are brightest. The ring is probably formed by clumps of ejecta superposed along the line of sight; individual clumps are clearly resolved in the center of the image. We calculated the average size and mass of the central clumps. By assuming the averages also apply to the unresolved clumps in the ring, we determined that the shell of Tycho's remnant contains some 400 clumps amounting to about one solar mass of material. Diffuse emission from the shell indicates the presence of another solar mass of unaggregated debris, making the total mass of the ejecta about twice the mass of the sun.

That is a bit high for a white dwarf, which would have no more than 1.4 solar masses. The discrepancy may be attributable to the uncertainties in our calculation. In particular the distance of Tycho's remnant (or of any of the other remnants) is not known precisely; we assumed a distance of 10,000 light-years. If the real distance is 15 percent less, the remnant's intrinsic X-ray luminosity would be smaller and the estimated mass of its stellar progenitor would fall to about 1.5 solar masses. Our results are therefore compatible with the theory that Tycho's remnant was produced by the complete disruption of a white dwarf. Although we cannot unequivocally rule

REMNANT	SUPERNOVA YEAR	SUPERNOVA TYPE	DISTANCE (LIGHT-YEARS)	ANGULAR DIAMETER OF X-RAY EMISSION (MINUTES OF ARC; FULL MOON = 30)	INTRINSIC DIAMETER OF X-RAY EMISSION (LIGHT-YEARS)	RELATIVE X-RAY INTENSITY
SNR 1006	1006	I?	3,300	30	28	6
CRAB NEBULA	1054	II?	6,500	1.5 × 1.7	2.8 × 3.2	1,000
SNR 1181	1181	II?	8,500	4 × 6	10 × 15	0.2
TYCHO'S SNR	1572	I	10,000	8	22	14
KEPLER'S SNR	1604	I	16,300 – 32,600	3	14 – 28	5
CASSIOPEIA A	~1680	II?	9,100	4	10	40

SIX SUPERNOVA REMNANTS in our galaxy are associated with stellar explosions of the past 1,000 years. None of the distance and size estimates are precise, but the estimates for Kepler's remnant are particularly uncertain. The date of the supernova associated with Cassiopeia A is unclear because there is no proof it was seen. The supernovas of 1006, 1054 and 1181 were recorded in China.

out the possibility that cool, unobservable ejecta, or even a neutron star, are hidden inside the hot shell, we think the weight of the evidence favors a low-mass explosion.

From the brightness of the shocked interstellar gas we have also estimated the mass of the swept-up material. It too is about two solar masses. Tycho's remnant appears to be in an intermediate stage of development in which the swept-up material is beginning to have a significant effect on the shell. In a few more centuries interstellar material will dominate the dynamics of the shell's expansion and will account for most of the radiation. As the material becomes thoroughly mixed with the ejecta, most of the information on the supernova explosion—the initial expansion velocity and the mass and heavy-element content of the star—will be lost. The mixing of ejecta and interstellar gas already makes it difficult to estimate the mass of the supernova of 1006. The remnant of the 1006 event, judging from its X-ray spectrum, is less enriched in the heavy elements than Tycho's remnant is, indicating that its stellar material has been more diluted.

The remnant of Kepler's supernova of 1604 presents another problem: it is probably twice as far away as Tycho's remnant, and on the Einstein image the shocked interstellar gas is not easily distinguishable from the ejecta. It might be possible anyway to estimate roughly the mass of the star, but we have not tried to do so. As we stated above, however, the absence of evidence for neutron stars in the 1006 and Kepler remnants supports the hypothesis that they were formed by exploding white dwarfs.

The absence of a neutron star in another shell-like remnant, one we have not mentioned so far, is more troubling. Called Cassiopeia A, it is the youngest remnant; the expansion velocity of the shell implies that the supernova occurred in the last decades of the 17th century. There is no conclusive evidence that the outburst was witnessed at all on the earth (although William B. Ashworth, Jr., of the University of Missouri has argued that John Flamsteed, the first astronomer royal of England, observed a "new star" in approximately the right location in 1680). Since the supernova did not arouse widespread interest, it certainly could not have been very bright, and this strongly suggests it was a Type II event. According to theory, the remnant should have a dense star at its center, but the X-ray image does not reveal the synchrotron nebula one would expect around a neutron star. In the absence of observational evidence one can only speculate that Cassiopeia A may contain a black hole.

In addition to being a luminous X-ray object Cassiopeia A is the brightest source of radio emissions in the sky; the other shell-like remnants can also be mapped in detail at radio wavelengths. (Cassiopeia A, Tycho's remnant and many older remnants were actually discovered as radio objects.) Whereas the X rays from these remnants are thermal, the radio emissions are not. They are synchrotron radiation produced by enormous numbers of highly energetic electrons moving in a magnetic field. The electrons cannot be as energetic as the ones in the Crab Nebula or they too would emit X rays. Nevertheless, they do achieve velocities approaching the speed of light. Given that the shell-like remnants appear to contain no neutron stars, where do the energetic particles come from?

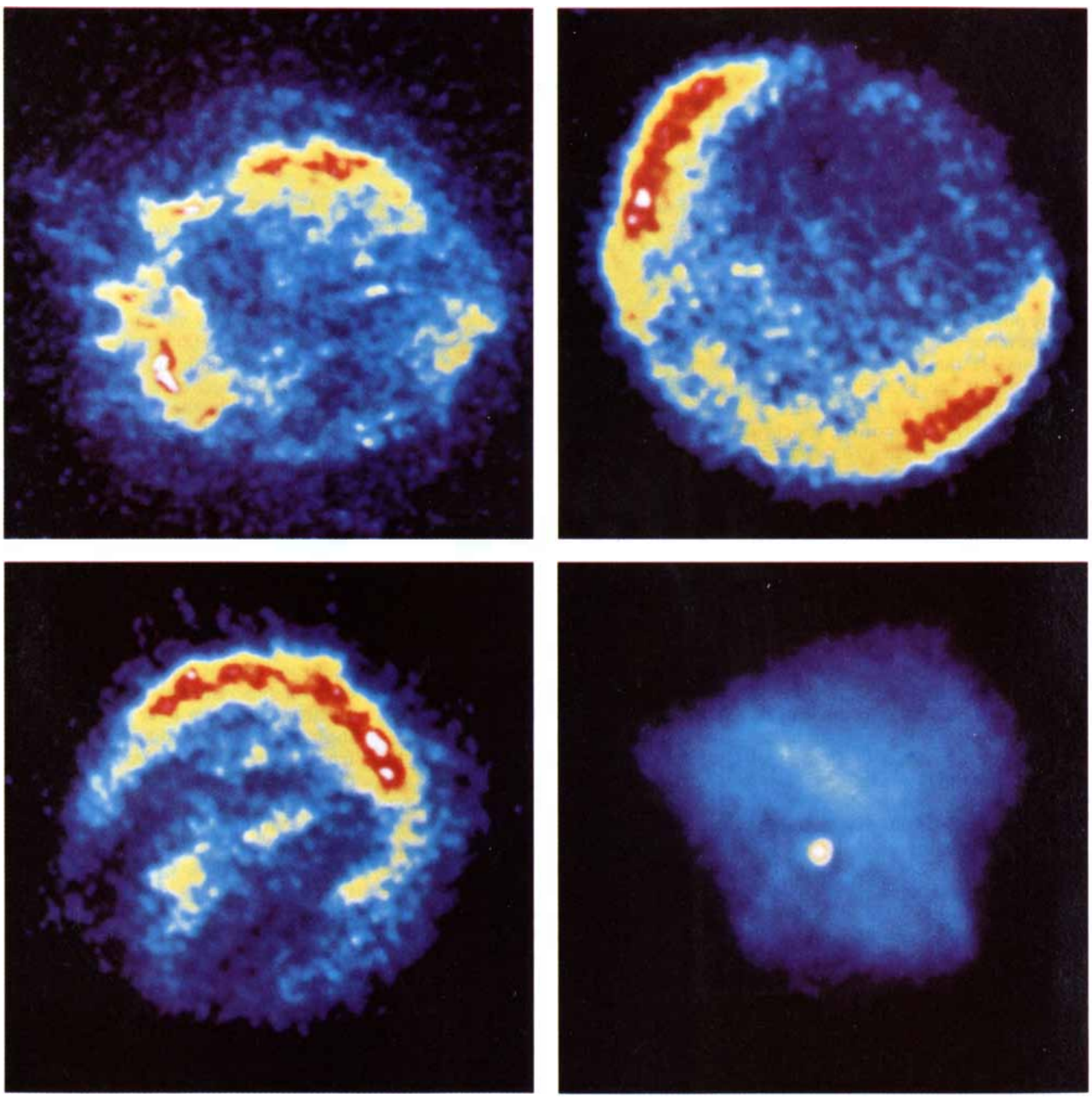
The question is of particular interest because a similar phenomenon seems to operate on a much larger scale in active radio galaxies and in quasars. In the case of supernovas the initial explosion cannot account for the presence in the remnant of high-energy electrons, because over the centuries the particles either would have escaped from the remnant or would have radiated most of their energy. Somehow the shock waves from the explosion must continue to accelerate electrons to high energies.

If the shock waves do account for the radio emissions, one would expect the radio and X-ray images of a shell-like remnant to be similar because most of the X rays come from shock-heated matter. Indeed, the correspondence between the two images is striking in all the young remnants we have examined. Combined radio and X-ray observations may in the future help to reveal the details of the shock-acceleration mechanism. According to one promising hypothesis, the ionized clumps of gas in the shocked ejecta form turbulent magnetic eddies; electrons in the remnant are accelerated by collisions with the clumps and emit

radio waves as they spiral around the magnetic field lines.

The discovery of Tycho's remnant with the Jodrell Bank radio telescope in 1952, some 378 years after the star had disappeared from view, was a milestone in astronomy. It was altogether appropriate that the value of the new technology should be demonstrated on an object associated with Tycho, since he was the first of a long line of astronomers who have devoted much of their energy—of which the flamboyant Tycho had a great deal—to the improvement of astronomical instruments. In the decades since the Jodrell Bank discovery X-ray images of remnants have added still another dimension to the study of supernovas.

Yet nothing learned from the observation of remnants matches what astronomers could learn about supernovas if they had the chance to see another bright explosion in our galaxy. Because nearby stars are by now well catalogued, it would be possible, for instance, to say whether the star that exploded was a white dwarf or a more massive object. When the next galactic supernova occurs, astronomers will be ready, thanks to the tradition begun by Tycho, with an arsenal of powerful instruments. Their observations will no doubt open up intriguing new avenues of investigation.



TWO TYPES OF SUPERNOVA REMNANT are illustrated by X-ray images from the Einstein Observatory. Like Tycho's remnant, Cassiopeia A (*upper left*), SNR 1006 (*upper right*) and Kepler's remnant (*lower left*) are shell-like: their X rays are thermal emissions from an expanding shell of stellar ejecta and swept-up interstellar gas. None of them seems to incorporate a central pulsar. In contrast, the X rays from the Crab Nebula (*lower right*) are synchrotron radiation from a small region around a bright central pulsar. The glare of the synchrotron nebula may obscure a surrounding shell; the Crab Nebula is 170 times as bright as SNR 1006 even though it is about the same age. (The true relative sizes and brightnesses of the remnants are given in the table on the preceding page.)

THE SEA OF SHAME.

The Gulf of Thailand is more than a body of water between Thailand and Vietnam. It's a gulf between human beings and their sense of humanity.

For almost ten years, it's been the scene of the atrocities dealt upon the Boat People. Since 1982 alone, more than 1,800 have been killed by pirates, 2,300 women have been raped, 850 more women have been abducted and remain missing. Still, the Boat People come.

The International Rescue Committee sees to it that those lucky enough to survive... those who have seen so little humanity along their journey... will at least find some once they arrive. For over 50 years, we have provided refugees with food, shelter, clothing, and medical care. We've helped them rebuild their lives and their hopes.

We can't do it alone. We need support from you. Please give to the IRC. Help those who cross The Sea of Shame find a cup of compassion.

IRC

People with the courage to start over should have the means to start over.

Send your tax-deductible contribution to: the International Rescue Committee at 386 Park Avenue South, New York, NY 10016.

Lawrence Robbins

The Antarctic Ice

On the average it is 2,200 meters deep; its climatic history is being learned through the analysis of deep cores. The bedrock of the continent has been mapped

by Uwe Radok

The words terra incognita (unknown land) that appear so often on ancient maps might reasonably be invoked even today to describe the vast sheet of ice that covers Antarctica and discharges more than a trillion (10^{12}) tons of ice each year into the seas surrounding the continent. Little was known about the ice until the findings made during the International Geophysical Year (IGY) were reported 20 years ago. Since then another international research effort—the International Antarctic Glaciological Project (IAGP)—has built up a further fund of information about the core of the region, the East Antarctic ice sheet. The findings are valuable for the knowledge they yield of a little-known area; they also help in assessing global pollution and the increase in atmospheric carbon dioxide resulting from the burning of fossil fuels. In this article I shall describe some of the findings of the project, with which I have been associated since its inception in 1969.

The study of the Antarctic ice sheet can be said to have begun when the Bellingshausen expedition dispatched by imperial Russia in 1820 confirmed the existence of an ice-covered southern continent. Subsequent expeditions gradually established the lateral extent of the ice sheet, but its average height remained totally unknown as late as 1956, at the beginning of the IGY. The only estimate of the average elevation had been made in 1911 by the German geographer Wilhelm Meinardus. He analyzed seasonal changes in barometric pressure and reasoned ingeniously that they result from air moving to and fro over the edge of the ice sheet. His calculations gave the average elevation of the ice as about 2,100 meters.

The first opportunity for checking the elevation came when data from numerous surface traverses made during the IGY were incorporated in the first factual maps of Antarctica. Even in these maps a large central area remained devoid of information, but the general shape of the ice sheet emerged. Its main features are the large East Antarctic ice mass and the much smaller West Antarctic lobe (which is itself as big as Greenland). The center of the ice sheet is not at the geographic pole but at 83° S, 53° E. That point is equally remote from all surrounding coasts and is therefore termed the Pole of Relative Inaccessibility. The mean elevation of the ice as recorded on the IGY maps is about 2,200 meters above sea level, quite close to the estimate made by Meinardus.

Progress since the IGY has resulted largely from technological advances such as satellite observations, remote-sensing devices and computer simulations. The organizational impetus from the IAGP has also been crucial. This project unites five major expeditions (of Australia, France, Japan, the U.S. and the U.S.S.R.) and a group from the U.K. in a concerted investigation of the East Antarctic ice sheet. It is run by a coordinating council of participating investigators who meet once a year to compare results and plan further operations, both separate and cooperative. Reports on the annual meetings appear in *Antarctic Journal of the United States* as a special contribution to the IAGP arranged by Guy Guthridge and Richard M. Cameron of the National Science Foundation's Division of Polar Programs.

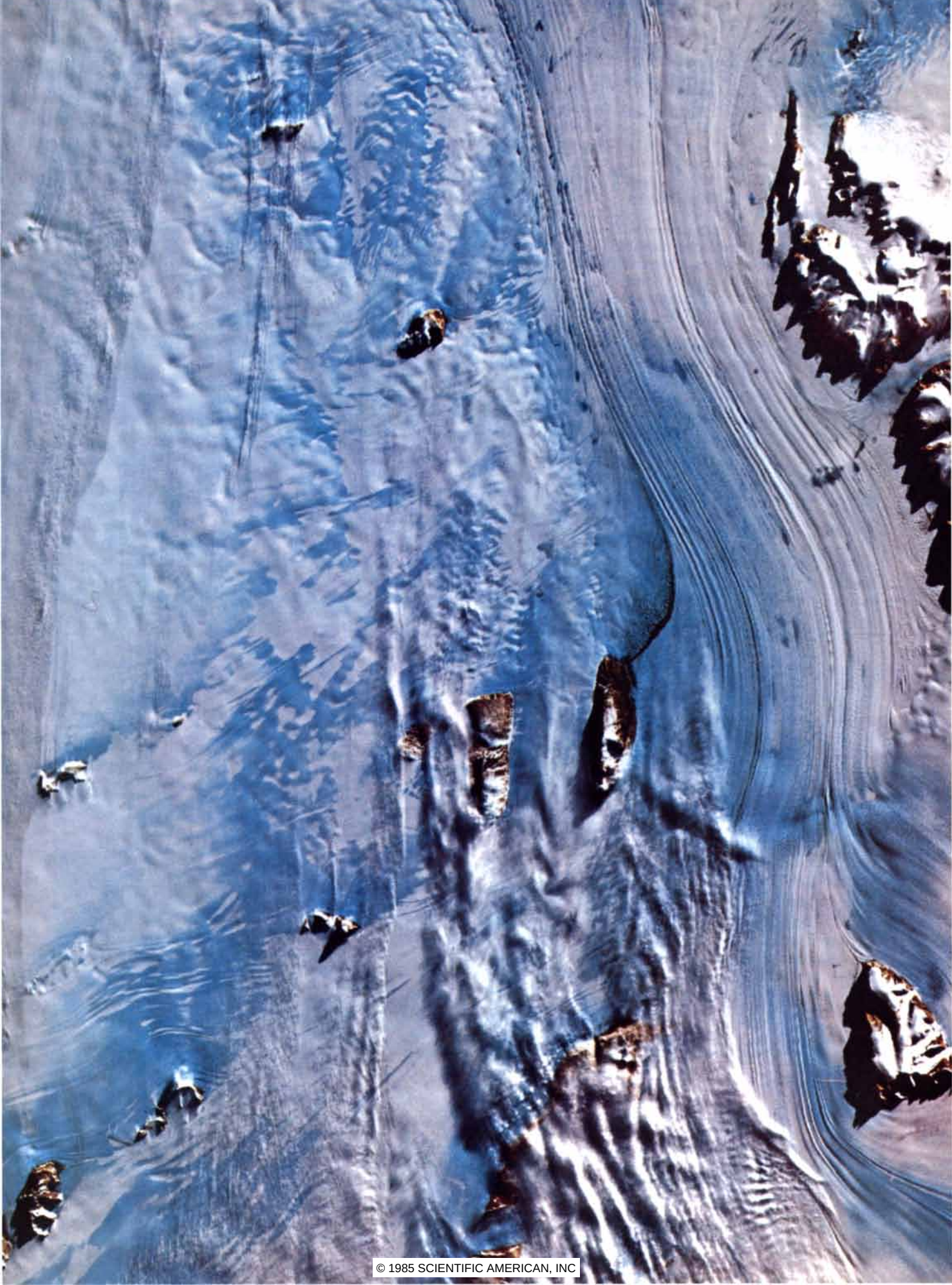
The ice sheet poses three broad research tasks. Its topography (surface and bottom) needs to be determined to yield a detailed description of the surface elevation and the thickness of the ice. The second task is to clarify what is called the regime of the ice: its temperature at various depths and the balance between what is gained as snow and what is lost as icebergs and meltwater. The third task is to determine the structural and chemical properties of the ice. Finally, all this information needs to be combined in models of the ice sheet that describe its history and predict its future.

During the IGY the surface elevations of the ice were determined along traverse routes by measuring the barometric pressure. The pressure changes with altitude at a rate that can be calculated from temperature readings in the atmosphere above the ice sheet. At that time and for some years afterward the thickness of the ice was determined by seismic and gravity methods. The seismic method involved interpreting seismic waves reflected from the ice bottom in response to explosions set off at the surface. The gravity method entailed measuring the value of gravity at numerous places, relying on the fact that the force of gravity at any point depends in part on the proportions of ice and rock under that point. (The altitude and latitude of the point are also factors.) The average thickness determined by these methods was approximately 2,000 meters. The latest estimate is 2,160 meters.

Eventually it was discovered that radar offers a better way of measuring both elevation and thickness. Its usefulness was revealed by difficulties with the radar altimeters of airplanes over Antarctica and quite independently by observations of the ionosphere during the IGY.

Pilots in the Antarctic had noticed that their radar altimeters occasionally

LAMBERT GLACIER, a prominent feature of the East Antarctic ice sheet, is the world's largest. The glacier flows into the Amery Ice Shelf, which is at the top of the picture. The satellite image covers an area of about 24,300 square kilometers (135 by 180 kilometers). This computer-enhanced image was prepared by the Image Processing Facility of the U.S. Geological Survey in Flagstaff, Ariz., from data obtained by a Landsat satellite.

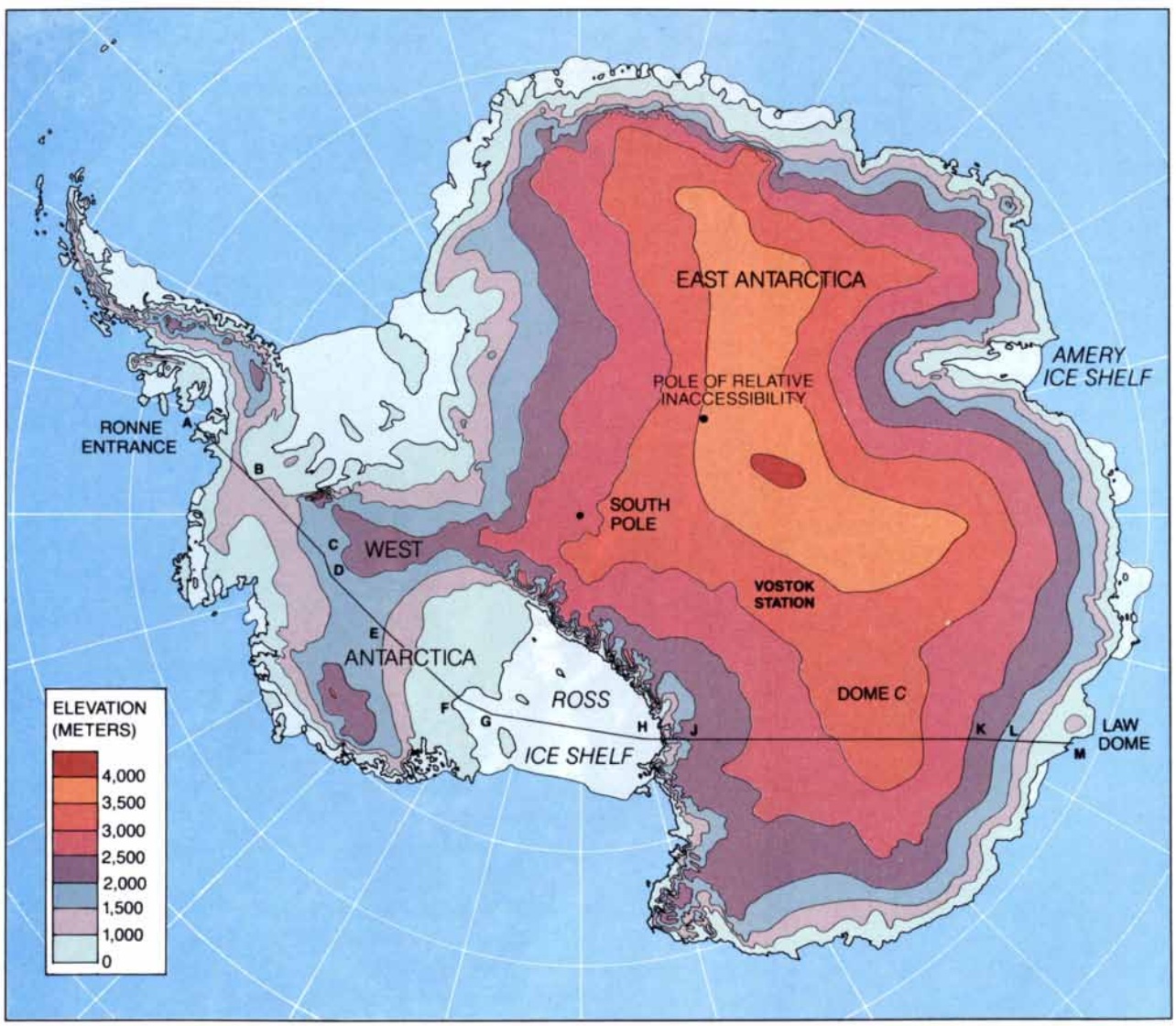


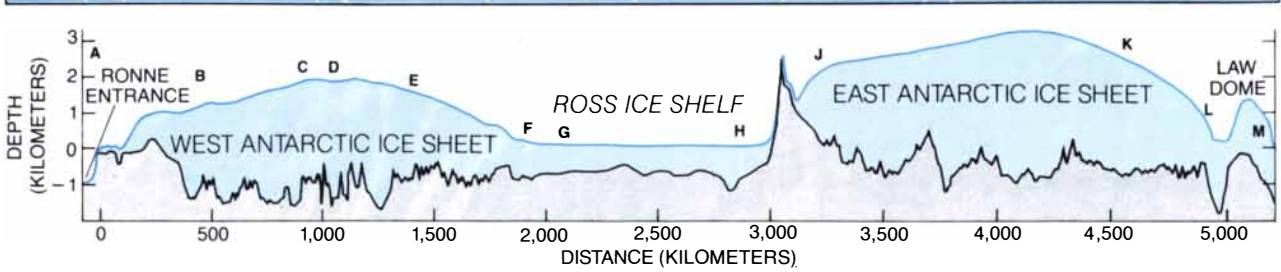
gave excessively high readings over the ice sheet. Indeed, the phenomenon had led to crashes and loss of life in "whiteout" conditions, when the surface cannot be discerned. The ionospheric phenomenon, which at first seemed to be quite unrelated, was that radar signals sent up from the ice sometimes failed to produce echoes even though the reflective conditions were good.

The explanation of both phenomena is that some of the radar signals had penetrated the ice and had been reflected upward from its base. The radar altimeters had then added the ice thickness to the true altitude, and the ionospheric echoes had been canceled by interference with the echoes from the rock below.

The first successful experiment in measuring the thickness of the ice by radar was made in 1957 by Amory H. Waite of the U.S. Army Signal Corps. The full development of what has become the discipline of radioglaciology was carried out at the University of Cambridge by Stanley Evans, Gordon de Q. Robin (the U.K. members of the IAGP's coordinating council) and David J. Drewry. As part of the IAGP the British group and the American, Russian and Australian expeditions have employed airborne radar to map the surface and bedrock elevations of a large part of Antarctica. Supplementary maps of the surface topography have been made with radar data collected by balloons circling the earth at a constant high level and by satellites.

The radar soundings have confirmed the existence of major mountain chains under the ice. They have also revealed the existence of deep ravines and extensive flat patches. The flat

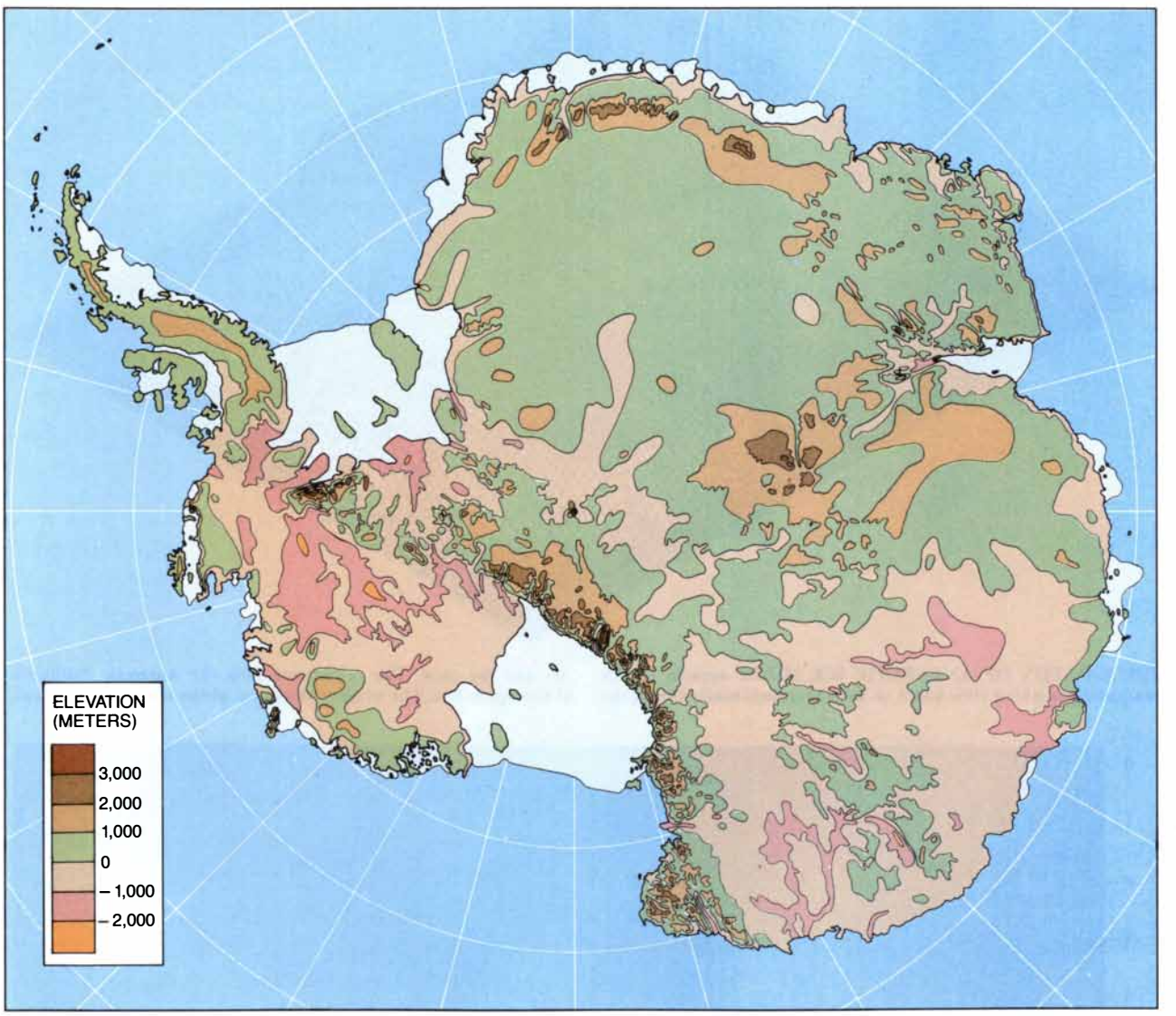




patches are presumed to be subglacial lakes, kept unfrozen by heat from the interior of the earth. Such topographic data are of prime importance in the selection of suitable sites for drilling into the ice to extract long cores that reveal the history of the ice over several hundred thousand years.

The increasing refinement and power of ice-radar systems has disclosed that most of the ice sheet is not uniform but instead is pervaded by internal layers. They reflect enough radar energy to be recorded in addition to the echoes from the surface and bottom. Such internal echoes could arise from a variety of irregularities in the ice. Among them are changes in density and crystal structure (which could result from melting or deformation), concentrations of impurities and layers of increased acidity resulting from ancient volcanic activity.

The internal echoes provide information on the history of the ice and therefore are the subject of intensive research involving ice cores and the ways in which electromagnetic waves behave and change as they travel through the ice. Recently Charles R. Bentley (the U.S. member of the IAGP's coordinating council) of the University of Wisconsin at Madison and his colleagues and students completed an extensive IAGP program of measurements in and around a 900-meter hole cored by French investigators on the ice rise known as Dome *C*. These measurements established firm values for the propagation velocity of radar signals in the ice sheet and explored for the first time the electrical resistivity of ice as an indicator of structural and chemical characteristics of the ice sheet. A finding of particular interest



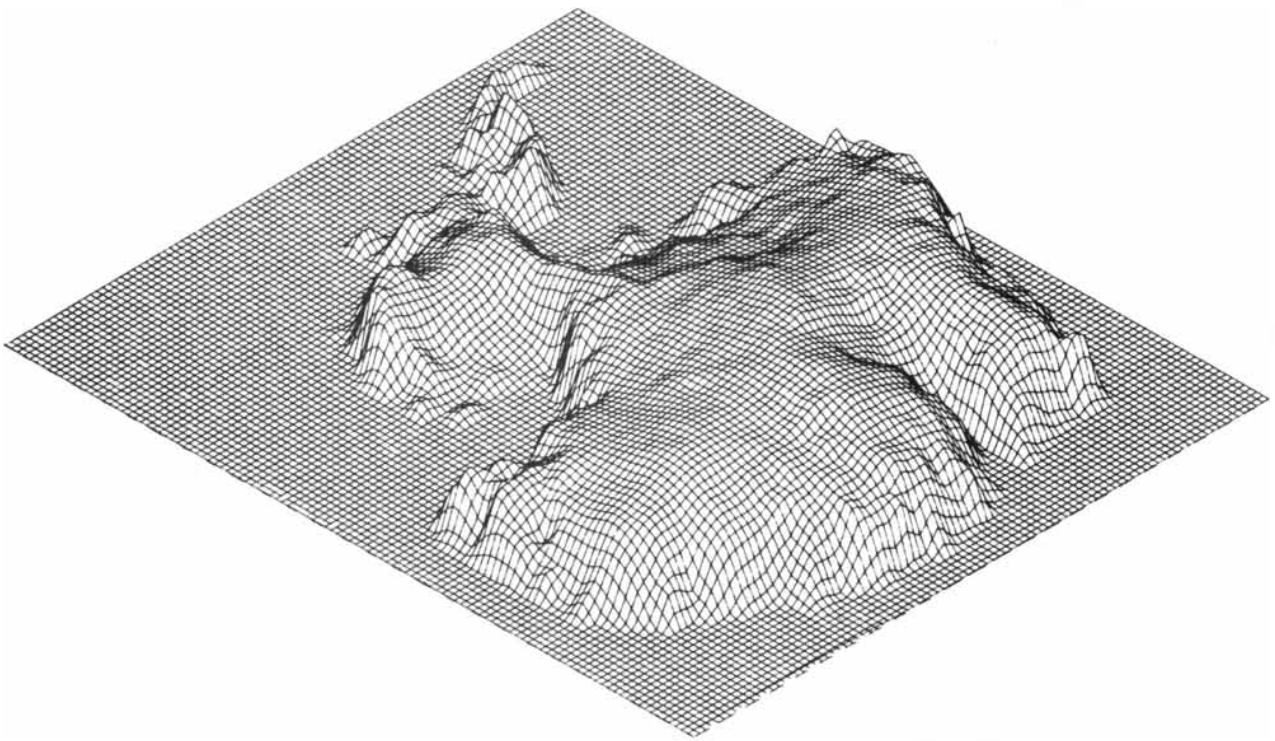
ICE AND LAND of Antarctica are mapped on the basis of data obtained during the International Geophysical Year (IGY, beginning in 1956) and since 1969 by the International Antarctic Glaciological Project (IAGP), which unites major expeditions of Australia, France, Japan, the U.S. and the U.S.S.R. and includes a group from the U.K. The map on the opposite page shows the ice surface. The Pole of Relative Inaccessibility (at 83° S, 53° E) is at the center of the ice sheet; it derives its name from the fact that it is equally remote from all the surrounding coasts. Below the map is a profile of the ice and bedrock made along the irregular line identified on the map by the letters *A* through *M*. The map above shows the bedrock of the Antarctic continent. The data on the bedrock are derived from radar soundings taken through the ice from the surface and from airplanes. The technique has been named radioglaciology.

is that the deeper ice, which was laid down during the most recent ice age, has a higher electrical resistivity than the rest of the ice, for reasons not yet fully understood.

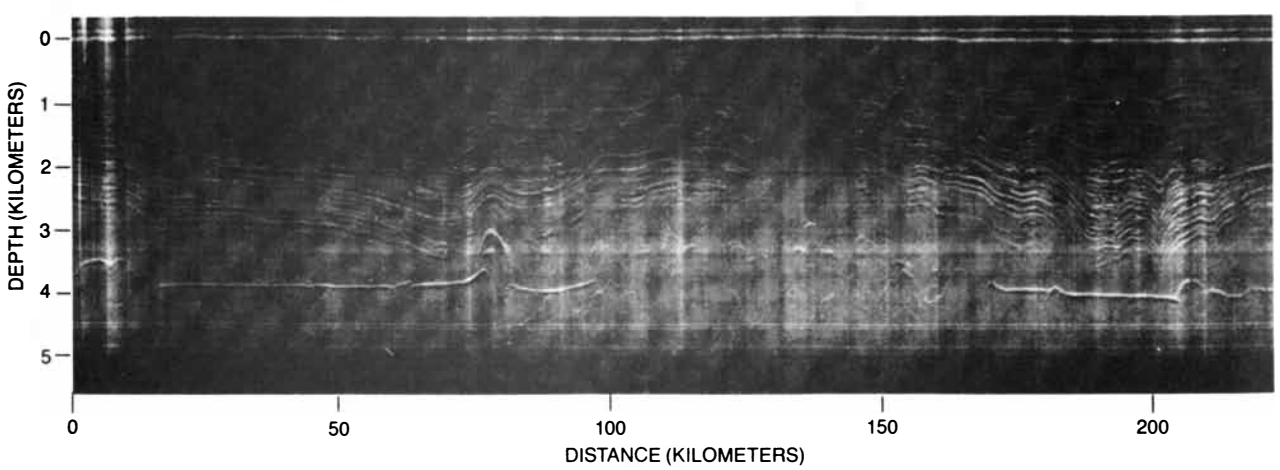
The surface slope and thickness of the ice jointly determine the direction and magnitude of the gravitational force that causes the ice to flow outward from its highest points. Until recently the motion was difficult to measure except in relation to a stable reference point, such as a rock outcrop. Vitali Bogorodskii of the U.S.S.R. and John F. Nye of the University of Bristol independently developed a technique for deducing motion in the absence of a stable reference point. It is based on subtle changes in radar echoes generated by even slight displacement of the ice over the unchanging irregularities of the bedrock.

A more powerful method now available relies on the Doppler effect: the change in frequency of a sound or an electromagnetic signal resulting from movement of the generating or reflecting object. Measurements of this effect on signals from satellites passing over a point on the ice sheet give the position of that point with a precision of a few meters. Repeated measurements of a given point reveal its displacement by the motion of the sheet even when the rate of movement is only a meter or two per year.

The work on the regime of the ice sheet—the temperature of the ice and the balance of mass between gain and loss—has advanced considerably less since the IGY. Indeed, all the rel-



TOPOGRAPHY OF ANTARCTIC ICE SHEET appears in this computer-generated view based on measurements made during the IGY and the IAGP. The orientation puts the Antarctic Peninsula at the upper left. The average thickness of the ice is 2,200 meters.



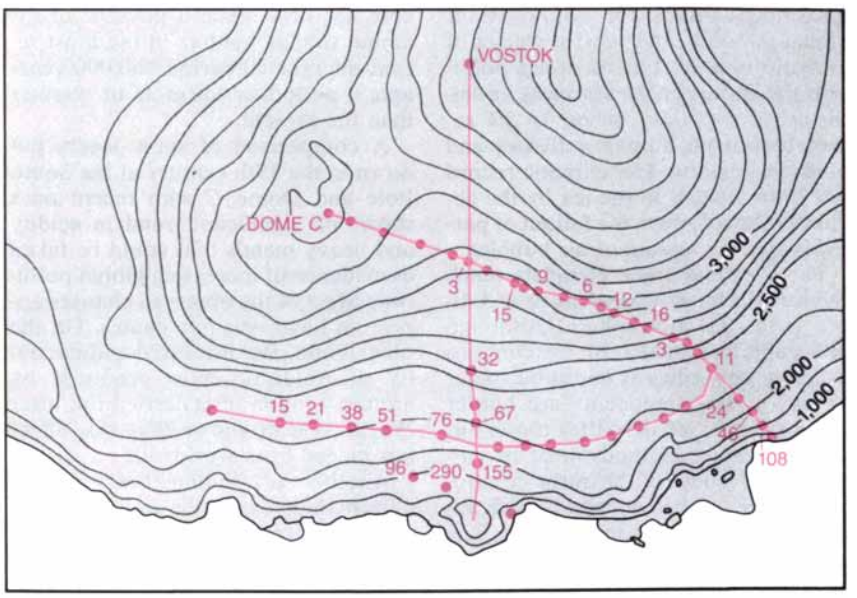
INTERNAL LAYERING of Antarctic ice is evident in this record of radar readings made from an airplane. The image reveals a subglacial lake, probably kept unfrozen by heat from the interior of the earth. The lake yields the flat radar echo that appears between 20 and 70 kilometers on the distance scale and again (because the airplane crossed its own track) between 170 and 200 kilometers. The record shows that radar echoes tend to fade near the base of the ice. This radar track was made in the vicinity of Vostok Station.

evant observations have come from traverses made during or soon after that year. The problem is the lack of new techniques for measuring from airplanes or satellites the temperature a few meters down in the ice and the rate at which the ice sheet builds up or diminishes.

H. Jay Zwally of the National Aeronautics and Space Administration has proposed a method for sensing remotely the rate at which snow accumulates. It is based on the electromagnetic radiation emitted naturally by the ice sheet at centimeter wavelengths. The emission depends on the ice temperature and on how much the radiation is scattered by ice crystals in the top 20 to 30 meters of the ice before reaching the surface. Therefore it should show changes in the snow structure related to the rate at which new snow is accumulating. With this technique, which is still under development, microwave sensors already in place on satellites could add greatly to the existing fund of information about the mass balance. Much of that information is embodied in the monumental Russian *Atlas of Antarctica* of 1966 and is being updated by the Arctic and Antarctic Research Institute in Leningrad under the direction of Evgeny Korotkevich and Grigory Averyanov.

In spite of inadequate data major advances in the understanding of the ice-sheet regime have flowed from IAGP measurements along contours of elevation. For example, Ian Allison and his colleagues of the Australian Antarctic group have measured velocities of ice flow along the 2,000-meter contour around the large indentation filled by the Lambert Glacier and the Amery Ice Shelf (the third-largest in Antarctica). The measured velocities are generally too low for the ice to be carrying off all the snow that falls annually upstream. This finding suggests the surface of the inland ice is rising at a rate of between one centimeter and eight centimeters per year.

Similar IAGP results have been reported by Australian glaciologists as well as by a Japanese team led by Kou Kusunoki (Japan's member of the IAGP's coordinating council) for some of the large outlet glaciers of Enderby Land, the western part of the IAGP study area. The glaciers appear to be roughly in balance near the coast but to be gaining mass 200 meters from the coast and farther inland. The fullest measurements of this kind result from several Australian and Russian traverses crisscrossing the eastern IAGP study area. In a forthcoming paper Trevor Hamley and Neal Young of the Australian Antarctic Division and Ian Smith of the University of Melbourne suggest (within a 10 percent range of uncertainty) that part of the Antarctic ice sheet is close to balance even though individual outlet glaciers may be out of balance. N. F. McIntyre of Cambridge has reached the same conclusion independently from a reanalysis of measurements on the Lambert Glacier and the Amery Ice Shelf.



FLOW OF ICE was measured at a number of points in the area studied by IAGP workers. The findings are represented on the map by colored numbers that give the velocity in meters per year. Measurements were made by the satellite Doppler technique. The technique relies on the fact that the frequency of an electromagnetic signal changes as the transmitter—in this case the satellite—moves past a receiver on the ice. The location of the point is thus established. Readings made some months later show how much the point has moved.

Data on the structural and chemical properties of the ice are derived mainly from cores drilled out of the ice at many locations. Most of the holes run to a depth of a few hundred meters; the 900-meter French hole on Dome *C* and a Russian hole (now more than 2,000 meters deep) aimed at the 3,700-meter bottom at Vostok Station are exceptions. Dramatic results from the Vostok hole, going back 160,000 years, are about to be published in *Nature* by a team that included Claude Lorius (the French member of the IAGP council), Korotkevich and Vladimir Kotlyakov (the Russian members), Nartsis Barkov of the U.S.S.R. and Jean Jouzel, Catherine Ritz and Liane Merlivat of France.

IAGP workers and others have studied the cores with a battery of macroscopic, microscopic and chemical methods. Macroscopic inspection reveals the presence of layers created by melting or by the fallout of dust; changes in crystal size can often be seen clearly too. Subtler changes can be detected only with instruments. An example is the small change in electrical conductivity that indicates an acidic layer resulting from the fallout of volcanic particles.

Another subtle feature is the orientation of the ice crystals. Establishing it entails examining thin sections of ice cores under polarized light. Samples from various locations and depths in the ice have revealed a characteristic sequence in the predominant orientation of crystals as one proceeds from the surface down to the bedrock. In the top and bottom layers the crystals have no systematic arrangement. Between these layers their principal axes tend to be aligned along directions determined by the deformation of the ice as it flows over large structures in the bedrock. Australian and French laboratory measurements of ice from Antarctic cores have shown that at a given temperature and stress the ice in the most anisotropic (nonrandom) layer deforms from five to 10 times more readily than the randomly oriented ice in the top and bottom layers. The finding has been confirmed in the field by measurements of how the continued flow of ice gradually makes a borehole tilt and contract after the core has been extracted.

The study of particles and trace elements incorporated in the ice has greatly expanded since the IGY. Lorius and his colleagues at the University of Grenoble have developed it into a major component of IAGP research. Their work and the coordinated IAGP studies

have implications for appraisals of global pollution and also for studies of climatic history. For assessing pollution the Antarctic ice serves as an undisturbed baseline, owing to its remoteness from human activities and volcanic regions. The climatic record has been created in the ice by the accumulation of snow, the fallout of particles and the sealing of air bubbles.

Particles and trace elements reach the ice within snow crystals or as fallout from the atmosphere. Although one cubic centimeter of ice contains as many impurities as one cubic meter of air, trace components are harder to measure in ice than they are in air, so that precise methods must be employed in studying the rarer components such as heavy metals and microparticles of continental origin. The analysis focuses on recent changes in concentrations (which should reflect volcanic events and possibly the effect of human activities on the global environment) and on changes during and after the most recent glacial period some 10,000 to 20,000 years ago (which elucidate the little-known environmental processes and conditions at that time). Moreover, the deep Vostok core has now made it possible to examine similar matters in the most recent interglacial period 160,000 years ago, a period as warm as or warmer than the present.

A comparison of snow layers put down in the 19th century at the South Pole and Dome *C* with recent ones shows no significant trend in acidity and heavy metals that could be taken as evidence of increasing global pollution. Most of the observed changes appear to have volcanic causes. On the other hand, the increased radioactivity of Antarctic snow produced by atomic-weapon tests carried out after World War II shows that this effect has indeed been worldwide.

Another key finding bears on the current debate over the environmental effect of increasing concentrations of carbon dioxide in the atmosphere as a result of the burning of fossil fuels. Analyses of the ice cores by French and Swiss workers have shown that concentrations of carbon dioxide in the preindustrial era were on the order of 260 parts per million. The finding lowers the hitherto accepted base value, suggesting that recent accumulations of carbon dioxide in the atmosphere may have built up faster than had been thought.

Valuable information on climatic history comes from the stable heavy isotopes of water: oxygen 18 and hydrogen 2 (deuterium). Water made up of these heavy molecules tends to condense and rain out of the atmosphere more readily than water incorporating the lighter isotopes oxygen 16 and hydrogen 1. When the remaining water falls as snow on the ice sheet, it is depleted in the heavy isotopes.

The amount of depletion reaches its final level when the snow crystal is formed and before it lands on the ice sheet. The depletion as read at the site of deposition is found to be closely related to the mean surface temperature there. In this way the isotopic concentrations measured along ice cores have revealed the major temperature changes in the Antarctic region during the interglacial period and at the end of the most recent ice age, together with the lesser fluctuations that have occurred since then. A striking finding is that the present warmer period in this region of East Antarctica appears to have begun 11,000 years ago. That is several thousand years earlier than the date generally accepted for the Northern Hemisphere.

The isotopic concentration also varies with temperature changes during the year. In regions with relatively high accumulations of snow these seasonal variations are preserved to considerable depths in the ice. They therefore offer a method of establishing the age of the ice in particular levels of a core by counting the layers with different concentrations of the isotopes, that is, the seasonal layers. A weakness of the method is that more than one warm and one cold layer can be formed in a single year and that entire annual layers can sometimes be carried off by strong winds. The age of some layers, however, can be established independently because the ice carries evidence of major events of known date. Examples are the radioactivity created in the snow by hydrogen-bomb tests in the atmosphere, which marked the Antarctic snow layers laid down in 1954–55, 1958, 1961–62 and 1964, and trace elements deposited in the snow after large volcanic eruptions such as that of Krakatau in 1883 and Gunung Agung in 1963. For estimating approximate dates going back much further, techniques for measuring minute quantities of the radioactive isotopes of lead, argon, silicon, carbon, chlorine and beryllium (with half-lives ranging from 20 years to 1.5 million years) are coming into use.

Another feature of the past has been deduced from the trapped air in differ-

THIN SECTION from a core of ice removed from the East Antarctic sheet is viewed in polarized light to reveal the orientation of ice crystals. Individual crystals can be differentiated because the color of the light that each crystal transmits depends on its orientation. The section is one millimeter thick and the crystals measure several millimeters across.

ent layers of ice cores. Air in the gaps between snow crystals is at the barometric pressure of the air at the top of the ice sheet. When the gaps between snow crystals are sealed off by the increasing weight of new snow, the trapped air retains its total mass. This phenomenon makes it possible, after the air has been recovered from a core, to calculate the original barometric pressure. By this means Ditmar Jenssen in Australia and Dominic Raynaud in France have reconstructed the surface elevations of the past.

Such measurements show that late in the most recent ice age the ice of the Adélie Land sector of East Antarctica was from 100 to 200 meters thinner in the interior than it is now but that it was from 400 to 500 meters thicker near the coast. A current thickening of ice streams inland and a thinning near the coast, both suggested by IAGP measurements, evidently continue a trend of long duration.

A further meteorological inference can be drawn from counts of dust particles of continental origin in ice cores. Ellen Mosley-Thompson and Lonnie G. Thompson of Ohio State University have found higher readings in ice-age layers of the cores than they have in more recent layers, a finding confirmed by French glaciologists. The finding suggests that wind speeds over regions of the globe that were deserts during the ice age were substantially higher than the wind speeds over the deserts of recent times.

From the start of the IAGP investigators recognized that the results of various studies would need to be fitted together to yield a comprehensive picture of the East Antarctic ice and its environment. Although a complete synthesis remains a distant goal, progress has been made through the construction (by computer) of simplified models of the ice sheet and of how it has changed in the course of time.

The development of models encompassing both the thermodynamics and the dynamics of the ice sheet was initiated by the Russian glaciologists Petr Shumsky and Igor Zotikov. In the IAGP this work has been carried forward above all by William F. Budd of the Australian Antarctic group, who has devised and directed the varied glaciological research of the Australian Antarctic research expeditions since the beginning of the IAGP and represents Australia on the IAGP council. Budd, Jenssen and I constructed the first detailed model of the East Antarctic ice sheet as an aid in planning the IAGP. We began by assuming that the ice flow exactly compensates for the accumulation of snow, so that the sheet is in a perfectly balanced state and does not change its shape with time. Basic principles of physics can then be factored in for the purpose of deriving a host of physical characteristics of the sheet.

In the next step Budd and Ian Smith expanded a model developed by Molly Mahaffy, then of the University of Colorado at Boulder, to simulate the raising or lowering of the ice sheet that must occur wherever and whenever the local gain in mass is in fact not balanced by the ice flow. (Melting can in general be disregarded.) The model then shows the growth of the sheet, beginning with an initially bare bedrock and taking into account the accumulation, flow and loss of ice thereafter.

If neither the accumulation nor the loss is allowed to change, the ice eventually reaches an equilibrium shape. In reality such a state is never reached. The growing ice sheet tends to cool the surrounding atmosphere, thereby reducing the rate of snowfall. Eventually the increasing load of ice depresses the bedrock, lowering the ice-sheet surface also and raising its temperature. Moreover, the almost concurrent development of ice sheets in the Northern Hemisphere would reduce the sea level, exposing submerged land over which the Antarctic ice could spread. These developments are known to have been coupled with changes in the radiation received from the sun resulting from variations in the earth's orbit and attitude. Simplified representations of all these processes have enabled the IAGP modelers to simulate the growth and decay of ice, both in Antarctica and on the continents of the Northern Hemisphere, over the past 500,000 years. The simulated pattern is in broad agreement with the geologic evidence and the findings from cores of ocean sediments.

For the modelers the accumulating data hold out the challenge of producing a history of the ice sheet that fits all the data. The models would then serve for predicting the future of the Antarctic ice. To do so the effects exerted on the climate of Antarctica by the remainder of the global climate system must be taken into account.

The problems are highlighted by the current debate on the growing atmospheric concentration of carbon dioxide, which could warm the polar regions and speed up the motion of the ice in the outer regions of Antarctica, particularly the continent's large floating ice shelves. The IAGP research is taking account of these concerns. It also looks beyond them to assess the wider role of the main Antarctic ice sheet in the history and future of the earth's environment.

Butterfly Engineering

Butterflies drink nectar, fly and bask in the sun to regulate their internal temperature. Underlying such life-sustaining functions are mechanisms that can be studied according to engineering principles

by Joel G. Kingsolver

The relations between a living organism and its environment are generally analyzed according to concepts drawn from biology. Biological organisms and machines are both subject to the same physical laws, however, and so organisms can also be analyzed by means of concepts derived from engineering. Underlying the functions needed to sustain the life of a living thing are a variety of physical mechanisms. The precise character of those mechanisms may determine the range of environments in which the organism can survive and reproduce. By simulating the effect of altering such physical mechanisms it is possible to find out how they affect the distribution and evolutionary success of an organism. I have applied these principles to a most unlikely engineering structure: the butterfly.

There are, to be sure, some fundamental differences between the engineering of a machine and the engineering analysis of a living thing. The designer of a machine begins with a set of functions and then formulates a design that carries out the intended functions with the maximum efficiency or the maximum output. A zoologist who takes an engineering approach reverses the process, inferring the details of function from the anatomical structures. Furthermore, unlike machines organisms are self-constructing and all new structures represent variations on existing ones. Instead of analyzing a structure in isolation, the engineering biologist must be concerned with evolutionary relations among organisms and with the developmental processes that produce the structure. An engineering analysis can elucidate both kinds of relations.

Three of the most significant functions for the survival and reproduction of the butterfly are feeding, regulating the internal temperature and flying. Flight, which has a central role in the biology of butterflies, is an energetically expensive means of locomotion that requires much energy-rich fuel. To acquire the fuel butterflies feed on nectar, a mixture of water and simple sugars made in the flowers of many plants. Most insects have several sets of mouthparts, but in the butterfly the separate parts are fused into a single structure called the proboscis: a long, flexible tube that is tightly coiled when the butterfly rests. During feeding the proboscis is extended and its tip is inserted in the flower.

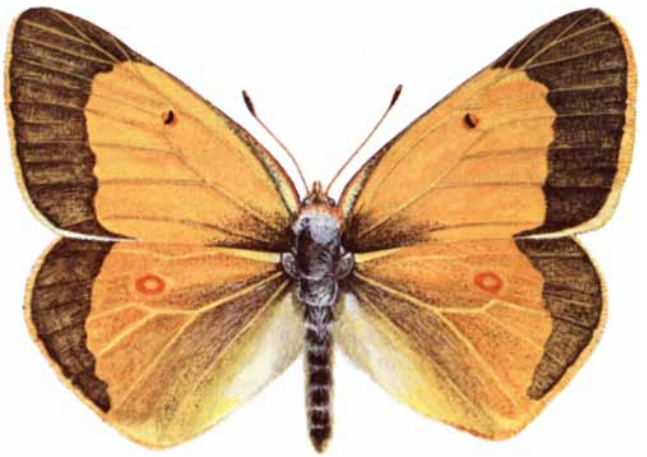
Inside the butterfly's head is a cavity whose size can be changed by muscles called the cibarial dilators. The contraction of the dilators creates a suction pressure that draws nectar up the long, narrow proboscis. From the standpoint of engineering, the butterfly's method of feeding is analogous to sucking sugar water through a long straw. Any fan of milkshakes knows at least two factors influence the efficiency of drinking through a straw. The first is that as the straw becomes longer or thinner the pressure needed to draw the liquid up increases. The second is that the pressure needed to raise the fluid also increases with the viscosity of the solution.

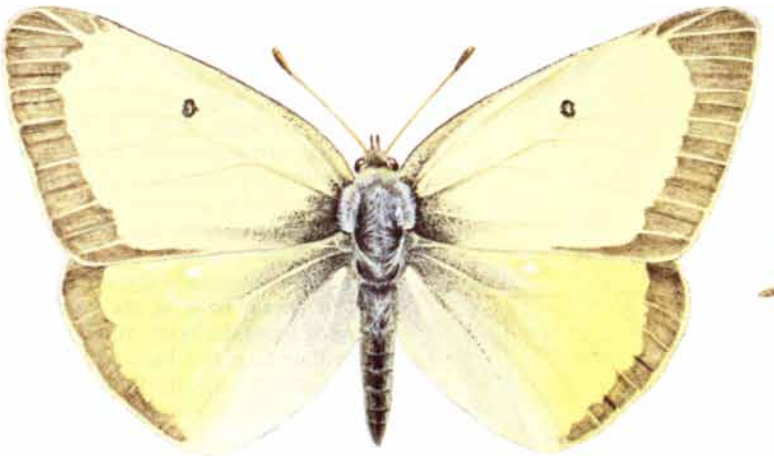
How can the butterfly assure itself of a high rate of energy intake? The sugar content of nectar varies considerably, and one might assume the insect can get energy fastest just by sucking the sweetest available nectar, but that is not so. It is true that sucking a nectar very low in sugar will yield energy slowly. A very sweet nectar, however, also yields energy at less than the maximum rate. The reason is that the viscosity of the solution increases rapidly with the sugar content. If the nectar is very sweet, it cannot be drawn up rapidly enough for the insect to obtain energy at the highest possible rate. Therefore the optimum nectar is neither the sweetest nor the least sweet.

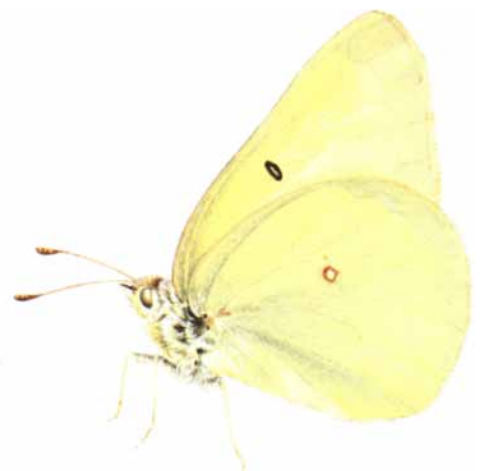
The precise optimum depends on the details of the butterfly's feeding mechanism. Thomas L. Daniel of the University of Washington, Peter G. May of the University of Florida and I analyzed two mechanisms: one in which the suction pressure exerted by the insect remains constant no matter what the sugar content of the nectar is and one in which the pressure varies with the sugar concentration. It turns out that if the suction pressure is constant, the optimum nectar is one containing between 20 and 25 percent sugar. The optimum does not depend on the size and shape of the proboscis. An increase in the diameter of the proboscis or a decrease in its length would increase the rate at which nectar flows up, but the flow rate would increase at all sugar concentrations and the optimum would not change.

Similarly, if the butterfly could suddenly double its suction pressure at all sugar concentrations, the rate of energy intake for a 40 percent solution would rise sharply, but the rate for a 25 solution would also rise by the same

THREE SPECIES of the genus *Colias* include populations that vary in wing color according to the elevation in the Rocky Mountains at which they live. The butterfly at the top is from a population of the species *Colias meadii* that is found at altitudes of from 3,300 to 3,600 meters. The one in the middle is from a population of *Colias alexandra* that is found from 2,700 to 2,900 meters. The one at the bottom is from a population of *Colias philodice eriphyle* that is found from 1,500 to 1,700 meters. *Colias* butterflies achieve the body temperature needed for flight by basking in the sun with the wings held together above the body. Solar radiation absorbed by the underside of the wing is conducted to the body. The region of the wing near the body makes the greatest contribution to warming the insect and, as can be seen in the painting, that region becomes darker with increasing elevation.







proportion. If, on the other hand, the suction pressure is doubled for a solution of 25 percent and tripled for a solution of 40 percent, the sweeter nectar might yield energy faster. Indeed, our work shows that when the suction pressure varies with sugar concentration, the optimum nectar is one containing between 35 and 40 percent sugar. Which mechanism most butterflies actually use is not known. Perhaps the only published data come from May, who studied two species and found they rely on a variable-pressure mechanism. Although the sucking mechanism that is generally used is not known, a high rate of return could be a significant factor in determining which flowers butterflies visit. Butterflies tend to feed on plants that produce a nectar containing from 15 to 30 percent sugar. In contrast bees tend to feed on plants yielding nectars with from 30 to 50 percent sugar.

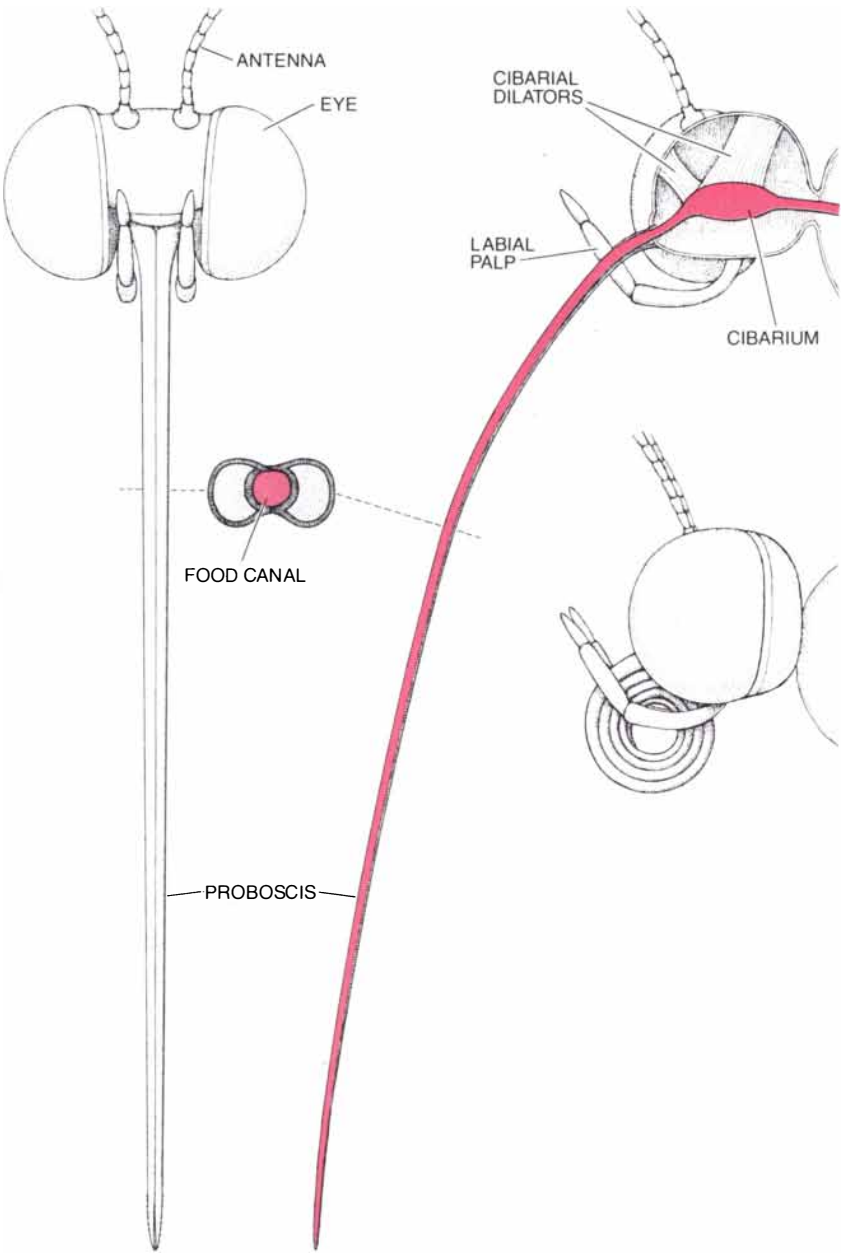
In butterfly feeding a small change in the underlying physical mechanism has significant implications for the relations between the organism and the environment. The same is true in the regulation of temperature. In order to fly many butterflies must raise their internal temperature to between 30 and 40 degrees Celsius, the range in which most mammals maintain their internal temperature. Mammals warm themselves by increasing the rate at which heat is produced metabolically and cool themselves by controlling the circulatory flow to the surface of the body. Butterflies regulate their internal temperature by quite different mechanisms. The insects do have an internal circulating fluid called hemolymph, but the flow of hemolymph is not precisely controlled. Furthermore, instead of raising its metabolic rate to warm up, the butterfly raises its internal temperature by basking in the sun and absorbing solar radiation until it can fly.

Butterflies bask in several different positions. Each species generally adopts one posture. The posture taken up by a particular species depends largely on the specific mechanism by which the body is warmed. When the butterfly basks, the wings convey thermal energy to the body in one of two ways. The wings can absorb the solar radiation and conduct the heat through the wing tissue to the body; the wings can also reflect the solar radiation directly onto the surface of the body, where it is absorbed.

Among the species that rely on absorption two attitudes are common: the dorsal posture and the lateral posture. In dorsal basking the wings are extended from the body much like the wings of an airplane. In lateral basking the wings are held together vertically over the body. No matter which attitude the butterfly takes up, in absorptive basking only the radiation striking the base of the wing helps to raise the body temperature. The light, thin wings are poor conductors of heat, so that radiation striking more than a few millimeters from the body is largely dissipated.

Because the color of a surface determines which wavelengths of radiation it absorbs, color has a significant role in butterfly thermoregulation. The color of the wing is due not to pigmentation of the wing itself but to pigmentation of the thousands of tiny, shingle-like scales that cover it [see "The Color Patterns of Butterflies and Moths," by H. Frederik Nijhout; SCIENTIFIC AMERICAN, November, 1981]. Each scale is colored by just one of the several pigments on the wing; the overall color is produced by the biological equivalent of the pointillism seen in the paintings of Georges Seurat.

Since the part of the wing near the body has the greatest effect on body temperature, the color there is more



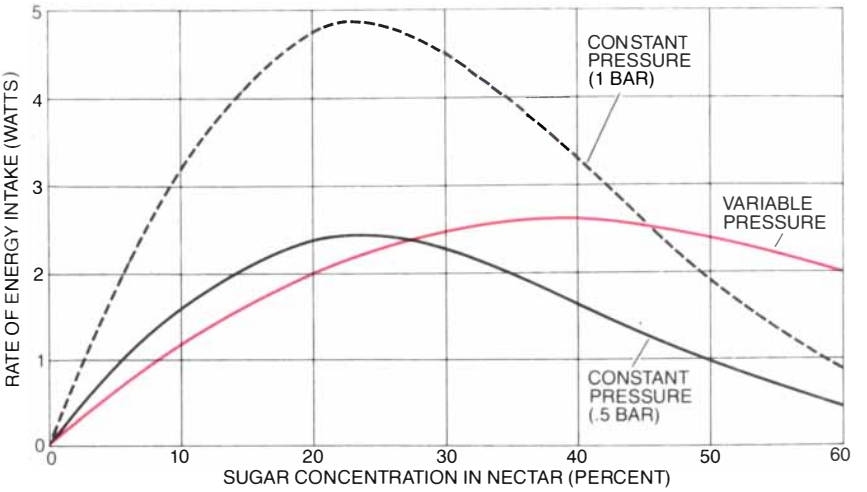
PROBOSCIS is a long, thin, flexible tube through which the butterfly feeds. In flight and at rest the proboscis is curled (*lower right*). On finding a flower containing nectar, the butterfly uncurls the proboscis and inserts it in the nectar (*upper left*). Inside the insect's head is a cavity that can be expanded by muscles called cibarial dilators (*upper right*). The expansion of the cavity creates a suction pressure that draws the sugary nectar up the food canal.

significant for temperature regulation than the color elsewhere on the wing. In most lateral or dorsal baskers the wing base is darkened by a black pigment called melanin; the amount of melanin determines how dark the wing base is. Ward B. Watt of Stanford University and I observed the wing-color pattern in several species of the genus *Colias,* which are often called sulfur butterflies. We found that sulfurs living at high elevations have darker wing bases than species living at low elevations. Darker wings are more efficient absorbers that enable the high-altitude species to make the best use of solar radiation in their cooler habitats.

Reflectance basking entails a completely different set of engineering problems. Absorption and reflection from a particular surface are inversely related. Dark surfaces, which are good absorbers, are poor reflectors. To effectively reflect sunlight onto the body the wings must be white or silvery. Indeed, the main group of reflectance baskers are called whites because of the color of the upper wing surface; among the whites are the genus *Pieris.* Reflectance baskers generally hold the wings at an angle above the body in a configuration resembling a V. The angle of the V has a considerable influence on how much reflected light reaches the body.

When a beam of radiation enters the V, it is reflected back and forth between the raised wings before striking the body or being reflected out of the V without reaching the body [*see lower illustration on next page*]. The beams that initially strike the inner region of the wing reach the body; those that strike the outer region do not. The angle at which the wings are held determines the proportions of the two regions. If the V is wide, only radiation reflected from a small region at the base of the wing is conveyed to the body. If the V is narrow, radiation is conveyed to the body from a much larger region, perhaps including even the tip of the wing. Each species of *Pieris* butterfly has a characteristic basking configuration and a corresponding pattern of wing color, in which the shade of the wing tip varies in part according to whether the tip has any influence on body temperature.

Although wing color is quite significant, there are at least two other anatomical factors that influence body temperature. The butterfly's body has three parts: the head, the thorax and the abdomen. The legs and wings are attached to the thorax; the thoracic temperature is more closely regulated than the temperature of the other body parts. Because there is a flow of hemolymph between the abdomen and the thorax, however, the temperature of the abdomen can affect that of the thorax. In dorsal baskers the abdomen is often placed in the shade of the wings, thereby reducing the thoracic temperature. Reflectance baskers can raise the tip of the abdomen above the wings to a position where it intercepts much reflected light and raises the thoracic temperature. In addition the butterfly's thorax has a hairy covering consisting of greatly modified scales. The covering is an insulator and can be quite thick on butterflies that live at high elevations or in northern latitudes.



OPTIMUM SUGAR CONCENTRATION in the nectar that nourishes a butterfly depends on the mechanism for drawing the liquid in. The optimum concentration is the one that yields the highest rate of energy intake. The graph shows how the rate of intake varies with sugar concentration for three slightly different methods of sucking. If the butterfly exerts a low constant pressure, the optimum concentration is between 20 and 25 percent (*solid black curve*). If the butterfly exerts a higher constant pressure, the rate of energy intake at all concentrations increases, but the optimum does not change (*broken black curve*). When the suction pressure increases with the sugar concentration, however, the optimum nectar contains about 35 percent sugar (*colored curve*). The constant-pressure curves are hypothetical. The variable-pressure curve is for *Agraulis vanillae,* the gulf fritillary, as measured by Peter G. May of the University of Florida. (The sweetest nectar is not optimum because viscosity increases with sugar concentration and viscous nectars are difficult to suck.)

The capacity to achieve the body temperature needed for flight is central to the biology of many butterflies. All the major activities in the insect's life cycle require flight, including feeding, finding a mate and laying eggs. Several recent studies in the U.S. and Europe have shown that reproductive success is limited partly by the time available for the female to fly in search of plants suitable for egg laying. One reason flying time is so crucial is that many butterflies have an adult life span of less than a week. Whether a butterfly can warm up enough to take off is determined largely by weather conditions and such conditions strongly affect the insect's success in reproducing.

Taking advantage of an engineering approach, Robert J. Moffat of Stanford University and I developed mathematical models relating the physical characteristics that affect temperature regulation, such as wing color, to the weather conditions in which a particular insect can fly. The relations can be shown graphically by means of the "flight space," a concept originated by Warren P. Porter of the University of Wisconsin at Madison and David M. Gates of the University of Michigan. On a plane with coordinates representing air temperature and wind speed, the flight space for a species is the area including all conditions under which that species can warm up enough to take off [*see illustration on page 112*].

By comparing the flight spaces of different species, one can observe how wing color affects the range of weather conditions in which the butterfly can take to the air. For example, the flight spaces of two of the *Colias* species Watt and I studied could be compared. Certain populations of the species *Colias philodice eriphyle* are found at altitudes of between 1,500 and 1,800 meters. Populations of the species *Colias meadii,* which has a considerably darker wing base, are found between 3,300 and 3,600 meters. Plotting the flight spaces of the two species shows that differences in wing color are among the means by which closely related butterflies adapt to local environmental conditions. Moreover, the mathematical model predicts that

the low-elevation species would not be able to fly at the higher altitude because it could not achieve the body temperature needed for flight. Experiments in which I transplanted *C. p. eriphyle* specimens to the high, cool mountain slopes proved the prediction correct.

The physical mechanisms underlying butterfly feeding and thermoregulation depend on physical principles that can be observed in the operation of many machines. Some of the mechanisms exploited by large butterflies in flying, on the other hand, appear to be quite different from those of flying machines. Only recently have laboratory investigations begun to shed light on the novel principles of butterfly flight. Many large butterflies fly using a "flap and glide" technique, in which periods of flapping alternate with periods of gliding. The flap-and-glide technique is not common among insects, although it is used to some extent by certain grasshoppers and dragonflies.

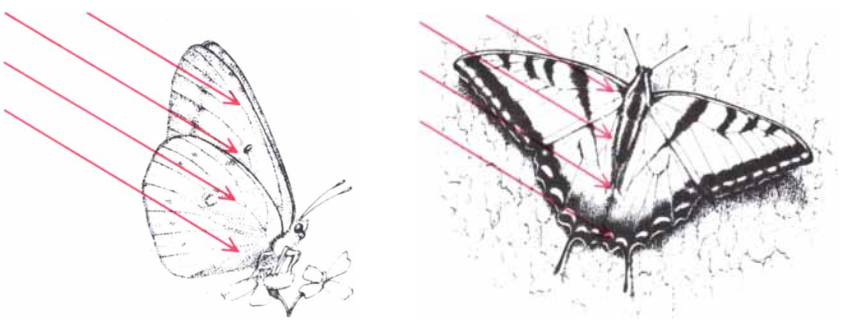
The mechanics of the gliding phase of butterfly flight are better understood than those of the flapping phase. Gliding is better understood in part because it bears a closer resemblance to the operation of a classical airfoil, such as an airplane wing. When a manmade glider descends, it moves along a flight path with the wings in a fixed position. The shape of the wing and the angle of the wing in relation to the air flow cause the air to move faster over the upper surface of the wing than it does under the lower surface. The difference in speed results in a pressure difference between the upper and lower surfaces. The pressure difference in turn yields a net force directed upward and perpendicular to the direction of motion of the craft. That force is called the lift.

The flow of air around the wing during gliding is quite complex. In order to analyze the flow it is sometimes helpful to describe it as a combination of two much simpler patterns: translation and circulation. The concept of translation describes a flow pattern in which the air moves in a direction opposite to that of the movement of the wing. Circulation is equivalent to a net spinning movement of the air around the wing. Translation and circulation are abstractions and the description of the airflow in terms of translation and circulation is a matter of analytical convenience. Actually there is no movement of air from the lower surface of the wing to the upper surface as there would be if there were in fact a spinning movement. Nevertheless it can be shown mathematically that the actual airflow pattern resembles what would be seen if the two simpler patterns were combined. As we shall see, translation and circulation are of great help in understanding hovering flight in butterflies.

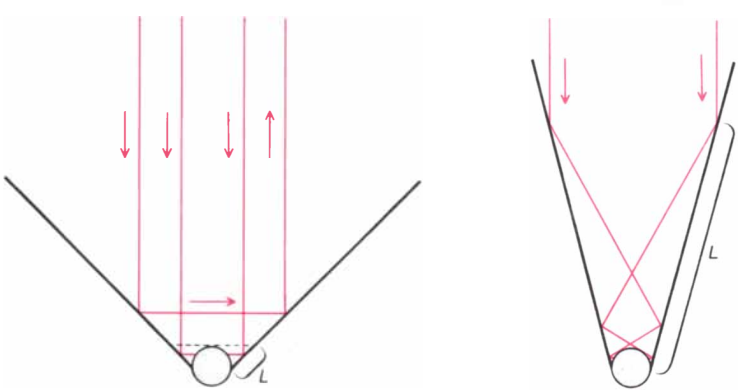
When the butterfly glides, its wings generate lift in much the same way as the wings of an aircraft do. There are, however, substantial differences between the aircraft and the organism. In the operation of a classical airfoil the main force opposing the forward motion of the wing is the drag. Drag results from a shearing between the wing surface and the air and from a pressure difference between the front and the back of the wing. The efficiency of an airfoil is measured by the ratio of lift to drag; manmade wings are carefully designed to achieve a high ratio.

The wing of a butterfly has a low ratio of lift to drag and in this sense is a rather poor airfoil. There is, however, another criterion for assessing the performance of a glider: the capacity to remain aloft for extended periods. By that measure the butterfly is a good glider. When butterflies glide, they descend slowly. The low rate of descent is due to the form of the insect's wing, which could be compared to that of a parachute. The concave shape of a parachute and the large surface area perpendicular to the direction of descent yield a substantial drag force opposing the force of gravity. Just as the drag force slows the descent of the person hanging from the parachute, so the drag on the butterfly's broad, flat wings serves to slow the insect's descent. Indeed, much of what appears to be gliding in butterflies could more accurately be described as parachuting. When the butterfly floats through the air, the large drag force produces a low ratio of lift to drag, but it also yields a low rate of descent.

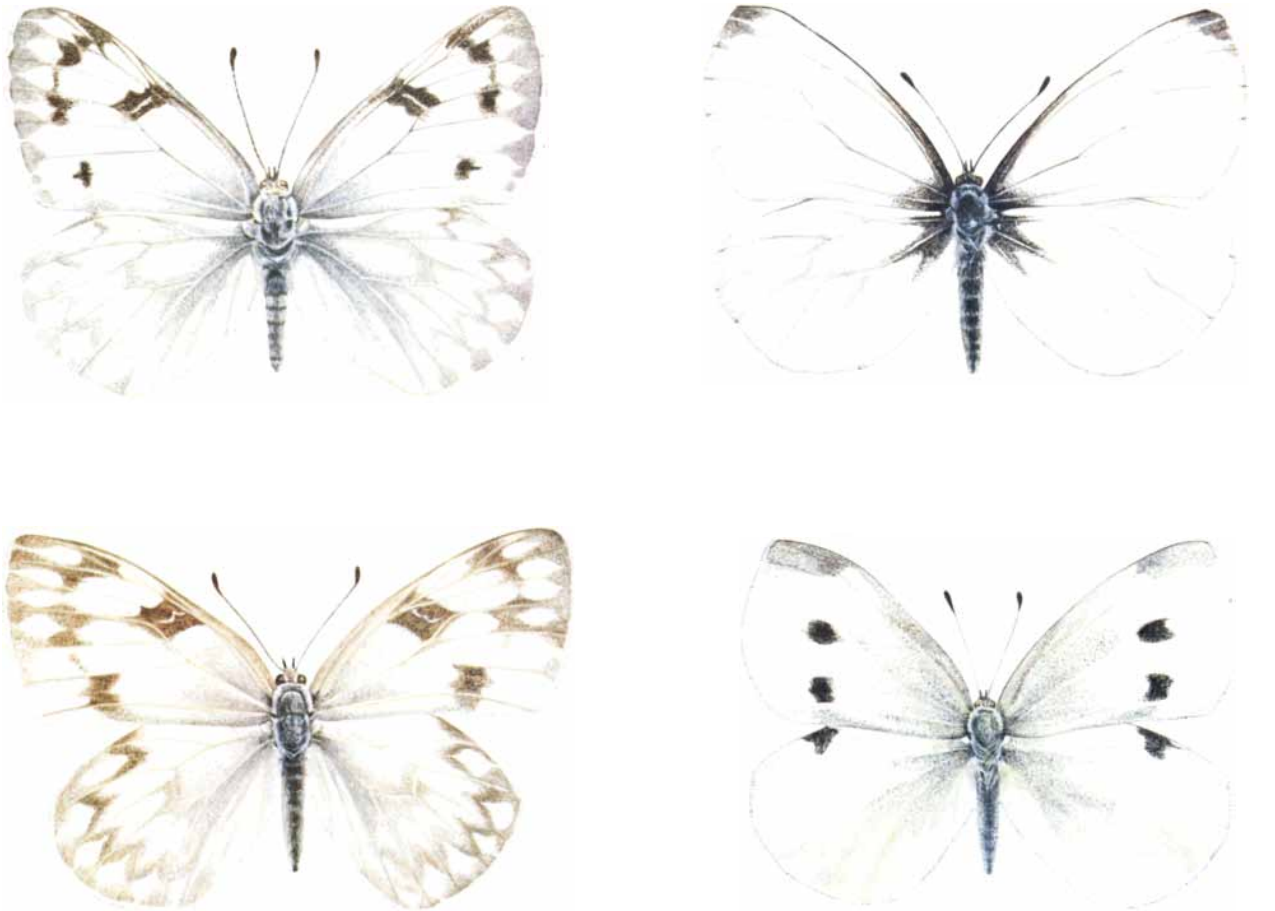
The flapping phase of butterfly flight is more complex than the gliding phase. The complexities are reduced



ABSORPTIVE BASKING, in which the wings absorb solar radiation and conduct thermal energy to the body, is done in two main postures. Lateral baskers hold the wings together over the back (*left*). Dorsal baskers extend the wings to the side (*right*). Each species generally adopts a single attitude for basking. Whether the color of the upper wing surface or the color of the lower one affects body temperature depends on which posture is adopted. The three species of *Colias* butterfly shown in the illustration on page 107 are lateral baskers.



REFLECTANCE BASKING depends on the reflection of solar radiation from the upper wing surfaces onto the body. In reflectance basking the wings are held above the body in the form of a V whose angle can vary considerably. Some species of reflectance baskers hold the wings in a broad V (*left*). If the V is broad, only the beams that initially strike a small region of the wing (*L*) are reflected onto the body. Other species hold the wings in a narrower V (*right*). As the V narrows, the region from which beams are reflected onto the body increases. If the V is very narrow, even the wing tips can reflect radiation onto the body.



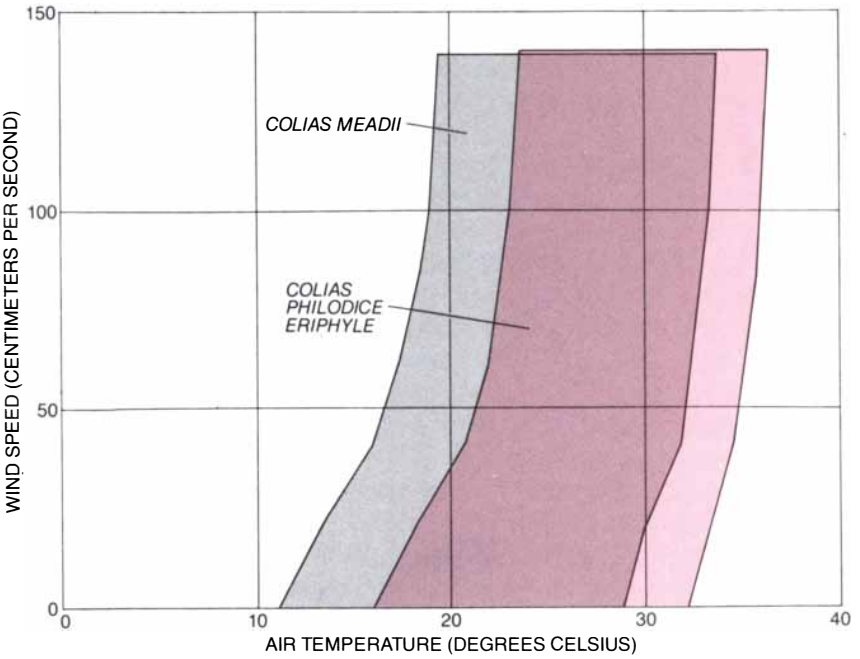
***PIERIS* BUTTERFLIES are reflectance baskers whose wing-color patterns are related to the climatic conditions and the wing angle assumed in basking. In reflectance basking light wing areas clearly provide more heat to the body than dark areas. The butterfly at the upper left is *Pieris occidentalis;* the one at the lower left is *Pieris protodice.* They are from the subgenus *Pontia,* whose members hold their wings in a broad V when basking. As a result of the basking posture the color of the outer margin of the wing does not affect body temperature and both *Pontia* species have extensive dark areas there. The butterfly at the upper right is *Pieris napi macdunnoughii;* the one at the lower right is *Pieris rapae.* They are from the subgenus *Artogeia,* whose members adopt a narrow V in basking. The color of the outer margin does affect body temperature and both *Artogeia* species have very light coloring there. The two upper butterflies come from Gothic, Colo., at an altitude of 2,900 meters; the two lower ones come from Delta, Colo., at 1,500 meters. The butterflies shown exploit absorptive as well as reflectance basking; hence the high-altitude species tend to have darker wing bases.**

somewhat by analyzing hovering. In hovering there is no net horizontal or vertical movement, although both the wings and the body may change position. Butterflies and moths are among a large group of organisms, including hummingbirds and many other insects, that are capable of hovering flight. When the flier hovers, the wings move through a complex cycle and the repetition of the cyclical movement keeps the bird or insect in the air.

If the rapid cyclical movement could be slowed down and viewed from the side, it would be seen that in most hoverers, including hummingbirds and many insects, the long axis of the cycle is nearly horizontal. The wing moves mostly back and forth and the amplitude of the vertical movement is small. When that is the case, the cycle is said to have a horizontal stroke plane. A cycle with a horizontal stroke can be analyzed as if the wing were an airfoil. The wing moves in a figure-eight pattern whose long axis is in the horizontal plane. While moving along that path the wing generates a lift force that can be computed at every instant of the cycle. By summing the lift force over the complete cycle it is possible to find the average lift that keeps the hoverer aloft.

If the cycle through which the butterfly's wing moves could be slowed and viewed from the side, a very different pattern would be seen. Recent work by Charles P. Ellington of the University of Cambridge shows that the butterfly's cycle entails vertical movements of large amplitude. Instead of a horizontal stroke plane the butterfly's motion has a vertical stroke plane. Although the cycle is largely vertical, it bears a resemblance to an oarsman's stroke. On the downstroke the wing is held perpendicular to the direction of its movement, whereas on the upstroke it is "feathered" so that it is parallel to the direction of movement. The butterflies are the only organisms so far known to employ a vertical stroke plane in the hovering cycle.

A cycle that has a vertical stroke plane, like its horizontally oriented counterpart, includes both translation and circulation. Yet a vertically oriented cycle cannot be analyzed as if the wing were an airfoil. In order to sum the lift generated by an airfoil over the entire cycle (as can be done when the stroke plane is horizontal) it is necessary to assume that the forces operating on the wing at each instant are independent of the forces generated in the preceding instant. The assumption works reasonably well for cycles where the amplitude of the vertical movement is small. In a cycle with a large vertical stroke, however, complex patterns of airflow are set up



FLIGHT SPACE represents in graphic form the climatic conditions under which a particular species of butterfly can achieve the body temperature required for flight. That a butterfly species derives an advantage from its wing-color scheme can be shown by plotting the flight spaces of two species. The gray area is the flight space of *Colias meadii*, the high-altitude species shown at the top of the illustration on page 107. The dark undersurface of the wings of *Colias meadii* enables the insect to fly in cool conditions. The colored area is the flight space of *Colias philodice eriphyle*, the low-altitude species shown at the bottom of the illustration on page 107. The lighter-colored wing undersurface of *C. p. eriphyle* yields a flight space well suited to the warmer climatic conditions that prevail at low altitudes.

around the wing. One result of the complex pattern is that the flow at a particular instant has a considerable effect on the flow at the next instant. The independence assumption is violated and the cycle is said to display "unsteady flow."

One mechanism for producing an upward thrust under conditions of unsteady flow has been described admirably by Torkel Weis-Fogh of Cambridge. That mechanism, which is called the fling, puts considerable emphasis on the rapid development of circulation [see "Unusual Mechanisms for the Generation of Lift in Flying Animals," by Torkel Weis-Fogh; SCIENTIFIC AMERICAN, November, 1975]. At the beginning of the fling the wings are held closed vertically over the back. The wings separate first at their leading edges, which are close to the head [*see illustration on page 113*]. The gap between the leading edges widens as the wings are flung open around the line defined by the trailing edges. The flinging motion generates a circulation from the lower wing surfaces around the leading edges to the widening gap between the upper surfaces. The circulation is maintained on the wing during the downstroke, generating an upward force.

Many small insects are known to use the fling, but it is rarely seen in larger species. Ellington showed that if the wings are large, the circulation tends to separate from the upper surface of the wing, which reduces its capacity to generate an upward force. There is, however, a way around the problem. If the wings are flexible enough to peel apart gradually from the leading edge, the circulation remains bound to the wing on the downstroke and the upward force results. Butterflies are the largest fliers that use the fling, and Ellington has demonstrated they employ the gradual peeling motion.

My own preliminary studies suggest the flexibility needed for the peeling motion is closely related to the pattern of veins in the butterfly's wing. The wing veins form a set of rigid struts over which a flexible membrane is draped. Butterflies have far fewer wing veins than most insects, particularly at the wing's outer edge in the region farthest from the body. The major veins radiate outward from the base of the wing. Of the main veins, the one at the leading edge is the stiffest. Except in the central region known as the discal cell there are few veins perpendicular to the ones radiating from the wing base. As a result of the distribution of veins the outer edge of the wing is quite flexible. Moreover, there is considerable flexibility from the leading edge to the trailing edge, which is called chordwise flexibility.

It appears that the stiff leading edge and the chordwise flexibility are essential in the peeling motion. There are no muscles in the wing itself and the peel must be accomplished by the contraction of the flight muscles at the base of the wing. The contraction of those muscles sends the stiff leading edge on its downward stroke. The chordwise flexibility of the remainder of the wing causes the movement of the leading edge to be translated into a gradual separation that proceeds from the leading edge backward to the trailing edge. Because of the great flexibility of the outer margin of the wing, there is also a peel that proceeds downward from the tip of the leading edge toward the body.

The reduced venation of the butterfly's wing makes it possible for the insect to perform the peel. Yet the same pattern of venation keeps the wing veins from being used as fluid conduits for regulating body temperature. The principle of combined benefit and detriment seen in the wing veins has a more general application. Each of the butterfly's major anatomical structures has a role in several functions, and a quality that is effective for one purpose may be quite limiting for another. In designing a machine with several functions, the designer sets out to achieve the most efficient combination of functions. By analogy with the machine, the butterfly wing could be considered the best compromise among several crucial functions.

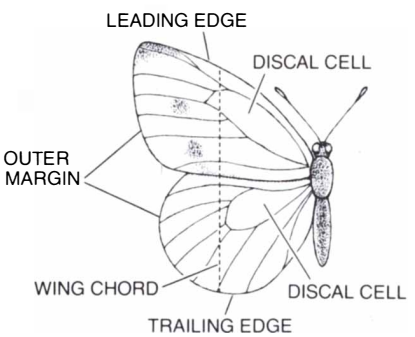
Since biological organisms are self-constructed and evolved, however, there is no reason to think they represent optimum designs. An alternative view is that the various anatomical structures establish constraints within which needed functions are fulfilled and evolution takes place. The effect of such constraints can be demonstrated by comparing two *Colias* species. Unlike the species described above, these are hypothetical species corresponding to adaptive extremes. Wing color in *Colias* results from a combination of two pigments: the black pigment, melanin, and a yellow-orange pigment, pteridine. The relative proportions of the two pigments determine the wing color of a particular species. Hence, along with body fur, the pigmentation sets the limits of the range of environments in which sulfur butterflies can survive.

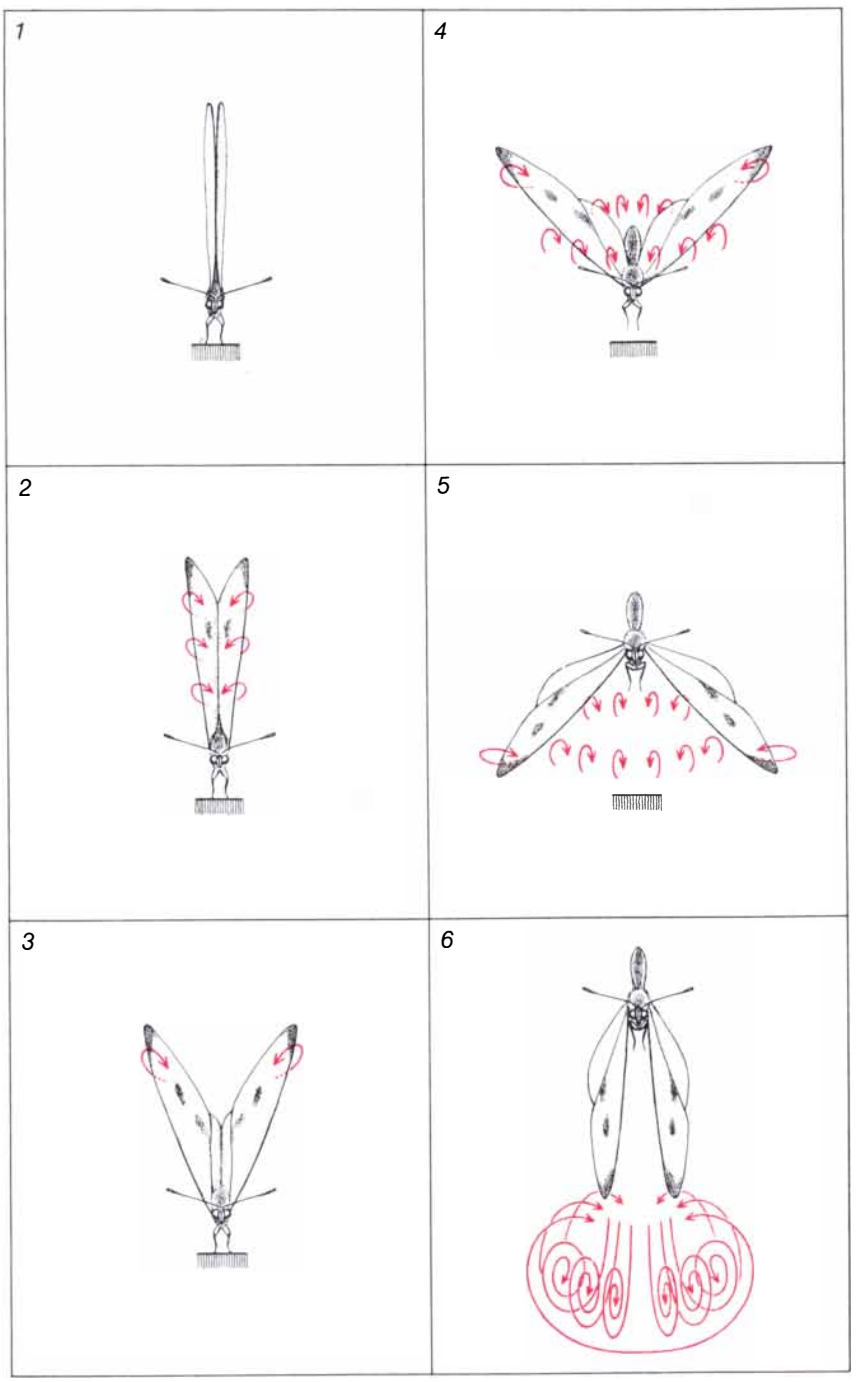
As I have explained, *Colias* butterflies are lateral baskers in which the degree of wing darkening increases with altitude. Therefore the hypotheti-

cal low-altitude species would have entirely yellow wings and no body fur; the hypothetical high-altitude species would have entirely black wings and a thick layer of fur. When the flight spaces of these two hypothetical species are plotted, they are found to differ far more than the flight spaces of the real species I compared above. Indeed, they are as different as two flight spaces can be within the constraints imposed by the two-pigment system.

Of course, some environments fall outside the limits dictated by the simple system of two pigments. To describe species that could survive there it is necessary to relax the constraints. Suppose a white pigment were added to the other two pigments as well as an additional form of thermoregulation, reflectance basking. Those additions extend the limits of the scheme to include the *Pieris* butterflies. With three pigments and three basking postures one can describe the flight spaces of all the butterflies in the family Pieridae, which includes the whites and the sulfurs. As the process of relaxing the constraints continues, broader taxonomic categories fall within the limits. By including some features not found in the Pieridae, such as metabolic heat generation before flight and regulation of hemolymph circulation, all the insects in the order Lepidoptera, which includes the butterflies and the moths, can be encompassed.

In this way the functional meaning of taxonomic categories can be uncovered. The taxonomic categories can in some instances be shown to differ in specific functional characteristics, as the genera *Pieris* and *Colias* were shown to differ in basking posture and wing pigmentation. Because characteristics such as wing pigmentation are genetically determined, the system of constraints also gives insight into the evolutionary relations among organisms. For example, the engineering analysis makes it possible to tell how much a mutation adding a third pigment to the wing of a *Colias* butterfly would affect the range of habitats in which the butterfly could live. By combining engineering analyses of organisms with studies of evolutionary relations, one can begin to understand how physical mechanisms constrain evolutionary change and thus shape the frail structure that enables the butterfly to flutter through the air.





PEELING FLING enables a *Pieris* butterfly to take off, as is depicted in a sequence of drawings based in part on the work of Charles P. Ellington of the University of Cambridge. The wings are initially held together over the body (*1*). As the fling begins the leading edges (near the head) peel apart (*2*). The peeling motion proceeds toward the trailing edges, generating a circulation around the wing as it does (*3*). When the peel has reached the trailing edges and the wings have separated, the downstroke begins (*4*). Throughout the downstroke the circulation is maintained on the wing, thereby yielding an upward force (*5*). At the bottom of the downstroke the circulation, which has been bound to the wing, is shed in the form of a vortex ring (*6*). The wings are then raised for another downstroke. The anatomy of the butterfly wing may make a significant contribution to the butterfly's capacity to perform the peeling fling (*above*). The pattern of wing veins results in a stiff leading edge and much flexibility along the wing chords.

THE AMATEUR SCIENTIST

Cooking outdoors with simple equipment demonstrates aspects of thermal physics

by Jearl Walker

Outdoor cooking can be a pleasant part of camping or a key to survival in an emergency. It can also provide a study in thermal physics: how heat can be transferred from a heat source to food. This month I analyze several ways of cooking food with flames, coals or charcoal briquettes. The techniques require little or no equipment.

A few fundamental concepts of thermal physics underlie all cooking procedures. One concept involves what is meant by heat and temperature. The atoms and molecules of a substance move randomly at any temperature above absolute zero. In a solid the motion consists in rotation and vibration. In a gas or a liquid the phenomenon also includes the random motion of atoms and molecules that are traveling in straight lines, colliding and then again traveling in straight lines.

When a substance is heated, the heat represents the additional energy imparted to the random motion. Temperature is a measure of the amount of energy in the random motion. Thus when the substance is heated, its temperature increases and the substance is said to be hotter. The heat of cooking increases the energy of the random motion of the atoms and molecules in the food, and the food thereupon cooks by undergoing certain chemical and physical changes.

Conduction, convection and radiation are the three primary ways of transferring heat energy. In conduction the heat is conveyed through some intermediate material such as a metal pan or foil by means of atomic collisions. As the outside surface of the metal warms, the energy in the random motion of the atoms there increases. They collide with atoms somewhat deeper in the metal, giving those atoms some of the kinetic energy derived from the heat source. Eventually atoms on the inside surface receive the energy and collide with atoms on the surface of the food, heating the food. Conduction continues for as long as the temperature of the heat source is above the temperature of the food.

Convection involves the ascent of a heated fluid, either air or a liquid. Heat increases the energy in the random motion of the fluid and decreases the density of the fluid. The surrounding cooler and denser fluid then pushes the heated fluid upward. As the hot fluid passes the food, the atoms and molecules in the fluid collide with those on the surface of the food and transfer energy to them.

Radiation involves the emission and absorption of electromagnetic waves. In cooking the source is light. The surface of a heat source such as burning coal emits light in the infrared and visible parts of the electromagnetic spectrum. Since light has energy, this emission is a radiation of energy. When the light is absorbed by atoms and molecules on the surface of the food, the energy of their random motion increases, as does the temperature of the surface. Heating by radiation therefore requires that the food absorb some of the light (primarily the infrared) emitted by the heat source.

Many campfire-cooking techniques draw on more than one of these primary means of transferring heat. For example, a fire might heat a metal pan by both convection of hot air and radiation of light. As the metal warms, energy is conducted through it to the food. As the surface of the food then heats up, conduction brings the heat into the food.

One of the easiest ways to cook food such as meat is to spear it with a stick or wrap it around the stick and then suspend it over the fire or coals. The food is heated by the convection of rising hot air and by the radiation from the heated surfaces of the wood and from the hot regions in the flame. You can save work by propping the stick over the fire or suspending it across the fire by means of two forked sticks driven into the ground on opposite sides of the campfire.

A large piece of meat suspended over the fire must be turned frequently, because only the side toward the fire gets the effect of the rising hot air and the radiation. The rig known as a dingle fan, probably from the logging-camp shed called a dingle, is helpful in this task. To make the apparatus attach a short chain to the upper end of a stick that is angled upward over the perimeter of the campfire. Suspend the meat from the chain by a string attached to a hook in one end of the meat. Tie a short stick to the string. One end of the stick holds a fan made of wire or branches wrapped in aluminum foil or leaves. To the other end attach a small rock to serve as a counterweight to the fan. Orient the plane of the fan somewhat off the vertical and arrange the entire assembly so that the fan is in the hot air rising from the campfire. The meat is not in that convection current but is exposed to the radiation from the fire.

The rising hot air pushes against the underside of the fan. The force moves the fan to one side, twisting the chain and rotating the meat. Once the fan is out of the convection current the chain untwists, rotating the meat in the opposite direction. It overshoots the original position, again twisting the chain. The cycle continues indefinitely, exposing about half of the meat to the radiation. After a while invert the roast and hang it from a hook on the other end to expose the other half of the meat to the radiation.

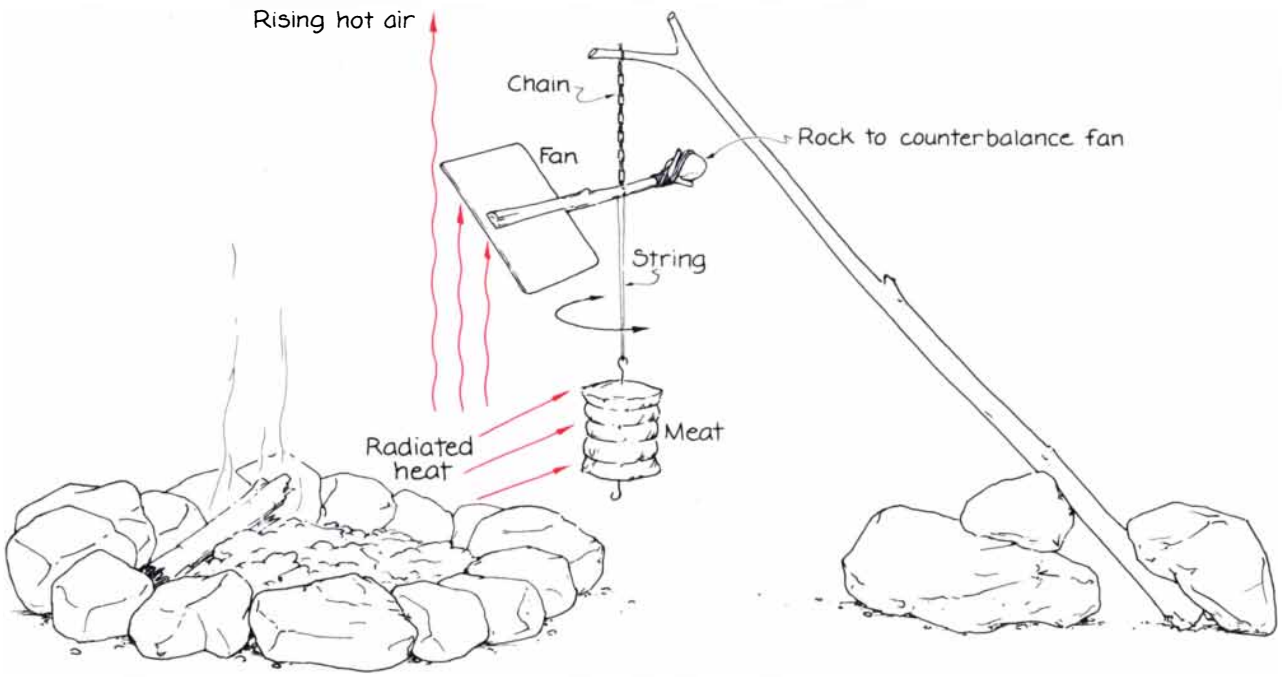
To fry food you can make a stove from an empty No. 10 can. Remove one end plate of the can and cut a flap at the open end. Bend the flap outward. Push the loose end plate into the can and against the other end plate. With a can opener (the kind that punches triangular holes) or a knife, cut flaps in the can in several places near the closed end. Push the flaps into the can and against the loose end plate to hold it near the fixed end one. Place the open end of the can over a heat source. The upper end plate serves as a surface on which eggs, bacon and other items can be fried.

The can functions as a chimney because cool air is sucked in through the open flap at the bottom to replace the hot air rising to the top and out through the holes there. The strong flow of air through the can fans the fire and keeps it burning briskly.

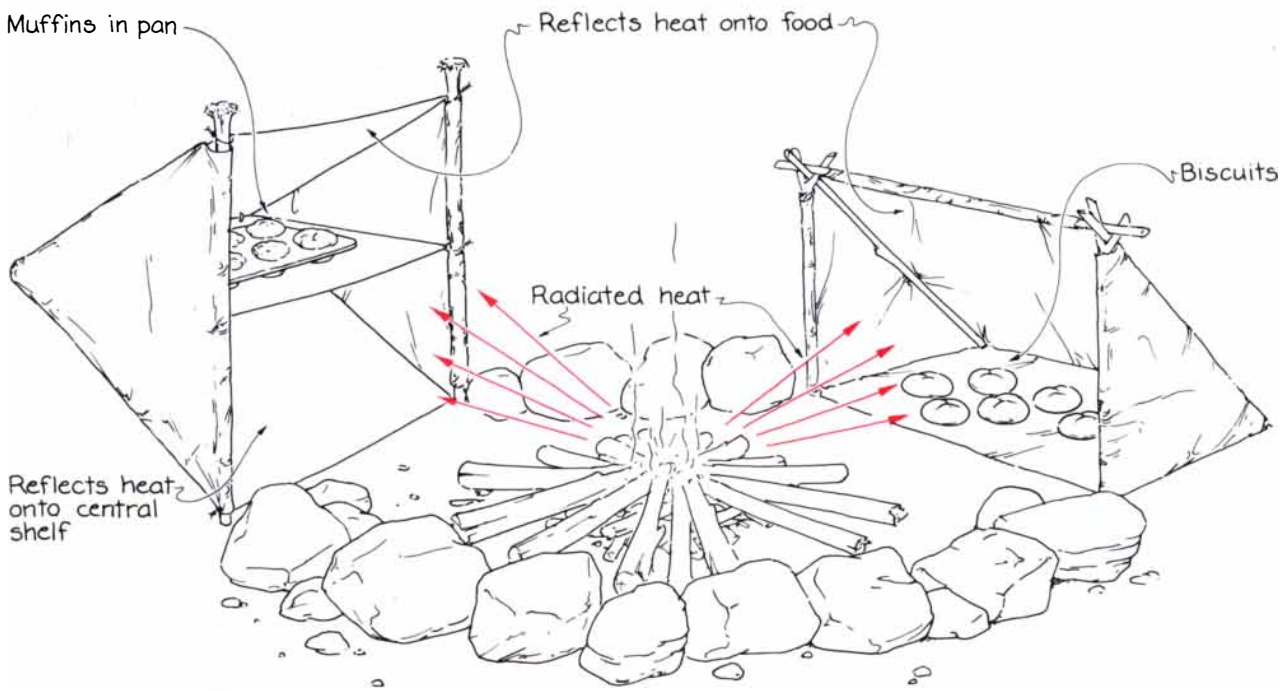
You could make the stove without the loose end plate held near the top.

That plate, however, helps to produce a nearly uniform temperature over the entire cooking surface. Without this plate the part of that surface directly above the heat source would be hotter than the rest of the surface because it receives more radiation from the source. The loose plate is intended to heat the small layer of air above it, transferring heat to the cooking surface evenly by the conduction and convection through the air.

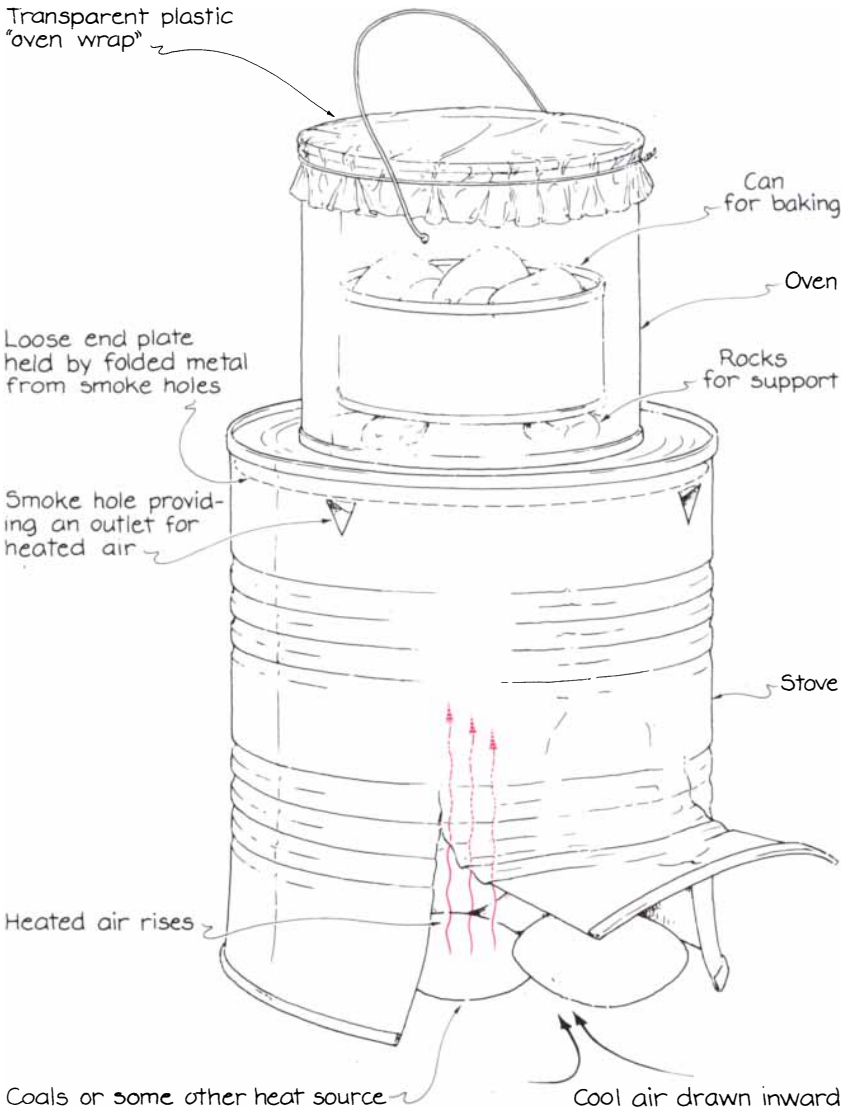
A popular heat source for the stove is a "buddy burner," a small can filled with corrugated cardboard over which hot paraffin has been poured. When the can is brought out for cooking, the paraffin is solid. A match melts and vaporizes some of it, and thereafter the vapor burns. The flame melts more paraffin, which is drawn to the top of the cardboard, where it vaporizes and burns. The cardboard also burns, but slowly, like the wick of a candle. A damper can be placed over part of the burner to slow things down if the stove gets too hot. Make the damper by fold-



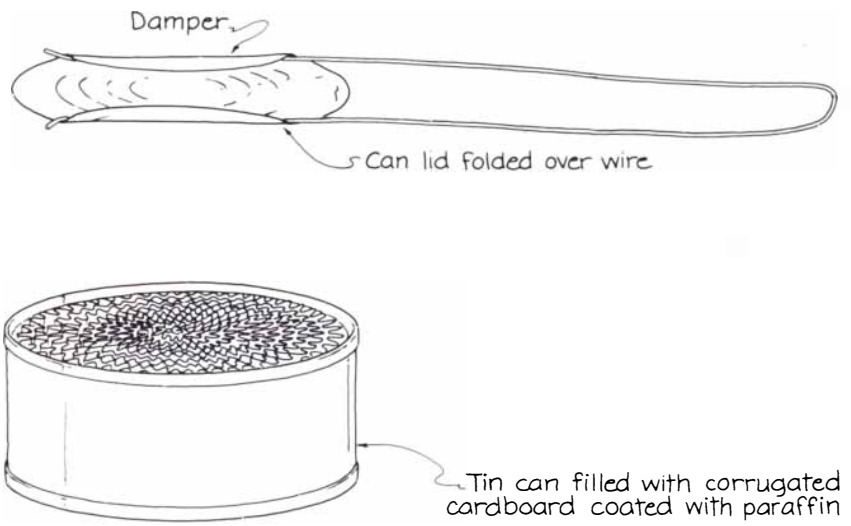
An arrangement for a "dingle fan"



Two designs for reflector ovens



An oven mounted on a can stove



A "buddy burner" and a damper

ing the lid from the can over a doubled piece of wire.

Charcoal briquettes also serve well as a heat source, but it is hard to light them. One easy solution is to put them in a can that works like a chimney. Cut both ends off a large can and punch flaps into the side near one end. Fold the flaps inward. Push large wads of newspaper against them so that they hold the paper in place. Prop the can against a rock with the paper downward. Leave an air space between that end and the ground. Put the briquettes in the can and ignite the newspaper. The flame pulls air into the bottom of the can to replace the hot air rising through the briquettes and out of the can. The strong flow of air makes the flames ignite the briquettes. When the newspaper has burned up and the briquettes begin to fall to the ground, lift the can off them. You now have your cooking fire.

The can stove can be converted into a small oven by means of two more cans and several flat rocks. One of the cans should be larger than the other but no wider than the top of the stove. Remove one end of the smaller can and both ends of the other one. Wrap transparent plastic around an open end of the larger can, holding it in place with a piece of wire. Tuck the ends of another piece of wire under the first one to form a handle. The plastic, commonly called oven wrap, is designed for oven cooking; it will not melt at typical oven temperatures.

Lay the flat rocks on top of the stove. Position the smaller can on them with its open end upward. Put the larger can over that with the plastic upward. This assembly is an oven. Food can be baked in the smaller can. The plastic serves as the oven window.

Several other simple baking rigs are available to the campfire cook. In one of them you put a small can on several flat rocks on the ground and put a larger can over it. (Take one end plate out of the larger can but leave the other one in place.) Heap coals or briquettes on the top plate and around the sides of the larger can. As that can gets warm the air trapped in it heats up too, forcing hot air to flow around the smaller can, which holds the food.

You can also bake or roast food by wrapping it in heavy aluminum foil and putting it directly on the coals. It can be helpful to pile dirt on the foil. The dirt cuts down the supply of oxygen, causing the coals to burn incompletely, but it also traps the heat for as long as they do burn.

One disadvantage of this technique is that some parts of the food will cook too fast and may burn. To eliminate

the problem wrap the food in two layers of foil with several layers of newspaper between them. The air trapped between the layers slows the transfer of heat; air is a poor conductor of heat and the pockets of air are too small to allow convection.

You can also protect the food from overrapid transfer of heat into the foil if at least part of the food is fairly moist. Then much of the heat goes into heating and vaporizing the water. The steam transfers heat to the rest of the food, cooking it.

The skin of an orange serves admirably as a baking dish. Slice off the top third of a large orange. Scoop out the rest of the orange and then partially fill it with dough. Replace the top of the orange and put the entire assemblage directly on the coals. Keep the orange upright so that the top stays on.

Heat conducted through the skin of the orange bakes the dough. Since the skin is thick and moist, it does not conduct the heat rapidly. Hence the dough is buffered against overheating at the places where the orange touches coals. Apples and bananas can be substituted for an orange. A large onion can be employed similarly to bake an egg or small pieces of meat.

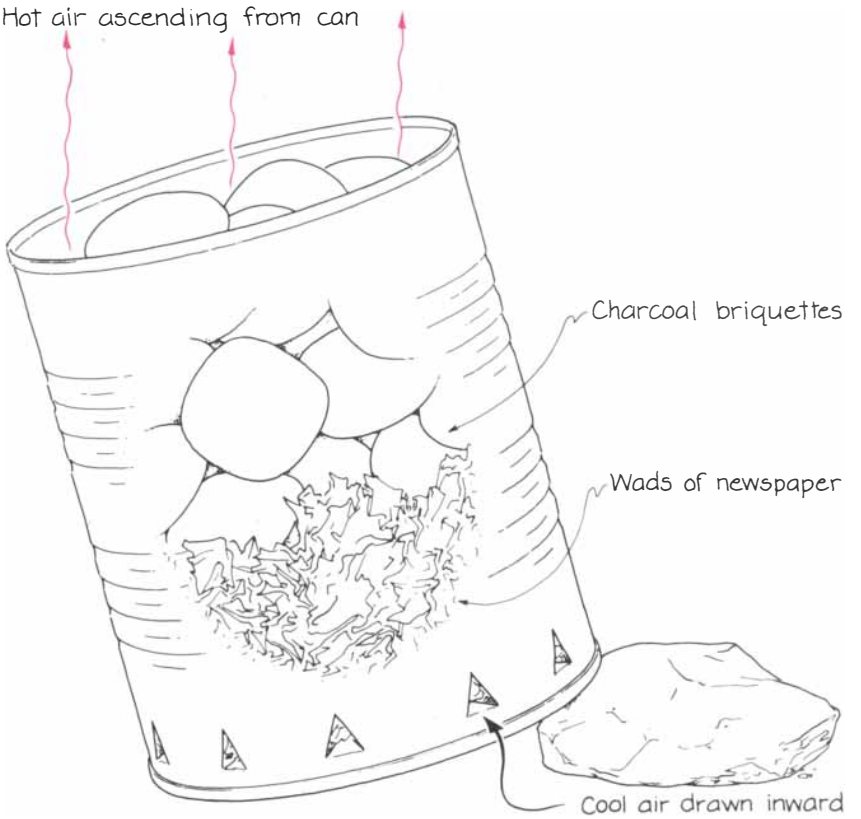
You can also bake with a reflector oven that utilizes the heat radiated by a campfire. This type of oven can be made of aluminum foil or metal plates (such as cookie sheets) supported by sticks. In the simplest version the food is put on a flat piece of foil. Another piece of foil set at an angle over the flat one reflects heat onto the food.

One side of aluminum foil is ordinarily shinier than the other. Position the shiny side of the reflecting layer toward the food. The foil holding the food should be placed with the dull side upward so that it absorbs some of the radiated heat instead of reflecting it all away.

A more ambitious reflector oven includes a second tilted layer of foil below the level one; it reflects radiation onto the bottom surface of the cooking layer. Here the level sheet should be a rigid layer of metal instead of foil, which would buckle under the weight of the food.

A crude reflector oven can be made out of a large can from which both ends have been cut. Place it on its side so that an open end faces the coals. The food goes inside. The coals radiate heat into the can, where it reflects into the food.

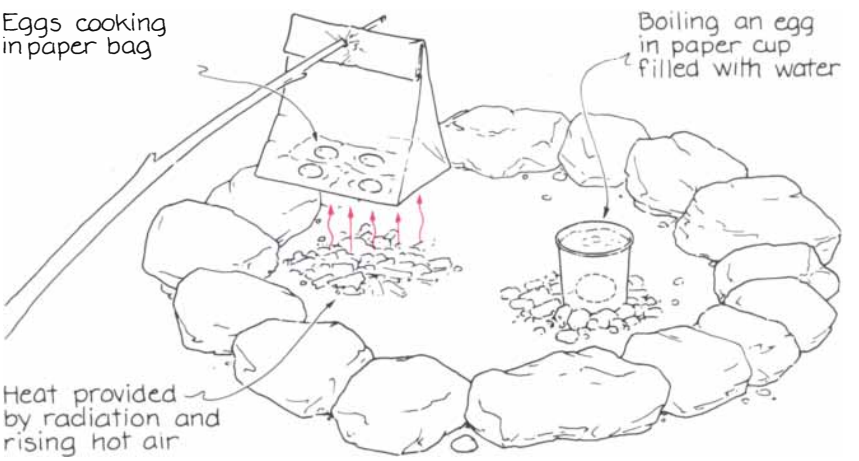
You can make a skillet for frying by wrapping heavy aluminum foil around the forked end of a stick. Wrap one layer of foil around the stick. Add several layers of newspaper and finally another layer of foil. The air trapped in the paper slows the transfer of heat, providing a more uniform temperature over the cooking surface.



How to ignite charcoal briquettes

In an emergency a flat rock put on the coals can serve as a skillet. Make sure that the rock is not wet or porous; water inside it may expand and blow the rock apart.

Some manuals for outdoor cooking describe how a whole chicken can be cooked with small rocks. Wrap the rocks in foil and put them on hot coals. When they are quite hot, transfer them to the interior of the chicken. Wrap the chicken in foil and place it in a container insulated with layers of newspaper or leaves. It will cook in an hour or more, depending on the temperature of the rocks. To avoid the possibility of salmonella and other types of food poisoning make certain that the chick-



Two ideas for cooking eggs

en cooks fully and is eaten soon afterward.

Another unusual cooking technique depends on the fact that liquid water can be no hotter than its boiling temperature of 212 degrees Fahrenheit. Relying on this fact, you can hard-boil an egg in a paper cup put directly on the coals. Paper will not burn until it reaches a temperature of at least 400 degrees F. The paper is so thin that the heat it receives from the coals is immediately conducted to the water.

Suppose you add coals below the cup or fan the coals already there so that heat is conducted faster into the water. Does the temperature of the water increase? No; the effect of the additional heat is to increase the frequency with which bubbles of water vapor form on the bottom. Because the water never gets hotter than 212 degrees F., the paper does not burn.

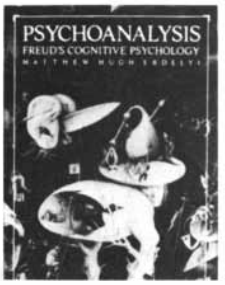
If the upper part of the cup is dry and receives enough heat by radiation from the coals or the rising hot air from them, it will ignite. To boil an egg in a paper cup you would do well to put enough water in the cup so that when the egg is in place, the cup is nearly full. The same principles apply when you boil water in containers made of canvas, bark, leaves or coconuts and placed directly on the coals.

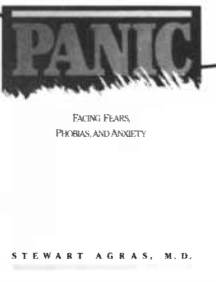
This fact about water accounts for another curious technique in cooking. Crack an egg (or more than one) into a small paper bag. Close the bag and fold the top down several times. Pierce the folded section with a stick. Prop the stick so that the bag is suspended above hot coals. The eggs will cook. The water in them keeps the bag from reaching its ignition point.

One drawback of this technique is that the eggs tend to adhere to the paper. My son Christopher found that the problem can be surmounted by coating the bottom interior of the bag with margarine or butter. (Some people solve the problem by lining the bottom of the bag with bacon. I avoid this procedure; if I cook the eggs for the proper length of time, the bacon is not cooked enough. Moreover, in order to kill the parasitic worm *Trichinella spiralis* that might reside in the bacon, all of the meat should be heated to at least 149 degrees F. The bacon in the bag is not that fully heated.)

You could look into a variety of other outdoor-cooking schemes. The most intriguing ones require only components that can be found in nature. You might also study the relative reflectance of thermal radiation by the two sides of aluminum foil or the rate at which heat is conducted through material such as an orange skin. Let me know what you find.

New and Captivating Reading from W. H. Freeman . . .





PSYCHOANALYSIS Freud's Cognitive Psychology

Matthew Hugh Erdelyi

The first scientific exposition of psychoanalysis to appear since the cognitive revolution in psychology, this long awaited, powerful, lucid, and engaging work sets forth Freud's often misunderstood ideas and brings them into play with modern developments in psychology and the neurosciences. Integrating a multitude of case histories with laboratory studies, visual art, poetry, jokes, and Freud's own writings, Erdelyi presents psychoanalysis as a quintessentially cognitive psychology, many precepts of which are only now being rediscovered and assimilated into mainstream psychology. He shows how, contrary to widespread myth, psychoanalysis is capable of scientific validation as well as falsification, and how many of its seminal insights have been borne out. What emerges is a Freud who is surprisingly modern and, in some respects, still visionary.

July 1985. 303 pages

THE BREAKING OF BODIES AND MINDS Torture, Psychiatric Abuse, and the Health Professions

Edited by Eric Stover
and Elena O. Nightingale, M.D.
Foreword by David A. Hamburg, M.D.
American Association for the Advancement of Science

"Everyone should read this book: former prisoners, future prisoners, and those who will never be prisoners, because everyone can and should struggle against the criminal use of science by state terrorism."
—Jacobo Timerman

Documenting the willing and unwilling participation of health professionals in torture and punitive psychiatry, the expert contributors to this timely new volume detail the historical and social contexts of these abuses, the ways in which they are medically "justified," and the efforts of other medical personnel to put a stop to them.

July, 1985. 309 pages

PANIC Facing Fears, Phobias, and Anxiety

Stewart Agras, *Stanford University*

Is panic learned, or inherited? The distinction, says Stewart Agras, is an artificial one. Nevertheless, advances in biology and learning theory have resulted in the current use of drug and exposure therapies which, in conjunction with social support, are proving effective in dealing with the Panic Syndrome—a disorder affecting 1.5% of the adult population. Examining closely related phobias and normal fears, Agras provides material for the reader's self-assessment in these areas and reveals how panic may well be traced to unresolved separation anxieties.

June 1985. 151 pages

PSYCHOANALYSIS
☐ Cloth ISBN 0-7167-1616-X.....$24.95
☐ Paper ISBN 0-7167-1617-8.....$14.95

THE BREAKING OF BODIES AND MINDS
☐ Cloth ISBN 0-7167-1732-8.....$21.95
☐ Paper ISBN 0-7167-1733-6.....$11.95

PANIC
☐ Cloth ISBN 0-7167-1730-1.....$19.95
☐ Paper ISBN 0-7167-1731-X.....$11.95

Postage & Handling.....$1.50
(NY, CA & UT residents please add appropriate sales tax.)

☐ I enclose a check made payable to W. H. Freeman and Company.
☐ MasterCard ☐ VISA Exp. Date: ____ / ____ Acct. # _____
Signature* _____
Name _____
Address _____
City _____ State _____ Zip _____
(All credit card orders must be signed.)

W. H. Freeman and Company
4419 West 1980 South, Salt Lake City, Utah 84104

0216

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

MATHEMATICAL RECREATIONS FOR THE PROGRAMMABLE CALCULATOR. Dean Hoffman and Lee Mohler. Hayden Book Co., 1982.

THE FRACTAL GEOMETRY OF NATURE. Benoit B. Mandelbrot. W. H. Freeman and Company, 1983.

DISCRETE ITERATIONS. François Robert; translated from the French by Jon Rokne. Springer-Verlag, in press.

A CRITICAL TIME FOR NUCLEAR NONPROLIFERATION

THE LAST CHANCE: NUCLEAR PROLIFERATION AND ARMS CONTROL. William Epstein. The Free Press, 1976.

WHY STATES GO—AND DON'T GO—NUCLEAR. William Epstein in *Nuclear Proliferation: Prospects, Problems and Proposals.* Annals of the American Academy of Political and Social Sciences, Vol. 430; March, 1977.

NUCLEAR PROLIFERATION TODAY. Leonard S. Spector. Vintage Books, 1984.

NON-PROLIFERATION: THE WHY AND THE WHEREFORE. Jozef Goldblat. SIPRI (Stockholm International Peace Research Institute), 1985.

THE ADRENAL CHROMAFFIN CELL

THE ADRENAL MEDULLA. Stephen W. Carmichael. Eden Press, 1979–84.

THE MOLECULAR ORGANIZATION OF ADRENAL CHROMAFFIN GRANULES. Hans Winkler and Edward Westhead in *Neuroscience,* Vol. 5, No. 11, pages 1803–1823; November, 1980.

THE CHROMAFFIN GRANULE. H. Winkler and S. W. Carmichael in *The Secretory Granule,* edited by A. M. Poisner and J. M. Trifaró. Elsevier Biomedical Press, 1982.

CALCIUM CONTROLS EXOCYTOSIS IN BOVINE ADRENAL MEDULLARY CELLS. Peter Baker and Derek Knight in *Trends in Neuroscience,* Vol. 7, No. 4, pages 120–126; April, 1984.

THE SCANNING TUNNELING MICROSCOPE

ENERGY GAP IN SUPERCONDUCTORS MEASURED BY ELECTRON TUNNELING. Ivar Giaever in *Physical Review Letters,* Vol. 5, No. 4, pages 147–148; August 15, 1960.

(111) FACETS AS THE ORIGIN OF RECONSTRUCTED AU(110) SURFACES. G. Binnig, H. Rohrer, Ch. Gerber and E. Weibel in *Surface Science,* Vol. 131, No. 1, pages L379–384; August, 1983.

REAL-SPACE OBSERVATION OF 2×1 STRUCTURE OF CHEMISORBED OXYGEN ON NI(110) BY SCANNING TUNNELING MICROSCOPY. A. M. Baró, G. Binnig, H. Rohrer, Ch. Gerber, E. Stoll, A. Baratoff and F. Salvan in *Physical Review Letters,* Vol. 52, No. 15, pages 1304–1307; April 9, 1984.

THE CLASSICAL VACUUM

A BRIEF SURVEY OF STOCHASTIC ELECTRODYNAMICS. T. H. Boyer in *Foundations of Radiation Theory and Quantum Electrodynamics,* edited by A. O. Barut. Plenum Press, 1980.

QUANTUM FIELD THEORY, HORIZONS, AND THERMODYNAMICS. D. W. Sciama, P. Candelas and D. Deutsch in *Advances in Physics,* Vol. 30, No. 3, pages 327–366; May–June, 1981.

DERIVATION OF THE BLACKBODY RADIATION SPECTRUM FROM THE EQUIVALENCE PRINCIPLE IN CLASSICAL PHYSICS WITH CLASSICAL ELECTROMAGNETIC ZERO-POINT RADIATION. T. H. Boyer in *Physical Review D,* Vol. 29, No. 6, pages 1096–1098; March 15, 1984.

THE MESOZOIC VERTEBRATES OF THAILAND

CYCLOTOSAURUS CF. *POSTHUMUS* FRAAS (CAPITOSAURIDAE, STEREOSPONDYLI) FROM THE HUAI HIN LAT FORMATION (UPPER TRIASSIC), NORTHEASTERN THAILAND, WITH A NOTE ON CAPITOSAURID BIOGEOGRAPHY. Rucha Ingavat and Philippe Janvier in *Géobios,* Vol. 14, No. 6, pages 711–725; 1981.

PALEOGEOGRAPHIC HISTORY OF WESTERN THAILAND AND ADJACENT PARTS OF SOUTH-EAST ASIA: A PLATE TECTONICS INTERPRETATION. Sangad Bunopas. Geological Survey of Thailand Paper No. 5, Department of Mineral Resources, Bangkok, Thailand, 1982.

TRIASSIC TURTLE REMAINS FROM NORTHEASTERN THAILAND. France de Broin, Rucha Ingavat, Philippe Janvier and Nares Sattayarak in *Journal of Vertebrate Paleontology,* Vol. 2, No. 1, pages 41–46; May, 1982.

GONIOPHOLIS PHUWIANGENSIS NOV. SP., A NEW MESOSUCHIAN CROCODILE FROM THE MESOZOIC OF NORTHEASTERN THAILAND. Éric Buffetaut and Rucha Ingavat in *Géobios,* Vol. 16, No. 1, pages 79–91; 1983.

THE PERMO-TRIASSIC "PALEOTETHYS" IN MAINLAND SOUTHEAST ASIA AND ADJACENT PARTS OF CHINA. D. Helmcke in *Geologische Rundschau,* Vol. 74, pages 215–228; 1985.

YOUNG SUPERNOVA REMNANTS

TYCHO BRAHE, A PICTURE OF SCIENTIFIC LIFE AND WORK IN THE SIXTEENTH CENTURY. J. L. E. Dreyer. A. & C. Black, Edinburgh, 1890.

THE HISTORICAL SUPERNOVAE. David H. Clark and F. Richard Stephenson. Pergamon Press, 1977.

SUPERNOVA REMNANTS AND THEIR X-RAY EMISSION. Edited by John Danziger and Paul Gorenstein. D. Reidel Publishing Co., 1983.

THE ANTARCTIC ICE

ANTARCTIC SNOW AND ICE STUDIES 2. Edited by A. P. Crary. Antarctic Research Series Vol. 16, American Geophysical Union, 1971.

FROZEN FUTURE: A PROPHETIC REPORT FROM ANTARCTICA. Edited by Richard S. Lewis and Philip M. Smith. Quadrangle Books, 1973.

THE CLIMATIC RECORD IN POLAR ICE SHEETS. Edited by Gordon de Q. Robin. Cambridge University Press, 1983.

BUTTERFLY ENGINEERING

MECHANICAL DETERMINANTS OF NECTAR FEEDING STRATEGY IN HUMMINGBIRDS: ENERGETICS, TONGUE MORPHOLOGY, AND LICKING BEHAVIOR. J. G. Kingsolver and T. L. Daniel in *Oecologia,* Vol. 60, pages 214–226; 1983.

THERMOREGULATION AND FLIGHT IN *COLIAS* BUTTERFLIES: ELEVATIONAL PATTERNS AND MECHANISTIC LIMITATIONS. J. G. Kingsolver in *Ecology,* Vol. 64, pages 534–545; 1983.

THE AERODYNAMICS OF HOVERING INSECT FLIGHT. C. P. Ellington in *Philosophical Transactions of the Royal Society, B. Biological Sciences,* Vol. 305, No. 1122, pages 1–181; February 22, 1984.

THE AMATEUR SCIENTIST

ROUGHING IT EASY. Dian Thomas. Warner Books, 1975.

BOY SCOUT FIELDBOOK. Boy Scouts of America. Workman Publishing Company, 1978.

THE U.S. ARMED FORCES SURVIVAL MANUAL. Edited by John Boswell. Rawson, Wade Publishers, Inc., 1980.

CENSORSHIP IN A FREE SOCIETY. IT'S A BAD MATCH.

Censorship is the greatest tragedy in American literature. It constricts the mind, teaches fear and leaves only ignorance and ashes.

Today, all over the country, books are being banned, burned and censored. Teachers, students, librarians, and book and magazine publishers are being harassed.

The attacks of these self-appointed censors are endorsed by our silence.

The freedom to read is one of our most precious rights. Do something to protect it.

Contact:
People For The American Way, P.O. Box 19900 Washington, D.C. 20036 or call 202/822-9450.

People For The AmericanWay

Stolichnaya
The Vodka

For gift delivery anywhere* call 800-528-6148 (Arizona 602-957-4923).

*Except where prohibited / Major credit cards accepted / Stolichnaya Vodka 80 and 100 proof, distilled from grain. Imported by Monsieur Henri Wines, Ltd. White Plains, NY 10604