SCIENTIFIC AMERICAN

NOVEMBER 1987
$2.50

Modern diet and Stone Age physiology raise the risk of cancer.
Now biologists have another powerful tool: manmade chromosomes.
A new class of supernovas has upset established theory.

***The Panda's** adaptation to a diet of bamboo obscured its ancestry; techniques of molecular biology have helped to solve the puzzle.*

Hundreds of times a second, microprocessors adjust its engine for peak performance.

Meanwhile, sensors on each wheel stand ready to "pump" the brakes faster than any race car driver could, to help prevent the wheels from locking in panic stops.

And throughout the cockpit, a veritable arsenal of computer technology informs the driver of everything from the moment outside temperatures near the freezing point, to the date when routine maintenance may be necessary.

In short, the 528e from the Bavarian Motor Works would qualify as one of the world's most advanced technology centers, were its true calling not something even more profound: a civilized, four-door luxury sedan whose astounding technology only renders it that much more civilized.

As such, the new, more powerful BMW 528e, in the words of Motor Trend magazine, provides drivers an experience akin to "piloting a precision aeronautical machine rather than a mere automobile."

And it relegates many of the automobiles that are currently trumpeting "technology" in their marketing strategies to the role of followers.

To these followers, the editors of Motor Trend have a suggestion: the 528e "provides a ride quality most manufacturers would do well to emulate."

To those in the market for a luxury performance sedan, we would suggest the choice is between the BMW 528e and those who would do well to emulate it.

Which is no choice at all, technically speaking.

THE ULTIMATE DRIVING MACHINE.©

SCIENTIFIC AMERICAN

November 1987 Volume 257 Number 5

DEPARTMENTS

THE COVER depicts a giant panda and its staple food (see "The Ancestry of the Giant Panda," by Stephen J. O'Brien, page 102). The evolutionary relationships of the panda have been obscured by the fact that it is mostly vegetarian and has certain anatomical and behavioral traits that enable it to subsist on a diet of bamboo. After nearly a century of debate the story of the panda's ancestry now appears to have been clarified by techniques of molecular analysis.

THE ILLUSTRATIONS

Cover painting by Marvin Mattelson

Page	Source
43	Carol Donner
44–48	Andrew Christie
51	John F. Kielkopf
52–57	Ian Worpole
58	European Southern Observatory (*top*), Ian Worpole (*bottom*)
63	Philip Hieter, Johns Hopkins University School of Medicine
64–68	Hank Iken
89	Lawrence Berkeley Laboratory
90–96	George Retseck
97	Keith Moffat, Cornell University (*left*); Janos Kirz, State University of New York at Stony Brook (*middle*); Corporate Research Laboratory, Exxon Research and Engineering Company (*right*)
98–99	George Retseck
102–103	Jessie Cohen, National Zoological Park, Washington, D.C.
104	Johnny Johnson
105	Patricia J. Wynne
106	Stephen J. O'Brien
107	Patricia J. Wynne
109	James Kilkelly and R. Bruce Laughlin
110–114	Jerome Kuhl
118–119	Ian Kinnes, British Museum
120–126	Tom Prentiss
128A	Nova Scotia Department of Government Services
128B-131	George V. Kelvin
135	Jearl Walker
136–138	Michael Goodman
141–144	Heinz-Otto Peitgen and Peter H. Richter, University of Bremen

SCIENTIFIC AMERICAN

Established 1845

EDITOR Jonathan Piel

BOARD OF EDITORS Armand Schwab, Jr. (Managing Editor), Timothy Appenzeller (Associate Editor), Timothy M. Beardsley (Washington), John M. Benditt, Laurie Burnham, David L. Cooke, Jr., Ari W. Epstein, Gregory R. Greenwell, John Horgan ("Science and the Citizen"), June Kinoshita, Philip Morrison (Book Editor), Ricki L. Rusting, Karen Wright

ART DEPARTMENT Samuel L. Howard (Art Director), Edward Bell (Assistant Art Director), Johnny Johnson

PRODUCTION DEPARTMENT Richard Sasso (Director of Production), Carol Eisler and Leo J. Petruzzi (Assistants to the Production Director), Carol Hansen (Electronic Composition Manager), Carol Albert, Nancy Mongelli, Jody Seward, William Sherman, Julio E. Xavier

COPY DEPARTMENT Sally Porter Jenks (Copy Chief), M. Knight, Michele S. Moise, Dorothy R. Patterson

PUBLISHER Harry Myers
ADVERTISING DIRECTOR Peter B. Kennedy
CIRCULATION MANAGER William H. Yokel
DIRECTOR OF OPERATIONS R. Vincent Barger

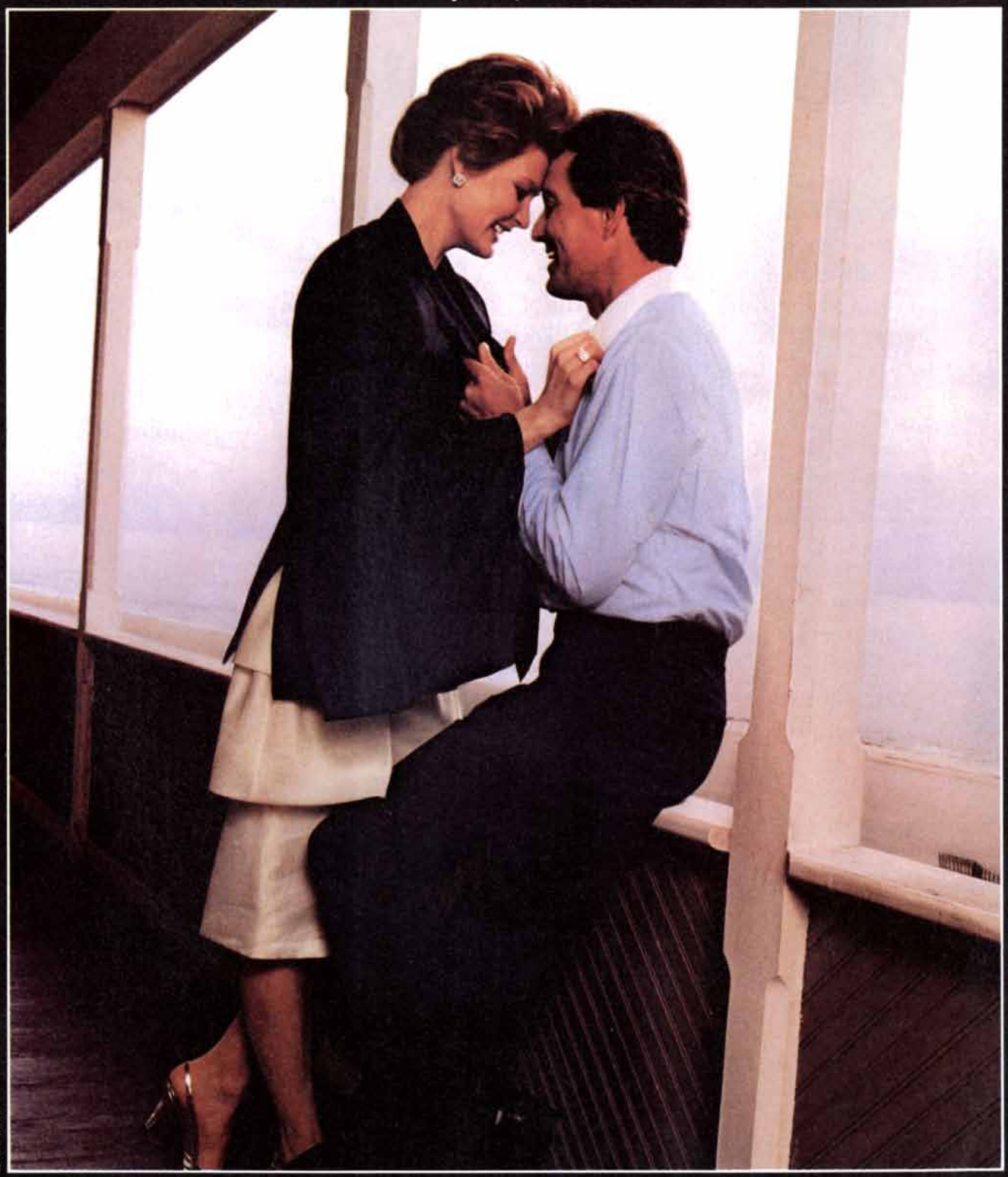
PRESIDENT Georg-Dieter von Holtzbrinck
EXECUTIVE COMMITTEE Claus-Gerhard Firchow, Executive Vice-President; Harry Myers, Senior Vice-President; R. Vincent Barger, Vice-President, Finance; Linda Chaput, Vice-President; Jonathan Piel, Vice-President; Carol Snow, Vice-President

CHAIRMAN OF THE BOARD Gerard Piel

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, P.O. Box 953, Farmingdale, N.Y. 11737-0000.

You always tell people how much you love your wife.
This year, tell her.



DeBeers

The Diamond Anniversary Ring.

A band of diamonds that says you'd marry her all over again.

A diamond is forever.

For anyone who's ever IBM presents personal

Now it's easier than ever to become a newsmaker in business, thanks to IBM's new desktop publishing system.

It's called the IBM® SolutionPac™ Personal Publishing System. And it comes complete with everything you need to start making headlines: the hardware, software, service and support.

The IBM SolutionPac for Personal Publishing lets you design and print near typeset quality documents.

Just one purchase, and you'll be on your way to designing and printing the kinds of presentations, reports and newsletters bound to open people's eyes.

The Personal Publishing System is as easy to set up as it is to buy. All the components are designed to work together, from the Personal System/2™ Model 30 to IBM's laser printer, the Personal Pageprinter.

We've also preloaded all the necessary software—including Windows,™

wanted to make headlines, publishing.

PageMaker® and PostScript®—on the system's hard disk. All you have to do is turn the system on, run through the tutorial program and begin composing your thoughts.

Because it's from IBM, you can transfer many of your IBM-compatible files directly into the Personal Publishing System—where you can then turn them into eye-catching documents.

There's good news for IBM PC XT™/286 and Personal Computer AT® owners, too. The IBM Personal Publishing Option offers you everything you need to upgrade your PC into a cost-saving print shop.

The IBM 3117 Scanner transfers artwork directly into the system.

To see that your newsworthy ideas are noteworthy, visit your IBM Authorized Advanced Products Dealer (ask about the IBM Credit Card), or call your IBM Marketing Representative.

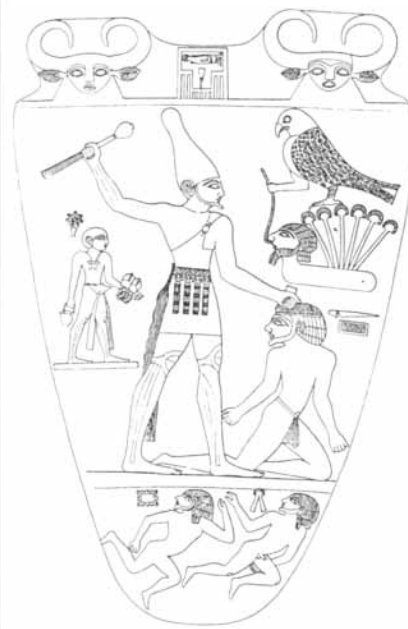
For the certified dealer nearest you, call 1-800-447-4700, ext. 48. (In Canada, call 1-800-465-6600.) IBM®

LETTERS

To the Editors:

As an egyptologist, I was struck by the aerial photograph of a desert kite in "Gazelle Killing in Stone Age Syria," by Anthony J. Legge and Peter A. Rowley-Conwy [SCIENTIFIC AMERICAN, August]. An almost identical form was illustrated by the ancient Egyptians in about 3100 B.C., on a votive palette recording the conquests of the First Dynasty king Narmer. Placed above the figure of a fallen enemy, the form is clearly intended to identify the tribe or people he represents. The device has traditionally been interpreted as some kind of plant form. But in 1955 the Israeli archaeologist Yigael Yadin pointed out its similarity to the shape of desert kites in aerial photographs. None of his comparative examples resembles the Narmer sign as closely as the photograph in the article does.

Beside the "kite" man lies a second foe, identified by a crenellated rectangle that is unmistakably the plan of a fort or fortified town. The two fallen foes seem to form a group, separated from the rest of the design by a thick horizontal line. Yadin made the plausible suggestion that these people had been defeated in military actions outside the Nile valley. Since he was most familiar with kite remains in northern Jordan, he assumed that this area had been the site of the conflict. It is, however, a very long march from Egypt. In their caption for the aerial photograph the authors mention traces of kites in the Sinai. This reference suggests that the Sinai is the likelier battleground, particularly since it shows many signs of Egyptian activity in the late fourth millennium B.C.



Narmer palette (Egyptian Museum, Cairo)

Whether the two defeated foreigners at the bottom of the Narmer palette represent regions, kingdoms or tribes, a fort and a kite would characterize them in terms of distinctly different ways of life. When Yadin made his suggestion, the kite was thought to be a protective enclosure for the flocks of nomadic herdsmen. If instead its sole use was in large-scale hunts (as described in the article), it would seem that the Egyptians—who surely knew the function of something they portrayed so accurately—intended to differentiate between fully agricultural populations in fortified communities and people who, whether nomadic or sedentary, still depended heavily on hunting.

Desert hunts were a feature of life in ancient Egypt, and they are often depicted. There is a symbolic and sometimes sporting element in these scenes, as well as a considerable degree of pictorial license (for example, the captured beasts sometimes include winged griffins and other imaginary creatures). But the representations clearly show that a method of hunting at least as important as tracking and the chase involved a large netted or fenced enclosure, within which the trapped animals were killed or captured. This construction is never fully portrayed. It has been variously interpreted as some kind of trap, as a holding pen in which to collect sparse desert fauna or even as a temporary game park for a great man's pleasure. Several features that are common to all representations of these enclosures, however—the inclusion of nongame animals, among them hedgehogs, the depiction of some animals mating or giving birth and the temporary appearance of the structures—suggest that they were set up for large drives, conducted annually at a certain season. If they were, it is likely that the Egyptian hunting enclosure was a flimsy form of desert kite.

The depletion and ultimate extinction of game animals in the deserts adjoining ancient Egypt have been attributed to both desiccation and overhunting. In Egypt as in other semiarid parts of the ancient Near East, the human contribution to the impoverishment of wildlife

BREAKTHROUGH: BOUNCE ENGINE NOISE OFF A CUSHION OF AIR TO QUIET AIRLINE CABINS.

Aircraft are being developed with large jet-driven fans that would give airlines great fuel economy. However, sound from the engines and the swiftly turning blades must be controlled to keep the cabin pleasant for the passengers.

To help preserve both cabin comfort and fuel efficiency, McDonnell Douglas researchers are investigating a special surface for the plane. Barely visible "bumps" on the aft fuselage create a thin cushion of turbulence that deflects the annoying sound. The noise is carried away with the wind, so passengers hear only the comforting hum of powerful engines at work.

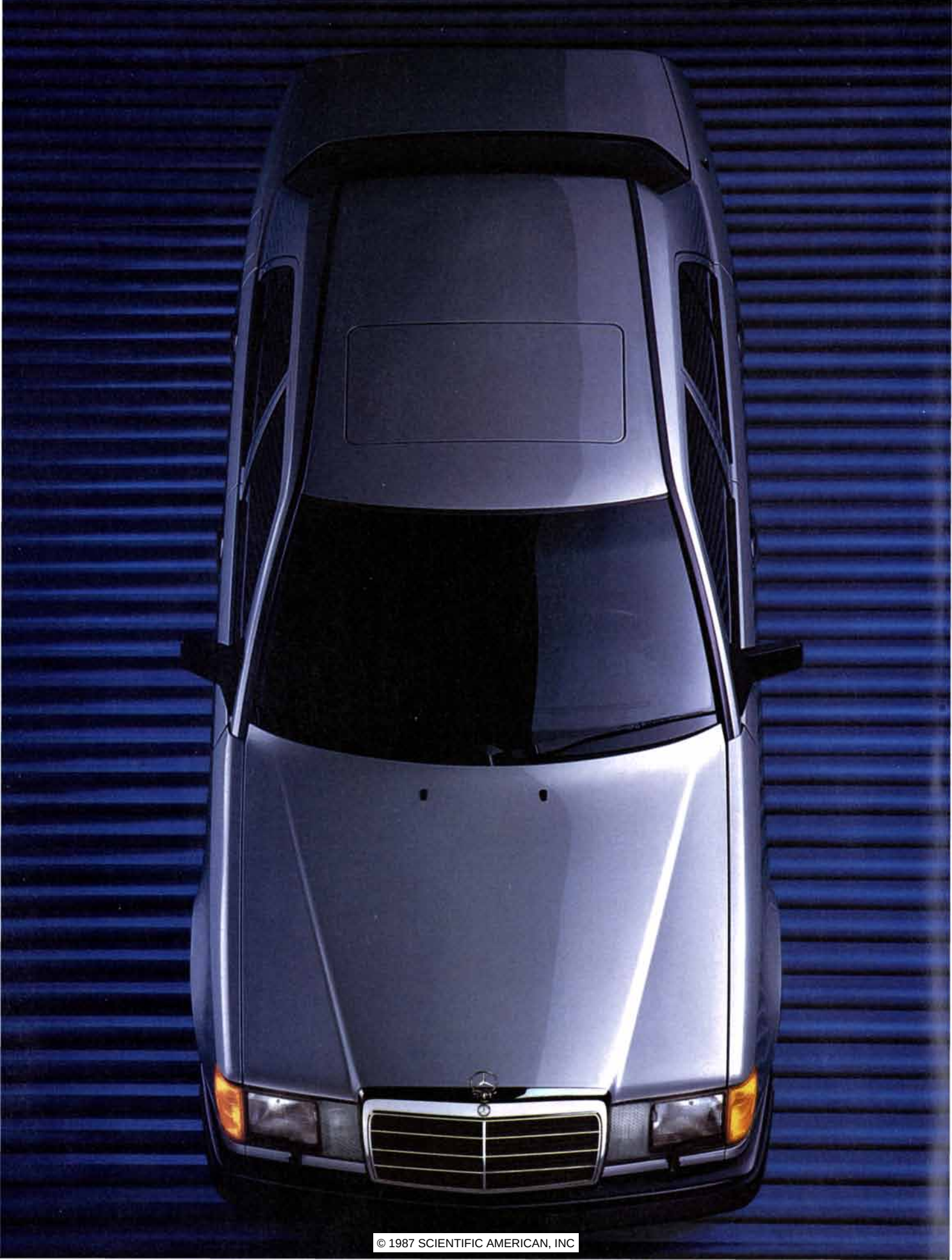
We're making breakthroughs that add value and quality for airlines and their passengers.

We're McDonnell Douglas.

For more information, write:
Noise Control, McDonnell Douglas,
Box 14526, St. Louis, MO 63178

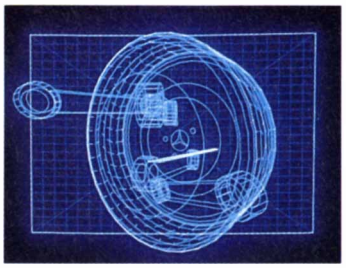
MCDONNELL DOUGLAS

COMMERCIAL & MILITARY AIRCRAFT INFORMATION SYSTEMS EQUIPMENT FINANCING SPACECRAFT & MISSILES HEALTH CARE HELICOPTERS



THE MERCEDES-BENZ 300 CLASS: TECHNOLOGY IS ONLY OPPORTUNITY KNOCKING– THE SECRET IS TO OPEN THE RIGHT DOORS.

A microchip might be able to think with unimaginable speed—but what does it think about? Computers, robots and laser beams might extend human reach—but to what end?



The answers to such questions, as delineated by the engineers of Mercedes-Benz, explain why the ambitions of the 300 Class seem to exceed those of other technologically intensive automobiles.

Thus no microchip in a 300 Class automobile thinks frivolously. One thinks about modulating brake pressure in a panic stop, to prevent skidding and preserve steering control. Another, about fine-tuning engine operation, seeking the ideal balance of power and efficiency under all conditions. Still others deploy a driver's-side air bag and front seat belt emergency tensioning retractors within milliseconds of a major frontal impact.

Advanced computers pretested 77 experimental variations of the 300 Class multilink independent rear suspension system. But only hard test driving and human feel could ultimately turn theoretical perfection into what *Automobile Magazine* has termed "the yardstick . . . because of its nearly perfect blend of handling, roadholding and comfort."

Extensive wind-tunnel testing shaped a 300 Class body that is aerodynamically slipperier than even the slippery Porsche 928. But human genius molded that shape into a spacious, civilized four-door sedan. Robots welded exotic steels into a lightweight yet rigid body structure. But the Mercedes-Benz safety philosophy demanded that the structure also *yield* in a severe front or rear impact—absorbing kinetic energy, cushioning the passenger shell within.

Synthesizing myriad diverse technologies, the creators of the 300 Class have achieved automobiles that merge 140-mph test-track performance with resolute safety-mindedness. That blend the instant satisfaction of driving pleasure with the long-term satisfaction of durability and reliability. They gave the 300 Class a superior automobile intelligence, in brief, serving the needs and well-being of driver and passengers as no automobile before.

An achievement that reflects not only the power of technology, but the power of the vision that makes a Mercedes-Benz.

Engineered like no other car in the world

may have included the all too deadly efficiency of desert kites.

EDNA R. RUSSMANN

Associate Curator
Department of Egyptian Art
The Metropolitan Museum of Art

To the Editors:

Gus W. Van Beek, in his excellent article "Arches and Vaults in the Ancient Near East" [SCIENTIFIC AMERICAN, July], does not mention what is perhaps the most fascinating aspect of these superb constructions: they are not circular arcs, as Roman and Renaissance arches are, but instead they approximate closely the funicular curve (the inverted curve of a hanging chain), which is structurally ideal for a uniformly loaded vault. The reader can verify this by pinning the cover painting upside down on a vertical surface and suspending a length of watch chain or limp string between the ends of the span; the length of chain or string can be adjusted until it hangs along the vault.

The Romans undoubtedly used the less than ideal semicircular form because they could more easily build centering (temporary support) and cut stone voussoirs (wedge-shaped blocks or bricks) for it; perhaps they also liked its simpler shape. But the builder of pitched-brick vaults learns instinctively from the tactile experience of construction that the semicircle is not the correct form and is considerably weaker structurally. Having built several pitched-brick vaults myself and with students, I can attest that the funicular shape emerges automatically, as if by magic, as one balances bricks against one another in the steeply sloping courses.

The introduction of voussoirs was perhaps not as much of a technological advance as the author suggests. The rectangular brick with wedge-shaped mortar joints is a universal, highly economical module for vaulting of any profile or span. The wedge-shaped brick does not in fact create a stronger structure, provided the mortar is as strong as the brick itself, and the problems of manufacturing and handling matched sets of voussoirs of varying shapes for a funicular vault are formidable.

EDWARD ALLEN

South Natick, Mass.

(*Van Beek's reply follows.*)

To the Editors:

Mr. Allen is certainly correct in stressing the strength and ease of construction of pitched-brick vaults when they are designed on the catenary (or funicular) curve, and his do-it-yourself demonstration of the catenary using the cover painting (which shows the vault at Ctesiphon) is imaginative and welcome. Hassan Fathy, the most prolific builder of pitched-brick vaults in the 20th century, uses catenarian vaults exclusively. I shall treat this subject in detail in my forthcoming book on mud-brick construction.

Yet it is misleading to overemphasize the use of the catenary curve in ancient pitched-brick vaulting. Most of the vaults at the Ramesseum are even more semicircular than the one that is shown in the article. Note also that the second preserved section of vaulting in the San Simeon monastery rises only slightly above a semicircular curve. The majority of the examples of pitched-brick vaulting surviving from antiquity are closer to the semicircular than to the catenary curve.

By allowing semicircular and even flatter vaults, voussoirs make possible the construction of an upper floor with a minimum of weight on the vaults. The reason is that space between vaults, or between a vertical wall and a vault, must be filled in with brick or some other device—such as the small, secondary air-shaft vaults in the San Simeon monastery—to support an upper floor. Vaults built on a catenary curve rise much higher than semicircular ones and therefore require a much more massive (and heavy) brick fill between vaults to carry an upper floor.

As for the strength of mud-brick voussoir vaults, with their flatter curves, it may be of interest to know that when my colleagues and I excavated the preserved vault in Room B of the Assyrian building (dating from about 675 B.C.) at Tell Jemmeh, we found an empty space—an air pocket—one meter wide, a little more than one meter long and about 35 centimeters high *under* the vault. Without support from below, this vault bore the weight of successive towns built on top of it during the next 450 years and of debris layers two meters thick for 2,200 more years. I suspect that few contemporary homes and functional office buildings, constructed with new materials and technologies, will equal this performance.

GUS W. VAN BEEK

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

NOVEMBER, 1937: "Astronomers often assert that what we human beings can see beyond the earth is only what the annoying earth's atmosphere permits us to see. However, by going in an airplane to an altitude of 25,000 feet (above most of the atmosphere) and photographing the sun's corona during an eclipse, Major Albert W. Stevens, a member of the Hayden Planetarium–Grace Eclipse Expedition, has revealed for the first time that the solar corona is not merely the familiar array of finger-like streamers from the sun but a globular layer a million miles thick surrounding it."

"In these days when everything gets bigger and bigger, scientific instruments are growing too. Champion for size, of course, is the 200-inch telescope, which on a weight-for-weight basis probably exceeds by several times the mass of any other instrument intended for pure scientific research. Its total weight will be in the neighborhood of 500 tons. One of the closest runners-up is the 'atom smasher' that research engineers of the Westinghouse Electric & Manufacturing Company are building at the company's East Pittsburgh research laboratories. It will stand 65 feet high, atop a two-story building, and will weigh about 100 tons."

"When soybean flour that has been completely freed of oil by solvent extraction is dissolved in water, the resulting solution can be whipped into a stiff white foam greatly resembling egg white. Since protein in the form of soy meal costs only about 1/10 as much as protein in the form of eggs or milk, which it resembles dietetically, research has been directed toward its use to replace these more expensive foods."

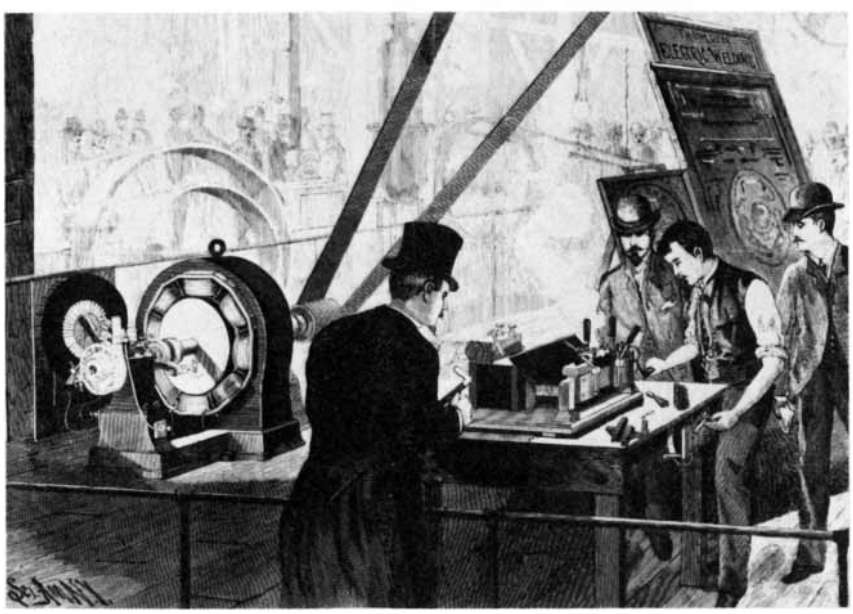
"A tireless automatic machine, which 'sees' better than a hawk, 'hears' better than a hare and has a sense of 'touch' infinitely acute, has been placed in operation at a Ford plant for inspection of the valve push rods—small, accurately finished engine parts. The uncanny device, which utilizes photoelectric cells and radio amplification to work its wonders, inspects valve push rods for hardness, hidden fissures and accuracy of dimension."

NOVEMBER, 1887: "The Russian paper *Caspian* gives the following interesting details of one of the largest naphtha fountains yet known, which has lately broken out near Baku and which threatens to inundate all Balakhani. The naphtha, owing to the pressure of the gases that accompany it, rises to a height of 280 feet to 420 feet and is carried away by the wind to a great distance, falling like fine rain at the more distant parts of the district, but near the fountain coming down in torrents that form rivers and streamlets. Farther on it falls like sleet and settles in a layer on all the buildings in the neighborhood."

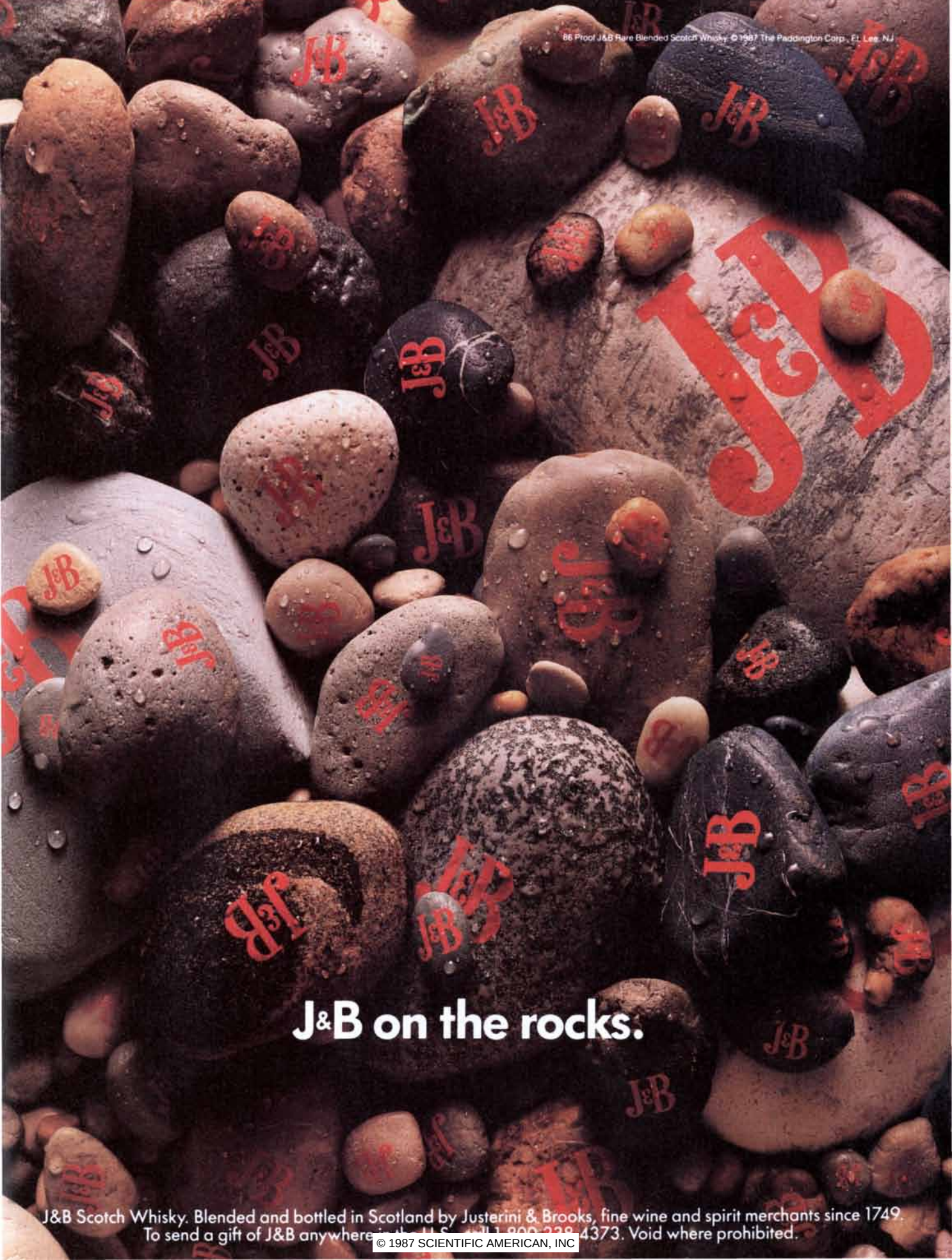
"At a recent meeting of the geological section of the British Association, a report was read on the present condition of the experimental heading for the channel tunnel between Dover and Calais, a distance of twenty-one miles, the completion of the work having been forbidden by the English government. A hole has already been bored seven feet in diameter, one mile and a quarter in length, nearly the whole of which is actually beneath the sea bottom. Most of the work was done five years ago, and as it has gone through a chalky formation needing no lining, it has remained perfectly dry."

"In these days, when dynamos and electric motors are everywhere met with, and when they are continually finding new fields for work, the production of a watch that is utterly unaffected by the strongest magnets is an improvement well worthy of special notice. When dynamos were first introduced, many watches were magnetized, and processes and machines were invented for demagnetizing them. Finally, the custom of leaving the watch outside when entering a room containing dynamos was adopted, to prevent the annoyance due to this trouble. A cure for the whole affair, however, is found in the invention of a Mr. C. A. Paillard of Geneva, Switzerland. He has applied palladium to the manufacture of watches, using it for those parts that are usually constructed of steel."

"The discovery of electrical induction, and the invention based on this discovery, by means of which mechanical energy is converted into electrical energy, has made electric illumination and a host of other commercial electrical applications possible. Among these applications, one of the most recent and interesting is that of the electrical welding of metals, invented by Professor Elihu Thomson of Lynn, Mass."



Electric welding at the American Institute Fair in New York



THE AUTHORS

LEONARD A. COHEN ("Diet and Cancer") studies interactions between foods and cancer at the Naylor Dana Institute for Disease Prevention of the American Health Foundation. After graduation from the University of Wisconsin at Madison in 1960, he studied at Brooklyn College and the Sloan-Kettering Institute for Cancer Research and received a Ph.D. in cell biology from the City University of New York in 1972. The following year he went to the Naylor Dana Institute, where his research has been focused on the relation between dietary fat and breast cancer; in 1984 he was made head of the section of nutritional endocrinology. Cohen serves on the board of scientific advisers of *Nutrition Action Newsletter,* published by the Center for Science in the Public Interest.

J. CRAIG WHEELER and ROBERT P. HARKNESS ("Helium-rich Supernovas") work together in the department of astronomy at the University of Texas at Austin. Wheeler is the Samuel T. and Fern Yanagisawa Regents Professor of Astronomy and chairman of the department. He got his B.A. (1965) in physics at the Massachusetts Institute of Technology and his Ph.D. (1969), also in physics, from the University of Colorado. After stints at the California Institute of Technology as a research fellow and at Harvard University as an assistant professor he joined the faculty at Austin, where he became full professor in 1980. Wheeler is the author of *The Krone Experiment,* a novel that was recently published by Pressworks, Inc. Harkness earned a B.Sc. (1977) in astronomy and physics at University College London and a Ph.D. (1981) in astrophysics from the University of Oxford. As a postdoctoral fellow at Oxford he pioneered numerical techniques for calculating the transfer of radiation within supernovas. At Austin, Harkness is associated with both the astronomy department and the university's Center for High Performance Computing as a research scientist.

ANDREW W. MURRAY and JACK W. SZOSTAK ("Artificial Chromosomes") have been constructing artificial yeast chromosomes together since 1982. Murray, who had received his B.A. at the University of Cambridge in 1978, came to Szostak's laboratory at the Dana Farber Cancer Institute and the Harvard Medical School as a graduate student in 1980. He got his Ph.D. in 1984 and worked with Szostak as a postdoctoral fellow until 1986, when he went to the University of California at San Francisco as a postdoctoral student. Szostak received his B.Sc. at McGill University in 1972, earned a Ph. D. from Cornell University in 1977 and investigated recombination in yeast as a postdoctoral student there. In 1979 he went to the Dana Farber Cancer Institute and the Harvard Medical School. He moved to the Massachusetts General Hospital's department of genetics in 1984 and was made assistant professor at the medical school.

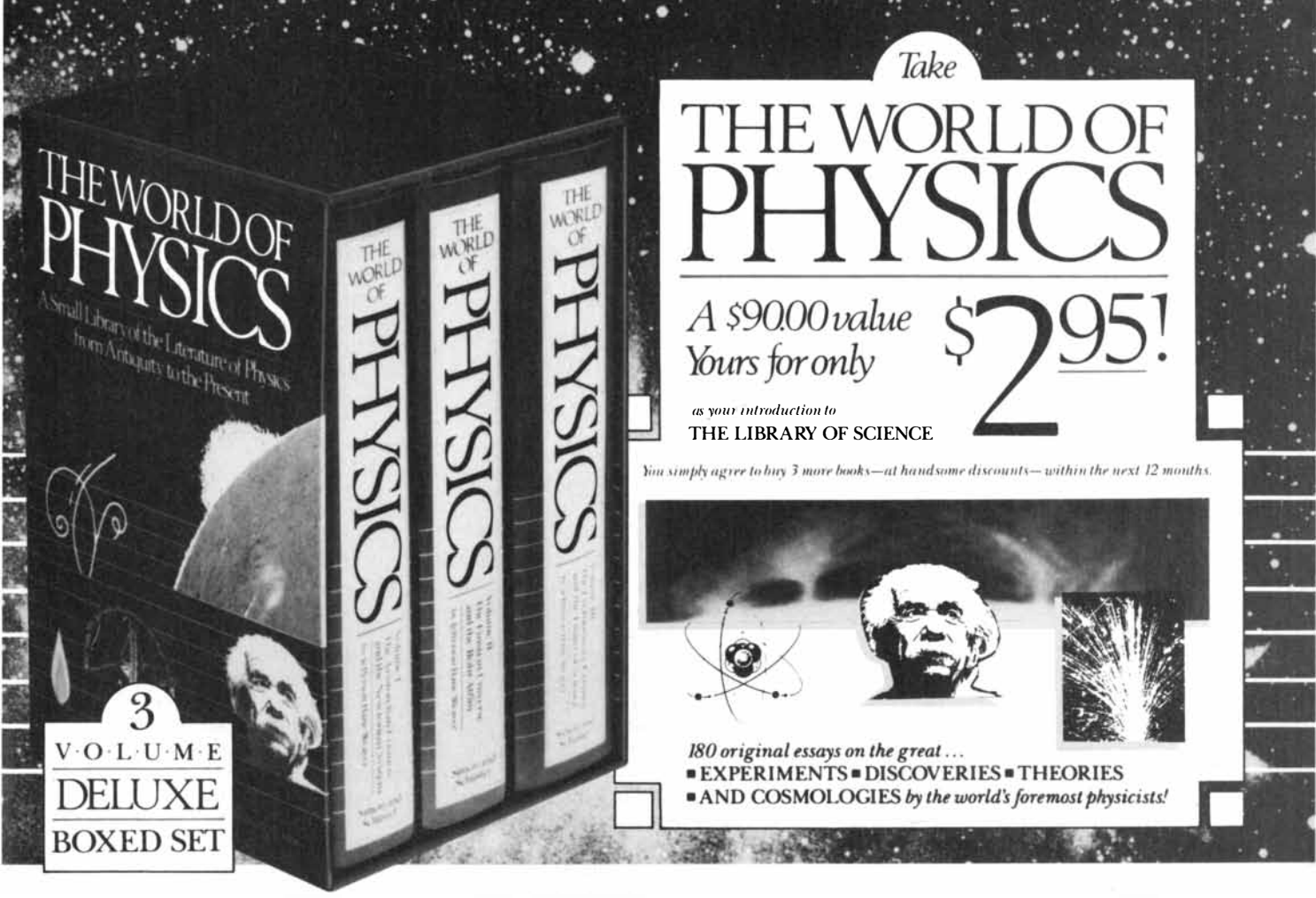
HERMAN WINICK ("Synchrotron Radiation") is professor in the applied-physics department at Stanford University and deputy director of the Stanford Synchrotron Radiation Laboratory. He did undergraduate and graduate work at Columbia University, where he obtained his doctorate in high-energy physics in 1957. He held positions first at the University of Rochester and then at Harvard University, where he worked as a staff scientist at the Cambridge Electron Accelerator. By 1973, when he left Harvard to go to Stanford, he was assistant director of the facility. The accelerator, he writes, "was also used (parasitically) as a synchrotron light source. I was appointed laboratory liaison with the synchrotron light users, to make sure they did not interfere with the high-energy-physics program. In the process I became very interested in the uses of synchrotron radiation." Winick recently won an award from the Department of Energy for his development of high-brilliance, high-flux beams.

STEPHEN J. O'BRIEN ("The Ancestry of the Giant Panda") has been investigating the subject of his article since he was asked to determine the paternity of a cub born to the celebrated Ling-Ling of Washington's National Zoological Park in 1983. He had specialized in mammalian somatic-cell genetics and molecular evolution at the National Cancer Institute after earning a Ph.D. from Cornell University in 1971. In 1973 he went to the NCI's Laboratory of Viral Carcinogenesis, of which he became chief in 1986. O'Brien is cochairman of the International Committee on Comparative Gene Mapping and editor of the journal *Genetic Maps,* and he was a coauthor of "The Cheetah in Genetic Peril" in the May 1986 issue.

CHARLES H. BENNETT ("Demons, Engines and the Second Law") is a member of the research staff at the IBM Corporation's Thomas J. Watson Research Center in Yorktown Heights, N.Y. After undergraduate education at Brandeis University he received a Ph.D. in chemical physics from Harvard University in 1970. He worked for two years at the Argonne National Laboratory and then went to IBM in 1972. There his continuing research interest has been in the mathematical theory of randomness and the relation between statistical mechanics and the theory of computation. Bennett was coauthor of "The Fundamental Physical Limits of Computation" in the July 1985 issue.

JOHN M. HOWELL ("Early Farming in Northwestern Europe") is a longtime student of the Neolithic period in Europe. He read prehistoric archaeology under Stuart Piggott at the University of Edinburgh before undertaking his doctoral research at the University of Oxford. Having done extensive fieldwork in western and eastern Europe, he has broadcast on the archaeology of those regions for the BBC and is the author of *Settlement and Economy in Neolithic Northern France.* Howell is honorary research fellow in prehistoric archaeology at the University of Liverpool, but "owing to the lack of university positions in the U.K." he also works as an international tax consultant with a London accounting firm.

DAVID A. GREENBERG ("Modeling Tidal Power") studies environmental issues associated with tidal power facilities as a research scientist with the Canadian Department of Fisheries and Oceans. He got his bachelor's and master's degrees at the University of Waterloo in Ontario in 1969 and 1970 and earned a Ph.D. in physical oceanography from the University of Liverpool in 1975 with a thesis on the subject of this article. That work has since been updated and expanded, first at the Department of Fisheries and Oceans in Ottawa and, since 1977, at the department's Bedford Institute of Oceanography in Dartmouth, Nova Scotia. Greenberg's recent interests include circulation in the Gulf of Maine and mechanisms generating the Labrador Current.



Take

THE WORLD OF PHYSICS

A $90.00 value
Yours for only **$2.95!**

as your introduction to
THE LIBRARY OF SCIENCE

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

180 original essays on the great...
▪ EXPERIMENTS ▪ DISCOVERIES ▪ THEORIES
▪ AND COSMOLOGIES *by the world's foremost physicists!*

Critical acclaim from today's leading physicists...

"THE WORLD OF PHYSICS is a well chosen selection of articles from 3,000 years of physics. It will be a valuable addition to the libraries of scientists and laymen alike."
—GERALD FEINBERG

"Puts, for the first time, a collection of simple and clear articles about the material world into the hands of the reader. I was edified and learned a lot."—HEINZ PAGELS

See the magnificent world of physics through the eyes of the greatest scientists. Discover revolutionary works that shattered conventional notions of space, time and matter by visionaries who mapped the course of scientific inquiry!

The breadth and authority of this 3,000 page physics library is simply unmatched. Together, the three volumes trace the evolution of major ideas in physics from classical Greece through Galileo, Kepler, and Newton to relativity and quanta and recent breakthroughs in high energy physics, GUTs, and creation itself.

Learn about radium from Madame Curie... the structure of atoms from Niels Bohr... quantum mechanics from Werner Heisenberg.

Read Albert Einstein's fascinating essay "How I Created the Theory of Relativity"... Erwin Schrödinger's seminal "What is Life?"... Sheldon Glashow's "Quarks with Color and Flavor"... Stephen Hawking on black holes... Hans Bethe on stars... Steven Weinberg on the first three minutes of the universe... and more.

This three-volume deluxe boxed set—a $90.00 value—*can be yours for only $2.95.*

So complete and return the postage-paid reply card today.

VOLUME 1
The Aristotelian Cosmos and the Newtonian System
Physics from its beginnings through the 19th century discoveries in energy, entropy and electromagnetism. Including—
▪ PLATO ▪ ARCHIMEDES ▪ PARACELSUS ▪ COPERNICUS ▪ ROBERT HOOKE ▪ HERMANN HELMHOLZ ▪ RENE DESCARTES ▪ JAMES CLERK MAXWELL

VOLUME 2
The Einstein Universe and the Bohr Atom
The golden age of physics and the landmark discoveries in relativity, quanta and high energy physics. Featuring—
▪ ERNEST RUTHERFORD ▪ ENRICO FERMI ▪ MAX BORN ▪ WOLFGANG PAULI ▪ PAUL DIRAC ▪ RICHARD FEYNMAN ▪ JAMES CHADWICK ▪ MURRAY GELL-MANN

VOLUME 3
The Evolutionary Cosmos and the Limits of Science
The latest hypotheses on the Big Bang, unity, symmetry, creation, and beyond the end of time in "superspace." Highlighting:
▪ CARL FRIEDRICH VON WEIZSACKER ▪ GEORGE GAMOW ▪ HERMANN BONDI ▪ ARTHUR S. EDDINGTON ▪ HARLOW SHAPLEY ▪ EDWARD R. HARRISON

MEMBERSHIP BENEFITS • In addition to getting The World of Physics for only $2.95 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 65% off the publishers' prices. • At 3–4 week intervals (16 times per year), you will receive the Library of Science News, describing the coming 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 is missing, please write to the Library of Science , Dept. 2-BT2-87667, Riverside, NJ, 08075, for membership information and an application.

SCIENCE AND THE CITIZEN

Ozone Watch

Additional data blame the "hole" on industrial chemicals

Atmospheric data from Antarctica continue to accumulate, shedding light on the peculiar stratospheric chemistry that is the probable cause each Antarctic spring of the ozone "hole," a temporary depletion of the ozone layer in the region (see "Science and the Citizen," August). No final proof is likely to be available for years that the hole is caused by chlorofluorocarbons (CFC's), manmade chemicals used as refrigerants and solvents. Data reported in early September in *Nature* seem, however, to have persuaded most investigators that it is easily the likeliest explanation. Preliminary results from new airborne measurements in Antarctica by the National Aeronautics and Space Administration suggest that unusual meteorology may be implicated along with the unusual chemistry.

The question is significant because a generalized depletion of the ozone layer would pose a significant threat: the layer protects human beings and other living organisms on the planet's surface from ultraviolet radiation. It has been established that ultraviolet radiation emanating from the sun can sharply increase the risk of skin cancer.

The ozone hole occurs inside the polar vortex, a large mass of air that remains relatively isolated over Antarctica during the long night of the polar winter and into the spring. The data reported in *Nature* by C. Bernard Farmer and his colleagues at the Jet Propulsion Laboratory show that while the ozone within the vortex is diminishing during the Antarctic springtime in September and October, the abundance of simple chlorine compounds such as hydrogen chloride and chlorine nitrate increases sharply.

Since laboratory experiments reveal that the chlorine from such compounds can destroy ozone, the JPL investigators say their results "are entirely consistent" with the notion that the atmospheric chlorine compounds, which come mainly from manufactured chlorofluorocarbons, cause the depletion. The latest NASA results confirm that at the same time there are local high concentrations of chlorine monoxide, which probably also destroys ozone.

The abundance of the chlorine compounds appears to be correlated with temperature. The JPL workers suggest that the compounds are trapped in some condensed form in polar stratospheric clouds during the Antarctic winter; as the sun rises during the spring and the temperature increases, they are released and do their damage. Hydrogen fluoride, whose only known source is manmade chemicals, is also present in the polar vortex in elevated amounts in the spring.

Joe C. Farman of the British Antarctic Survey, who first described the ozone hole, thinks that CFC's rise into the stratosphere in equatorial latitudes and that the chlorine compounds resulting from their breakdown are then sucked down into the polar vortex as it cools during the Antarctic winter. An alternative theory to explain the hole, now losing ground, is that it is caused by an upwelling of ozone-poor air from lower altitudes. Unequivocal confirmation of one or the other view will be difficult to achieve because many crucial measurements have to be made with light from the sun. New instruments are being developed that will avoid this limitation and so allow expanded measurements during the polar night. How to make measurements in moving air masses during the Antarctic winter, when travel on the continent is impossible and balloons cannot be recovered, remains an unsolved problem.

Many governments have apparently decided not to wait for definitive proof. In September countries that produce 80 percent of global CFC's signed a protocol to the Vienna Convention for the Protection of the Ozone Layer. The protocol obliges them to cut consumption of CFC's and related ozone-destroying compounds to half the 1986 levels by 1999. There will be some margin for "economic rationalization." Developing countries are given an extra 10-year grace period, and targets can be modified. —*Tim Beardsley*

Odyssey

Dig on Ithaca seeks the dwelling of Homer's famed wanderer

The *Odyssey* describes the journey of the Greek chieftain Odysseus from the fallen city of Troy back to his kingdom on Ithaca. For almost 200 years archaeologists have searched modern Ithaca—a 17-mile-long, figure-eight-shaped island to the west of Greece—for evidence that Odysseus was modeled after a king who ruled the island in the 13th century B.C., during the Mycenaean period. The last attempt to find his dwelling place, undertaken by a British team in the 1930's, ended as all others had—in failure. Three years ago Sarantis Symeonoglou of Washington University in St. Louis resumed the quest, which even he admits is somewhat quixotic.

"There is one faction of scholars who say that Homer's stories are myths and have no relation to reality," he remarks. "Others say there was a historical presence. It really is a matter of faith." Symeonoglou believes that Homer, the blind poet who supposedly composed the *Odyssey* in the 8th or 9th century B.C., visited Ithaca, and that his tale incorporates his impressions of the island and the historical accounts of the islanders. The archaeologist contends his "faith that Homer described something real, not just imaginary," is bolstered by a growing body of circumstantial evidence.

Symeonoglou is directing excavations on a mountain slope, dotted with thorn oaks and thistles, near the narrow middle of the island. In at least three respects, he maintains, the site matches Homer's descriptions of the city ruled by Odysseus: it has ample sources of underground

128 formulas at the speed of light.

The Casio FX-5000F Formula Calculator. To save you the time and trouble of looking up or memorizing many of the most important mathematical and scientific formulas, we put 128 of them in our FX-5000F Formula Calculator. And you can call them up in a flash.

The formulas are numbered and cover the fields of *math, statistics, physics, electronics* and *mechanics.* Plus you can input 12 of your own. Just key-in the appropriate number and the formula you need appears instantly. The alpha numeric display with 10 digit mantissa plus 2 digit exponent is easy to read and scrolls to over 70 characters. Its two-line display shows both the formula and the answer simultaneously.

And it doesn't stop there. Once you've recalled the formula, the calculator prompts you to input the values of the variables and automatically calculates the result.

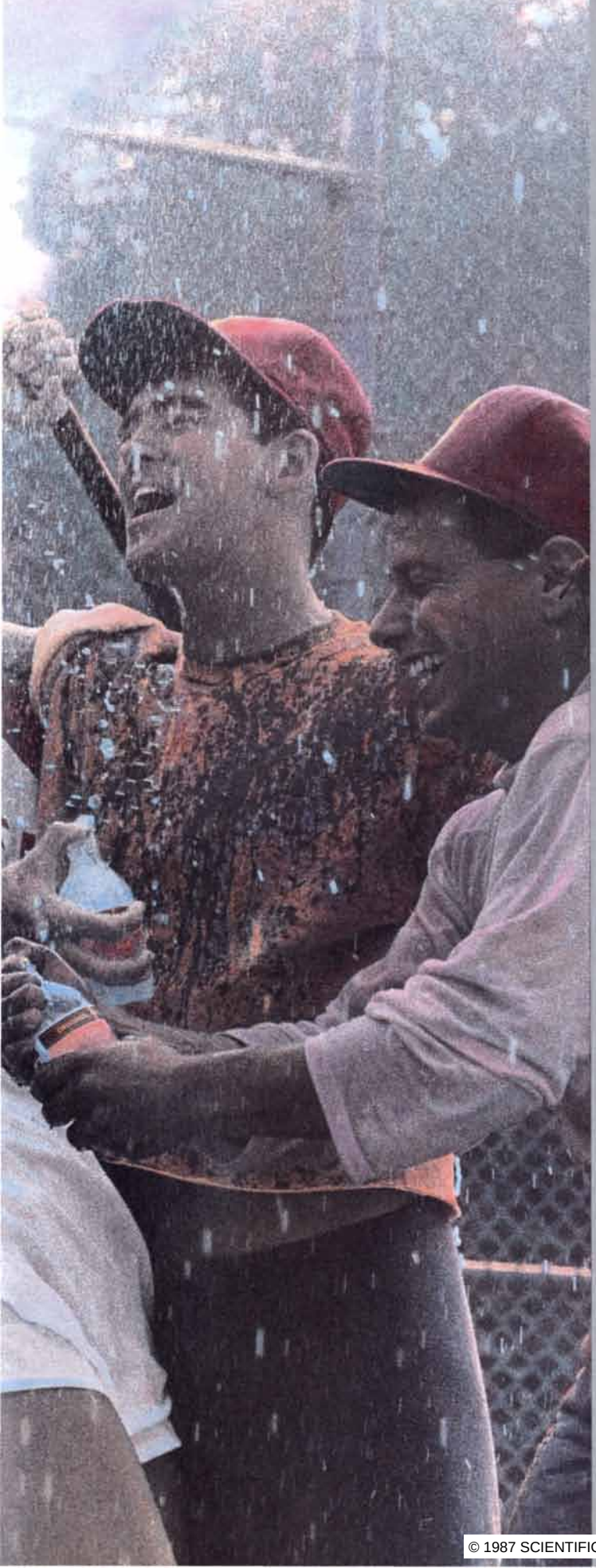
The FX-5000F's 675 steps of program memory allow you to create some pretty sizeable programs and store them until needed. While an instant replay feature lets you review and edit formulas at the touch of a button.

Adding to its usefulness are an additional 160 powerful functions, for a combined total of 288 functions and formulas.

To get the list of 128 formulas, as well as more information about how the Casio FX-5000F delivers them with the speed of light, call 1 800-626-2916 EXT. 89.

CASIO
Where miracles never cease

Casio, Inc. Consumer Products Division: 570 Mt. Pleasant A



"The only thing that used to bring this company together was corporate softball.

Even then we weren't exactly championship material. You put people in little cubicles, in separate buildings, using separate phone systems and you're bound to end up looking minor league. Something had to change. And it did. When we plugged in that communication system from AT&T. We began cutting costs on redundancies. And started working like a team of pros. Anticipating each other's moves, covering bases. As for our scorecard, would you believe a 9 and 1 season?

AT&T comes through."

The AT&T System 85

With complete networking transparency, AT&T's premier voice and data switch can help bridge the gap between scattered outposts – across the street or across the country.

A single System 85, with remote modules, can serve up to thirty locations in a nineteen square-mile ring. While multiple systems can be linked to better serve far-flung operations.

Full feature integration means dispersed offices can share central resources, including AT&T's AUDIX message-recording system. So the investment you make today, will be protected tomorrow.

To learn more, contact your AT&T Account Executive. Or call 1 800 247-1212. And ask about financing through the AT&T Credit Corporation.

From equipment to networking, from computers to communications, AT&T is the right choice.

AT&T
The right choice.

water, it commands a view of the entrance to the island's main harbor and it once held an outdoor shrine dedicated to the god Apollo.

Evidence for the shrine, Symeonoglou says, derives from several sources. In the 1930's the British workers, after digging up religious artifacts dating from between 900 and 600 B.C., proposed that a shrine once stood on the site but did not ascribe it to any particular deity. In 1979 Greek investigators combing through a museum in Athens came on more telltale objects discovered at the site in the late 1800's by farmers: a bronze figurine of Apollo, a miniature bronze lyre (commonly associated with the god) and a fragment of pottery inscribed with his name. The final piece of evidence, Symeonoglou says, is a stone foundation dating from about 650 B.C. On the basis of his excavations, he thinks the foundation was once a temple of Apollo built to replace the outdoor shrine.

The oldest building uncovered in Symeonoglou's excavation dates from the Homeric period. Workers have found fragments of pottery, however, with the crude, hand-painted drawings characteristic of the Mycenaean period. They have also discovered potsherds with designs characteristic of the so-called protogeometric period (from 1050 to 900 B.C.), when artisans began using tools to draw precise geometric patterns, and the geometric period (from 900 to 700), when they perfected their methods. These findings suggest to Symeonoglou that the site was occupied continuously from the Mycenaean period through the Homeric period. "The tradition of Odysseus was not invented by a poet. It survived because of this continuity."

STEM from a Mycenaean cup more than 3,000 years old was found in a dig on Ithaca directed by Sarantis Symeonoglou.

Symeonoglou notes that the city he is searching for was probably a small one, consisting of several hundred people living on a few acres of land. His excavations, however, encompass only about 450 square yards; many more years of digging, he says, will be necessary to establish clearer links between the site and the *Odyssey*.

Peter M. Warren of the University of Bristol maintains that establishing a connection between any archaeological finding and Homeric mythology is extremely difficult, if not impossible, and that the real value of such excavations derives from what they reveal about ancient cultures. "The glamour lies in relating this kind of research to the story of Odysseus," he says, "but at the heart of the matter Homer has nothing to do with it." —*John Horgan*

PHYSICAL SCIENCES

Double-Beta Decay

Predicted decades ago, a rare radioactivity is finally seen

After 13 years of trial and error, nuclear physicists at the University of California at Irvine have finally observed the radioactive decay of an isotope of selenium. The wait was not long, considering that an atom of the isotope has only a 50-50 chance of decaying over 1.1×10^{20} years, a period some 10 billion times longer than the estimated age of the universe. The wait was also worthwhile. The physicists—Michael K. Moe, Alan A. Hahn and Steve R. Elliott—have established the longest atomic half-life ever measured in the laboratory; more important, they have made the first direct observations of an extremely rare form of radioactivity predicted half a century ago: double-beta decay.

Ordinary beta decay occurs when a neutron in the nucleus of an atom spontaneously changes into a proton, an electron (known as a beta particle) and a neutrino. The proton is retained in the nucleus; the electron and the neutrino are ejected. Since the atom gains a unit of positive charge, it moves up one slot in the periodic table. By this process carbon 14 (the number refers to the total number of protons and neutrons in an isotope's nucleus) gradually disintegrates into nitrogen 14 and tritium (hydrogen 3) changes into helium 3.

In ejecting an electron and a neutrino an atom that undergoes spontaneous beta decay always loses some of the energy binding its nucleus together. For this reason some elements cannot undergo beta decay, because the elements immediately following them in the periodic table have nuclei with higher binding energies. In 1935 theorists speculated that there might be a way around the restriction: an atom of such an element might occasionally decay two neutrons at a time, thereby moving up two slots in the periodic table to become an atom with a lower binding energy. Numerous detectors designed to record the phenomenon failed, however, mainly because they could not distinguish between particles emitted during the decay and spurious particles.

Moe began his search in 1974 by selecting selenium 82 as a target. Selenium, which has 34 protons, is followed in the periodic table by bromine, which has 35; since bromine has a higher nuclear binding energy, selenium 82 can decay only into krypton 82, which has 36 protons and a lower nuclear energy. Investigators at the Max Planck Institute for Nuclear Physics in Heidelberg had derived a tentative estimate of the half-life of selenium 82 from analyses of traces of krypton 82 in billion-year-old selenium ores. "We had a target half-life we could shoot for," Moe explains. Selenium 82 was also thought to have a high transition energy, so that electrons from its decay would be relatively easy to distinguish from background radiation.

The Max Planck experiments had shown that any one atom of selenium 82 has a vanishingly small chance of decaying in a given period. These odds could be overcome only by assembling many atoms of the rare isotope. Moe obtained 14 grams—about 10^{23} atoms—from the Oak Ridge National Laboratory, which had produced selenium 82 while processing materials for nuclear weapons. This sample was the largest ever produced in the U.S.

At first Moe put the selenium in a detector designed for other particle-physics experiments. The detector

Merrill Lynch offers all kinds of mutual funds to meet all sorts of goals. Plus someone to help you choose the best ones for you.

It's easy to find a mutual fund–today there are more funds than stocks on the New York Stock Exchange.

But, like stock, you've got to pick the right one. Narrow down the field to the steady performers. And then, go farther–search out the fund that is right for you. The fund that seeks to meet your investment goals: what you're hoping to achieve and the type of investment that makes you comfortable.

After that, the most important thing is to develop a relationship with someone who understands your goals and can help you reach them.

A professional like a Merrill Lynch Financial Consultant.

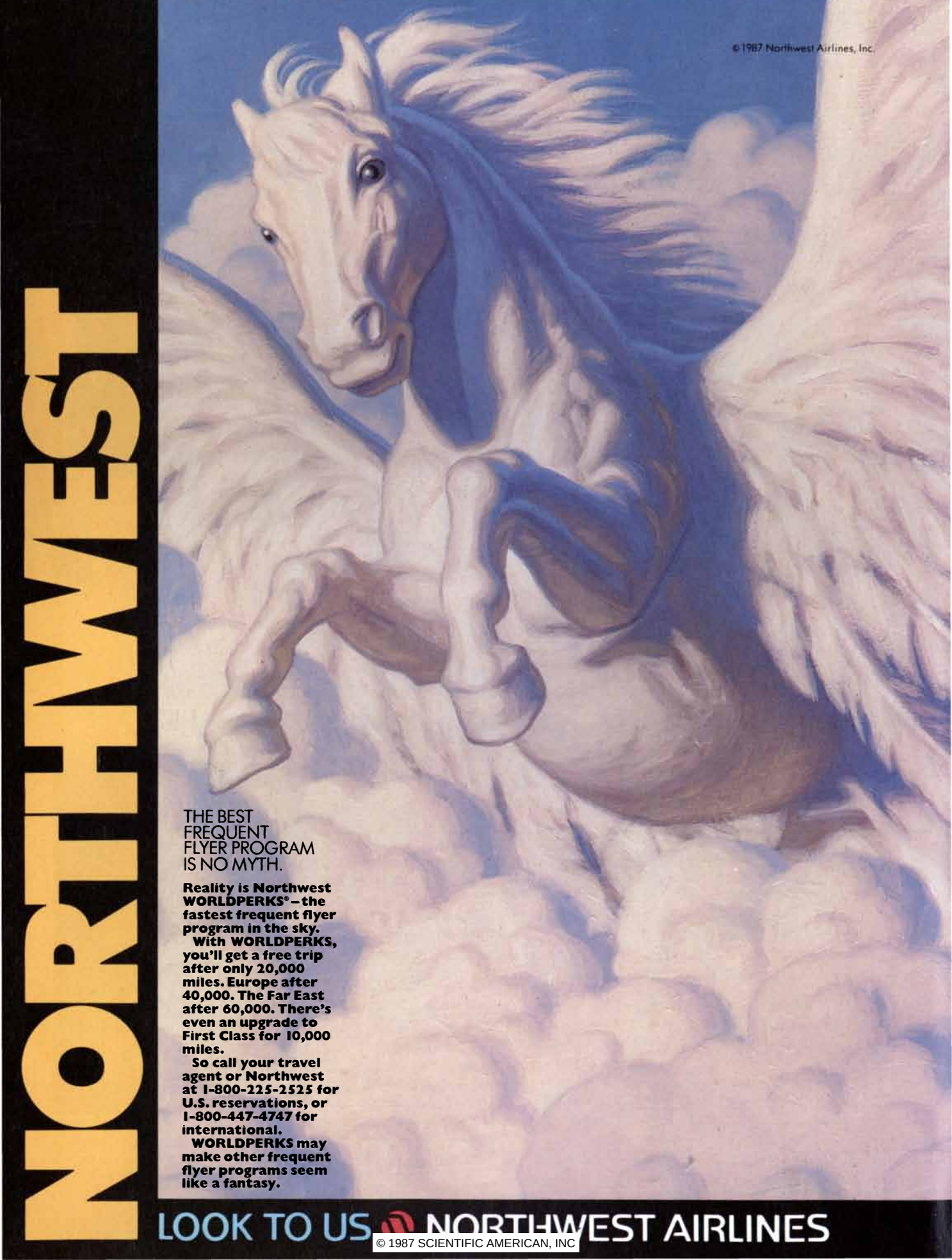
A Merrill Lynch Financial Consultant can help you analyze your situation and the degree of risk you're willing to assume. And then help you select from the many different mutual funds we have to offer.

Today Merrill Lynch Asset Management, Inc. manages over $80 billion in more than 50 different mutual funds.

And that's just the beginning of what we can offer you. Find out more with our *Smart Investor's Guide to Successful Money Management.* For your free copy, call toll-free 1-800-637-7455, Ext. 6279, or visit your local Merrill Lynch office.

Merrill Lynch Financial Consultants like Burt Haskins understand your changing financial needs, and have access to unmatched resources to help you reach your goals.

Merrill Lynch



kept recording spurious events, and after six years Moe abandoned it and teamed up with Hahn and Elliott to build a new detector called a time-projection chamber. It is filled with helium; when electrons traverse the gas, they leave a trail of ions, which induce pulses in wires lining the sides of the chamber. By analyzing signals from the wires the investigators determine the point of origin, the path and the energies of electrons passing through the chamber. The emission of the two neutrinos from each decay is inferred from the behavior of the electrons.

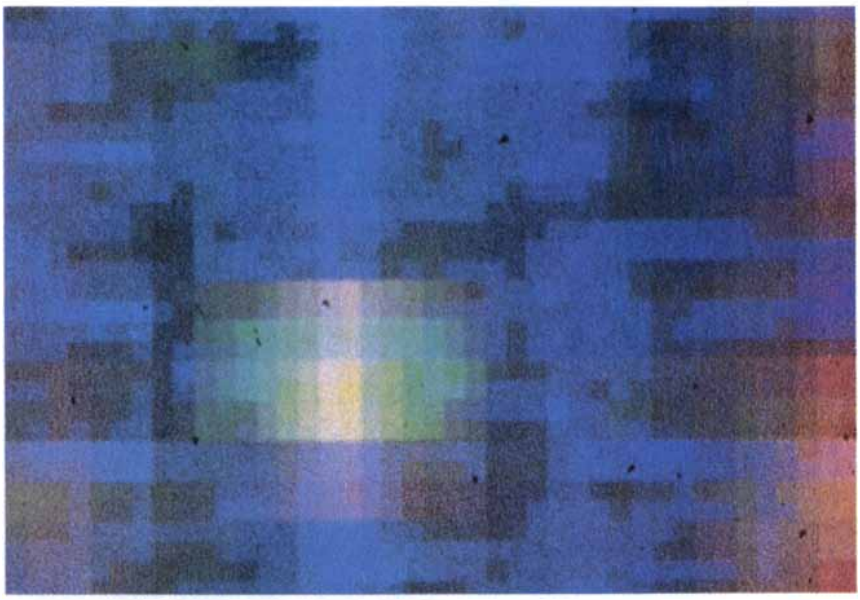
It took several years to reduce background noise and then, early in 1986, the workers began recording double-beta decay regularly. The events occur about once every three or four days: given this frequency, together with the total number of atoms in the selenium-82 sample, the workers derived their 1.1×10^{20} half-life value, which roughly matches the one obtained from the Max Planck group's analyses.

The precise, definitive nature of the findings gives nuclear theorists a benchmark against which their models can be tested, according to Hahn. In particular, he notes, the data may guide the many investigators (including those at Irvine) looking for an exotic form of double-beta decay that produces no neutrinos. Zero-neutrino double-beta decay could occur in at least two modes: either the neutrinos ejected from the decaying neutrons annihilate each other, creating a burst of energy that is imparted to the two electrons, or the neutrinos converge into a hypothetical particle known as a majoron.

Detection of either mode of zero-neutrino decay, Hahn says, would indicate that the neutrino has mass and so would have dramatic consequences for all physics. Theorists seeking to explain gravity, the electroweak force and the strong nuclear force with a so-called grand unified theory, for example, have proposed the existence of a neutrino that has mass. Neutrino mass also plays a key role in speculations about whether the universe has enough "dark matter" to halt its present expansion.

About a dozen experiments designed specifically for spotting zero-neutrino double-beta decay are under way around the world. Frank T. Avignone III of the University of South Carolina and Ronald L. Brodzinski of the Pacific Northwest Laboratory in Richland, Wash., are monitoring a crystal of germanium 76 (which should decay into selenium 76) isolated in a gold mine in South Dakota. Early this year they reported having found tentative evidence of the majoron form of double-beta decay. The report has been rebutted by other investigators, however, and Avignone acknowledges that "we need cleaner data." The Irvine team's results, he notes, will help him and other workers, both by revealing more about the nucleus of selenium and by demonstrating the process of double-beta decay, at least in its most conventional form. —*J.H.*

A tiny, frozen planet has yielded its secrets to a heat-sensitive spacecraft from the earth



INFRARED RADIATION from Pluto and its moon Charon stands out against the cooler background of space in this image made by the Infrared Astronomical Satellite *in 1983.*

Pluto's Poles

Infrared observations suggest they are capped by methane ice

In the past decade much of the mystery shrouding Pluto, the smallest and usually the most far-flung planet of the solar system (it occasionally swings inside Neptune's orbit), has dissipated. In 1976 workers found that methane ice covers at least part of its surface; in 1978 others spotted a moon, later named Charon, circling the planet, and in 1980 still others suggested that methane gas envelops the planet. Recently, because the plane in which the planet and its moon orbit each other has been edge on to the earth, they have been observed eclipsing each other frequently. By measuring the fluctuations in light during these eclipses investigators have estimated the diameter of Pluto and Charon—about 2,290 and 1,284 kilometers respectively—and their density. The density estimates suggest that both objects consist primarily of rock mixed with water ice.

Now Mark V. Sykes, Roc M. Cutri and Larry A. Lebofsky of the University of Arizona and Richard P. Binzel of the Planetary Science Institute in Tucson have brought Pluto into even sharper focus, providing new estimates of the degree to which methane ice and gas surround the planet. They derive their findings, which they report in *Science,* from images made four years ago by the *Infrared Astronomical Satellite* (*IRAS*). Sykes found the images last fall in the vast *IRAS* archives at the Jet Propulsion Laboratory in Pasadena, Calif.

The *IRAS* images do not show Pluto and Charon as distinct objects but only the infrared radiation, or heat, that they both radiate. Since other observations have indicated that Charon is too small to retain an atmosphere and that it has no methane ice on its surface, the workers were able to estimate and then subtract the moon's contribution to the infrared images and focus on modeling Pluto.

They calculated that if methane ice

Sometimes the best way

It's a very important lesson we learned early in our 65-year history.

Leadership in the worldwide insurance industry isn't won by imposing our culture on local operations. It's won by becoming part of the social, cultural and moral fabric of the countries in which we do business.

That idea is one reason for our current success

to stand out is to blend in.

in providing a broad range of commercial and personal insurance to businesses and individuals around the world. To learn more, contact AIG, Dept. A, 70 Pine St., N.Y., N.Y. 10270

Thinking more like our customers than other insurance companies is how we won our stripes.

Insurance Companies That Don't Think Like Insurance Companies. AIG

covered all of Pluto, the planet would be too cold to produce the *IRAS* images. By reducing the amount of ice on the planet, the workers obtained a model that matched both the observations of methane ice and the *IRAS* data. The model shows ice extending from the poles of the planet to a maximum latitude of 45 degrees; a band free of methane ice circles the planet's middle.

Based on this model, the investigators propose that Pluto has an atmosphere of methane gas, sublimed from the ice, about 1,000 times less dense than the earth's atmosphere at sea level. That figure is lower than other workers have recently proposed. "We're talking about a very small amount of gas," Sykes points out. The group estimates that the average temperature on the planet varies from about 59 degrees Kelvin (−214 degrees Celsius) at the equator to 54 degrees at the poles.

Sykes notes that ice caps have been evoked in the past to explain why Pluto has dimmed gradually over the past few decades. The planet appeared to be brighter previously, according to the hypothesis, because one ice-covered, highly reflective pole was pointed toward the earth; in recent years more of the darker equatorial band has been facing the earth. Another variable, Sykes adds, could complicate this hypothesis: since the axis of Pluto is tilted with respect to the sun, the methane ice caps probably wax and wane during its orbital cycle, which takes 248 earth years to complete. —J.H.

Goethe *v.* Newton

Shedding light on the poet's "refutation" of Newton's Opticks

The German poet Johann Wolfgang von Goethe (1749–1832) wrote a number of scientific treatises, including an analysis of light and color called *Zur Farbenlehre,* published in 1810. In it he argued that the model of light propounded in Isaac Newton's *Opticks* in 1704 was incorrect. Light cannot consist of differently colored constituents, Goethe contended, because it is essentially pure and immutable; it only appears to an observer to take on different colors when it interacts with various mediums. Among those mediums is darkness, which Goethe considered not the absence of light but an entity in its own right.

Ever since, Goethe has been ridiculed by scientists for having allowed mysticism to cloud his better judgment. Now, however, Michael Duck of the United Kingdom Atomic Energy Research Establishment at Harwell has reproduced one of the poet's key experiments. He maintains in *American Journal of Physics* that Goethe's model, while incorrect, describes how people actually perceive the spectrum more accurately than Newton's model does.

Whereas Newton refracted a beam of light through a prism onto a screen, the experiment reproduced by Duck entails holding a prism to the eye and looking at a white sheet against a black background. From close up the observer sees colored fringes emanating from two opposite edges of the sheet: one fringe is yellow, orange and red and the other is blue, indigo and violet. As the observer steps back from the paper, still holding the prism to an eye, the fringes broaden, and finally green appears in the middle of the sheet.

A Newtonian would say that the white sheet, when viewed from a short distance, actually reflects a number of spectra. Because the spectra closely overlie each other in the middle of the sheet, one sees only white light there, but at the edges of the sheet the outermost colors of the spectrum stand out and so one sees them as fringes. If one steps back far enough, however, the entire sheet comes into focus as a single spectrum. The spectrum should now appear unchanged as the observer moves back still farther.

Goethe disagreed. He maintained that looking through the prism causes the darkness and light to seem to overlap; as one steps back the colored fringes created by the overlapping broaden. Green finally appears in the middle not because it is inherent in the white light all along but because the inner borders of the two fringes, one yellow and the other blue, have merged and created green. Goethe believed his theory would be proved if, as one continues to retreat, the yellow and blue keep waning until they are entirely replaced by green.

"It will no doubt come as something of a shock to Newtonians," Duck writes, "that this is precisely what does happen; as the observer retreats, yellow and blue wane, and eventually do indeed disappear: Goethe would seem to be vindicated, and *Newton* proved wrong!"

Actually, Duck explains, Goethe had stumbled on a subtle physiological effect named (after the two German scientists who first described it in the late 1800's) the Bezold-Brücke phenomenon. The observed hue of a colored light source, these scientists found, can vary according to its brightness. As the intensity of the source decreases, the eye's receptors become more sensitive to red, green and violet than to yellow and blue. The effect can be observed by simply examining the night sky through a prism. The moon, Venus and a few bright stars display all the colors of the spectrum. In fainter stars blue and yellow are not visible.

Duck says Newton may have observed the Bezold-Brücke phenomenon and chosen to ignore it since he could not explain the effect in a way that agreed with his basic theory. Newton is suspected of having occasionally modified his experimental data to make them match his hypotheses. He "proved" in one experiment, for example, that it is impossible to build an achromatic lens—one that does not distort the color of images. "Newton may not have been entirely honest," Duck remarks, "but then, who is?" —J.H.

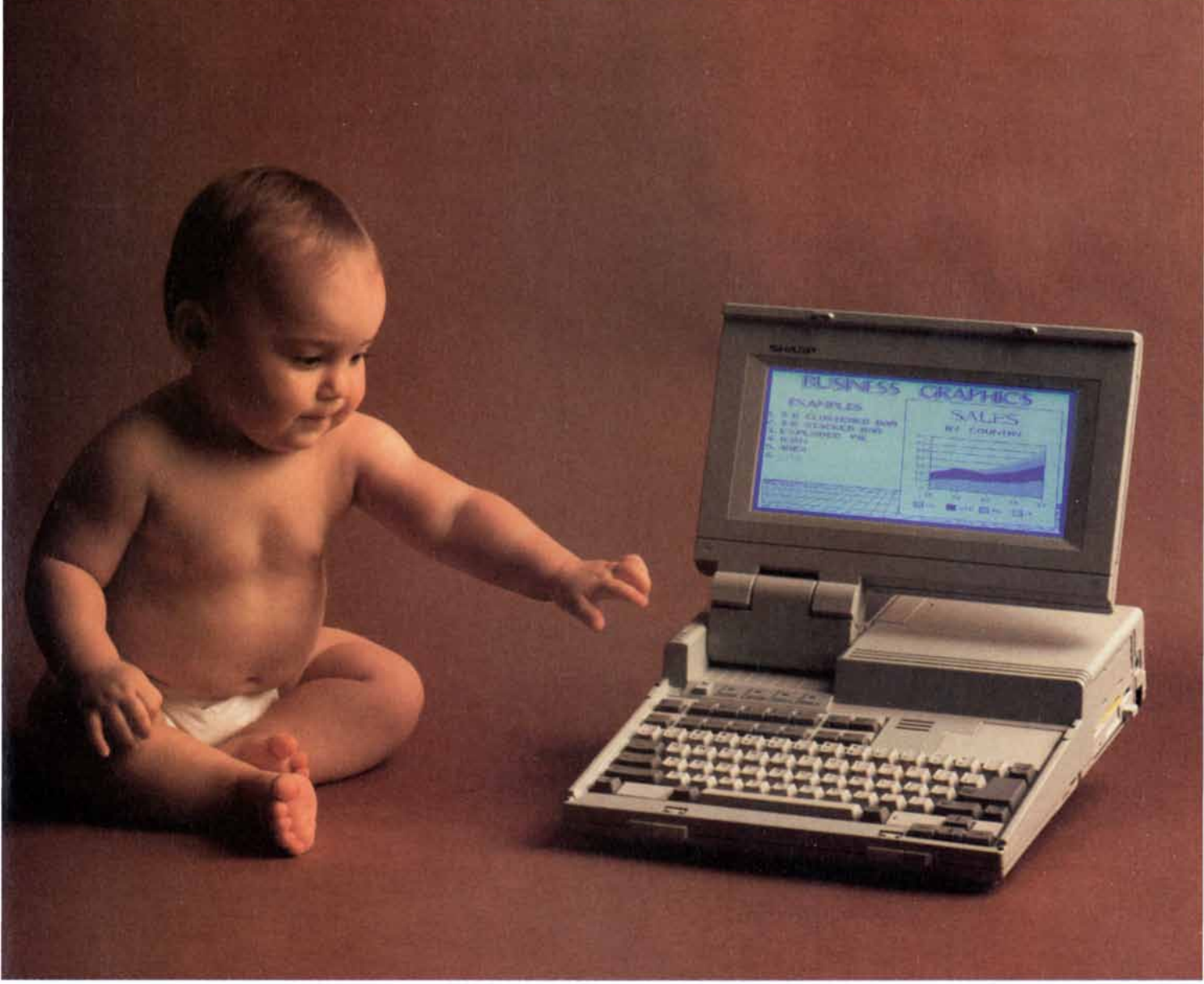
BIOLOGICAL SCIENCES

DNA Unveiled

Tunneling microscope offers a closer look at the stuff of life

The scanning tunneling microscope, invented in 1980, is the most powerful imaging tool now in existence. A resolution of less than one angstrom unit (one ten-billionth of a meter) enables the microscope to show the outlines of individual atoms, which are typically three angstroms or so in diameter. The microscope is already widely used by workers in microelectronics who are studying conducting, semiconducting and superconducting materials.

In the last few years investigators have begun to probe far more complex materials: those that constitute living things. Now a group of workers that includes one of the inventors of the scanning tunneling microscope has produced the first images



THE FIRST LAPTOP DESIGNED TO BE YOUR FIRST LAPTOP.

Sharp proudly announces the arrival of our laptop computer.

We've named it the PC-4501. You'll call it the ideal first laptop.

Because it packs the power of a full-size desk top in a convenient laptop bundle. It sports features like 256KB dynamic RAM. 3½″ 720KB floppy disk drive. Flip-up, easy to read Supertwist LCD display.

You also get a laptop that's IBM® compatible, expandable to 640KB RAM, and an optional second 3½″ 720KB floppy disk drive. And, since it can run on batteries, it's completely portable.

One more thing you'll enjoy: a price that will pamper your wallet.

For first hand information on the new PC-4501, see your Sharp dealer. Or call 1-800-BE-SHARP.

SHARP®

FROM SHARP MINDS COME SHARP PRODUCTS™

Create a new world where nature "thinks" like a farmer.

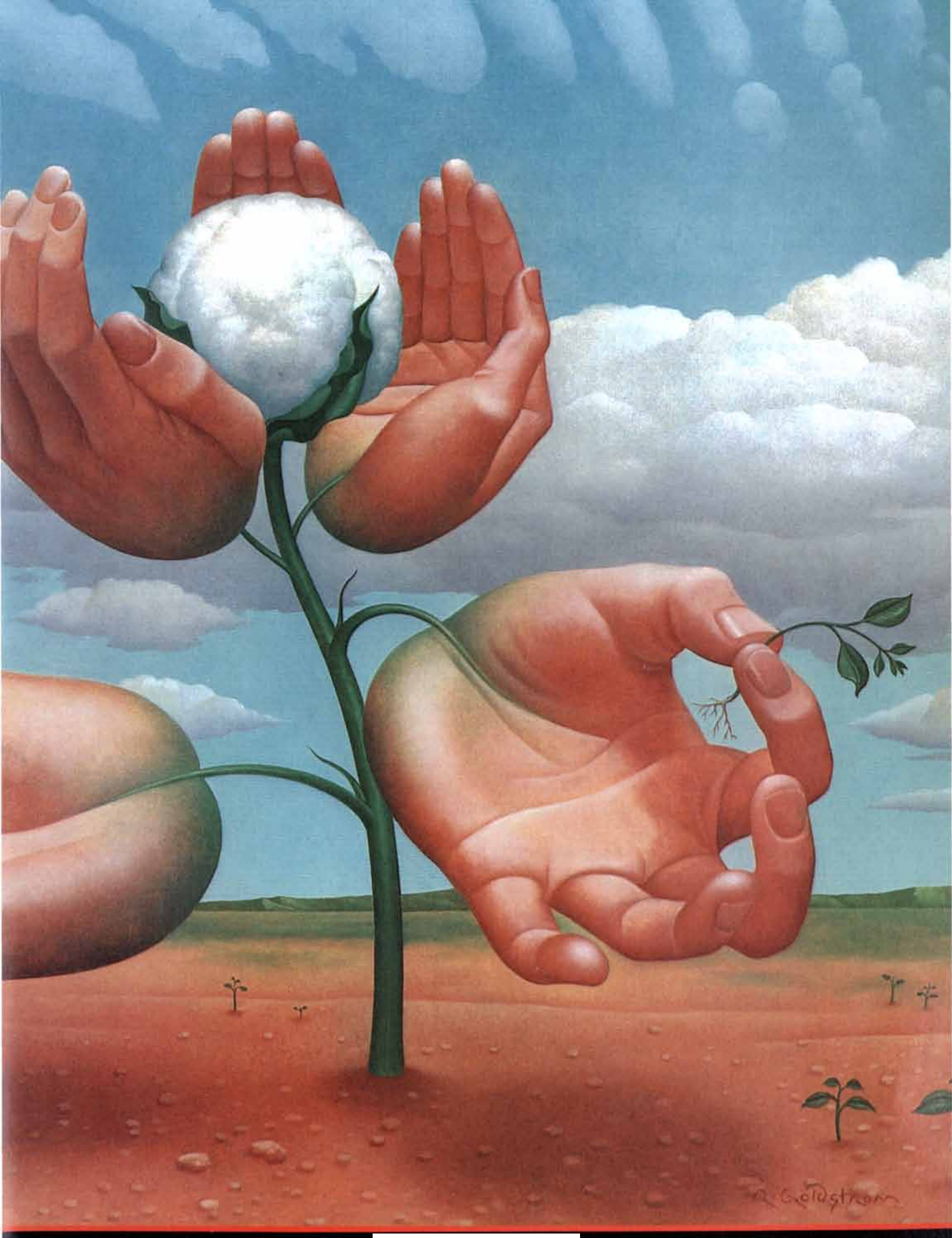
A quarter of North America's potential cotton crop is ruined each year by weeds, insects, disease.

At BASF, we looked at the problem in new ways and saw the cotton plant's environment as an ally. The result was a BASF breakthrough: Pix® Plant Regulator. Pix causes nature to rethink the plant's growth. Incorporating natural defenses, the plant and farmer work together toward a healthier, more bountiful harvest.

In one industry after another, from agribusiness to textiles, our broad-based technologies help us create new worlds by seeing in new ways.

The Spirit of Innovation

BASF



that depict directly the helical twists and turns of individual strands of DNA. The observations were made by Heinrich Rohrer, who shared the Nobel prize in physics in 1986 for his development of the scanning tunneling microscope, and Giorgio Travaglini, both of the IBM Corporation's Zurich Research Laboratory; they collaborated with Matthias Amrein and Heinz Gross, both biologists at the Swiss Federal Institute of Technology in Zurich.

Unlike either optical or electron microscopes, the scanning tunneling microscope does not form an image of a surface from reflected particles. Instead it traces contours with a needle that scans the clouds of electrons bound to atoms at the surface. The term tunneling refers to the process whereby the electrons, as a consequence of their wavelike properties, randomly jump high enough above the surface to be detected.

The Swiss workers prepared their specimens by allowing a drop of water containing DNA molecules to evaporate on a substrate made from a crystal of graphite. Graphite crystals were chosen because they are extremely flat; they usually have protuberances of no more than two or three atoms.

Apparently because DNA does not bind well to the graphite's smooth surface, the investigators found the strands only where they had become lodged in "steps": regions where the substrate rises abruptly by two or more atoms. This positioning limited somewhat the degree to which the microscope could scan the DNA strands. Nevertheless, the workers were able to measure the diameter and distance between successive spirals (two and 3.5 nanometers respectively) of individual strands.

Unlike the electron microscope, the scanning tunneling microscope can scan surfaces not only in a vacuum but also in air and liquid. This capability allowed the Swiss investigators to scan the DNA molecules under normal atmospheric conditions. They are now searching for a substrate that, unlike graphite, will bind closely to DNA in water and so will allow the molecules to be scanned in their natural environment.

An article by Rohrer and Travaglini describing this research appeared in *Spektrum der Wissenschaft*, the German-language edition of SCIENTIFIC AMERICAN, in August. —*J.H.*

Unadvertised Receptivity

Why did "concealed ovulation" evolve in the vervet monkey?

Most mammals mate only when the female is in estrus: a brief period around the time of ovulation during which the female advertises sexual receptivity by distinctive visual, behavioral or pheromonal signals. A few primates, including human beings, have developed a contrasting pattern, in which females do not show obvious external signals of ovulation and are sexually receptive even when they are not ovulating. Looking for clues to the evolutionary origins of concealed ovulation and its adaptive value for females, Sandy J. Andelman of the University of Minnesota examined the mating system of the cat-size African monkeys called vervets.

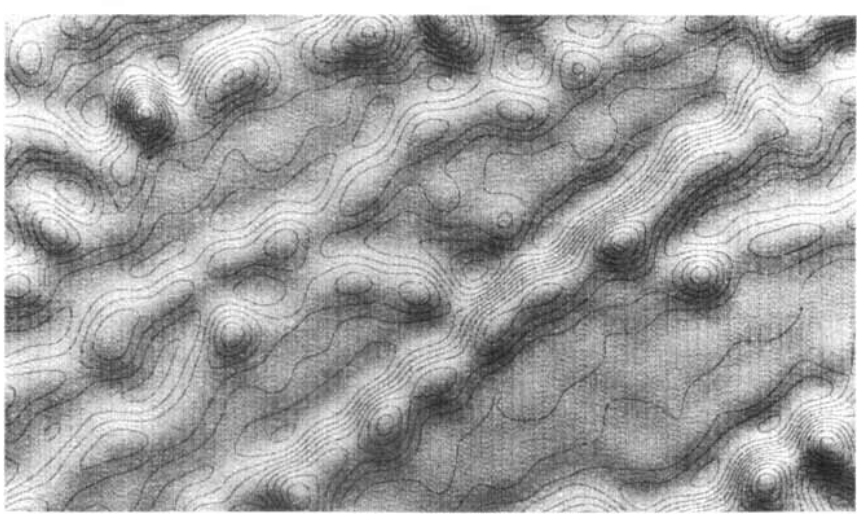
In the *American Naturalist* Andelman describes behavioral and physiological data she collected over six years from three groups of semiarboreal vervets in a national park in Kenya. Her team recorded attempted and successful copulations by all males in the groups. Because animals in game reserves must not be disturbed, the team had to monitor the females' ovarian function by collecting urine samples with a plastic funnel from monkeys perched in trees.

Reproduction among the vervets was highly seasonal, with 86 percent of all births occurring from October through December (the major rainy season on the savanna). Breeding behavior was also seasonal, but it was not limited to the period of ovulation: it began in April—about two months before the onset of normal ovarian function—and extended through October. The vervet females gave no signals of ovulation, and 60 percent of the males in a group mated with more than 80 percent of the group's females.

Such a mating pattern gives the female vervet a number of advantages, Andelman says. Primate males, particularly recent transfers from other groups, are known to kill infants they did not sire, and male transfer occurs at a high rate among vervets during the mating season. Because vervet females copulate with most of the males in the group and show no signs of estrus, however, a male cannot know when conception has taken place or which infants are his. Moreover, males are more likely to gather food for females with which they have copulated, and to defend them from predators. Because vervets fill an opportunistic niche in marginal savanna areas, the females are often undernourished and in particular need of such assistance.

Sarah Blaffer Hrdy of the University of California at Davis thinks it is "puzzling" to find hidden ovulation in a multimale social system. When a number of males are continuously present, the combination of estrus and ovulatory signals is advanta-

Magnified 3.8 million times, the twisted topology of DNA is bared by tunneling-electron microscopy



STRANDS OF DNA rise above a substrate of graphite in this scanning tunneling micrograph. The bumps testify to the periodicity of the spirals in the helical molecules.

The difference between a shattered windshield and a shattered life.

On November 17, 1986, the car in which Tara Meyer was riding was in an accident. Like 150,000 other victims each year, her head hit and shattered the windshield.

But unlike many of the others, Tara suffered no facial cuts or lacerations.

The difference?

A new idea developed by a group of Du Pont engineers. They believed that a thin layer of plastic along the inside of the windshield would keep the razor-sharp, shattered glass on the outside—away from the occupants.

Their idea helped make more than just a safer windshield. . . it made a difference.

And that's something we try to do in everything we do.

At Du Pont, we make the things that make a difference.

Better things for better living. DU PONT

REG US PAT & TM OFF

geous: it tends to compress the copulatory period and so reduce the energy expenditure associated with mating and increase the time available for foraging. Assuming that female primates prefer a wide range of partners, the situation is different for females in systems involving a single mate. For them a prolonged period of receptivity and hidden ovulation should indeed be advantageous, because they increase the likelihood of mating with a male from outside the group. —*Elizabeth Collins*

The behavior of the female vervet monkey offers clues to the evolution of human sexuality

VERVET MONKEYS socialize on the plains of Amboseli National Park in Kenya in this photograph made by Sandy J. Andelman of the University of Minnesota.

Hope for Blighted Bays

The reduction of nitrogen inflow may benefit coastal waters

The ecological problems plaguing estuaries and bays along the East Coast are well known: precipitous declines in fish catches, dying oyster beds, murky water and algal blooms. New insights into the causes are emerging from studies of Chesapeake Bay, the largest estuary and bay system in the continental U.S. The results are suggesting steps that could benefit other coastal waters that suffer from oxygen depletion caused by excessive plankton.

The root of the problem is well understood: the burgeoning population in the area and the growing use of fertilizers on lawns and croplands have increased inputs of nutrients, principally nitrogen and phosphorus. Sewage-treatment plants also contribute to the problem. The nutrients stimulate the growth of phytoplankton, which cloud the water. The remains of the plankton nourish bacteria in the bottom sediment that consume dissolved oxygen, and lack of oxygen kills aquatic plants and fish. The process is intensified by the pres-

ence of a bottom layer of saltier water that is largely isolated from the oxygen-richer surface layers.

Uncertainty about the precise sequence of ecological events has hampered efforts to find a solution. Until recently most studies on the effects of nutrients on plankton growth have involved freshwater systems. They showed phosphorus to be the crucial factor limiting the growth of algae, and therefore the logical target of control; other nutrients were thought to be more abundant in relation to the plankton's needs, so that larger and more costly reductions in their concentrations would be necessary to be effective. Maryland and Virginia have accordingly banned phosphorus in detergents, and sewage-treatment plants on the Potomac River incorporate expensive chemical processes that reduce phophorus levels in their effluent.

Christopher F. D'Elia of the University of Maryland and others have shown, by adding nutrients to samples of Chesapeake Bay water, that in estuaries and bays phosphorus is not always the limiting factor. Estuaries have complex circulation patterns and chemistry. Phosphorus is released from bottom sediments in large amounts during late summer and fall, when nitrogen input from agricultural runoff is at its lowest, and so nitrogen becomes the limiting factor at this time over much of the bay. The Scientific and Technical Advisory Committee (STAC) to the joint Federal/state Chesapeake Bay Program strongly recommended last year that wastewater treatment in the bay region should include controls on inputs of nitrogen.

New results from computer simulations of the bay's chemistry and flow patterns are confirming that assessment. A review of the modeling exercise conducted in August by STAC concluded that "there is no question that parts of the estuary are nitrogen-limited and parts are phosphorus-limited," according to STAC chairman Maurice P. Lynch. The Environmental Protection Agency which has been reluctant to recognize the importance of nitrogen, now appears to have acquiesced: Charles S. Spooner of the EPA notes that the agency "recognizes the advantages of nitrogen controls, at least in the southern part of the bay." Spooner says the recent exercise shows that adding controls on nitrogen to the existing phosphorus controls might increase dissolved oxygen in parts of the bay by 50 percent.

Sewage plants account for about a third of the bay's nitrogen input, and reducing the amount of nitrogen in their effluent requires a very expensive chemical process. Several plants in the bay's catchment area are nonetheless incorporating such chemical treatments: carbon is added (in the form of methanol) as food for sewage-digesting bacteria. The bacteria proliferate and convert nitrates into gaseous nitrogen, which escapes into the atmosphere.

The cost of control may come down. Innovative treatment processes tested in two pilot plants by Clifford W. Randall of the Virginia Polytechnic Institute and State University have shown that much nitrogen and phosphorus can be removed together. Randall's three-stage "biological" technique depends on carbon initially present in the sewage, rather than methanol, to feed the nitrate-converting bacteria.

Independent studies have shown that Randall's process, which is extremely fast, can reduce operating costs by up to 20 percent, and Randall says a plant to operate his process would cost only 3 percent more than a conventional plant. Citizens' groups are going to court to force local authorities to incorporate the

1988 CADILLAC ELDORADO
New look. New V8 power. Cadillac style.

The restyled, longer Eldorado is that rare contemporary automobile that still makes a very personal statement about the man or woman who drives it.

Unmistakably Eldorado.
From its bolder grille, through its blade fenders, to its hallmark vertical taillamps, this is unmistakably Eldorado. These distinctive elements are crafted onto a contemporary aerodynamic profile.

Powerful new 4.5 liter V8.
At Eldorado's heart is a larger, more powerful V8 that is even smoother than its predecessor. With 240 lbs.-ft. of torque—more torque than any other front-drive engine sold in the U.S.—response is lively, instantaneous. With 155 horsepower—nearly 20% more th[illegible] pass-ing are even more effortless.

Now available—ABS.
Eldorado's four-wheel disc brakes can now be fitted with a sophisticated anti-lock braking system. It can help take the panic out of panic stops by controlling the brakes up to 15 times per second—faster and more accurately than is humanly possible. This enhances driver control under

LET'S GET IT TOGETHER...BUCKLE UP.

most traction conditions.

Renowned Cadillac comfort and convenience. With six-way power seats for driver and front passenger, and numerous other amenities, this driver's coupe is, above all else, a Cadillac.

Backed by Eldorado's 6-year/60,000-mile warranty. Every new Eldorado comes with the long-term peace of mind of a 6-year/60,000-mile warranty.*

Your personal invitation. Discover the new Eldorado at your Cadillac dealer.

CADILLAC ELDORADO. THE ONLY WAY TO TRAVEL IS CADILLAC STYLE.

*See your dealer for terms and conditions of this limited powertrain warranty.

process in new sewage-treatment projects. Spooner says the EPA is persuaded that Randall's process is a good one, and a large new plant incorporating it will soon come into operation in Virginia.

One of the first consequences of the recognition of the role of nitrogen is a draft agreement between the governors of states adjoining the bay that is expected to be signed in December. It commits the signatories to reduce inputs of nitrogen and phosphorus to the bay by 40 percent by the year 2000, primarily by controlling sewage-plant outflows. —*T.M.B.*

MEDICINE

Stressed Out

Learned helplessness in rats sheds light on human depression

Investigators who work with animals are familiar with a psychological state known as learned helplessness. It is induced by unavoidable stress; its characteristics include poor appetite, loss of interest in surroundings, psychomotor dysfunction and inability to perform previously learned skills or tasks. In the 1970's Martin E. Seligman of the University of Pennsylvania noted that learned helplessness is much like depression in human beings and that drugs relieving depression also relieve symptoms of learned helplessness in animals. He proposed that the animal syndrome could serve as a model for depression in humans. Seligman's insight served as the foundation for much useful psychological study.

Now neurobiologists have begun to build on this work by relating the behavioral and psychological manifestations of learned helplessness to physiology: neurotransmitters, hormones and their receptors in the brain, the endocrine glands and the immune system. A typical experiment involves three groups of rats. One group experiences no stress. Members of the second group are subjected to stress, such as electric shock, but are allowed to escape through some self-initiated course of action. These rats remain healthy and alert. Animals in the third group are not allowed to escape the noxious stimuli. Many of them learn that their responses to stimuli have no effect: they learn helplessness. Changes in the concentration of various neurotransmitters in different parts of the brain can then be related to the animals' behavior.

Just what conclusions should be drawn from these results is a matter of considerable and productive debate. The neurotransmitter norepinephrine, for example, figures in a number of hypotheses, which differ on many points. Among them is the issue of whether changes in the levels of the hormone, which helps to mediate "fight or flight" reactions and other functions of the adrenergic component of the nervous system, are a primary or a secondary effect of learned helplessness.

Thomas G. Reigle of the University of Georgia sees a complicated interaction between norepinephrine and the endogenous opiate system, a system that produces neurochemicals resembling opiates that inhibit the ability to feel pain and induce a sense of well-being. Rats that have learned helplessness show relatively high levels of endogenous opiates and are relatively unresponsive to pain. If rats receive naloxone, a drug that blocks opiate activity, before exposure to stress, they are less likely to learn helplessness than untreated rats.

Changes in opiate levels, Reigle says, lead to changes in norepinephrine activity in the hippocampus, a brain region implicated in emotional response. Animals that have learned helplessness show decreased levels of metabolites of norepinephrine, indicating that the adrenergic system is working inefficiently. Jay M. Weiss of the Duke University Medical Center says there is a larger drop in norepinephrine (30 or 35 percent as opposed to some 10 percent in the hippocampus and cortex) in the locus ceruleus, which supplies the cortex and hippocampus with the transmitter. He has been able to affect the behavior of stressed animals by manipulating norepinephrine levels in the locus ceruleus.

Interactions between the adrenergic system and the endogenous opiate system could also explain how the immune system is impaired by learned helplessness. Tumors grow faster in animals subjected to uncontrollable stress, and *T* cells, a major immune-system component, respond less quickly to infection. Several investigators have postulated that the opiates bind to receptors in the immune system, causing a partial shutdown.

Workers caution that whatever is found out in animals is not ipso facto transferable to humans. "Learned helplessness is one route into depression," says Steven F. Maier of the University of Colorado at Boulder. "It is not synonymous: there is no one-to-one relationship." Learned helplessness, he points out, affects every system in the brain, bringing about a potent and widespread cascade of effects that are expressed in a range of behaviors.

It is also unclear precisely how changes in levels of neurotransmitters and neurotransmitter metabolites are related to specific patterns of behavior in the rats. It follows that the same ignorance extends even more completely over their role in human emotion and behavior. At this stage, Maier says, the broad spectrum of effects noted in learned-helplessness experiments means different theories may not necessarily be incompatible. —*E.C.*

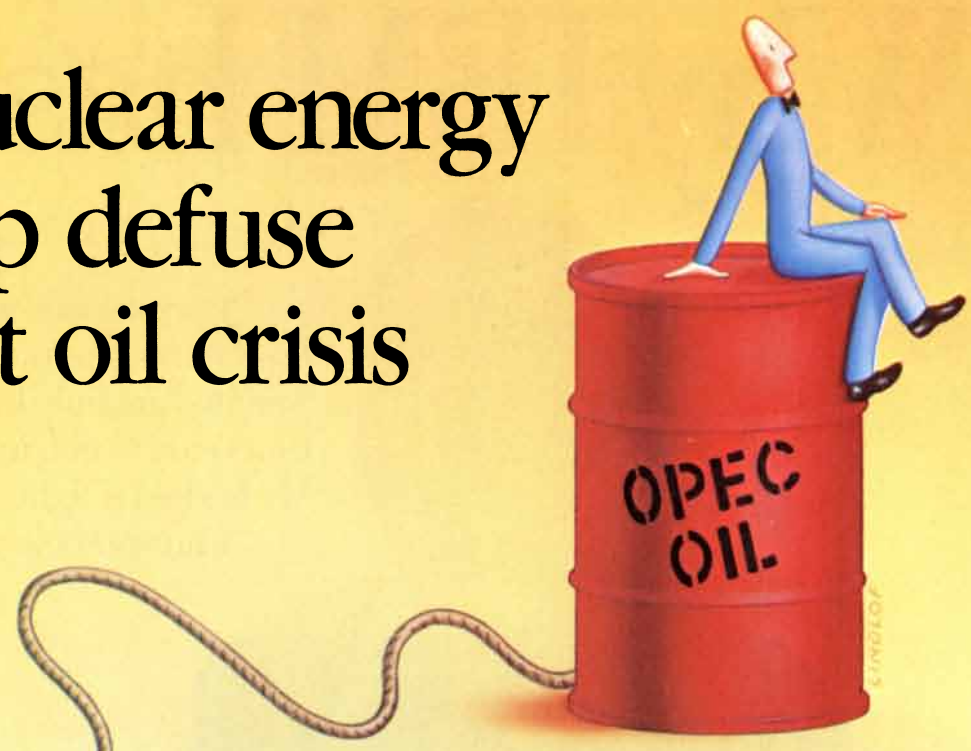
Stimulating Recovery

Wound healing is a promising field for biotechnology

Modern medicine has not yet devised any widely accepted treatment that actively promotes the healing of wounds. Rather, by closing wounds and keeping them moist and sterile physicians can only try to make it as easy as possible for nature to take its course. That may soon change: several companies are now exploiting recombinant-DNA technology to produce in large quantities substances that occur naturally in the body and have a potent stimulatory effect on cell migration and cell division, two processes central to wound healing.

Some of the substances, collectively called growth factors, have been known for many years, and more are being discovered. They are polypeptides, secreted by certain cells, that can bind to special receptors on the surfaces of other cells, or sometimes on the same cells. When a growth-factor molecule binds to a receptor, it sends a biochemical signal into the cell's interior that can (in ways that are far from being fully understood) stimulate or inhibit cell growth or movement.

How nuclear energy can help defuse the next oil crisis



Nuclear-generated electricity, still the fastest-growing major energy source in America, may be our best defense against another oil crisis.

More and more energy experts are asking the same question: How long before another oil shock torpedoes our economy and threatens our national security?

Oil turmoil

Signs of the next energy crisis:

- U.S. oil imports soared last year, costing the country $27 billion. This year, America's foreign-oil bill is expected to grow even bigger.
- Many oil analysts are saying that in three years or less, as much as 50% of all the oil used in the U.S. will have to be imported. That's a higher percentage than we have ever imported before, even during the oil crises of the 1970s.
- A whopping two-thirds of the world's oil lies under the sands of OPEC nations.

The need for nuclear

Nuclear electricity is a *domestically produced alternative to foreign oil.* Not just at the power plant, where nuclear energy is used instead of oil to generate electricity, but wherever Americans choose electricity (instead of oil) to heat their homes or run their factories.

The 1987 special report on U.S. energy security, ordered by the President and prepared by the U.S. Department of Energy, states that without electricity from nuclear energy, the United States "would be using more oil, paying more for each barrel of it, and feeling much less secure about its energy outlook."

The more we use our own nuclear electricity, the less we'll have to rely on energy from unstable regions of the world.

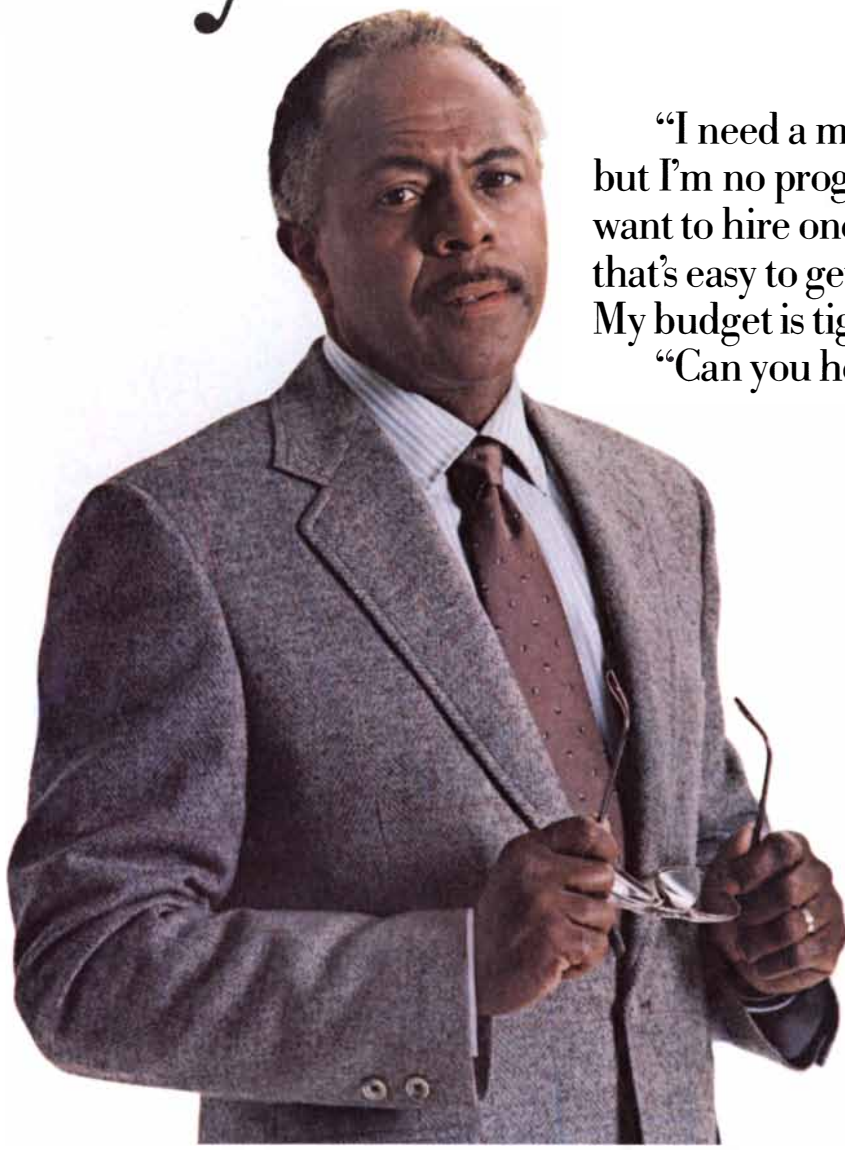
Nuclear energy for a secure future

With over a hundred operating plants in the U.S., nuclear energy is now our second leading source of electricity. But in spite of all that we have accomplished, the threat of foreign oil dependence remains. Difficult choices still need to be made, but one fact is clear: the more we develop our own energy sources, the more we can control our own destiny.

For a free booklet on energy independence, write to the U.S. Council for Energy Awareness, P.O. Box 1537 (FQ20), Ridgely, MD 21681. Please allow 4-6 weeks for delivery.

Information about energy America can count on
U.S. COUNCIL FOR ENERGY AWARENESS

Why IBM has a full



"I need a mid-size computer, but I'm no programmer and I don't want to hire one. I want software that's easy to get, install and learn. My budget is tight.
"Can you help me?"

Like the north and south poles, people who buy mid-range computers can be alike, yet opposite.

Their needs may be worlds apart, but each asks, "Can you help me?" And IBM answers, "Yes."

The fact is, IBM can help more kinds of customers, more ways, than anybody. We offer an extraordinary range of computing power, and all of it is available to any of our mid-range customers.

But there's one thing we don't have, and neither does anybody else: a "one-size-fits-all" design that can satisfy everyone without compromise. The needs of a small law firm, a retail chain and a multi-national corporation, for example, are too diverse for one architecture to serve each equally well.

So IBM offers a choice of mid-range systems, System/3X and System/370. Neither is "better" than the other, they're just better suited for different kinds of jobs. You may even want both.

And since we give you a choice, we'll help you make the right one. Together, we'll analyze your needs for software, connectivity and training. We'll look ahead to future growth, and we'll work to protect your previous investments.

We'll fit our system to your needs, not vice versa.

Power to start with, and grow with.

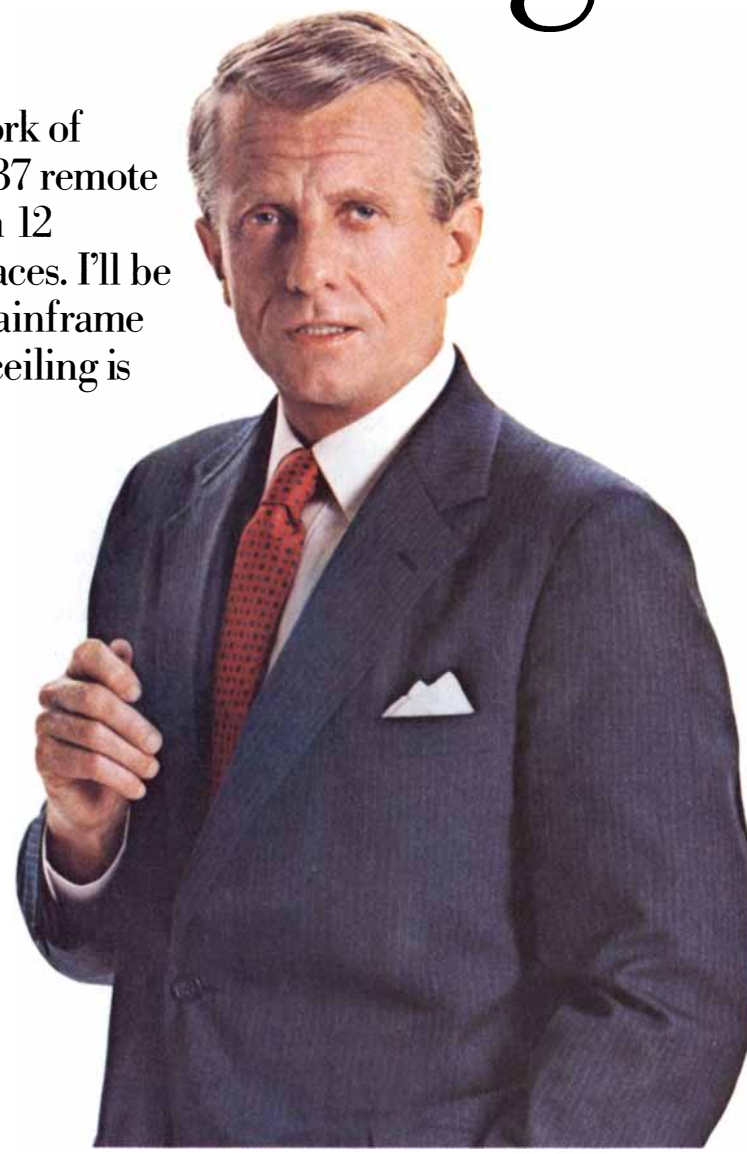
The System/3X family includes the System/36 and System/38.

The IBM System/36 is our most affordable and easiest-to-learn system. With over 4,000 programs available, it's ideal as a small company's first computer, with plenty of room for growth. It's also

-range mid-range.

"I'm building a network of mid-range computers in 37 remote locations connecting with 12 mainframes in 5 other places. I'll be sending both data and mainframe applications. My growth ceiling is pretty much unlimited.

"Can you help me?"



extremely connectible. So a larger business can build a network of System/36s, or include them in existing networks with PCs and mainframes.

The IBM System/38 makes it easy to develop applications, is renowned for its versatile data base, and connects with PCs, hosts and System/36.

Mainframe architecture for smaller neighborhoods.

Our System/370 architecture brings large system computing to mid-size environments with the IBM 9370 Information System and the newly updated IBM 4381 Processor family.

The 9370s are surprisingly compact, yet they not only connect with IBM's biggest mainframes and other systems (through a variety of networks), they can run thousands of mainframe programs.

For sending mainframe power out to departments, or for communicating between departments, 9370s are a terrific choice.

IBM's new 4381s bring similar benefits, plus additional ones: added performance, added growth, and the ability to run MVS/XA, the most powerful IBM mainframe operating system.

The IBM system of choice.

But regardless of architecture, the most important considerations for any system are the people who'll be using it and the challenges they face.

And that's why IBM's full-range mid-range is the right idea. No matter where you fit into it, the fit feels good.

IBM®

Most of the published research on the tissue-regenerating properties of growth factors consists of in vitro tests or work on the eye, where growth-factor treatment may be particularly valuable because the cornea lacks a blood supply. Investigators at the Washington University Medical Center in St. Louis have recently pushed the effort forward to an important new stage.

They write in *Science* that the factor called transforming growth factor-beta (TGF-B) is a "potent pharmacologic agent that can accelerate wound healing in rats." The workers found that when human TGF-B extracted from blood platelets was applied in a slow-release suspension to standard incisions in the skin of rats, it speeded healing by about three days; five days after wounding, incisions treated with TGF-B were more than twice as strong as those receiving no TGF-B. The compound hastened the influx of new cells and stimulated cell division and the deposition of collagen, a fibrous protein. One of the investigators, Thomas A. Mustoe, says it is reasonable to infer some similar effect in humans.

The work was carried out in collaboration with Amgen, Inc., a biotechnology company that is also investigating the tissue-regenerating effects of other growth factors in several animal models. Now Amgen has raised $40 million to spend on clinical testing of tissue-regeneration compounds. The clinical trials should start within a year, according to a spokesman. (Although it is true that aberrant regulation and processing of growth factors may underlie transformation of some cells to a cancerous state, there is no evidence suggesting that simply applying a growth factor would transform cells.)

If effective preparations incorporating growth factors can be developed—and much work remains—they could be particularly valuable in the treatment of severe burns, skin-graft donor sites and injuries to the cornea. Surgeons may in the future also use sutures impregnated with growth factors, an Amgen spokesman speculates.

The Chiron Corporation of Emeryville, Calif., has a product incorporating epidermal growth factor already in clinical trials; more than a third of eye-injury patients treated with the product show complete healing, according to Edward E. Penhoet, president of Chiron. Bedsores and diabetic ulcers, which are hard to treat, might also be attacked with synthetic growth factors. The stakes are high: Penhoet points out that the annual market for wound-treatment products can probably already be measured in the billions of dollars. —*T.M.B.*

One-Stop Medicine

More and more physicians sell prescription drugs

Who would turn down $50,000 in extra income if it could be earned with a modicum of extra effort and minimal capital investment? Physicians are being offered supplementary earnings in this range by companies that sell them drugs prepackaged for the patient. After writing a prescription, the physician offers the drug for sale directly, saving the patient a trip to the pharmacy and allowing the physician to charge a dispensing fee. About 5 percent of doctors are already routinely selling drugs, and the figure seems likely to increase.

The trend worries some academic clinicians and others concerned with medical ethics, who fear that the practice establishes a conflict of interest: it may be hard for some doctors to remain impartial in their prescribing if they stand to gain personally by extending treatment for an extra week or by choosing a drug they have in stock instead of another that might be slightly better.

A bill introduced in the House of Representatives by Congressman Ron Wyden, Democrat of Oregon, addresses such concerns by forbidding doctors to sell orally administered drugs at a profit unless there are emergency medical reasons or the patient would find it hard to get the drugs from a pharmacy. Doctors in rural areas would be exempted from the prohibition. "Medicine should not be like buying bananas in the supermarket," according to Wyden, who says that overprescribing does occur and that drugs have been sold by physicians with markups of more than 200 percent.

The pressure to earn a supplementary income by selling drugs is likely to worsen as the number of physicians continues to grow, increasing the competition for patients. Arnold S. Relman, editor of the *New England Journal of Medicine,* has argued in support of Wyden's bill that markets serve the interests of consumers only when they "are able to make their own purchasing decisions and can judge the value and quality of what they buy."

The companies that sell "repackaged" drugs are predictably opposed to Wyden's bill; pharmacists, who stand to lose out to the repackagers, support it. The American Medical Association is also lobbying against the bill. Ray W. Gifford, Jr., an AMA trustee, says the association believes medicine should continue to be regulated by states rather than by Federal statute. Although the association disapproves of doctors who sell drugs without special reason, it thinks Wyden's bill could unduly limit the individual physician's judgment. Wyden replies that Federal legislation is necessary because the Federal Trade Commission has reversed the position it took a few years ago and is urging states not to interfere with the practice. Jeffrey I. Zuckerman, director of the commission's Bureau of Competition, has written to legislators in Maryland and Georgia that dispensing by physicians "increases service and price competition among practitioners."

Immediate action on Wyden's bill is not expected; indeed, there is no companion bill in the Senate yet. The House bill has been reported out of committee with a comfortable majority, however, and Wyden says "the word's getting around." —*T.M.B.*

TECHNOLOGY

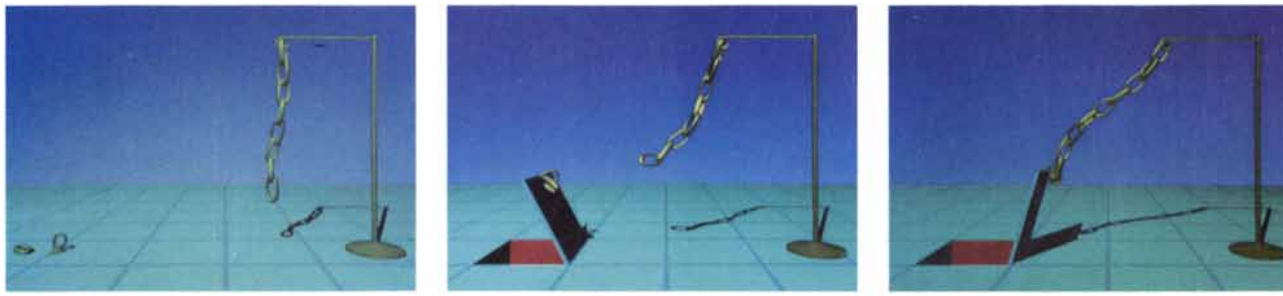
Faces, Couches, Cats...

Such "complex" objects stymie creators of computer imagery

No longer content with sending brightly colored, crisply shadowed spheres and cubes spinning through surreal Euclidean landscapes, computer-graphics researchers are trying to reproduce the subtle shadings and motions of reality. With varying success, they are simulating such mundane but complex phenomena as the wrinkling of fabric, the swinging of a chain, the glow of a lamp and the smile on a human face.

The pursuit of these goals is changing the way computer images are made. Whereas in the past workers would depict a specific phenomenon

Graphics based on real-world laws may be more lifelike, less predictable



CHAIN AND TRAPDOOR illustrate a new animation technique developed at the California Institute of Technology. The objects, once given a goal, automatically pursue it—here connnecting the chain's last link to the trapdoor—by following physics-based rules.

on a screen in the most efficient way possible, without regard to its real-world analogue, the new programs build on rules of physics like those that govern reality. Such programs require huge amounts of computer time and still cannot accurately depict many basic phenomena. But as hardware becomes faster and programmers become more sophisticated, the versatility of "reality-based" programs, Alan H. Barr of the California Institute of Technology says, should bring about "the end of the ad hoc approach to graphics."

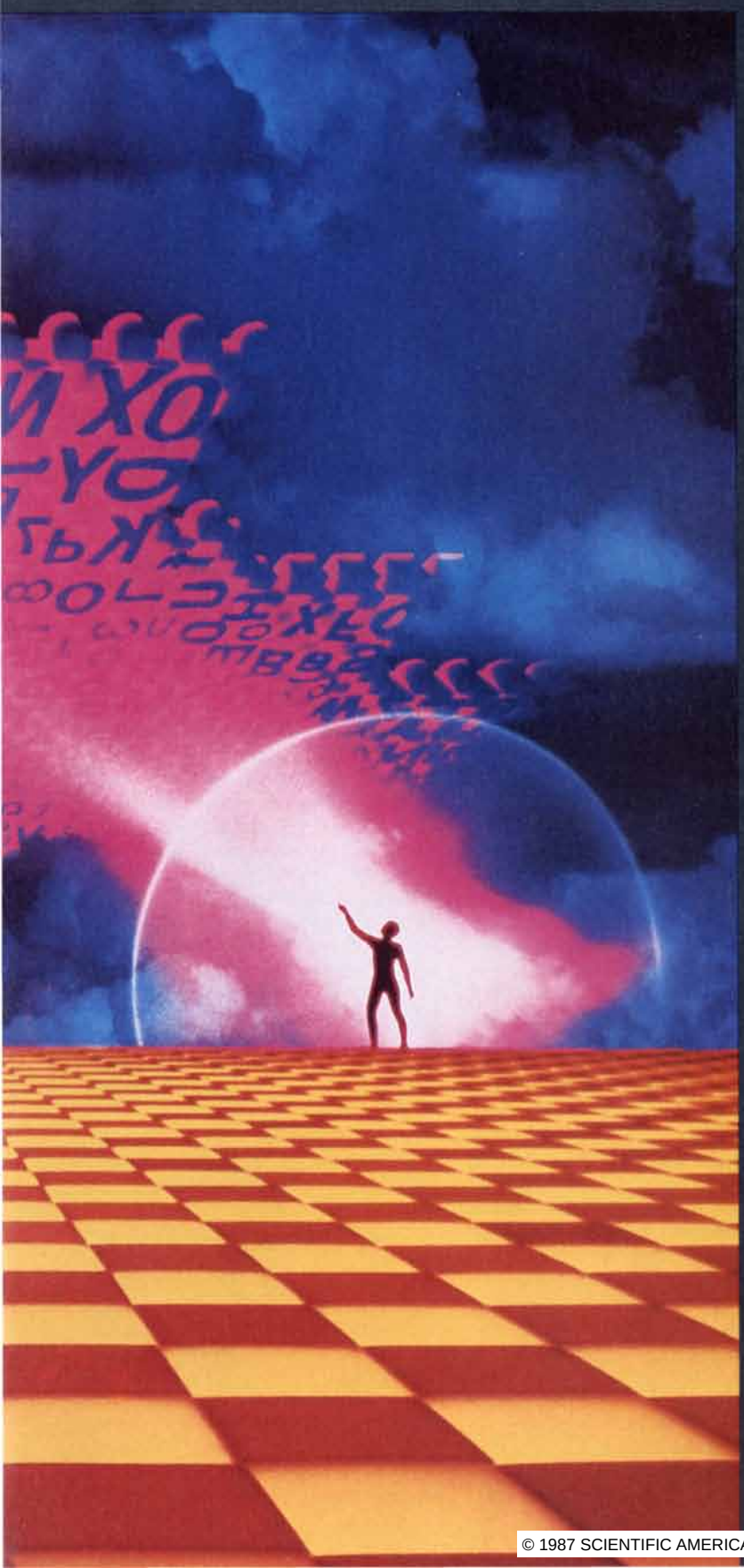
The effort to mimic reality goes forward on three often overlapping fronts: modeling, the depiction of three-dimensional objects and the space they inhabit; animation, the movement of these objects through space, and rendering, which captures the "look" of the objects and their environment on a two-dimensional screen.

The art of rendering leaped forward in the early 1980's with the development of a technique known as ray tracing, which represents the effects of individual rays of light reflecting off smooth surfaces and entering the eye of the observer. The major limitation of ray tracing is that it imparts a somewhat artificial, plastic sheen to objects. Recently workers at Cornell University, after extensive research into optics, created algorithms that combine ray tracing with a new technique called radiosity. The algorithms accurately depict how light is diffused by dull or roughly textured objects and by translucent mediums such as frosted glass, fog, smoke and fire.

James T. Kajiya of Caltech has de-

WE BELIEVE THAT THE GENIUS OF THE FUTURE LIES NOT IN TECHNOLOGY

ALONE, BUT IN THE ABILITY TO MANAGE IT.



Information.
Overpowering Force? Or Powerful Resource?

In today's fast-paced, high-tech business world information comes with no guarantees. It's either use it, or lose it. So how well you manage information can very well be the difference between business as usual and unusually good business. And that's where we come in. With the kind of end-to-end expertise most companies can't offer.

At Bell Atlantic, we can show you how to put information to work. How to gather it. Sort it. Organize it. How to get it to the right people. Apply it to the tough problems. And use it to gain the edge. With everything from local telephone to mobile communications, voice/data networks to software and computer maintenance, information systems to equipment leasing and financing, we can put it all together.

We're an integrated family of companies offering a broad range of solutions to your information and communications needs. And with our unmatched record of innovation, reliability and experience in moving information, we can put the full power of information within your grasp.

For more information about the Bell Atlantic companies, call toll-free 1-800-621-9900.

Bell Atlantic

WE MAKE TECHNOLOGY WORK FOR YOU.®

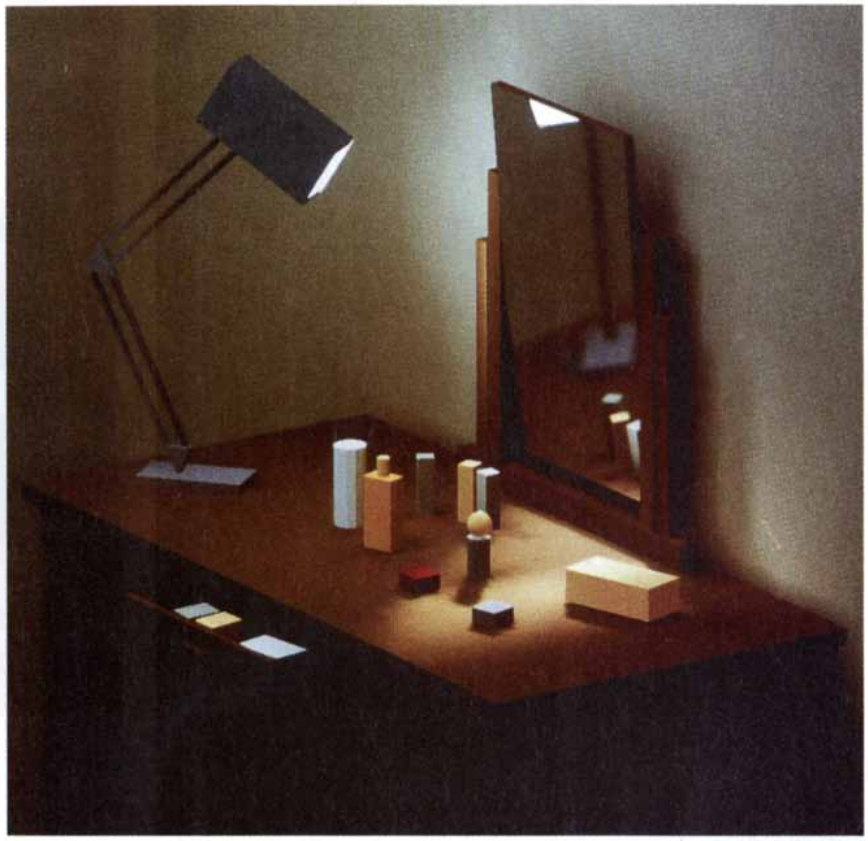
LOCAL TELEPHONE AND MOBILE COMMUNICATIONS
Bell of Pennsylvania • C&P Telephone • Diamond State Telephone • New Jersey Bell • Bell Atlantic Mobile Systems

INFORMATION AND COMMUNICATIONS SYSTEMS
Bell Atlanticom™ • Bell Atlantic Business Supplies • A Beeper Company • CompuShop® • Bell Atlantic Business Center • MAI Canada • Telecommunications Specialists, Inc.

COMPUTER MAINTENANCE AND SOFTWARE
Sorbus® • Technology Concepts, Inc. • Electronic Service Specialists, Ltd.

DIVERSIFIED FINANCIAL SERVICES
Bell Atlantic TriCon Leasing • Bell Atlantic Systems Leasing • Bell Atlantic Properties

Although computers paint more deftly than ever, they still cannot render a "dirty old couch"



DRESSING-TABLE SCENE rendered at Cornell University shows the interplay of the specular reflections from the mirror and the diffuse reflections from dull surfaces.

vised what he calls a "rendering equation" that achieves the same effects as both ray tracing and radiosity and also shows how light is refracted by a transparent medium onto another surface. This last effect—examples include the shimmering patterns that sunlight makes on the bottom of a swimming pool and the brilliant beam produced by a magnifying glass—had been extremely difficult to render.

Compared with rendering, computer animation is still relatively primitive: most computer-graphics programs still move an object by leading it along a prespecified path step by step. Barr and his colleagues at Caltech, however, have been able to develop a technique that gives objects a kind of autonomy. After creating an object, such as a chain, the workers establish a set of rules, based on Newtonian physics, that govern it. Then they specify a goal—swinging to a certain point, for example—and the object responds automatically. "I don't know in advance how the object will achieve the goal," Barr says.

Investigators can only render and animate what has already been realistically modeled, and the repertoire of modeling, Kajiya points out, is limited. The last major breakthrough in modeling occurred when workers found that fractals could simulate such natural phenomena as mountains and waves. There are many down-to-earth phenomena, Kajiya notes, that fractal-based algorithms cannot model. "You never see a computer image of an old beat-up couch, or a cat, or a person," he says. "These are things that have texture." Modeling faces is now a major endeavor of researchers, "but the models are very crude," Kajiya adds. The objects that most workers render and animate are rigid "geometric primitives" constructed from polygons.

At the AT&T Bell Laboratories, however, Gerald I. Weil is trying to model an almost infinitely malleable object: cloth. Like Barr, Weil has created programs based on "simple Newtonian physics." By calculating the forces exerted on individual threads in a piece of fabric, the program depicts how the fabric might deform when it falls through the air or is wrapped around a mannequin. Weil is also trying to model the appearance of hair, curly or combed.

Who wants this kind of realism? More specifically, who will be willing to pay for it? Perhaps the military, which puts a premium on realism in flight-training simulators, or Hollywood, which is always seeking flashier special effects. Donald P. Greenberg of Cornell maintains that the rise in speed and drop in cost of computer hardware will ultimately make the new programs available to designers "dealing with aesthetic evaluations, whether it's a car or a building." Weil offers a different view: he hopes the ability to simulate reality, once it is mastered, will eventually "lead back to surrealism and fantasy," in which real-world things behave in otherworldly ways. —*J.H.*

Aquatic Astronomy

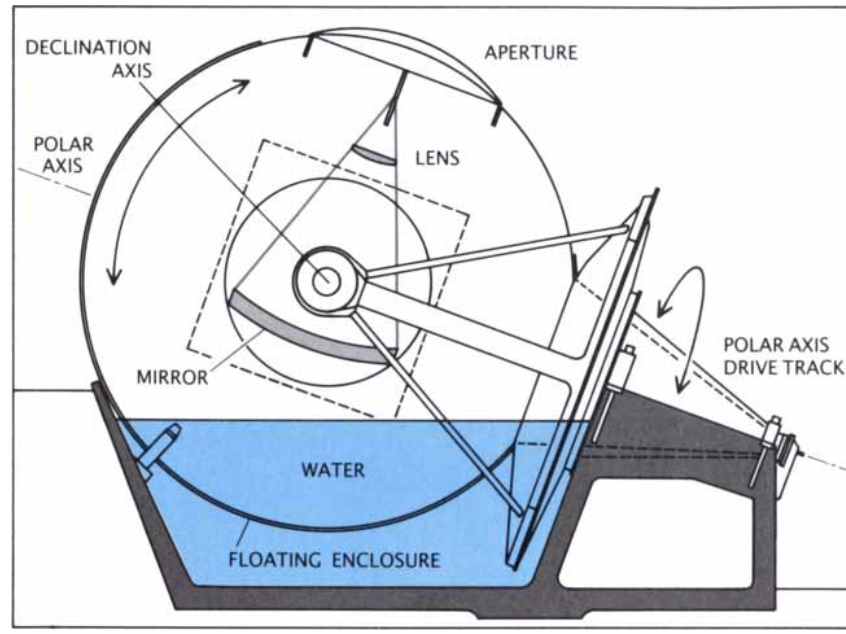
Floating telescope may solve some mounting problems

A large fraction of the investment in a major telescope pays for massive mounts that are needed to support the heavy primary mirror while it tracks an astronomical object with sub-arc-second precision. Conventional mounts give rise to problems: sometimes they flex, and their oil-pad bearings emit heat that produces troublesome convection currents in the air if the oil is not cooled. A radically new design proposed by Coast Steel Ltd. of Port Coquitlam, British Columbia, is said to be cheaper and to eliminate such problems. The company wants to build a telescope mounted in a steel sphere that floats in a tank of water.

Large telescopes have traditionally had "equatorial" mounts, in which one axis of rotation is parallel to the earth's polar axis. The design has one definite advantage: to track a celestial object the telescope can simply be rotated about the polar axis at a constant speed to compensate for the earth's rotation. Such a mounting is expensive, however. More recently built telescopes have "alt-azimuth" mounts. Like a gun turret, an alt-azimuth mount has one vertical and one horizontal axis. The bearings are cheaper, but in order to track an object the telescope must be driv-

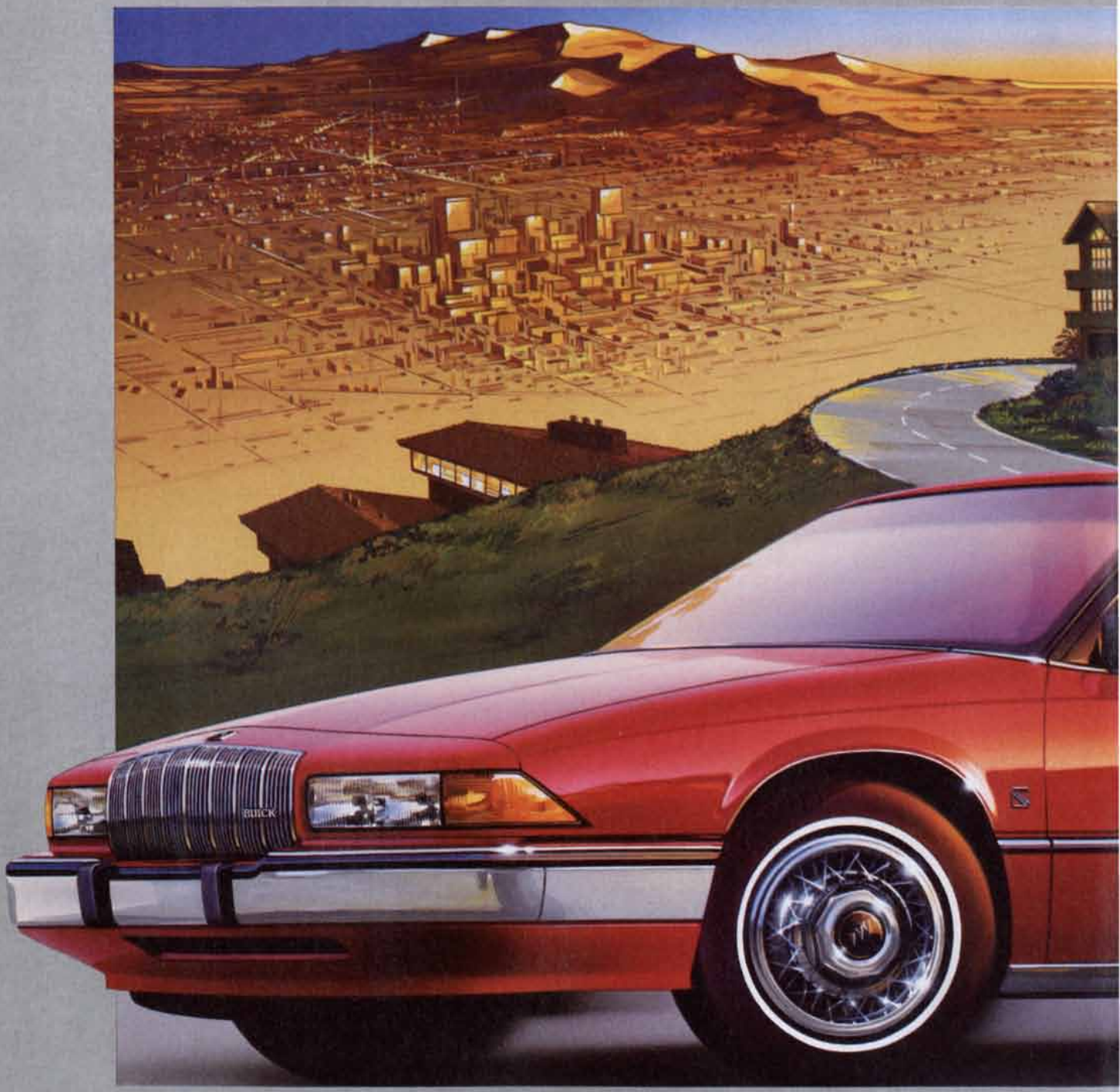
en at continuously varying speeds in the two planes. Furthermore, according to E. Harvey Richardson of the Dominion Astrophysical Observatory in British Columbia, a consultant for Coast Steel, the images produced by alt-azimuth telescopes rotate during tracking, so that instruments must also be rotated. Richardson thinks such mounts might have other disadvantages too: they expose the primary mirror to convection currents by placing it directly over bearings.

A Coast Steel floating boule telescope would be fixed within a steel cube that in turn fits into a steel sphere floating in a large tank. Most of the weight would be supported by the sphere's displacement of water, and so the equatorial mount that points the sphere and telescope together would not have to be very strong. David J. Halliday of Coast Steel points out that the engineering tolerances for building the sphere are less demanding than those for a conventional mount, so that costs would be lower. Computer modeling has shown the idea is feasible, Halliday says. An eight-meter mount is estimated to cost only $5.4 million, compared with $14 million for a conventional design. The company has some standing in the field: it is currently completing the dome of the 10-meter segmented-mirror Keck telescope on Mauna Kea in Hawaii. Halliday is expecting to receive government support for a small prototype boule and hopes to interest groups that are planning new telescopes. A particularly promising prospect is Saudi Arabia, which plans to spend $250 million on a multitelescope observatory. —*T.M.B.*



FLOATING TELESCOPE would be mounted in a spherical chamber cushioned by a tank of water. The design would reduce pressure on bearings on which the chamber rotates.

The new Buick Regal. There's nothing like it on the American road.



The 1988 Buick Regal. Not just a new car, a new Buick. One specifically designed for America's great roads — like Mulholland Drive, which winds through the Hollywood Hills to the San Fernando Valley in Southern California.

Beautiful to behold. Regal's styling is beautifully Buick. It's also exclusively Buick. There's nothing like it. On any road. Anywhere.

Sculpted in the wind tunnel, Regal's .305 coefficient of drag is the lowest in Buick's 85-year history. It is over 30% better than its predecessor, and is among the best in the world.

Beautiful to drive. Regal has a 2.8-litre V-6 engine with multi-port fuel injection . . . front-wheel drive . . . 4-wheel independent suspension . . . 4-wheel power disc brakes. And Regal is the first 6-passenger coupe in the world to have all these features — standard.

Beautiful to be in . . . and buy. Inside, Regal is also beautifully new and richly Buick, with room for six adults, and a generous 15.7-cubic-foot trunk. There is an impressive list of comfort, convenience and luxury features. Even an electronic digital speedometer is standard. And drivers who truly appreciate beauty and value will find Regal ownership pleasantly within reach.

Buckle up and see your Buick dealer for a test drive of the all-new 1988 Buick Regal today. There's nothing like it on the American road.

BUICK

GM

The Great American Road belongs to Buick

SCIENTIFIC AMERICAN November 1987 Volume 257 Number 5

Diet and Cancer

Recommendations aimed at reducing the incidence of cancers associated with nutrition are based on limited but suggestive evidence from epidemiological studies and animal experiments

by Leonard A. Cohen

In 1982 the National Research Council, an arm of the U.S. National Academy of Sciences, issued provisional guidelines intended to reduce the risk of cancer related to food. The main recommendations were to lower the intake of fat from the current U.S. average of 40 percent of total calories to 30 percent; to eat more fiber, fruits and vegetables; to increase the consumption of complex carbohydrates (such as the starch in flour and potatoes), and to decrease the consumption of pickled, salted or smoked foods as well as of simple carbohydrates (such as refined sugar).

These recommendations have generated a considerable amount of controversy and confusion among the public, dietary specialists, Government agencies, public health officials and the food industry as well. The controversy arises mainly from the fact that although the evidence on which the recommendations were based was highly suggestive, it was in no case definitive. The question is therefore whether it would be prudent to observe the recommendations or whether they should be regarded as premature and perhaps ill-founded.

One clear point is that the human diet has changed profoundly over a short period of recent evolutionary time. Anthropological studies of the diets of hunter-gatherer societies that have survived into the 20th century, such as the Bushmen of the Kalahari Desert in southern Africa, provide a valuable evolutionary perspective on modern dietary practices and their possibly detrimental effects. S. Boyd Eaton and Melvin J. Konner of Emory University, who analyzed data from several such societies, concluded that the fat intake of prehistoric people living in temperate climates was about 20 percent of total calories—half the current intake in the U.S.—with an appreciably higher ratio of unsaturated to saturated fat. Fiber intake appears to have been about 45 grams per day, compared with 15 grams or less in the U.S., and the intake of ascorbic acid was four times the U.S. level.

Assuming that modern human beings (*Homo sapiens sapiens*) evolved some 30,000 to 50,000 years ago, the species has subsisted for more than 90 percent of its history on just such low-fat, high-fiber diets rich in ascorbic acid and calcium. Hence modern people are served by metabolic and digestive systems that evolved to deal with dietary patterns very different from the current ones. The hunter-gatherer consumption pattern persisted (with minor changes when agriculture developed perhaps 10,000 years ago) until about 250 years ago, when the Industrial Revolution transformed the way many people live: fat intake rose steadily, fiber intake declined and consumption of refined sugars rose as consumption of complex carbohydrates fell.

In a sense, then, modern people are living in a biological time warp: their Stone Age physiology contends daily with a 20th-century diet to which it is poorly adapted. The changes in diet, together with a more sedentary way of life, are thought to be responsible not only for the increased size of modern people but also for the prevalence of obesity, early maturity and such chronic diseases as cardiovascular disorders and cancer (which were less common even in older individuals in Western societies of the 18th and 19th centuries and are uncommon today in hunter-gatherer societies).

The primary reason for the uncertainty about the dietary recommendations of the National Research Council lies in the complex nature of the processes that give rise to cancer. Historically two approaches have dominated investigations of the causes of cancer. One, which focuses on specific agents in the environment, derives from the suggestion by Sir Percival Pott in 1775 that the cancers of the scrotum and nasal cavity that were commonly found among chimney sweeps in England resulted from excessive exposure to soot. Years later laboratory experiments showed that certain polycyclic hydrocarbons formed during combustion are indeed carcinogenic in animals. Still later other chemicals, certain viruses and X rays were also found to cause cancer in animals.

The second approach, which examines a broader range of less specific factors, originated in the work of the Italian physician Bernardino Ramazzini. In 1700 he noted that breast cancer was more prevalent among nuns than among women who led conventional lives. Ramazzini suggested that the contrast might be due to the different experience of the two

populations with regard to pregnancy and lactation. Since characteristics such as reproductive life encompass a wide range of physiological events, rather than being carcinogens in the conventional sense, they are called risk factors.

These two approaches were subsumed within a single conceptual framework in the 1940's with the demonstration, by Isaac Berenblum of the Weizmann Institute of Science and his colleagues and by Roswell K. Boutwell of the University of Wisconsin at Madison, that the carcinogenic process has at least two distinct stages: initiation and promotion. It is now widely believed that cancer develops in discrete stages, each of which is regulated independently at different times by different agents. Initiation, as it is currently understood, involves a brief and irreversible interaction between a carcinogen and the genetic material of its target tissue. The reaction results in a molecular lesion, or mutation, which may transform some cells to an abnormal state but does not generate a clinically observable tumor unless it is acted on by another class of agents, called promoters.

A promoter can cause transformed cells to proliferate and form a tumor, but in itself it is neither mutagenic nor carcinogenic. Moreover, it must be applied continuously to have a biological effect; if the promoting stimulus is removed, its effects are reversible. (That is not the case for a carcinogen, and the difference is important when it comes to considering preventive strategies.) The time between an initiation event and the appearance of a recognizable tumor is called the latent period. In human beings it ranges from 10 to 20 years.

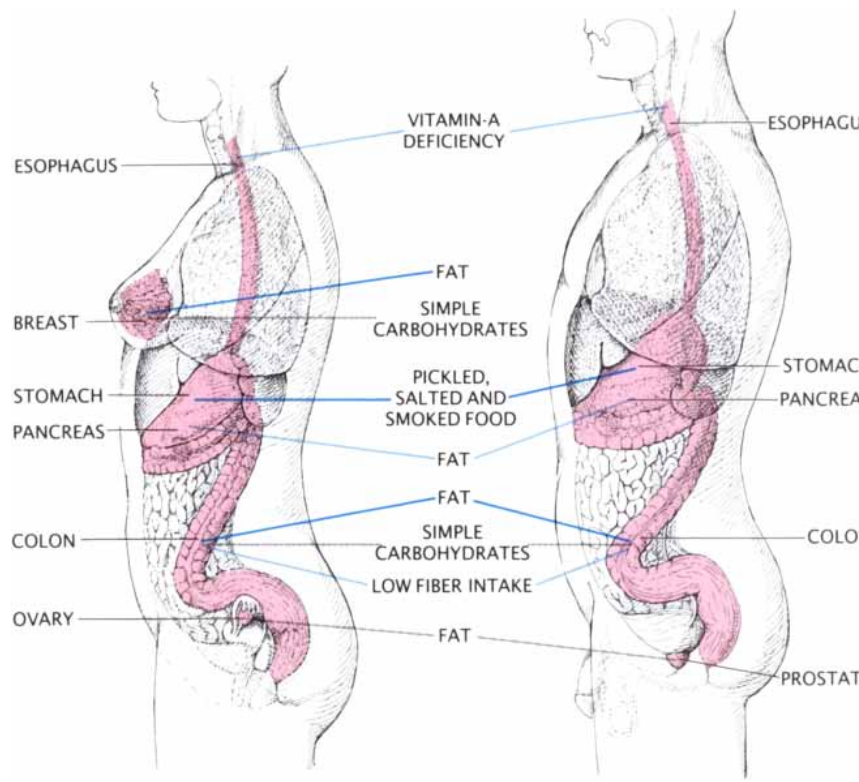
If initiating factors could be unequivocally identified, the incidence of cancer could be reduced by eliminating them or reducing exposure to them. Only a few human cancers, however, have been associated with specific causative agents—for example mesothelioma of the lung with asbestos and Burkitt's lymphoma with the Epstein-Barr virus. Both cancers are rare. With regard to the major cancers afflicting Western populations, only lung cancer has been associated with a somewhat specific agent, namely inhaled tobacco smoke. As for the other major cancers—cancers of the breast, colon, prostate, pancreas, uterus and ovary, which account for some 48 percent of all deaths from cancer among females and 28 percent of those among males—little is known of the initiating agents.

In any case, the identification and elimination of initiating chemicals, which are presumably at low concentrations in the environment, is not a feasible means of preventing these common cancers. The most effective available strategy is either to identify the elements in the environment that promote these cancers and then reduce or eliminate exposure to them, or to identify the elements in the environment that act as antipromoters and then increase exposure to them. If a promoting agent can be removed or modified during the long latent period, a lesion either will remain permanently in the latent phase or will progress to the tumor stage at a much lower rate than would ordinarily be the case.

At present several factors in the diet are seen as tumor promoters or antipromoters. The item most clearly established as a promoter is dietary fat. The possible antipromoters include dietary fiber; vitamins A, C and E; the trace element selenium, and certain compounds in such vegetables as broccoli, cabbage and cauliflower, which are called cruciferous vegetables because they are in the Cruciferae family of plants. Evidence from clinical studies and laboratory experiments suggests that some of the vegetable compounds may act at more than one point in the carcinogenic sequence, affecting enzymes that detoxify initiating carcinogens and also serving as antipromoters. Another class of potential dietary antipromoters has been studied in detail by Walter Troll of the New York University School of Medicine. Found in beans and plant seeds, they are protease inhibitors: they counteract the effects of enzymes that digest proteins and presumably help a tumor to invade neighboring tissue.

There are two major sources of evidence suggesting a role for dietary factors in cancer: epidemiological studies and laboratory experiments with rodents. The former can be done at the national population level or among individuals.

Comparisons of cancer death rates in different national populations have provided important clues to the causation of cancer. One of the most informative comparisons is between the U.S. and Japan. Both countries have similar levels of industriali-



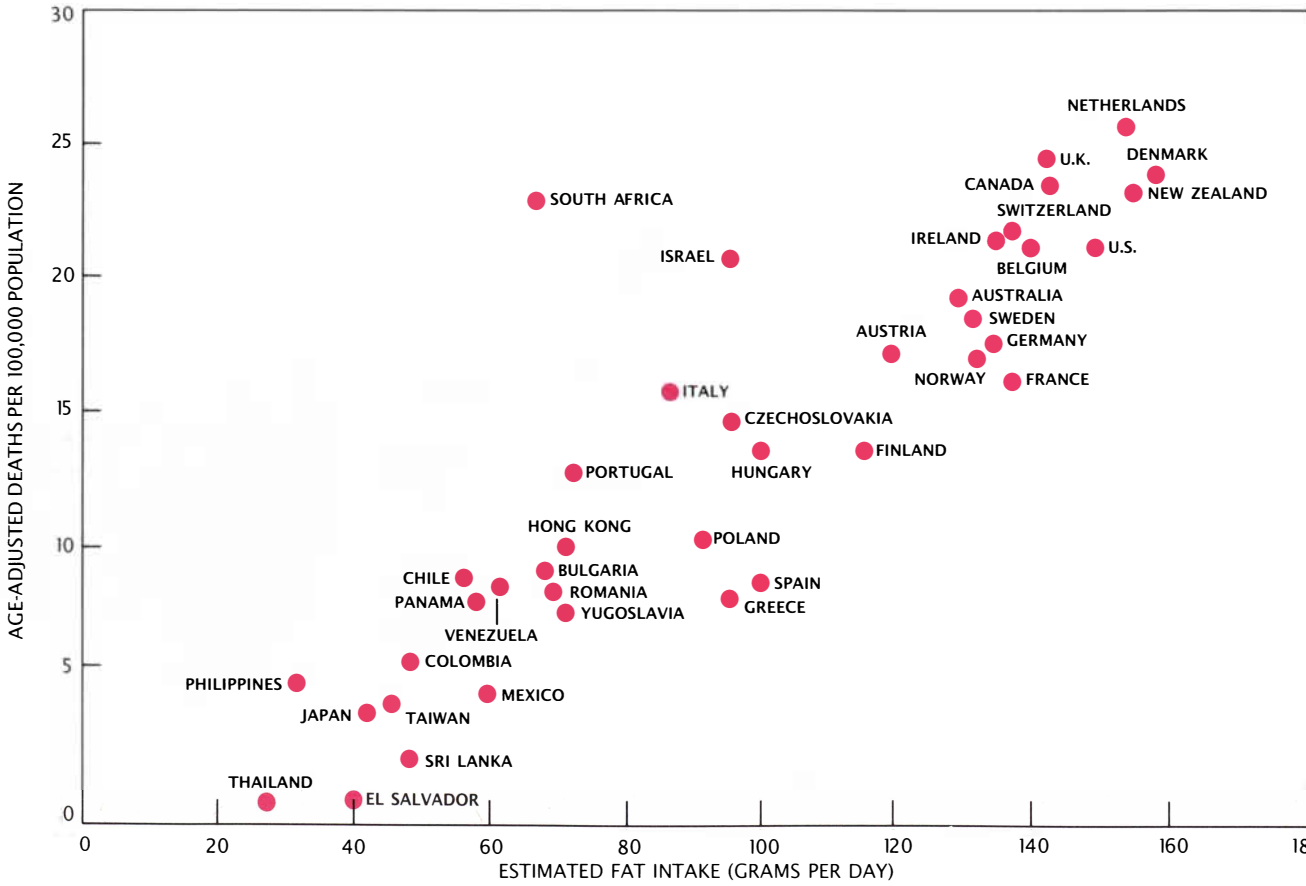
COMPONENTS OF DIET are linked more or less closely to the risk for particular cancers on the basis of epidemiological clues and experiments with rodents. The strongest links are shown in dark blue, weaker ones in pale blue and the weakest ones in gray.

zation and education, high medical standards and good vital statistics. Surprisingly, although the overall cancer rates are similar, the two countries yield mirror images when specific types of cancer are compared. For example, cancers of the breast, colon and prostate are common in the U.S. but rare in Japan. Conversely, cancer of the stomach is common in Japan but rare in the U.S. Are the environmental factors enhancing breast, colon and prostate cancer more prevalent in the U.S., or are factors that protect against such cancers more prevalent in Japan?

One might answer that the differences are based on genetic differences between the two populations. That hypothesis is refuted by the data on migration from one country to another. They show that the worldwide geographic variation in organ-specific cancer rates can be attributed to environmental factors rather than genetic ones. For example, within two generations Japanese immigrants in Hawaii and California had breast-cancer rates approaching those of the Hawaiian white population and significantly higher than those of native Japanese. Conversely, gastric-cancer rates became significantly lower among Japanese Hawaiians than among native Japanese. Increases in breast-cancer rates were also found in populations emigrating from Poland, where the risk is low, to the U.S.

Such shifts naturally focus attention on the possibility that differences in diet may be a factor. The idea that diet may have such a role was suggested more than 50 years ago by the English epidemiologist Percy Stocks, who associated differences in cancer rates in various parts of England with differences in the consumption of certain foods. In the 1950's Ernest L. Wynder of the American Health Foundation concluded that the wide differences between the breast-cancer rate in Japan and that in the U.S. could not be explained on the basis of such established risk factors as family history or reproductive characteristics; he proposed that an environmental characteristic such as diet was the key determinant. In the 1960's Kenneth K. Carroll of the University of Western Ontario published a series of geographic comparisons showing a strong correlation between the intake of dietary fat and breast-cancer mortality rates in 39 countries. Similar but weaker correlations were shown for fat intake and cancer of the colon and prostate.

Yet another epidemiological clue came from the analysis of long-term trends in the incidence and mortality rates of cancer. The data show that in a given country the mortality rate for a particular cancer often varies with time. The changes may be related to widespread changes in the environment, particularly in the population's dietary habits. For example, in developing countries as much as 80 percent of total calories comes from cereals and grains, which consist of complex carbohydrates. With industrialization and economic advance the caloric intake shifts toward fats derived from meat and vegetable oil. The consumption of sugar (simple carbohydrate) also increases. Such changes are taking place now in a number of countries, including Iceland, Italy, Greece and Japan. As the changes have occurred, rising rates

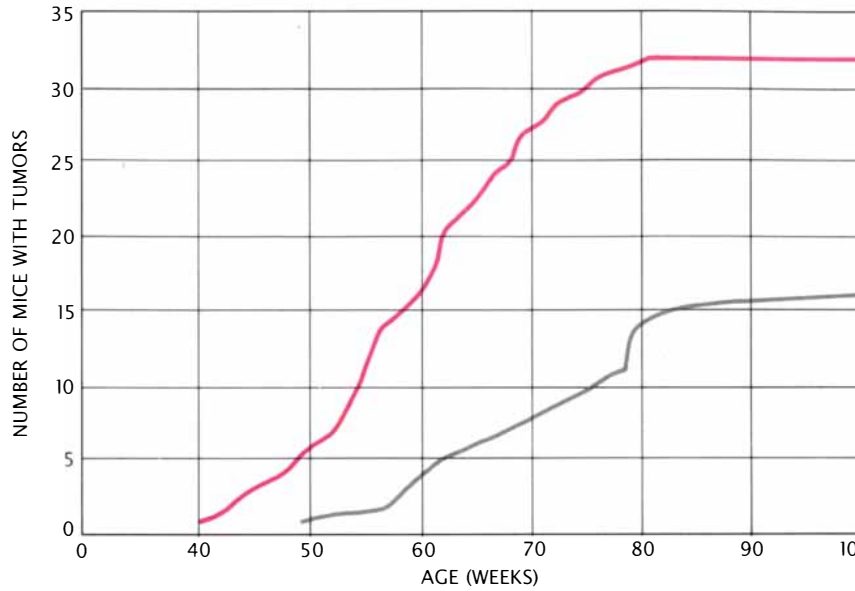


STRONG LINEAR CORRELATION is seen when data for dietary fat and deaths from breast cancer are plotted. (Fat intake was derived by dividing total fat consumed in a country by population; waste and consumption by animals were not taken into account.)

of breast and colon cancer (and of prostate cancer in certain places) have been observed.

The same changes have been observed in the U.S. Since 1900, for example, the intake of animal and vegetable fats has risen by 40 percent while the consumption of potatoes has declined by 60 percent and of flour by 50 percent. Since 1930 the incidence rates of breast, colon and prostate cancer have gradually increased. On the other hand, deaths from gastric cancer have fallen dramatically. The drop has been attributed to the fact that refrigeration has largely replaced salting, pickling and smoking in the preservation of food.

Epidemiological studies involving individuals also provide evidence of links between diet and cancer. Retrospective studies entail comparisons of cancer patients with healthy people. By obtaining detailed dietary histories one can determine whether the cancer patients differed from the controls in what they ate earlier in their life. The results tend to be somewhat inconsistent, however, mainly because people cannot always remember their past diets in detail and because their subjective reports are heavily influenced by their current dietary practice.



CLASSIC EXPERIMENT on female laboratory mice, done by Albert Tannenbaum of the Michael Reese Hospital in Chicago, showed a strong link between a high-fat diet and the cumulative incidence of spontaneous mammary tumors. Hydrogenated cottonseed oil made up 12 percent of the high-fat diet (*color*) and 3 percent of the control diet (*gray*).

One must therefore ask how good the epidemiological evidence is in support of a cancer-enhancing effect for dietary fat and a cancer-inhibiting effect for dietary fiber, vitamins A, C and E and selenium. (The evidence for compounds in cruciferous vegetables comes largely from experiments with animals rather than from epidemiology.) For most of the substances the evidence is limited. Although many and varied epidemiological studies have been made, none of them fully satisfies the criteria epidemiologists employ in drawing inferences about causes from statistical associations.

The association between dietary fat and breast cancer comes closest to fulfilling the criteria. The link between fat intake and colon cancer, although weaker, is significant. The association between dietary fat and prostate or ovarian cancer can be demonstrated only at the national-population level.

The idea that dietary fiber—the part of the plant cell that is not digested in the human small intestine—has a protective effect on large-bowel cancer originated about 15 years ago in the findings of the English epidemiologist Denis P. Burkitt and others. Burkitt, serving as a medical officer in Africa, proposed that the high fiber intake of certain African populations protected them from several alimentary-tract diseases, including appendicitis, diverticulitis and colon cancer. Efforts to test this hypothesis by making international comparisons have been hampered, however, by the lack of general agreement on how to quantify the intake of fiber. Early studies examined food groups known to contain high levels of fiber—cereals, grains, fruits and vegetables—rather than the nutrients themselves. With these groups as indicators of fiber intake it was found that mortality from colon cancer varied inversely with consumption of cereals and grains but not with consumption of fruits and vegetables.

In 1978 Alison A. Paul and D. A. T. Southgate of the Agricultural and Food Research Council's Food Research Institute in Norwich, England, published chemically defined values of fiber in various foods, giving epidemiologists better tools for determining fiber intake and its relation to colon cancer. What emerges from these studies is the finding that "fiber" is a general term encompassing such chemically and physiologically distinct substances as cellulose, gums, pectins and lignins. As a case in point, Sheila A. Bingham of the University of Cambridge and her colleagues reported that mortality from cancer of the large bowel was inversely related to intake of one particular class of fibers, the pentosans, but not to total fiber intake.

The epidemiological evidence for anticancer effects from vitamin A (or its dietary precursor, beta-carotene) comes almost exclusively from the comparison of cancer patients with healthy people. The findings pertain mainly to lung cancer; to a lesser extent they also encompass cancer of the colon, stomach, bladder, esophagus and oral cavity. There is literally no evidence for a protective effect of other vitamins, with the exception of a possible association between the intake of vitamins C and E and reduced rates of gastric cancer.

An anticancer role for selenium was originally proposed by Raymond J. Shamberger of the Cleveland Clinic Foundation. He based the suggestion on findings that levels of selenium in soil, forage crops and blood were inversely related to the risk of breast and colon cancer in various cities and regions of the U.S. International studies revealed a similar relation. The strength and consistency of these associations, however, are weak in some cases. For example, New Zealand has one of the lowest levels of selenium intake (50 micrograms per day) and a breast-cancer mortality rate of 30 per 100,000 population per year, whereas the U.S. has a selenium intake of about 120 micrograms per day and almost the same rate of deaths from breast cancer (27 per 100,000). Such inconsistencies do not rule out a role for selenium, but

they suggest that other factors may be more dominant.

As the above discussion indicates, epidemiologists essentially look at the results of experiments in nature and work backward to causes. Because only relatively crude methods are available for measuring food consumption, particularly what people have eaten in years past, the best an epidemiologist can do is to demonstrate strong and consistent correlations or associations, such as the one between dietary fat and breast cancer. Statistical associations can only imply causation, however; they cannot prove it.

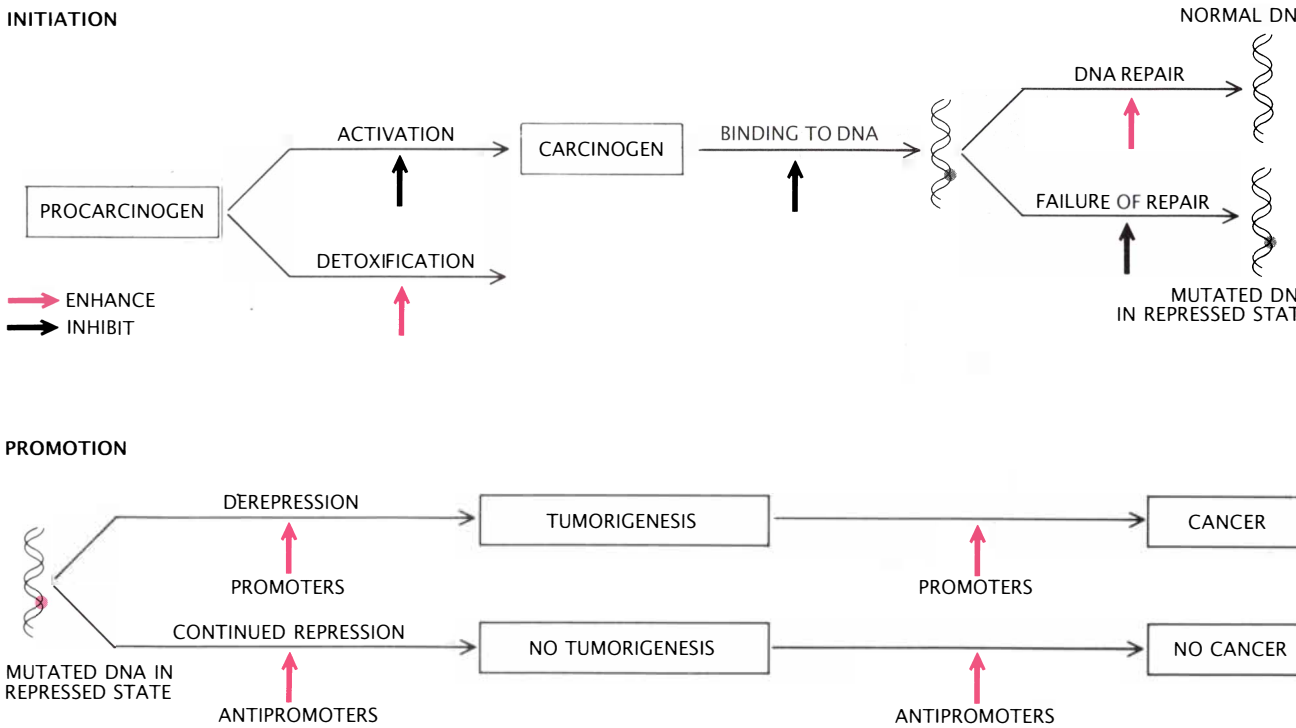
An experimentalist, on the other hand, works forward from causes to results and thereby provides the second major source of evidence bearing on the relation between diet and cancer. A drawback of this work is that the results of laboratory experiments involving rodents must be treated with caution, particularly when the findings are being extrapolated in an effort to assess effects on human beings.

Experimental studies going back nearly 60 years have indicated that the development of cancer in laboratory mice can be modulated by nutritional means. Again the strongest correlation is between dietary fat and mammary cancer. (What is called breast cancer in women is called mammary cancer in animals.) This association was first demonstrated by Albert Tannenbaum of the Michael Reese Hospital in Chicago in the 1920's. In the late 1960's Carroll, studying a chemically induced mammary tumor that enabled him to distinguish between initiation and promotion, demonstrated that fat acted as a promoter. Since then the tumor-promoting effect of dietary fat has been demonstrated in mammary tumors induced by chemicals, X rays and viruses. A diet high in fat has also been shown to promote the development of experimentally induced cancers of the colon and pancreas.

Animal studies have also shown that the type as well as the amount of dietary fat governs the development of mammary, colon and pancreas tumors. High-fat diets rich in certain fatty acids promote the formation of tumors, whereas similar diets high in other fatty acids do not. The difference seems to be in the chemical structure of the fatty acids. High-fat diets rich in linoleic acid (found in corn, safflower, sunflower and other common plant oils) act as promoters; similar diets rich in oleic acid (from olive oil) and eicosapentaenoic acid (from fatty fishes and marine mammals) do not act as promoters.

These findings in animals may help to explain several epidemiological enigmas, including the low rates of breast and colon cancer among the Eskimos of Greenland. They have a diet extraordinarily high in fat (60 percent of calories), but the fat is derived almost entirely from fish and marine mammals. Similarly, the moderate rates of breast cancer in such countries as Greece and Spain, where fat intake is fairly high (32 percent of calories), may be due to the fact that olive oil is a major source of fat in Mediterranean diets.

How might dietary fat promote cancer of the breast, colon and pancreas? Because the mammary gland is responsive to hormones, the earliest hypotheses envisioned an indirect effect by way of the endocrine system. The effect could be mediated either by a central process at the lev-



INFLUENCE OF DIETARY FACTORS on the initiation and promotion of carcinogenesis is outlined schematically here on the basis of current evidence. Cancer can be initiated when a procarcinogen (dietary or other) is activated and binds to DNA, giving rise to a mutation in the DNA (presumably often affecting an oncogene). The mutation may or may not be repaired. Steps in activation, binding and DNA repair may be enhanced (*color*) or inhibited (*black*) by such dietary components as vitamins A, C and E, indoles from cruciferous vegetables and the trace element selenium. An unrepaired mutation may transform cells to an abnormal state, but it remains repressed—a tumor does not develop and there is no progression to cancer—except in the presence of promoters, such as a high-fat diet. Antipromoters such as vitamin A tend to maintain repression and inhibit cancer growth.

el of the pituitary gland or higher centers, which regulate the secretion of hormones such as prolactin, or by a peripheral process, at the level of the target organ, involving the binding capacity of specific hormone receptors.

Another hypothesis focuses on prostaglandins, a class of biologically active lipids synthesized by the body from fatty acids (primarily linoleic acid) ingested in food. In a short series of metabolic steps the acid is converted into the potent hormone-like prostaglandins, which are involved in such functions as the regulation of smooth-muscle contraction, inflammatory reactions, blood coagulation and the immune response.

Mammary tumors and breast cancers produce abundant amounts of prostaglandins, which may act by suppressing the host's immune response. Their synthesis can be manipulated by dietary means, suggesting that experimental high-fat diets may exert their effects by altering the amount and direction of prostaglandin synthesis. Such a mechanism would help to explain not only the finding that a certain amount of dietary linoleic acid is required in experimental diets to achieve the high-fat effect but also the inhibitory role of marine oils: they interfere with the conversion of linoleic acid into prostaglandins.

The most popular current hypothesis about the link between a high-fat diet and colon cancer is based on findings that a high fat intake results in increased secretion of bile acids (which help to digest fat) and alters the bacterial population in the large bowel, thereby enhancing the conversion of primary bile acids into secondary ones. Secondary bile acids in turn are thought to be promoters of lesions that are already present in the large bowel.

Experimental evidence from rodents has also contributed to what is known about anticancer agents. Several studies have shown that administration of vitamin A or its analogues (collectively termed retinoids) inhibits the development of cancer induced experimentally in skin, the respiratory tract, the mammary gland and the bladder. The amount of retinoid necessary to show an effect, however, is some 20 times the recommended daily allowance for the vitamin. Side effects such as reduced weight gain, increased fragility of bones and liver dysfunction have been reported at these high doses. Selenium has been shown experimentally to reduce the frequency of chemically induced mammary and colon tumors in mice, but only at between 30 and 40 times the recommended daily allowance.

The laboratory results with retinoids and selenium as dietary supplements point to three cautionary conclusions. First, supraphysiological doses must be given to obtain the desired effects. Second, there is a thin line between a therapeutic dosage level and a harmful one. Third, neither compound can completely prevent the development of a tumor.

Several animal studies have shown that increasing the amount of fiber in the diet from the usual 5 percent to 15 percent or more reduces the frequency of chemically induced tumors of the colon in mice. The data are not consistent, however, particularly with respect to the specific fiber tested. For example, wheat bran acts to depress the incidence of colon tumors whereas apple pectin, alfalfa and oat bran either have no effect or stimulate the formation of colon tumors, depending on the experimental conditions.

A number of mechanisms have been proposed to explain the apparent protective effects of fiber on colon cancer. Studies by O. M. Jensen of the Danish Cancer Registry and Bandaru S. Reddy of the American Health Foundation indicate that populations at low risk for colon cancer excrete smaller amounts of bile acid than populations at high risk and that the low-risk groups consume more dietary fiber. Decreased bile acid content may result because fiber increases the bulk of feces, thereby diluting their content of bile acid per unit of volume, or because fiber enhances mechanisms that move bile acid from the large bowel into the circulation. Increased fiber also reduces the time between eating and excretion, giving potential mutagens in the feces less time to act on the colon.

Laboratory studies also suggest an anticancer effect for flavones, indoles and isothiocyanates (nonnutritive compounds often called secondary plant constituents) derived from cruciferous vegetables. At present the effect has been shown only for chemically induced tumors of the forestomach, lung and mammary gland in rodents. The compounds apparently act by blocking the activation of hydrocarbon agents in the host. (Most carcinogens are consumed as inactive procarcinogens and must be activated by enzymes in the liver before they can react with the DNA of target cells.)

The effects of salting, pickling and smoking foods have also been explored experimentally. These methods of preservation create conditions under which nitrates can be converted readily into reactive nitrites; these in turn have been shown to combine with appropriate substrates (in the test tube) to form carcinogenic nitrosamines and nitrosamides. The cancer-causing effect of such compounds has been well established in laboratory experiments done by Peter N. Magee of the Temple University School of Medicine. Sidney S. Mirvish of the Eppley Institute for Research in Cancer in Omaha has shown that nitrosation reactions do not proceed in the presence of ascorbic acid (vitamin C), which blocks the conversion of nitrates into nitrites.

Several foods consumed by populations at high risk for gastric cancer—among them smoked fish (in Japan) and fava beans (in Colombia)—have been shown to yield mutagens after nitrosation. These same populations consume small amounts of fresh fruit and vegetables and therefore take in low levels of vitamin C, suggesting that gastric cancer may result from the formation in the stomach of carcinogenic nitrosamines or nitrosamides from dietary nitrate in the absence of vitamin C. In no case, however, have the hypothetical "natural" nitrosated substances to which gastric cancer has been attributed been isolated in people or in experimental animals.

The experimental evidence supporting the notion that complex carbohydrates are better in the diet than simple sugars is quite limited. Several reports have indicated that laboratory rats fed a diet containing starch following the intake of a chemical carcinogen showed a significantly lower frequency of mammary tumors than rats fed a similar diet containing dextrose (a simple sugar) in place of starch. Aside from the fact that complex carbohydrates are digested more slowly than simple sugars and therefore exert a less pronounced effect on the insulin-secreting cells of the pancreas, little is known about the mechanisms that give rise to the effects of these two classes of carbohydrates on experimental carcinogenesis.

What does all the experimental evidence add up to? In general it is consistent with the evidence from epidemiological studies, particularly in the

case of the link between dietary fat and breast cancer. Many animal studies show that high-fat diets (40 percent of calories) stimulate the development of mammary tumors and low-fat diets (10 percent of calories) do not. What is most interesting about these findings is the fact that the quantities of fat in the high-fat and the low-fat experimental diets are much the same as those consumed by people in high-risk countries such as the U.S. and in low-risk countries such as Japan respectively. What remains to be determined is whether the risk for breast cancer varies steadily with fat intake or whether there is instead a threshold below which increased fat intake does not increase risk. Recent studies in my laboratory at the Naylor Dana Institute for Disease Prevention suggest there is a threshold somewhere between 25 and 32 percent of total calories.

In the case of studies on colon cancer too, the experimental diets are similar to the diets of human populations (although in this case the effect of fiber intake must also be taken into account). In contrast, experimental studies on retinoids and selenium entail quantities far exceeding recommended daily allowances. As potential anticancer agents these substances are more like prescription medicines than like components of a normal diet.

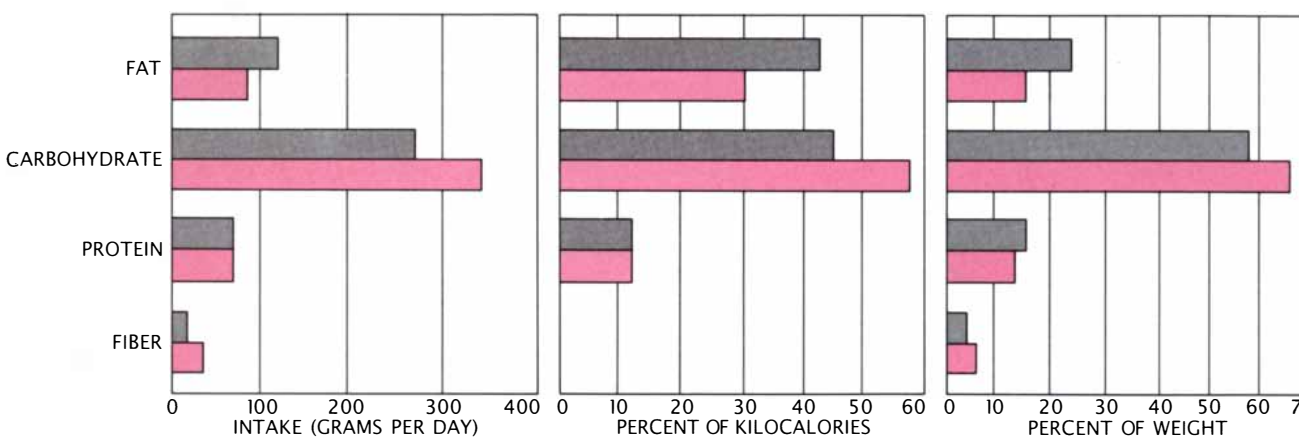
In the end it must be understood that no matter how suggestive epidemiological or experimental studies may be, they cannot provide unequivocal proof that a given person would be wise to cut down on fat in the diet, take more vitamins, eat more fiber and cruciferous vegetables and so on. In no case has it been directly demonstrated that implementing dietary changes in a given individual inhibited the onset of cancer or kept an established cancer from spreading.

Definitive data could be obtained from experiments involving people. Randomized dietary-intervention trials could compare the effect of different diets on the appearance of disease in different groups of healthy people. Such trials would require tens of thousands of subjects, monitored for dietary compliance over at least 10 years. The logistics would be formidable and the cost astronomical, and the outcome might not be known for from 10 to 20 years.

As an alternative a number of clinical trials are being done on people who are at high risk for particular cancers, so that results might be clear within five or 10 years. For example, a study sponsored by the National Cancer Institute involves healthy women deemed to be at high risk for breast cancer because they have a family history of the disease or have had benign breast disease. Individuals chosen at random are asked either to maintain their usual high-fat diet or to switch to one low in fat. Even this trial is expected to involve some 30,000 women, to last for 10 years and to cost more than $100 million. Other such trials are under way. One, in the U.S., will assess the effect of increased fiber on the incidence of colon cancer in individuals with benign colon polyps (a high-risk group); another, in Italy, will study the effect of retinoid supplementation on the appearance of second breast cancers in women who have had breast cancer.

The results of the clinical trials now in progress are not likely to be available until the 1990's or later. Hence one must ask whether it is better to ignore the indirect and circumstantial evidence linking diet and cancer and not make specific dietary recommendations or to act on a reasonable best guess and make interim or provisional recommendations. Many experts argue on ethical grounds that public health recommendations should be based only on unequivocally proved findings. Others say that with more than 400,000 deaths per year from cancer in the U.S., even a small reduction could save many lives, so that dietary recommendations aimed at reducing the risk of cancer should be made in the interest of public health even if they are provisional.

The National Research Council has taken the second course, issuing interim dietary guidelines based on the available epidemiological and experimental data. As more evidence becomes available the recommendations will surely require revision; at present they reflect the current state of knowledge. They form the basis of a diet that is not likely to have harmful effects and has a good possibility of reducing the risk of the major cancers induced by diet.



DIET IN U.S. would change appreciably if recent recommendations of the National Research Council and other agencies were followed. The current diet (*gray*) and a recommended one (*color*) are compared in terms of each component's quantity (in grams per day) and its share of the diet's total kilocalories and total weight. The aim is to reduce the intake of food items thought to initiate or promote cancer and increase the intake of those thought to nullify or moderate the cancer-initiating or cancer-promoting effect of other items. Note that when total caloric intake is held constant at 2,400 kilocalories, food intake by weight increases in the recommended diet because calorie-dense fat is replaced by less calorie-dense carbohydrate and by fiber.

How a luxury car should behave.

Lincoln Mark VII LSC.

It should respond to your every command. The 1988 Mark VII has been built to do just that.

Its engine is a more powerful 5.0-liter HO V-8, rated at 225 horsepower developing 300 ft-lbs of torque. When you want it to go, it goes. Very quickly. With a 0-60 acceleration that will make the most serious driver smile.

Mark VII has the most advanced braking system in the world. The Anti-lock Brake System. So when you want Mark VII to stop, it stops. Very quickly. Without the wheels locking. It helps you brake so well, it can shorten stopping distances up to 40% on slick surfaces. Even under the heaviest braking. Yet, while Mark VII exhibits exhilarating road behavior, it never forgets it's a luxury car.

As *Car and Driver* said, "If it's peace and quiet blended with excellent roadability you're after, look no further."

Mark VII LSC. The spirit and credentials of a sports coupe. And as you'd expect, all the luxury of a Lincoln.

LINCOLN.
What a luxury car should be.

For more information, call 1 800 822-9292. Buckle up—together we can save lives.

Lincoln-Mercury Division Ford

Helium-rich Supernovas

Computer modeling suggests that they result when the denuded core of a massive star collapses. These "peculiar" supernovas are close cousins of this year's bright event

by J. Craig Wheeler and Robert P. Harkness

The Rev. Robert Evans of Hazelbrook in New South Wales, Australia, must have learned to suffer tasteless jokes about his "direct line to God." When he can spare the time from his parish duties, he repairs to his telescope to scan other galaxies in search of supernovas: rare, catastrophic explosions in which, for a few weeks, a star flares to the brightness of its host galaxy. Whether it is by dedication and persistence or by a special talent for pattern recognition, Evans has compiled a record of discovery that is rapidly becoming legendary. Of the 66 supernovas spotted between 1983 and the middle of 1986, Evans discovered 10, the majority in relatively nearby galaxies, where they appear bright and can be studied easily by professional astronomers.

Evans' findings are distinguished by more than just their number. Precise, digitized spectra of light from two of his events, coupled with computer models of exploding stars, have forced the recognition of a new class of supernovas. Supernovas have traditionally been placed in two classes according to their spectra. Each class was thought to be associated with a distinct kind of explosion in a specific star type: the thermonuclear explosion of an old, dwarf star or the violent collapse of a bloated, massive star rich in hydrogen. Evans' discoveries and other, related events—all apparently marking the demise of massive stars previously stripped of hydrogen—have unsettled the spectral classification system, suggesting that one or both of the explosion mechanisms can give rise to either kind of spectrum.

It was cloudy in New South Wales this February 23, depriving Evans of the opportunity to spot SN 1987A, the brightest supernova since the time of Kepler and a major focus for supernova research for decades to come. Instead the first to notice the explosion, which appeared in a nearby, irregular galaxy known as the Large Magellanic Cloud, were Ian Shelton of the University of Toronto, a staff astronomer at the Cerro Tololo Inter-American Observatory in Chile; Oscar Duhalde, a night assistant there and, independently, Albert Jones, another Australian amateur. Yet the study of Evans' events had prepared the way for SN 1987A. The explosion has shown features closely related to those of the new class of supernovas, as well as a host of unconventional characteristics.

The existing classification scheme has been in place since the 1930's, when Fritz Zwicky of the California Institute of Technology started to catalogue supernovas. The first three dozen he observed were all of a kind. Their spectrum was incomprehensible, different from any star's but similar from event to event. Their light curve, representing the change in brightness with time, was also characteristic: the brightness peaked after about two weeks, declined steeply for about two more weeks and later faded exponentially, with a half-life of about 50 days.

Then Zwicky found a second kind of event. These were intrinsically less bright, which may account for their late discovery, and their light curve varied. Unlike the first group, these displayed spectra with a recognizable feature: spectral lines of hydrogen. The lines represent wavelengths at which excited hydrogen atoms emit photons as they undergo transitions between energy states. Rudolph Minkowski, working with Zwicky, distinguished supernovas showing the characteristic but uninterpretable spectrum and those showing plain spectral evidence for hydrogen in the ejected material as Type I and Type II respectively.

Theoretical advances pointed to two distinct explosion mechanisms, which were described by Fred Hoyle of the University of Cambridge and William A. Fowler of Caltech in the 1950's. The mechanism for Type I supernovas envisions, quite simply, a star-size bomb, powered by runaway thermonuclear reactions. The setting is thought to be a white-dwarf star: a compact, ancient remnant of a parent star no more than about eight times as massive as the sun that has burned most of its hydrogen and helium to the intermediate-mass elements oxygen and carbon.

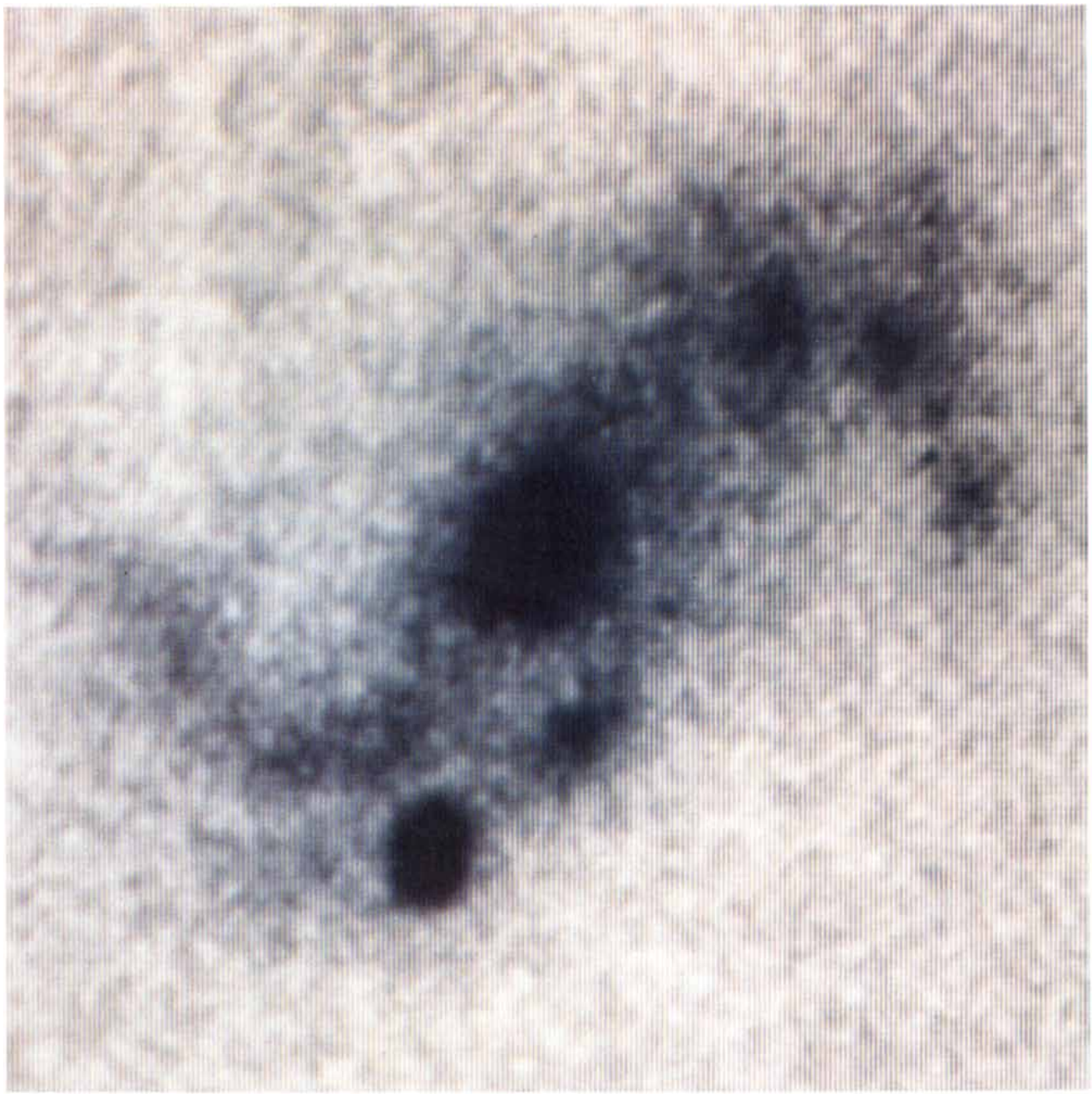
A binary companion star triggers the explosion. Earlier in the evolution of the star the companion may have wrested away the white dwarf's outer layer of hydrogen, baring its dense oxygen-carbon core. Later the white dwarf draws in matter from the companion. The density and temperature of the core increase sharply, and carbon there begins to fuse into heavier elements, among them radioactive nickel. A burning front rips outward, completely destroying the star, but the expanding cloud of ejecta glows brightly for many weeks as the nickel decays into cobalt and then iron.

A more exotic mechanism powers Type II supernovas. Zwicky himself had proposed that the formation of a neutron star, a sphere of matter as dense as an atomic nucleus but several kilometers in diameter, might release enough energy to trigger a supernova—a notion that lies at the heart of the mechanism now recognized for Type II events. In the core of a star of more than eight solar masses thermonuclear burning proceeds rapidly through a series of fusion reactions that end in nickel and iron, whose nucleus is the most tight-

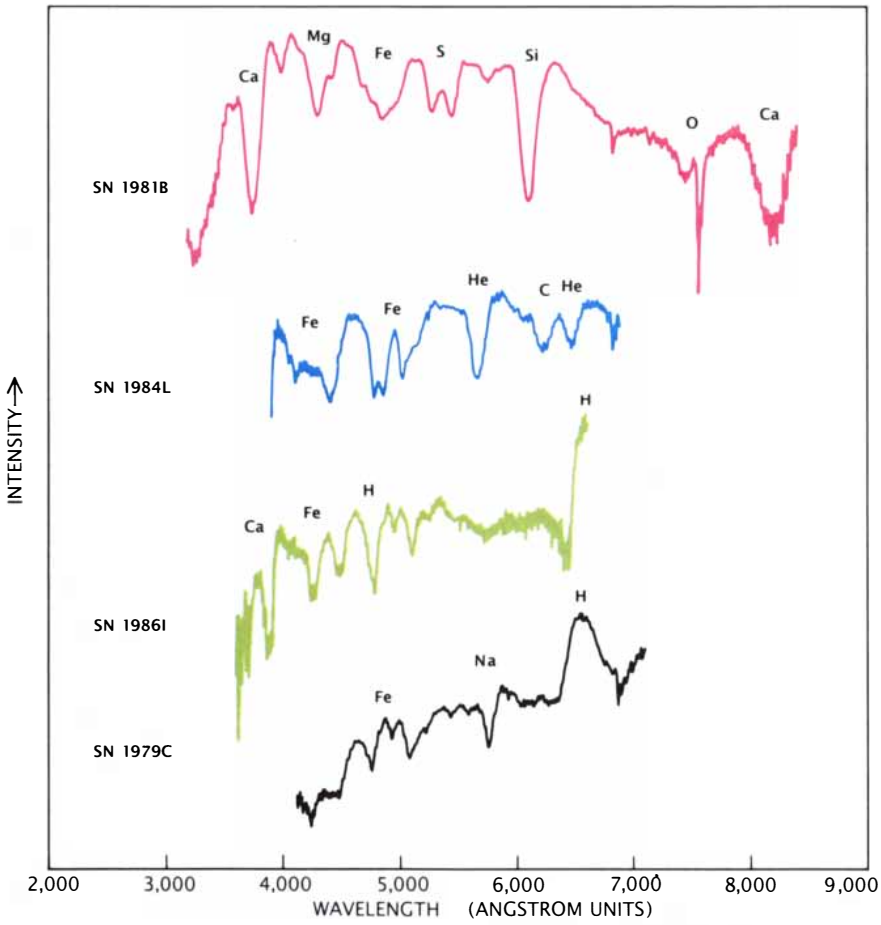
ly bound of any element; further fusion would consume rather than release energy. As heat production in the core ends, the star collapses precipitously and the protons and electrons of atoms in the core are forced together. They combine into neutrons, with the release of a neutrino for each neutron formed, and the collapsing ball of neutrons halts and bounces when it is squeezed down to nuclear densities [see "How a Supernova Explodes," by Hans A. Bethe and Gerald Brown; SCIENTIFIC AMERICAN, May, 1985].

Although the details of the process are not yet clear, the collapse liberates an enormous amount of energy. Most of it is lost in a burst of neutrinos, to which the outer layers of the star are transparent; the energy deposited in the star by the neutrinos, combined with the shock wave generated when the collapsing core rebounds, suffices to blow off the outer layers of the star. The ejecta include the star's extended hydrogen envelope, which accounts for the hydrogen lines of a Type II spectrum. The nuclear matter is left behind as a neutron star at the center of the expanding cloud of ejecta.

Two exceptions to this dichotomy of spectral types were noted more than 20 years ago by Francesco Bertola of the Astronomical Observatory



PECULIAR SUPERNOVA 1983I flares in the galaxy NGC 4051. The supernova (evident as a black spot in this negative image) seems to coincide with an arm of the galaxy, which suggests it originated in one of the massive stars that are common in galactic arms. The event was confirmed as a member of Type Ib, a new class of exploding massive stars, by computer models showing that a variant of the star type that has been implicated in other events belonging to Type Ib could account for the spectrum of this supernova too. The photograph was made by John F. Kielkopf at the Moore Observatory of the University of Louisville.



SPECTRA from supernovas of the various types indicate light intensity across a range of wavelengths. Peaks and troughs in the spectra identify elements in the supernova atmosphere, which absorb or emit photons at specific wavelengths. SN Ia (*red*) lack hydrogen but show strong evidence of silicon. SN Ib (*blue*) show helium but no hydrogen or silicon, and SN II are rich in hydrogen; the two subclasses of SN II, SN II-p (*green*) and SN II-l (*black*), vary both in details of their spectra and in the way their brightness changes over the course of time. The spectra were obtained by Beverley H. Wills, Derek Wills and Russell M. Levreault at the McDonald Observatory of the University of Texas.

in Padua. Bertola reported that the spectra of two supernovas, SN 1962L and 1964L, showed no evidence of hydrogen, which marked the events as Type I. The spectra of classical SN I at maximum light, however, consist of lines of intermediate-mass elements formed in the explosion—oxygen, magnesium, silicon, sulfur and calcium—all Doppler-shifted by the rapid expansion of the supernova ejecta. Bertola's exceptions, which he called "peculiar" Type I supernovas, lacked one of the strongest of the lines: a 6,150-angstrom-unit absorption line that is attributed to ionized silicon. (An element's absorption lines are seen when the atom or ion lies in front of a bright background of radiation, which it scatters out of the line of sight by absorbing and reemitting photons.)

As with so many unexplained "exceptions," Bertola's observations attracted little attention, other than an occasional reference in the literature, included for the sake of honesty. The situation changed dramatically with Evans' discovery of supernovas 1983N and 1984L. Bertola had to rely on photographic spectra, from which it is difficult to judge the absolute intensity of individual spectral lines. New digital spectrographs, in contrast, not only show whether a line is present but also yield quantitative data on the intensity and shape of spectral features. Now the loose threads represented by SN 1962L and 1964L led to the raveling of a curtain and the opening of a new vista of supernova research.

In the broadest terms, the spectra of 1983N and 1984L marked them as Type I: they showed no evidence of hydrogen. They also lacked the 6,150-angstrom silicon line, which linked them with Bertola's peculiar Type I events. Nevertheless, spectra obtained at maximum brightness showed the same cluster of lines between 4,000 and 6,000 angstroms, probably resulting from excited states of iron, as spectra of three-week-old Type I events do. The peculiar events, according to a short-lived aphorism, were "born old." This catchy phrase suggested that the peculiar events were variants of classical Type I events but generically related to them.

The apparent relation eroded as we proceeded to study the spectra. We found that the expanding gases from the peculiar supernovas contain helium, an element never definitively detected in the spectrum of a classical Type I or Type II event. The spectra showed a strong absorption line at 5,700 angstroms, exactly the frequency at which the 5,876-angstrom line of neutral helium falls when it is Doppler-shifted by 7,500 kilometers per second, a typical expansion velocity for the atmosphere of a supernova. The same Doppler shift also brought three other helium lines into coincidence with observed spectral features. The new events (termed SN Ib) were a breed apart from classical Type I events.

Other features of the events suggested that they in fact bear some relation to Type II explosions. Classical Type I events (now SN Ia) often are observed in elliptical galaxies, which are populated mainly by small, old stars—progenitors that require the thermonuclear-explosion model. Type II events, in contrast, usually flare in so-called H II regions: patches of ionized hydrogen, common in the arms of spiral galaxies, that represent the birthplace of stars and include many young, massive stars. SN 1983N, SN 1984L and Bertola's two events all seemed to fall in or near an H II region.

The events shared a second feature of SN II: detectable radio signals. Using the Very Large Array of radio telescopes in New Mexico, Richard A. Sramek of the National Radio Astronomy Observatory, Nino Panagia of the Space Telescope Science Institute and Kurt W. Weiler of the National Science Foundation observed radio emissions from both SN 1983N and SN 1984L and traced the radio "light curve" at various frequencies as the

power decayed. Such emissions, according to a plausible explanation put forward by Roger A. Chevalier of the University of Virginia, are generated when a supernova shock wave catches up to a circumstellar "wind" that is composed mainly of hydrogen blown off a star before the explosion. Massive stars, which are very bright, naturally expel such a wind by radiation pressure; the explanation therefore requires that the exploding star be massive or have a massive companion.

How could one make sense of the traits SN Ib displayed—the seeming links to both Type I and Type II and the evidence for helium, which sets the new category apart from both classical types? What kind of explosion was occurring, and in what kind of star? To interpret the evidence we turned to a new analytical tool: computational models that predict the spectrum of light emitted from a supernova with given characteristics. The expanding supernova atmosphere is modeled as a set of concentric shells, each of them containing matter with a specified composition, density, temperature and radial velocity. The computation then traces radiative transfer through the shells: how light at each frequency is generated, scattered or absorbed by excited atoms or ions as it flows toward the visible surface of the supernova. The result of the computation is the predicted spectrum of light from the explosion.

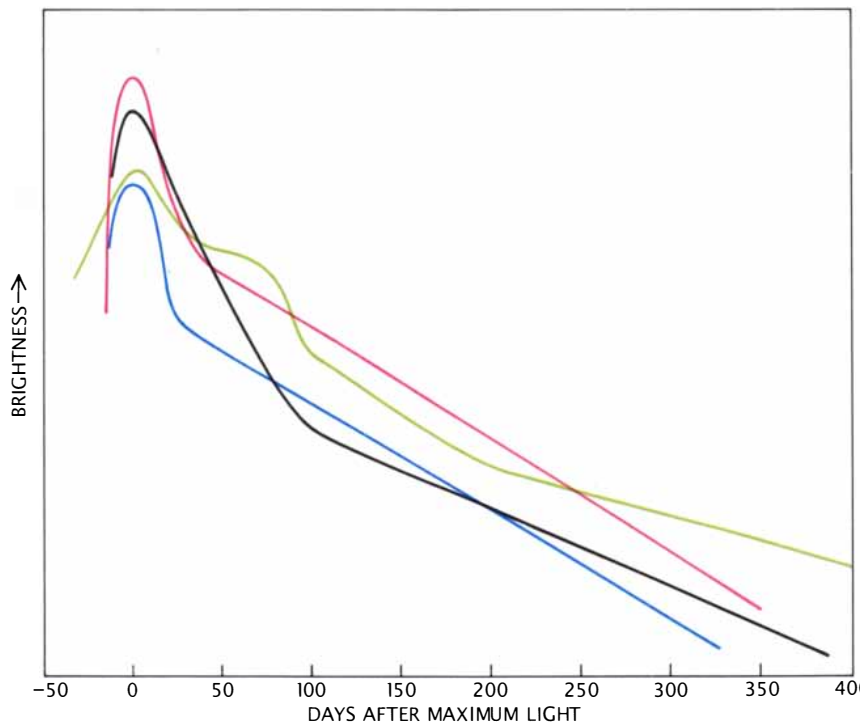
The calculations were pioneered by David Branch of the University of Oklahoma. In current versions radiation is assumed to be transported through matter in local thermodynamic equilibrium, that is, matter in which the proportion of each species of atom or ion in an excited state depends only on the energy of the state and the local temperature. The computations also take into account the effect of the supernova's rapid expansion on the coupling between radiation and matter. A large Doppler effect, for example, arises because a supernova's expansion velocity increases with distance from the center. From the point of view of any atom in the ejecta, matter is receding in all directions, with the result that radiation coming from all directions is red-shifted (Doppler-shifted toward the red, or longer-wavelength, end of the spectrum). Moreover, the ejected matter is moving at some fraction of the speed of light, which gives rise to relativistic effects.

To allow for these effects of rapid expansion, the equations of radiative transfer are solved in a reference frame that is moving with the matter. Calculating the interplay of matter and radiation in a "comoving frame" gives the frequencies and intensities that are transferred outward. A final transformation of the results, from the moving frame of the supernova ejecta to the stationary frame of the observer, yields the predicted spectrum. Such calculations tax the memory and speed of the fastest supercomputers.

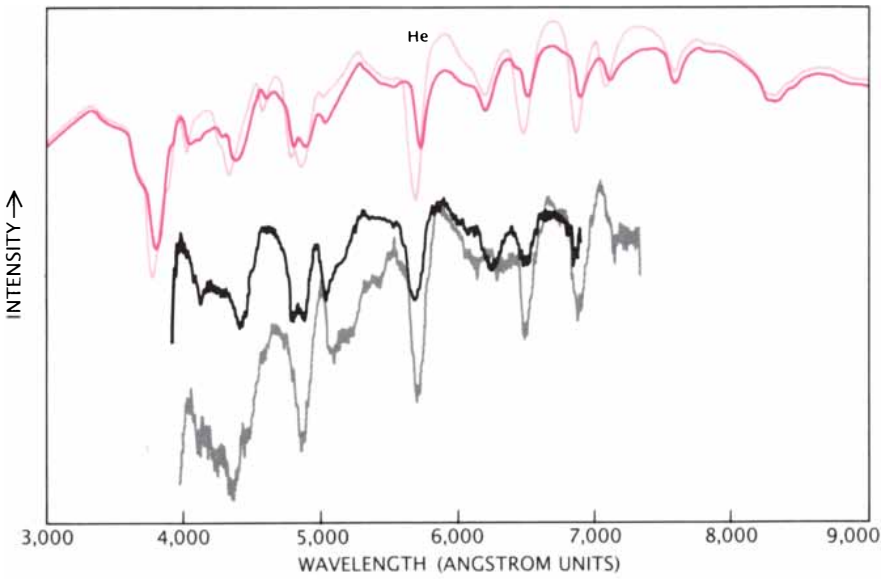
Attempts to construct simple model supernovas that reproduced the observed spectrum of SN Ib events gave insight into their composition and mass. The models specified plausible profiles of temperature and density and a composition consisting of a large amount of helium and smaller amounts of oxygen and carbon, elements that are thought to be abundant in the core of massive stars. A Cray XMP/24 at the University of Texas' Center for High Performance Computing, supplemented by various other supercomputers at the national centers supported by the National Science Foundation Office of Advanced Scientific Computing, provided the computing power.

The spectrum calculated for a model composed of 90 percent helium, 1 percent carbon, 9 percent oxygen and small proportions of heavier elements (a composition thought to be typical of massive stars) gave some resemblance to the observed spectrum. The helium absorption line—the most crucial feature—was too faint, however. Since even a very high proportion of helium could not reproduce the observed line, the helium in the SN Ib atmosphere must be excited beyond thermodynamic equilibrium. Multiplying by 100 the number of helium atoms predicted to occupy excited levels at equilibrium gave a theoretical spectrum that closely matched the spectrum of SN 1984L at its peak.

What could generate anomalously high helium excitation in an actual supernova is not known, but the ex-



CHARACTERISTIC LIGHT CURVES of supernova types trace their change in brightness (plotted on a logarithmic scale). The light curve of SN Ia (*red*), on the average the brightest kind of supernova, drops steeply after maximum light and then begins a smooth, exponential decline that has a half-life of about 50 days—the behavior that is expected if such supernovas shine by radioactive decay. SN Ib (*blue*) are dimmer and their light curve is less well understood, but radioactive decay is thought to power the curve's tail. Type II-p events (*green*) are so called because their brightness levels out in a plateau soon after its peak, whereas Type II-l supernovas (*black*), which probably contain less hydrogen and may have a different explosion mechanism, fade linearly.



THEORETICAL SPECTRA calculated for an exploding massive star composed of 90 percent helium, 9 percent oxygen, 1 percent carbon and traces of heavier elements are compared with the observed spectra of a Type Ib supernova (SN 1984L) near its maximum brightness (*black*) and three weeks later (*gray*). The resemblance suggests that SN Ib originate in helium-rich massive stars; to match the strong helium absorption line at 5,700 angstrom units, however, the calculations had to assume that the fraction of excited helium atoms is many times higher than the temperature of the supernova atmosphere would warrant. In the model that gives the best match to the maximum-light spectrum (*dark color*), helium is excited by a factor of 100; a factor of 10^4 gives the best match to the spectrum of the supernova three weeks after maximum light (*light color*).

tremely low density of the ejecta and the resulting infrequency of atomic collisions may play some role. Collisions, which tend to knock excited atoms or ions back into lower energy levels, act as a restoring force in maintaining thermodynamic equilibrium. Favoring the possibility that the strong helium line results from an imbalance between excitation (by radiation from deeper layers of the ejecta) and deexcitation (by collisions) is the fact that the helium lines grow stronger as the supernova ages and continues to thin out. Whereas an enhancement of 100 in the excited-level populations was sufficient to match the spectrum of SN 1984L at maximum light, an enhancement of 10^4 was needed for a theoretical spectrum that fitted the supernova's spectrum three weeks later.

The radiative-transfer calculations thus showed that the spectral features of SN Ib could well originate in the explosion of a helium-rich massive star. Varying the mass of the model star to achieve the best match with the observed spectra showed that the mass of the exploding material had to be roughly several times that of the sun.

The models also confirmed that two more events belong to the class SN Ib. SN 1983I and 1983V, discovered respectively by D. Yu. Tsvetkov of the Sternberg Astronomical Institute and by Evans, attracted little notice at the time. J. H. Elias and his colleagues at Caltech identified SN 1983I as Type Ib from the similarity of its light curve at infrared wavelengths to light curves of SN 1983N and 1984L. Spectra also suggested a relation to SN Ib: they lacked hydrogen features and the 6,150-angstrom silicon line. The characteristic 5,700-angstrom helium line, however, was apparent only as a shallow dip in intensity.

The relation of these events to SN Ib was strengthened by the demonstration that a variant of the theoretical model reproducing the original SN Ib spectra could also account for the new spectra. The variant specified an atmosphere enriched in carbon and oxygen and depleted in helium. Oxygen and carbon readily lose electrons to become ionized, and free electrons scatter photons of all wavelengths. Electron scattering reduces the contrast between an absorption trough and the surrounding wavelengths in a spectrum; in this case, the model showed, the electron scattering due to the increased oxygen and carbon further diluted helium lines already weakened by the reduction in helium. A model containing 10 percent helium, 10 percent carbon and 80 percent oxygen, again with small admixtures of heavier elements, gave an adequate match to the observed spectra. SN 1983I, 1983N, 1983V and 1984L apparently all belong to the class SN Ib, but the class is heterogeneous, encompassing events with a wide range in the ratio of helium to oxygen.

These theoretical models of atmospheres help to make sense of the early days of a supernova: the first month or so after the explosion, when the ejected matter is optically thick and each photon interacts many times with electrons or ions in the high-velocity gas. The detected radiation reveals the properties of the supernova only at the surface of the optically thick region—the depth from which photons can escape directly. Ever deeper layers are probed as the matter expands and thins out. Finally, after a month or so, the ejecta become so optically thin that radiation from any of the expanding matter can reach the observer directly. This tenuous remnant of a supernova has been called its supernebular phase, and it holds its own clues to the nature of SN Ib.

A supernova's spectrum changes as it enters the supernebular phase. In the early days after the explosion the spectrum is composed of "permitted" lines: lines that represent the absorption or emission of photons in the course of quantum-mechanically favored energy transitions. Only such transitions, which occur quickly, can leave their signature on the spectrum of the expanding atmosphere; transitions that take place more slowly are forestalled by atomic collisions, which dissipate the energy of an excited atom or ion before it can be radiated. In a spectral plot of intensity against wavelength each permitted line is marked by an emission peak juxtaposed with a trough (indicating absorption) at a slightly shorter wavelength. Matter expanding directly toward the observer produces the trough by scattering photons from deeper, opaque regions of the supernova.

In the supernebular phase, in contrast, "forbidden" lines dominate the spectrum. Such lines represent transitions taking place so slowly that they are never seen in laboratory gases. They are favored in the thin supernebular gas, however, where the interval between collisions is so long that each atom or ion can ab-

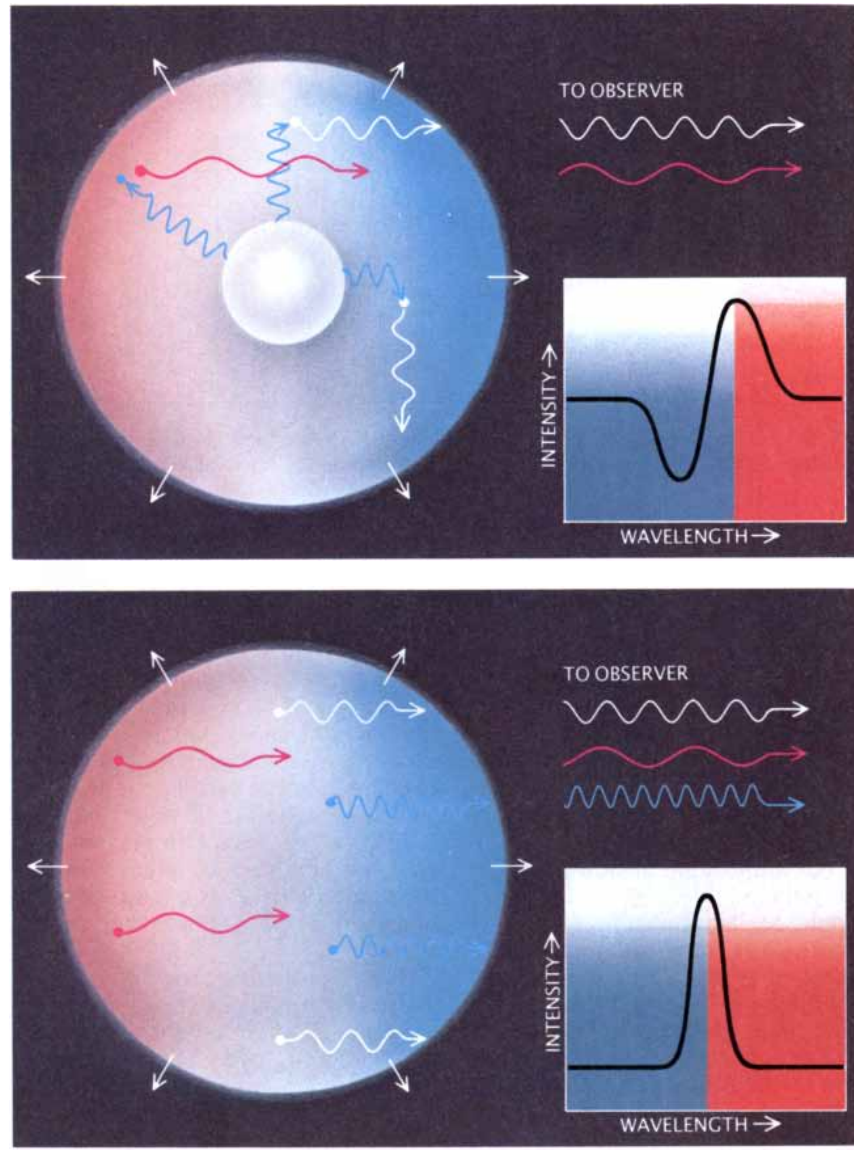
sorb and release energy virtually in isolation. Because a photon released in this optically thin environment is extremely unlikely to be scattered, forbidden lines are evident only as emission peaks.

Not only does a given element show different spectral features in the supernebular phase but also the supernova as a whole manifests a new composition, enriched in heavier elements that were hidden earlier in the explosion. Timothy S. Axelrod of the Lawrence Livermore National Laboratory has shown, for instance, that the supernebular spectrum characteristic of SN Ia events can be reproduced by overlapping forbidden emission lines from an expanding cloud of ionized iron and cobalt—the expected products of the runaway thermonuclear burning thought to power a Type Ia supernova. James Graham, Peter Meikle and their colleagues at the Imperial College of Science and Technology in London have identified smaller quantities of iron in the supernebular phase of a Type Ib supernova (SN 1983N).

The supernebular phase of SN Ib events also shows an entirely distinctive feature: evidence of large quantities of oxygen. This feature was established indirectly, beginning with the discovery of SN 1985F by Alexei V. Filippenko of the University of California at Berkeley and Wallace L. W. Sargent of Caltech. The supernova was discovered in its supernebular phase, many months after maximum light. Its supernebular spectrum was unprecedented, displaying very intense emission lines of oxygen.

A few months later a spectrum of SN 1983N that had been obtained about eight months after maximum light was finally corrected for instrumentation bias, and it proved to be strikingly similar to the spectrum of SN 1985F. The similarity led immediately to two remarkable conclusions. The first was that Type Ib supernovas evolve into a unique, oxygen-dominated supernebular phase quite unlike the supernebular phase either of SN Ia or of SN II (whose supernebular phase, like earlier epochs of the explosion, is dominated by hydrogen lines). The second was that SN 1985F was not unique but a member of the class SN Ib.

A great deal of effort is now being spent to interpret the supernebular spectrum of SN Ib and model this phase of the explosion. Presumably the spectrum reveals oxygen-enriched inner layers of ejecta, but what is the source of the oxygen?



AGING OF A SUPERNOVA changes its spectral characteristics. In the early phase of the explosion (*top*) light is emitted from the surface of a hot, opaque central region and is scattered in the surrounding semitransparent material. The supernova is expanding; photons from matter moving toward the observer are blue-shifted (Doppler-shifted toward shorter wavelengths), photons from matter moving laterally remain at their rest wavelength and photons from matter moving away are red-shifted. Elements along the line of sight to the core scatter photons away from the observer, and so a blue-shifted absorption dip marks each spectral feature. An emission peak accompanies each trough; the peak results as photons emitted away from the line of sight are scattered toward the observer. After a month or so, in the supernebular phase (*bottom*), all the material is transparent and light can reach the observer directly from any point. Because very few photons are scattered, each feature forms a pure emission peak.

Both a white dwarf (in some models) and a massive star are expected to contain oxygen-rich regions. The strength of the supernebular oxygen lines, which seem to indicate several solar masses of oxygen, suggests to us that the interior of a disrupted massive star is the likelier origin.

Thus the spectral evidence, the radio observations and the association of SN Ib events with H II regions all suggest that supernovas of this new class originate as massive stars. Considerations of stellar populations and their evolution, supplemented with further theoretical modeling, lead to a more specific hypothesis about the progenitors.

First of all, the explosions are not likely to occur in only the most massive stars—stars of more than 30 solar masses. Such stars are too rare to account for the growing list of SN Ib

SUPERNOVA	LOCATION IN HOST GALAXY	SUPERNEBULAR OXYGEN	HELIUM
1954A	?	HIGH?	?
1962L	H II REGION	LOW	?
1964L	SPIRAL ARM	LOW	?
1982R	?	LOW?	?
1983I	SPIRAL ARM	LOW	?
1983N	H II REGION	HIGH	HIGH
1983V	H II REGION	LOW	?
1984L	H II REGION	HIGH	HIGH
1985F	H II REGION	?	HIGH
1986M	?	LOW?	?

TYPE IB SUPERNOVAS, suspected and confirmed, are listed together with their defining traits. Early SN Ib candidates have been included on the basis of archival spectra.

events [*see table above*], which by now suggests that they may explode about as often as SN Ia and SN II. Light would diffuse slowly out of a very massive supernova, moreover, giving SN Ib light curves a very different shape from what is observed. Finally, evolutionary models suggest that the strong stellar wind of a very massive star would ultimately remove not only all the star's hydrogen (which would be in keeping with the observed composition of SN Ib) but also all its helium.

Could SN Ib represent exploding white dwarfs, contrary to the indications that the progenitor stars are young and massive? Atmosphere models show that spectral lines from the abundant iron in an exploding white dwarf would swamp any helium features, which makes this possibility also seem quite remote.

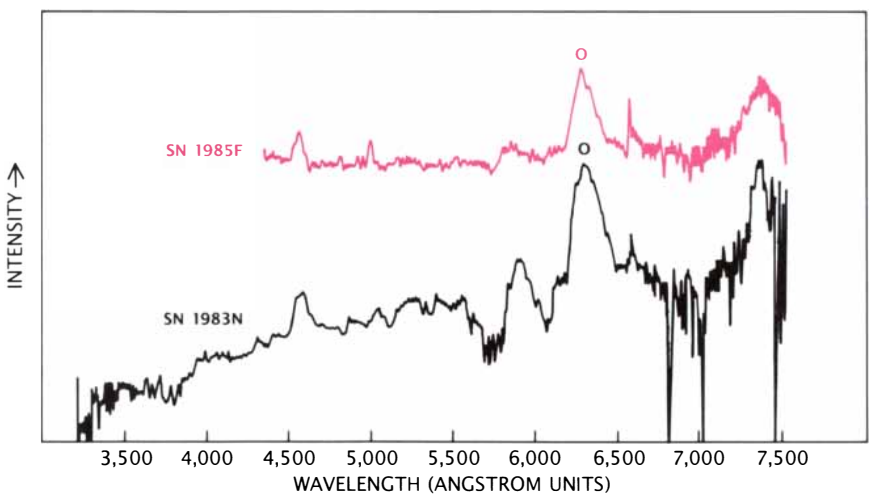
There is a middle ground: the cores of stars that once had a mass of between eight and 25 solar masses but have somehow lost their hydrogen envelope. Stars in that mass range are thought to account for most SN II events, and SN Ib would presumably have the same explosion mechanism: a core collapse that would compact some of the core into a neutron star and generate a shock wave, blowing off the rest of the star's mass. First the helium-rich outer layers of the progenitor star would be exposed and then its oxygen-rich intermediate layers.

How would the progenitor have lost its hydrogen envelope? Stars of less than 25 solar masses are not luminous enough to eject their envelopes spontaneously in a strong stellar wind. They probably must rely on a companion star to pull off their hydrogen layer, exposing the helium and heavier elements. In this picture SN Ib would differ from many Type II events only in that the latter come from massive stars having either no companion star or a very distant one with which they never interact; SN Ib, on the other hand, occur among the fraction (perhaps half) of the stars in the same mass range that happen to have a close companion.

This suggestion may solve a puzzle of stellar evolution posed by a major class of binary X-ray sources that consist of a neutron star orbiting a rather massive companion. (Matter falling from the companion onto the neutron star emits the X rays as it is compressed and heated.) Presumably such X-ray sources originated as binary systems containing two massive stars, and the more massive partner evolved and collapsed to form the neutron star. Given that the abrupt loss of most of the star's mass in the explosion could well have destabilized the binary system, the star probably had already lost its outer hydrogen envelope (and therefore most of its mass) into space or to the gravitational pull of its companion. Hence the stellar core that ultimately exploded would have been deficient in hydrogen. The binary X-ray sources, in other words, may be relics of Type Ib supernovas.

More generally, the possible close relation between SN Ib and SN II events means that the new kind of supernova can offer insight into classical core-collapse explosions. Attempts to understand the process by which core collapse in a massive star can lead to an explosion have consumed many man-years of work and hundreds of hours of computer time, but they have been hampered by a lack of observational detail with which to compare the theory; the thick hydrogen envelope of Type II events masks spectral features from deeper layers. Computer models applied to the direct observations of the supernova interior available in SN Ib explosions should make it possible to glimpse the inner workings of core collapse.

The realization that core collapse can generate a spectrum resembling the spectrum of classical Type I events, with their very different explosion mechanism, has focused attention on the classification of supernovas. Clearly the physics of supernovas are not constrained by any spectral classification scheme. Just as either explosion mechanism can produce a hydrogen-deficient spectrum,

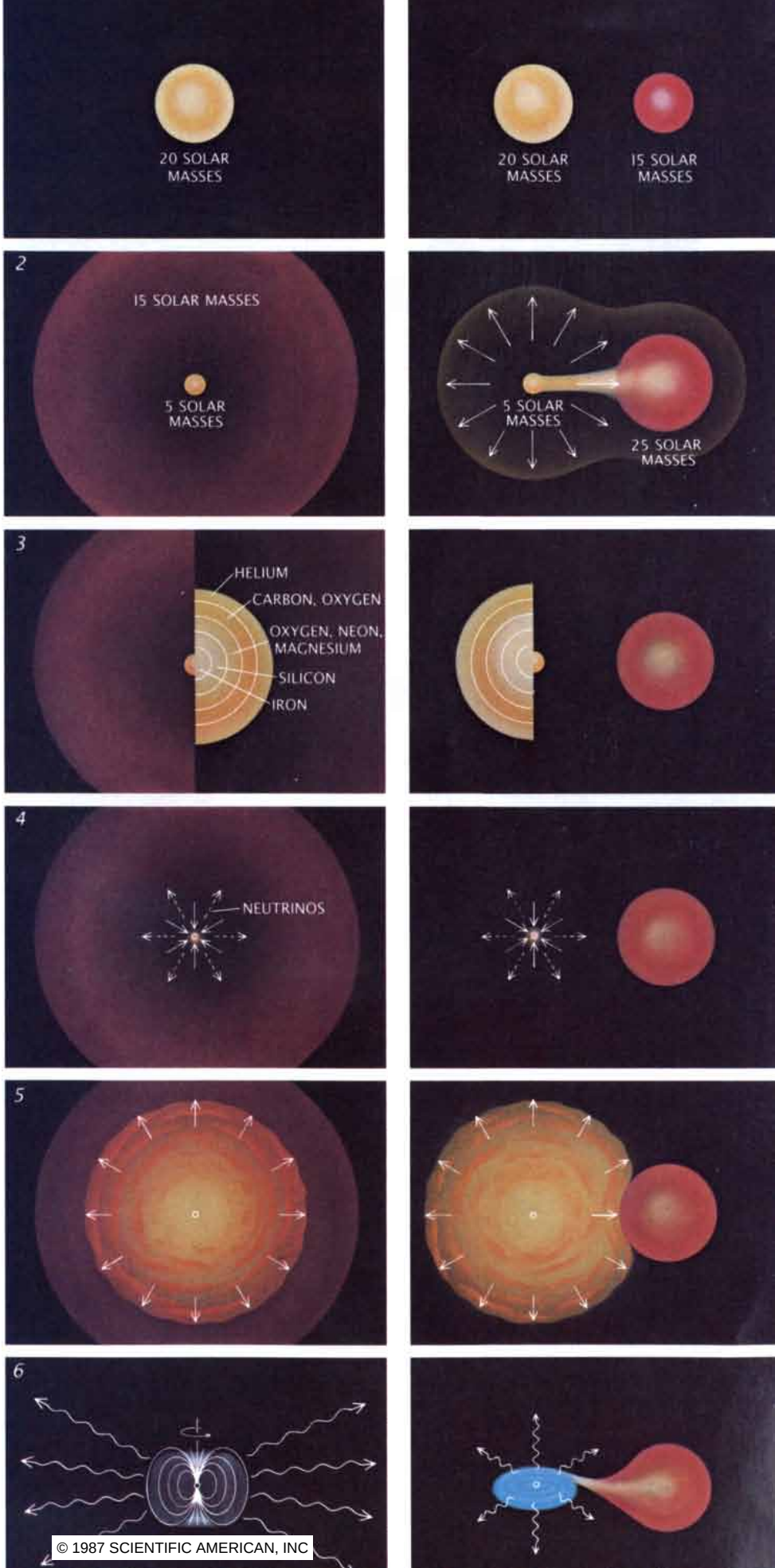


SUPERNEBULAR SPECTRA of two Type Ib supernovas show an emission peak suggesting the presence of oxygen. The element, first seen in the supernebular phase of SN 1985F and later in archival spectra from SN 1983N, probably originates in oxygen-rich inner layers of a progenitor star. The spectra were obtained at the McDonald Observatory by Harriet Dinerstein and Donald Garnett (1985F) and by Martin Gaskell (1983N).

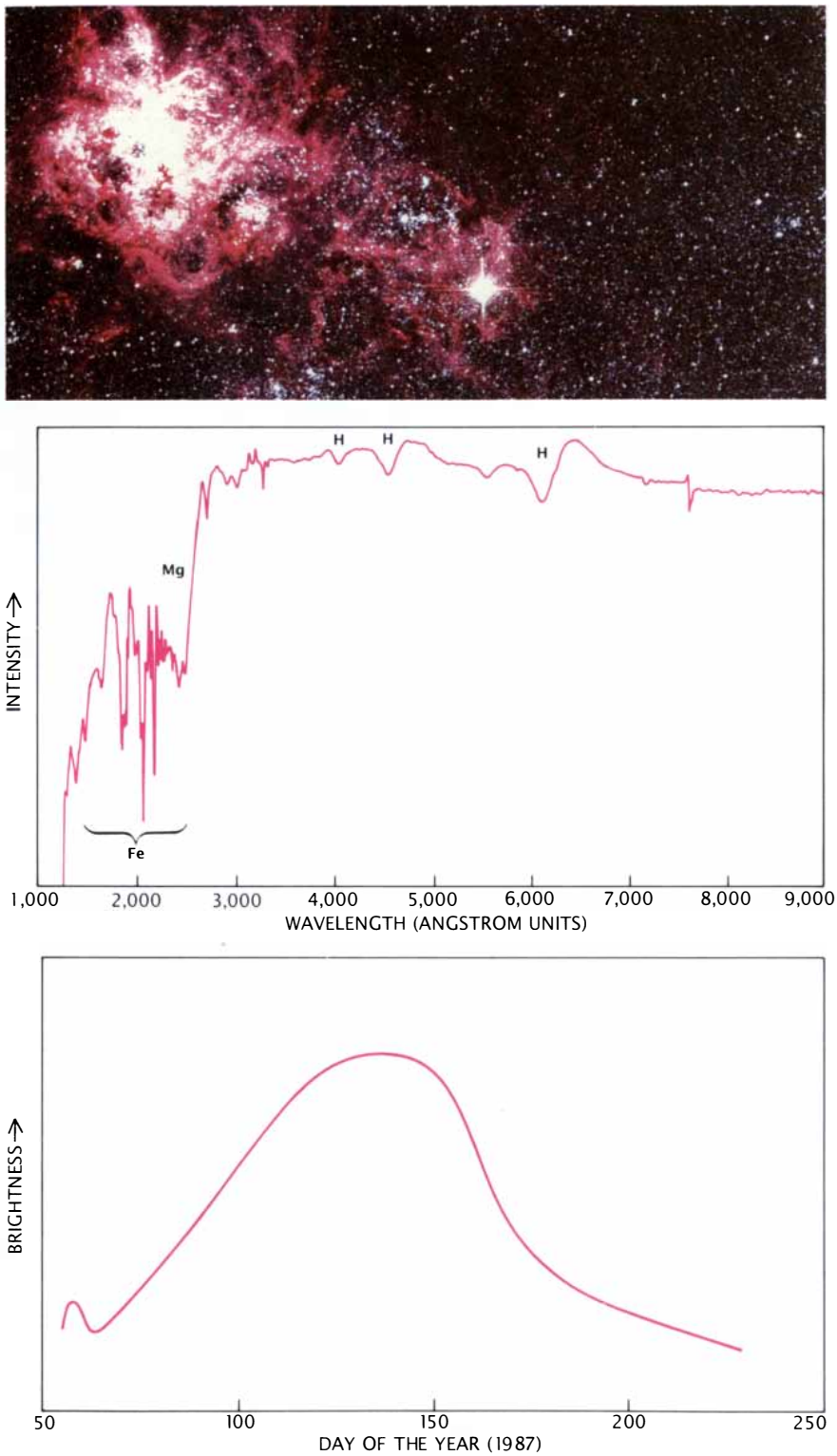
so there are now serious suggestions that some Type II events may represent a completely disruptive thermonuclear explosion within a hydrogen envelope. As our understanding of supernova explosions evolves, the existing classification scheme will inevitably become less important.

An appreciation of this new complexity was emerging when SN 1987A flared in the night sky. Data are still being gathered and feverishly analyzed, but it is already clear that the new event represents the collapse of a massive star in still another configuration. In one sense the explosion provided a spectacular vindication of theory: the burst of neutrinos from the collapse of a stellar core into a neutron star was observed directly for the first time. Detectors in Japan and Ohio—giant baths of water, specially instrumented and placed deep in mines to shield them from cosmic rays—registered the arrival of the neutrino burst just a few hours before the first light from the supernova reached the earth.

In another sense the new supernova presents a theoretical challenge. The spectrum shows clear evidence for hydrogen, which marks the event in general terms as an SN II. Yet a massive star at the end of its short lifetime, when it is susceptible to a core collapse, has usually bloated into a cool, red supergiant. The progenitor star of SN 1987A has been identified from preexisting images as a blue supergiant: a more compact and therefore hotter stage in the career of a massive star, never before



RELATED SCENARIOS may give rise to both Type II supernovas (*left*) and Type Ib events (*right*). A Type II explosion can begin with a massive star (*1*) that has aged and bloated into a red supergiant—a star consisting of an extended hydrogen envelope (*2*) and a layered core of heavier elements (*3*). The core collapses to form a neutron star, releasing neutrinos (*4*) and generating a shock wave that blows off the star's outer layers (*5*). Often the explosion leaves behind a pulsar: a powerful source of radio waves produced as the neutron star spins rapidly within its own magnetic field (*6*). The SN Ib progenitor star may differ only in having a massive binary companion. As the star bloats, some matter is expelled from the system and some is drawn into the companion, leaving a bare core that collapses and explodes just as in a Type II event. In the aftermath of the supernova the neutron star might accrete material from the companion star. The material would emit X rays as it was compressed and heated.



SUPERNOVA 1987A gleams next to a star-forming nebula known as 30 Doradus in the Large Magellanic Cloud (*top*). The supernova's spectrum (*middle*) combines a visible-light spectrum provided by Mark M. Phillips of the Cerro Tololo Inter-American Observatory and an ultraviolet spectrum from the *International Ultraviolet Explorer* (*IUE*) satellite, provided by Robert P. Kirshner of the Harvard-Smithsonian Center for Astrophysics. Hydrogen lines mark the event as Type II. Several features of the event, however, are anomalous, among them the shape of the hydrogen absorption lines, the ultraviolet spectrum and the shape of the light curve (*bottom*). The curve, made by Kirshner and George Sonneborn of the Computer Sciences Corporation from *IUE* data, shows an anomalous decline in brightness soon after the explosion, an unusually slow climb to maximum light and an exponential decline probably driven by radioactive decay.

known to explode. It probably had a mass of between 15 and 20 solar masses—the same mass range expected for SN Ib and some SN II.

The evolution of the explosion itself calls for a compact progenitor: the neutrinos, released at the moment of core collapse, arrived three hours at most before the first light, generated when the shock wave reached the surface of the star. From the lag and the estimated shock-wave velocity, W. David Arnett of the University of Chicago, Stanford Woosley and his collaborators at the University of California at Santa Cruz and groups in West Germany, the Soviet Union, Japan and elsewhere have determined that the radius of the supernova progenitor must have been about 30 times the radius of the sun—some 25 times smaller than a red supergiant.

The compactness of the progenitor probably explains another notable feature of SN 1987A: although it looked extremely bright thanks to its relative closeness, it did not reach the absolute brightness typical of an SN II or SN Ib. In any core-collapse supernova some of the energy deposited by the shock wave goes into expanding the star and making it diffuse enough to radiate efficiently. The smaller the progenitor star is, the more heat is converted into motion and the less energy is ultimately available to make the supernova shine. (SN Ia, which originate in white dwarfs, would hardly shine at all if it were not for the delayed release of heat by radioactive decay.)

SN 1987A dimmed anomalously during the first week after it was sighted and then brightened continuously for 10 weeks, in an unprecedentedly long climb to maximum light. It later settled into an exponential decline in brightness, which is the pattern expected if the source of this late light is the decay of radioactive cobalt and nickel produced in the explosion. Radioactive decay is now suspected of contributing to the brightness not only of SN Ia events but also of SN Ib and some SN II.

One thing is clear: continuing observations of SN 1987A will both alter and extend our comprehension of collapsing massive stars. Still more progress will come from an integration of all our knowledge. Comparisons of Type Ia, Type Ib and Type II events with SN 1987A will yield a wealth of information and a deeper understanding of the myriad manifestations of stellar catastrophe.

ONE GREAT TRUCK LEADS TO ANOTHER.

Drive home a 1988 full-size Sierra Pickup from GMC Truck, and it could be the start of something bigger than both of you.

That's because Sierra, like every vehicle from GMC Truck, offers so much that it demands to be part of a collection of similarly gifted vehicles. And Sierra is a great place to start, with its ample room for three adults, standard V-6 or optional V-8 power and the ability to carry 4' X 8' plywood sheets between the wheelhousings on long-beds.

And Sierra's standard equipment list includes a rear-wheel anti-lock brake system (operates only in two-wheel drive), electronic fuel injection on gas engines, power steering, power brakes and more.

TRAVEL FIRST CLASS.

Safari offers a similar degree of refinement in a very versatile van. Select a Safari with the newly available Touring Package and enjoy luxurious upholstery, rich carpe[ting], ample storage pockets and compartments, and supreme comfort for up to eight passengers with available seating. Here is very grand touring in a very manageable size.

Why not buckle up now and visit your nearest GMC Truck dealer? There's never been a better time to start your GMC truck collection — or a better choice of fine GMC trucks to choose from.

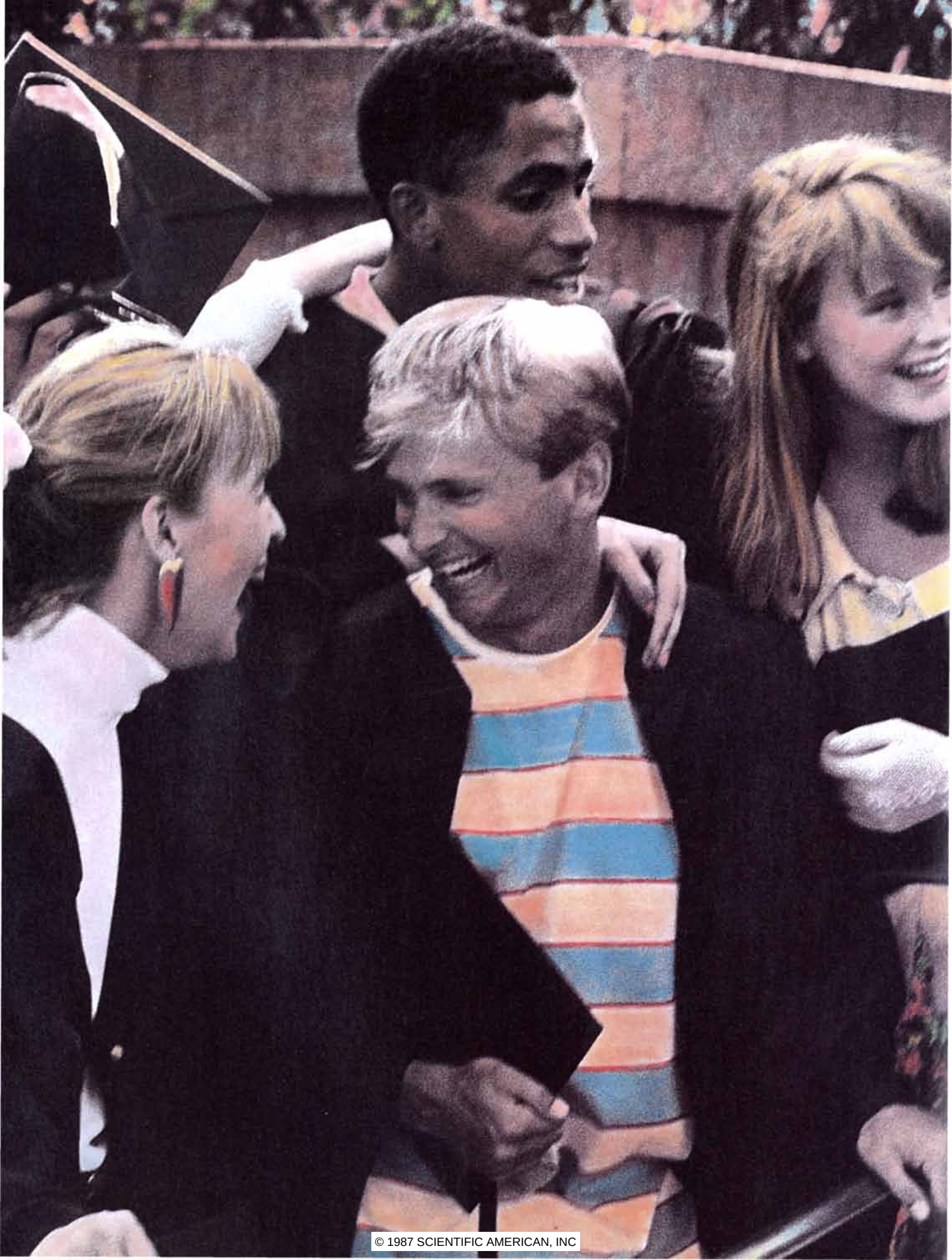
See your dealer for the terms and conditions of this limited powertrain warranty.

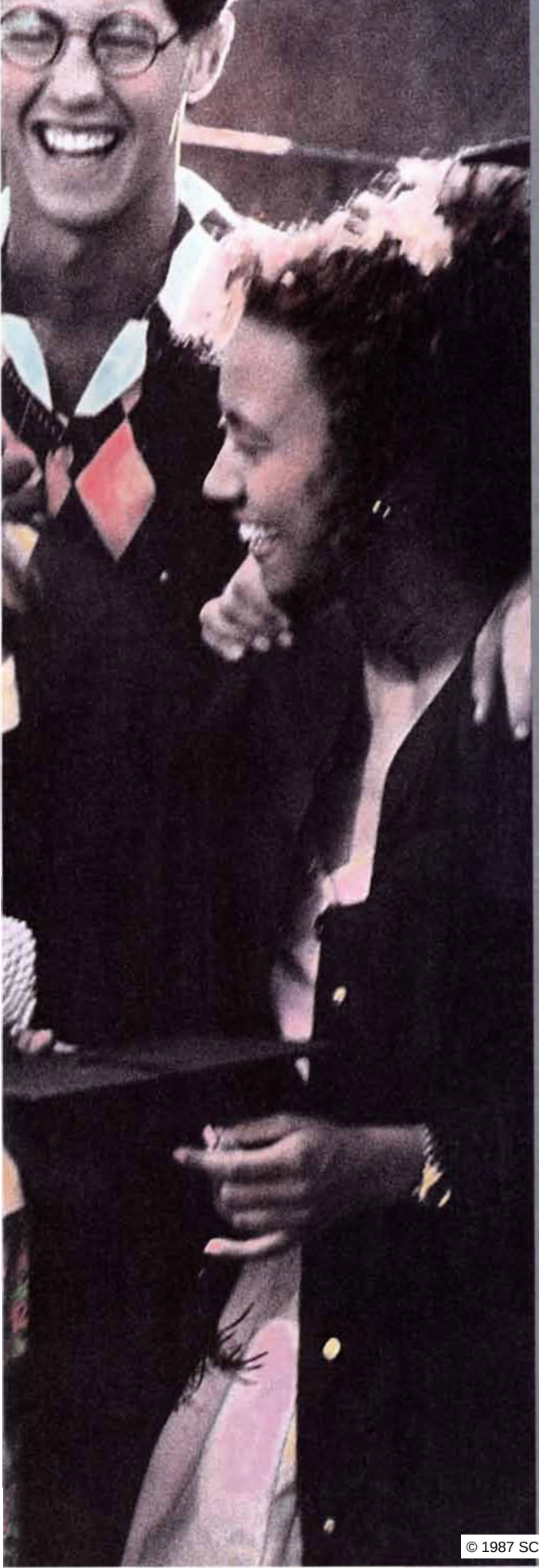
GM 6/60 QUALITY COMMITMENT PLAN

For a catalog and the name of your nearest dealer, call 1-800-VALUE88 (1-800-825-8388).

GMC TRUCK

IT'S JUST A TRUCK ANYMORE.





"You think they look happy? You should see the guys in admissions.

Good as the Bradley name has always been, it's not yet on everyone's lips. That's always made the competition for students and faculty harder for us. But now Bradley's changing the way we educate by changing the way we communicate. We've built a futuristic, electronically unified campus around AT&T voice and data networking. Now offices, labs, libraries and dorms are linked. And as for results, well, an extra 50 members of the class of 1991 can tell you,

AT&T Comes Through."

– Dr. Martin G. Abegg, President
Bradley University

The AT&T Intelligent Campus

In place of multiple phone systems and incompatible computers, Bradley University now has a single fiber optic network, controlled by an AT&T System 85 voice and data switch and an AT&T Information Systems Network. All financed through the AT&T Credit Corporation.

Now mail and messaging for faculty and staff move electronically. Students can tap nine host computers at five locations from their dorm rooms via AT&T PC 6300s. And executives at nearby companies can take graduate courses via AT&T Teleconferencing.

What's more, reselling AT&T WATS on campus reduces payback time on the whole system by a full year.

To learn more contact your AT&T Account Executive or call 1 800 247-1212.

From equipment to networking, from computers to communications, AT&T is the right choice.

AT&T
The right choice.

Artificial Chromosomes

In its infancy genetic engineering was confined to the manipulation of individual genes. Now the same strategies can be used to create whole chromosomes in order to investigate chromosome behavior

by Andrew W. Murray and Jack W. Szostak

The genes of higher organisms are borne on chromosomes, ribbons of DNA much longer than any single gene itself. The cells of human beings have 46 chromosomes each, the cells of a pea plant 14 and those of a fruit fly eight. In this era of genetic manipulation, individual genes are the primary focus of attention and chromosomes may seem to be little more than vehicles for moving genes around during cell division. Actually that is what chromosomes are meant to do. Their task is not simple, however, and their mechanism of action is far from apparent. How do chromosomes know where to go during cell division?

To address that question, we embarked five years ago on an effort to construct chromosomes from their raw materials. Because we were the architects of these molecules, we could deprive them of various functional elements and take note of the chromosomes' behavior in the absence of such elements. We have succeeded in making artificial chromosomes that replicate faithfully in a living cell and that are evenly distributed between the two "daughter" cells when the "mother" cell divides. We have also made many chromosomes that do not distribute well at all, and in the process we have learned a few things about the factors governing chromosome inheritance.

Our interest in chromosome behavior is more than academic. The existence of life itself depends on the accuracy of the genetic transmission mediated by chromosomes. The consequences of errors in chromosome distribution can be disastrous. Most often cells that have inherited the wrong number of chromosomes die, and those that do not die are usually abnormal. The human birth defect Down syndrome, for instance, is caused by an error in chromosome distribution that leaves every cell in the victim's body with an extra copy of chromosome 21.

Artificial chromosomes may also become a useful tool in genetic engineering. Current means of introducing foreign DNA into a cell cannot accommodate large genes or gene clusters. Our artificial chromosomes have a much greater capacity. Already other laboratories have been able to propagate human genes in yeast cells by inserting them in artificial chromosomes.

In order to make artificial chromosomes that function like their natural counterpart, we needed to be able to manipulate the elements involved in chromosome behavior. The raw material of all chromosomes consists of two helical DNA strands wrapped around each other in a kind of spiral staircase. The strands are made up of bases called *A, G, T* and *C*. The bases on one strand pair with the bases on the opposite strand to hold the two strands together, and the pairing is strictly ordered: *A* binds with *T* and *G* binds with *C*. Hence the strands are said to be complementary, and from the base sequence of a single strand one can deduce the sequence of its partner.

Actually that is how the cell is able to duplicate its set of chromosomes prior to mitosis, or cell division. During DNA replication the two strands of the double helix are separated and each acts as the template, or guide, for the synthesis of a complementary strand. In this way two chromosomes arise where only one existed before. Then in mitosis the two identical "sister chromatids" are pulled apart by slender spindle fibers that radiate from opposite poles of the dividing cells. This segregation ensures that each daughter cell contains a full set of chromosomes on completion of cell division.

It is thought that three chromosomal elements are necessary to accomplish replication and segregation. One is the origin of replication, a site on a chromosome at which synthesis of the new DNA begins. Because chromosomes are very long and replication is very slow, there are in fact many origins of replication on a single chromosome and many enzymes work in concert to produce the sister chromatids. Another element necessary for proper chromosome function is the centromere, a segment of DNA to which the spindle fibers attach during mitosis. The centromere is often but not always at the center of a chromosome. Finally, repetitive DNA sequences at each end of a chromosome are needed to protect terminal genes from degradation; these are called telomeres (from the Greek *telos*, meaning "end").

In this article we shall discuss recent progress in understanding how these elements contribute to the faithful inheritance of chromosomes. Most of the relevant studies were done in yeast, a single-cell organism that is easily manipulated and in many respects resembles the cells of higher organisms. Inheritance errors are rare in yeast, occurring only once in every 100,000 cell divisions.

Yeast is also the only organism known so far in which the centromeres, telomeres and origins of replication are short, well-defined segments of the DNA molecule. Yeast cells have 16 chromosomes that vary in length from 300,000 to more than two million base pairs. (Human chromosomes are much larger: their average length is some 100 million base pairs.) The functional elements of yeast chromosomes range in length

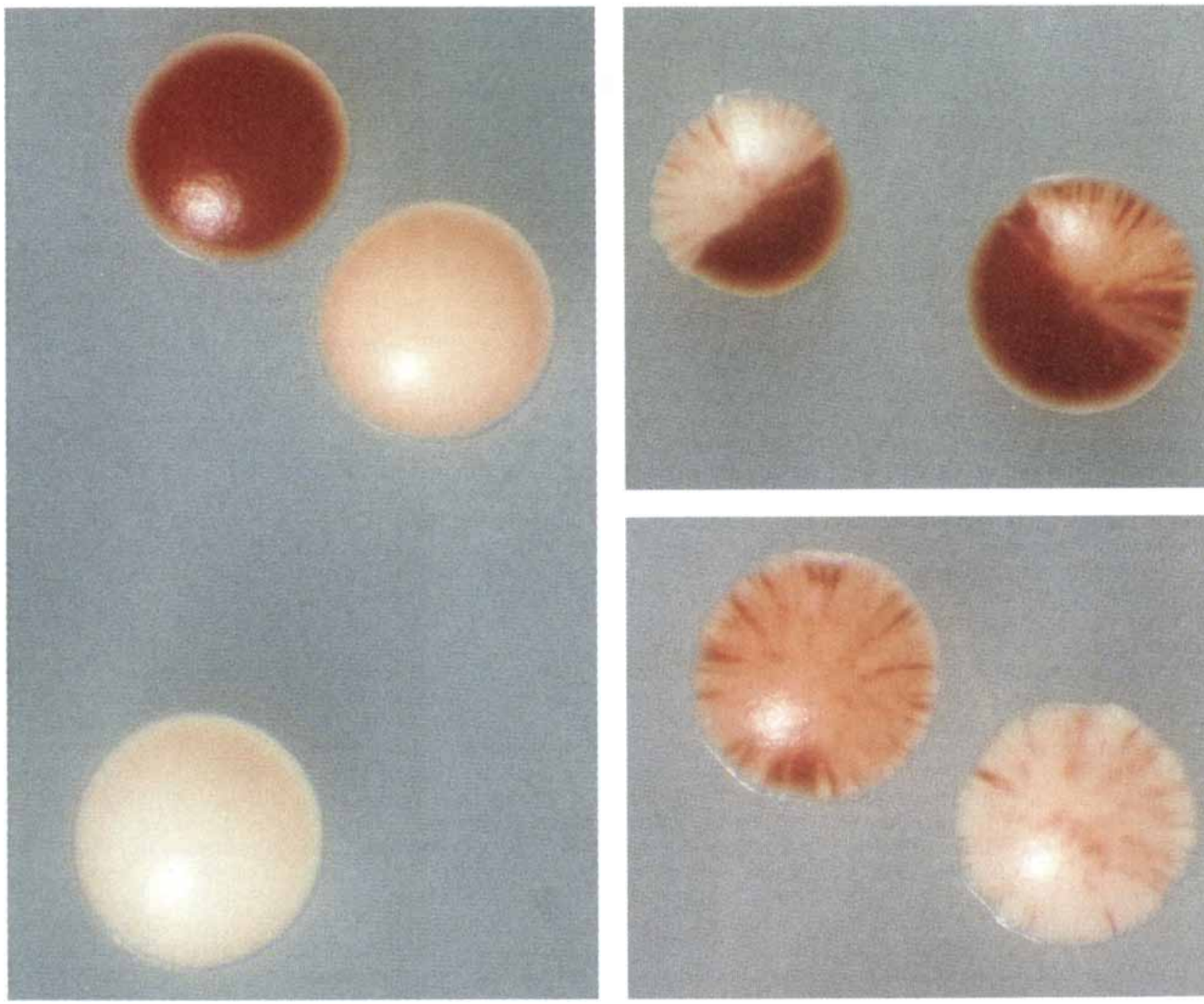
from 100 to 1,000 base pairs. In contrast, the centromeres of other organisms (guessing from what little is known about them) may be as many as one million base pairs long.

How were these chromosomal elements harnessed for our experiments? Origins of replication were the first to be cloned. Eight years ago Kevin Struhl, Dan T. Stinchcomb, Stewart Scherer and Ronald W. Davis, then all at Stanford University, found they could get pieces of yeast DNA to replicate independently of the yeast cell's own chromosomes if the pieces were combined with certain other segments from yeast chromosomal DNA. The DNA sequences that confer the ability to replicate free of the chromosomes are probably the same sequences that act as origins of replication on chromosomes.

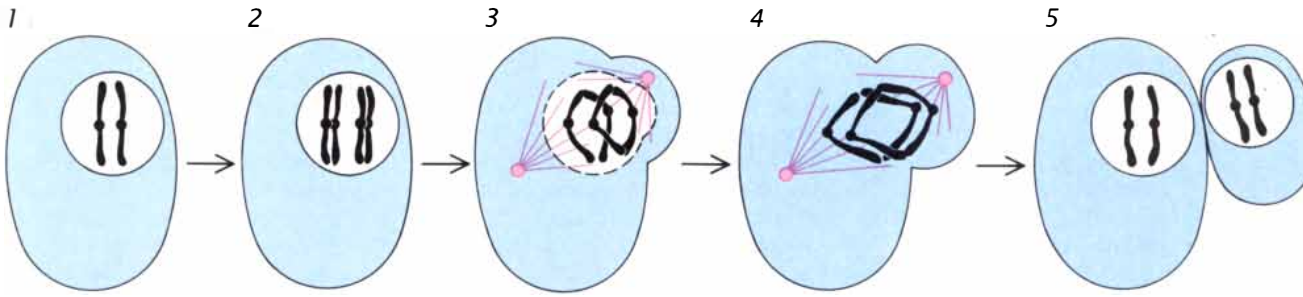
DNA molecules that can replicate freely are known as plasmids. Most plasmids, including those the Stanford investigators constructed, are circular, rather than linear like most chromosomes. In spite of their ability to replicate, these plasmids do not behave like normal chromosomes when it comes time to segregate. Often both copies of the plasmid are retained in the mother cell during mitosis. This phenomenon reinforces the idea that additional chromosomal elements are necessary to direct the inheritance of DNA molecules after replication.

In 1980 Louise Clarke and John A. Carbon of the University of California at Santa Barbara cloned the DNA that constitutes the yeast centromere. The centromere sequences exert a remarkable effect on plasmids into which they are inserted: such plasmids segregate correctly in 99 percent of cell divisions. The centromeres from 12 of the 16 yeast chromosomes have now been cloned and studied in detail, and their sequences are quite similar. It is thought these regions are the sites of attachment for microtubules, tiny rods that make up the spindle fibers. Chromosomes without centromeres probably cannot attach to the mitotic spindle and will therefore segregate randomly at best (that is, half of the time sister chromatids will segregate to different cells and half of the time they will stay in one cell).

Telomeres were the last chromoso-



TRICOLOR YEAST COLONIES reveal the inheritance of artificial chromosomes. Because the chromosomes contain a gene inhibiting pigmentation, cells with one copy are pink, cells with two copies are white and those with no copies are red. Each colony of 10 million cells is derived from a single ancestral cell. Hence if artificial chromosomes are properly meted out during cell division, all the cells in a colony will be the same color (*left*), but errors in inheritance produce sectors of color (*right*).



MITOSIS is the process by which one cell divides into two identical daughter cells. The behavior of chromosomes in mitosis includes replication to produce sister chromatids (*1, 2*), which segregate into separate cells (*3–5*). Segregation is accomplished with the aid of spindle fibers (*red*) that radiate from the poles of a dividing cell and pull sister chromatids in opposite directions.

mal element to be cloned from yeast; their terminal position precludes manipulation by standard methods of genetic engineering. For many years natural telomeres were studied in a type of protozoan in which the elements are particularly abundant. In 1982 one of us (Szostak) and Elizabeth H. Blackburn of the University of California at Berkeley devised a way to use these protozoan telomeres to build a DNA molecule that could replicate in yeast cells and remain linear. Soon after that we extended our technique to clone the telomeres of yeast. At about the same time Ginger M. Dani and Virginia A. Zakian of the Fred Hutchinson Cancer Research Center in Seattle also reported cloning yeast telomeres.

Because telomeres are necessary to preserve the integrity of a linear chromosome in a cell, the cloning of telomeres allowed us to consider making an entire artificial chromosome. We could now manipulate all the elements that were thought to control chromosomal replication and segregation; with these elements at our disposal the construction was fairly simple. In 1983 we made our first chromosome. It was quite short: with only 11,000 base pairs, it was less than 5 percent as long as the shortest natural yeast chromosome.

Construction of an artificial chromosome proceeds stepwise, beginning with a bacterial plasmid that is commonly used as a cloning vector. At each step we cut the plasmid DNA with enzymes, creating a gap into which we splice a gene or one of the three chromosomal elements. First we insert genes conferring distinctive qualities that will indicate which yeast cells have picked up the chromosome; these genetic signals may turn yeast cells a certain color or may allow the cells to grow in the absence of some normally essential nutrient. Then we add origins of replication and a centromere. Finally two fused telomeres are spliced into the plasmid. When the plasmid is put into a yeast cell, enzymes clip the fused termini into two functional telomeres, turning the circular molecule into a linear one.

What did we hope to learn from the behavior of the artificial chromosome? We had three objectives. First, we wanted to know whether centromeres, telomeres and origins of replication are the only specialized DNA sequences required for normal chromosome behavior. Second, we wanted to learn more about the mechanism by which the mitotic spindle separates chromosomes during cell division.

Finally, we wanted to learn how errors in segregation arise for both artificial and natural chromosomes. There are two types of segregation error. One type, called 1:0 segregation, results in one daughter cell's having one copy of the chromosome and the other's having none at all. This error can occur if the chromosome fails to replicate or if one of the sister chromatids gets degraded. The other type of error, 2:0 segregation, results in both sisters' going to the same daughter cell. The 2:0 errors are also called nondisjunction errors, because the sister chromatids fail to disjoin during mitosis.

The most convenient way to distinguish between these two types of error involves an assay based on the color of yeast colonies. When we construct our DNA vector, whether plasmid or chromosome, we include a gene that suppresses the production of a red pigment manufactured by some yeast strains. A cell that does not inherit the chromosome will retain its natural red color, but if a cell contains one copy of the gene (and therefore of the chromosome), it will be pink. If a cell inherits two copies, it is white. Assuming we start with pink yeast cells bearing one copy of the chromosome, a 2:0 segregation error will produce a white cell and a red cell; a 1:0 segregation error will result in a pink cell and a red cell. Two pink cells result from normal segregation. This assay was developed by Philip Hieter of the Johns Hopkins University School of Medicine and Carl Mann, Michael Snyder and Davis of Stanford. Douglas Koshland of the University of California at San Francisco and John C. Kent and Leland H. Hartwell of the University of Washington devised a similar system.

We can quantify the frequencies of error types by growing yeast colonies from single pink cells and observing the color patterns these colonies exhibit. If an error occurs in the very first cell division only, half of the colony will be one color and the other half another color. If errors occur in subsequent divisions, the round colonies will be splintered into different hues like a pie cut into pieces of three different colors. Thick "slices" indicate that a segregation error happened early in the colony's history and was perpetuated; thin slices of color result from later errors. In other words, colonies with few segregation errors have few segments and colonies in which errors are frequent look highly segmented.

We have used this and other assays to compare the segregation accuracy of plasmids, artificial chromosomes and natural chromosomes. The behavior of our artificial chromosomes surprised us. Unlike circular plasmids with centromeres, our small linear chromosomes segregated randomly. This unexpected result left us with two questions: Why did linear molecules fail to segregate

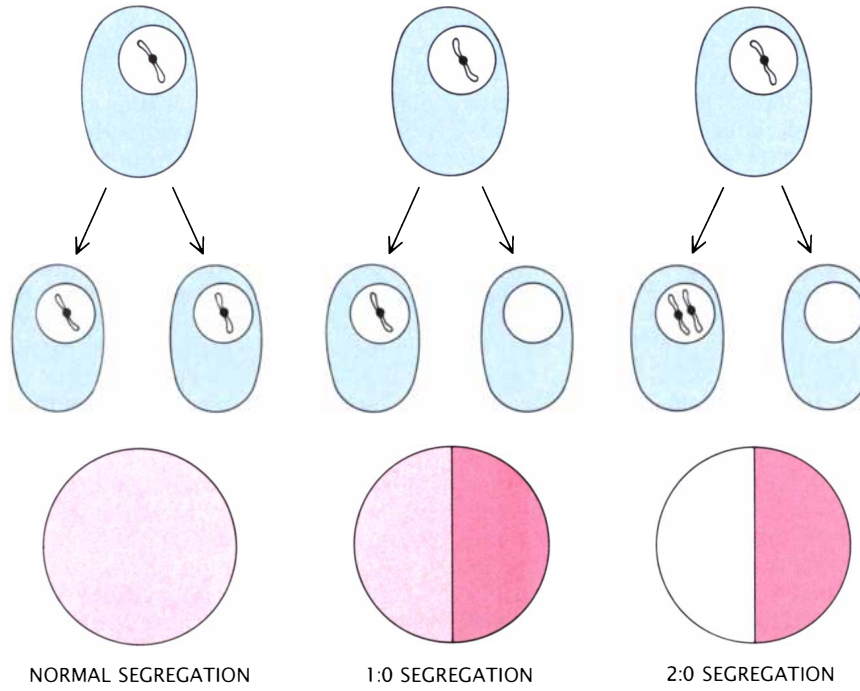
correctly where circular molecules had succeeded in doing so? And what changes were necessary to make linear artificial chromosomes behave more like natural chromosomes?

We tested a number of specific ideas that could account for the poor segregation of our artificial chromosomes. At first we thought the telomeres, which were derived from protozoa rather than yeast, might be at fault, but we found that they functioned exactly the way yeast telomeres do. We thought the centromere might have been damaged during cloning, but it turned out to be intact. We were left with the explanation that the topology of the linear molecule itself must somehow be responsible for the difference in segregation patterns. In spite of the fact that our chromosome was merely a shorter version of a natural chromosome, it segregated more poorly than a circular facsimile. Could it be that the chromosome's length itself was the problem?

We examined this possibility by constructing an artificial chromosome 55,000 base pairs long. The extra DNA came from a bacterial virus we used as the foundation for the larger chromosome; we assume this DNA is inert in yeast. In all other respects the molecule was the same as our 11,000-base-pair chromosome.

The longer artificial chromosome shows errors in segregation in only 1.5 percent of cell divisions. Of these errors 90 percent are due to chromosome loss (1 : 0 segregation) and only 10 percent to nondisjunction (2 : 0 segregation). This frequency is still 1,000 times greater than that of natural chromosomes, but when we later doubled the length of the chromosome to 104,000, the frequency of segregation errors fell to .3 percent. Hieter and his colleagues report similar findings in their laboratories. These data clearly demonstrate that chromosome length exerts a strong influence on the accuracy of chromosome segregation.

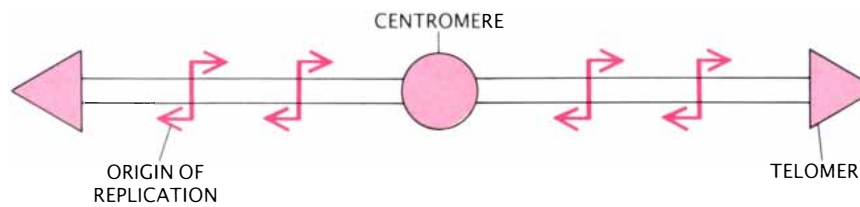
How could length be involved? We entertained two related possibilities. Most obviously, the overall physical length of the molecule could in itself have some bearing on the mechanism of segregation. Alternatively, the separation between the centromere and the ends of the molecule, which would vary with chromosome length, might affect segregation. In order to test the latter possibility we modified the structure of a natural yeast chromosome so that its centromere, which is normally situated at the center of the molecule, was shifted toward the end. The alterations reduce the accuracy of segregation only marginally; the chromosome still segregates 10,000 times more accurately than short artificial chromosomes do.



CHROMOSOME SEGREGATION determines the color of the yeast colonies shown on page 63. The founding cell in each colony has just one copy of an artificial chromosome and is therefore pink. If during mitosis that chromosome replicates and segregates normally, both daughter cells will receive one copy and will also be pink (*left*). Sometimes, however, a sister chromatid is not synthesized or is lost (*middle*). This 1 : 0 segregation error gives rise to a pink cell and a red cell. If no errors occur in subsequent divisions, half of the resulting colony will be pink and half will be red. In 2 : 0 errors (*right*) both copies of the chromosome go to the same daughter cell; that daughter is white and the other is red. Again the colony will be colored in halves if no errors occur after the first cell division. When additional segregation errors do occur, they show up as smaller segments of color. Both halved and highly sectored colonies are pictured on page 63.

It appeared that overall length was the determining factor. If it was, we reasoned, short fragments of natural chromosomes ought to behave just as poorly as artificial chromosomes during cell division. We were able to produce such fragments and found that the accuracy of segregation did indeed decline as the length of the fragments decreased. The segregation accuracy of the fragments starts to drop below that of full-length chromosomes at sizes between 100,000 and 150,000 base pairs.

Although the total frequency of segregation errors is lower for fragments of natural chromosomes than it is for artificial chromosomes of comparable length, most of the dif-



FUNCTIONAL ELEMENTS of chromosomes are specialized DNA sequences essential to chromosome performance. The centromere is the site of attachment for the spindle fibers; telomeres protect the ends of the chromosome from degradation. Origins of replication, where DNA duplication begins, are scattered throughout the chromosome.

ference can be accounted for by a higher incidence of 1 : 0 segregation errors in the artificial vectors; the frequency of 2 : 0 errors is similar. We assume these findings arise from a preponderance of replication failures in the artificial chromosomes rather than failures during cell division. Hence the overall length of a chromosome rather than some secondary feature affected by length does seem to be a crucial determinant of segregation accuracy.

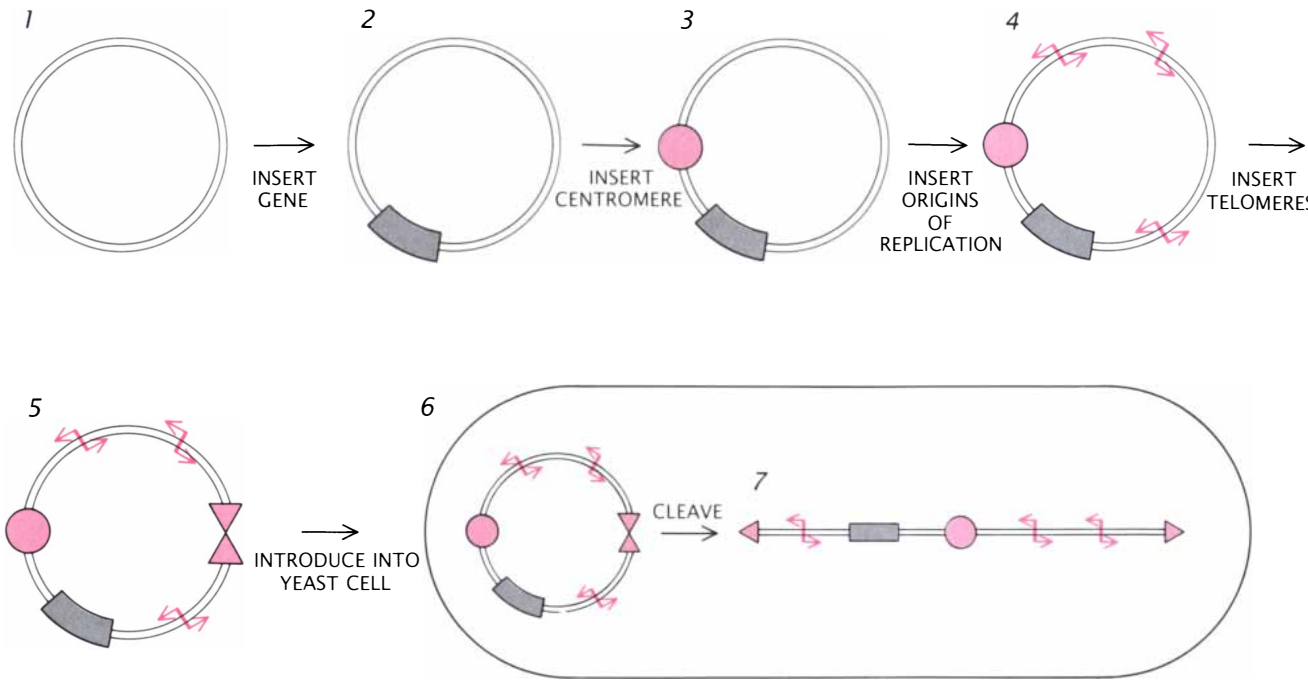
We have outlined a general model to explain the effects of length and topology on the properties of artificial chromosomes. The model rests in part on the previous observations of investigators regarding the forces that control chromosome movement during cell division. In the 1970's R. Bruce Nicklas and his collaborators at Duke University did an elegant series of experiments demonstrating that the attachment of the centromere to the spindle fibers is stable only under tension. In other words, if there is no force resisting the pull of the fibers, the centromeres essentially fall off the spindle apparatus. The force arises because sister chromatids are somehow bound together.

No one is quite certain what holds the two chromatids together, but we have proposed a tentative model. We believe the linkage is caused by the catenation, or intertwining, of the two double helixes of the sister chromatids. Catenation occurs in many places along the chromatids during replication, when enzymes that are supposed to keep DNA from tangling get squeezed out by the synthetic machinery [*see illustration on page 68*]. The regions of catenation probably resist separation during mitosis, thereby providing the opposing force that stabilizes centromere attachment to the spindle fibers. Later in mitosis another enzyme allows one double helix to pass through the other so that the two chromatids can eventually segregate—but not, perhaps, until the stability of chromosome attachment to the mitotic spindle has been established.

We think catenation might be the operative physical linkage between sister chromatids because it so neatly explains our observations with artificial chromosomes. Plasmids, being circular, always give rise to two linked sisters that cannot drift apart; hence no matter what size a plasmid is, its sister chromatids will resist separation during mitosis. In contrast, catenation between a pair of linear molecules can be resolved simply by twisting one molecule around the other. Even in the absence of enzyme activity two linear sister chromatids can become disentangled, and if they are detached, they will not provide each other with the tension required to accomplish accurate segregation.

The shorter a chromosome is, the more easily its sister chromatid unwinds and the less likely the molecules are to segregate properly. Likewise, if a chromosome fails to be replicated or if one member of a pair of sister chromatids is lost, the single chromatid will have no opposing partner and will not be able to attach to the mitotic spindle. Both scenarios agree with our observations of 1 : 0 and 2 : 0 segregation errors.

Our research has also suggested that cells have mechanisms for monitoring the success of chromosomal segregation. We have found, for instance, that in yeast strains carrying a short artificial chromosome many dividing cells look morphologically abnormal. In such divisions the daughter cell grows to the same size as the mother before the two cells separate; normally the daughter cell pinches off from the mother when it is roughly half the mother's size. When we examine these strains for cell divisions that involve segregation errors, we see that more than 90 percent of such divisions display this kind of abnormality. We believe the unusual appearance of these cells reflects the fact that they have been



CONSTRUCTION OF AN ARTIFICIAL CHROMOSOME begins with a circular piece of DNA called a plasmid (*1*). The authors insert an indicator gene such as one affecting pigmentation (*2*), then add functional elements one by one (*3–5*). After inserting two fused telomeres they put the plasmid into a yeast cell (*6*). There the telomeres are cleaved, making the molecule linear (*7*).

Using new technologies, an advanced solid-state laser prototype has been produced that is more efficient and more readily scaled from low to high power than currently available models. The Hughes Aircraft Company-built prototype uses optical phase conjugation, ensuring that all light waves emitted are in phase, compensating for aberrations and distortions in a laser beam. Also, the new laser material used, co-doped gadolinium scandium gallium garnet, approximately doubles the efficiency and energy storage capacity of the laser. A follow-on contract has been awarded to Hughes for the second and third stages of the U.S. Air Force's Medium Energy Source (MES) program. Future applications of the new laser include communications, range finding, and target designation.

The United Kingdom Infrared Telescope (UKIRT) in Hawaii was the first to use a new infrared focal plane array, which has caused a technological revolution in infrared astronomy. The Hughes-built microchip "sandwich" provides sharp, fast infrared images of our solar system and the galaxies. Astronomers can now obtain a better look inside mysterious clouds of dust and gas, known as nebulae, to learn more about the life cycle of stars. The array also produces, for the first time, fine-grain infrared images of objects within nebulae that were previously hidden.

A processor utilizing advanced microchips will offer processing power equal to 200,000 desktop personal computers. This new programmable signal processor (PSP) uses very large scale integrated (VLSI) circuits called gate arrays to pack this power into a 65-pound box, which measures just one cubic foot in volume. By comparison, these 200,000 personal computers would weigh roughly 3.4 million pounds, without any add-on memory cards, disk drives, or monitors. If stacked 20 units high, the 200,000 personal computers would fill a room approximately 144 x 120 feet. The PSP is being built by Hughes for the APG-70 radar system to be used aboard the U.S. Air Force F-15.

The nation's newest weather satellite has completed the U.S. weather-watch system. The Geostationary Operational Environmental Satellite (GOES) H, launched into orbit over the Atlantic seaboard, provides data for meteorologists to predict and monitor storm fronts threatening the East coast. Designed and built by Hughes for the National Oceanographic and Atmospheric Administration, GOES H carries two experimental payloads: a space-environment monitor (SEM) for solar-wind measurements, and a receiver designed to aid international search and rescue missions.

Hughes Research Laboratories needs scientists for a spectrum of long-term sophisticated programs, including: AI knowledge-based systems, planning, and computer vision; space plasma sources; pulsed power switches; free electron lasers; applications of focused ion beams; masked ion beam lithography; liquid-crystal materials and displays; nonlinear optics and phase conjugation; submicron microelectronics; computer architectures for image and signal processors; GaAs device and integrated circuit technology; optoelectronic devices; infrared detection and superconductivity applications. Send your resume to Professional Staffing, Hughes Aircraft Company-Research Laboratories, 3011 Malibu Canyon Road, Malibu CA 90265. Equal opportunity employer. Proof of U.S. citizenship required for some positions.

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

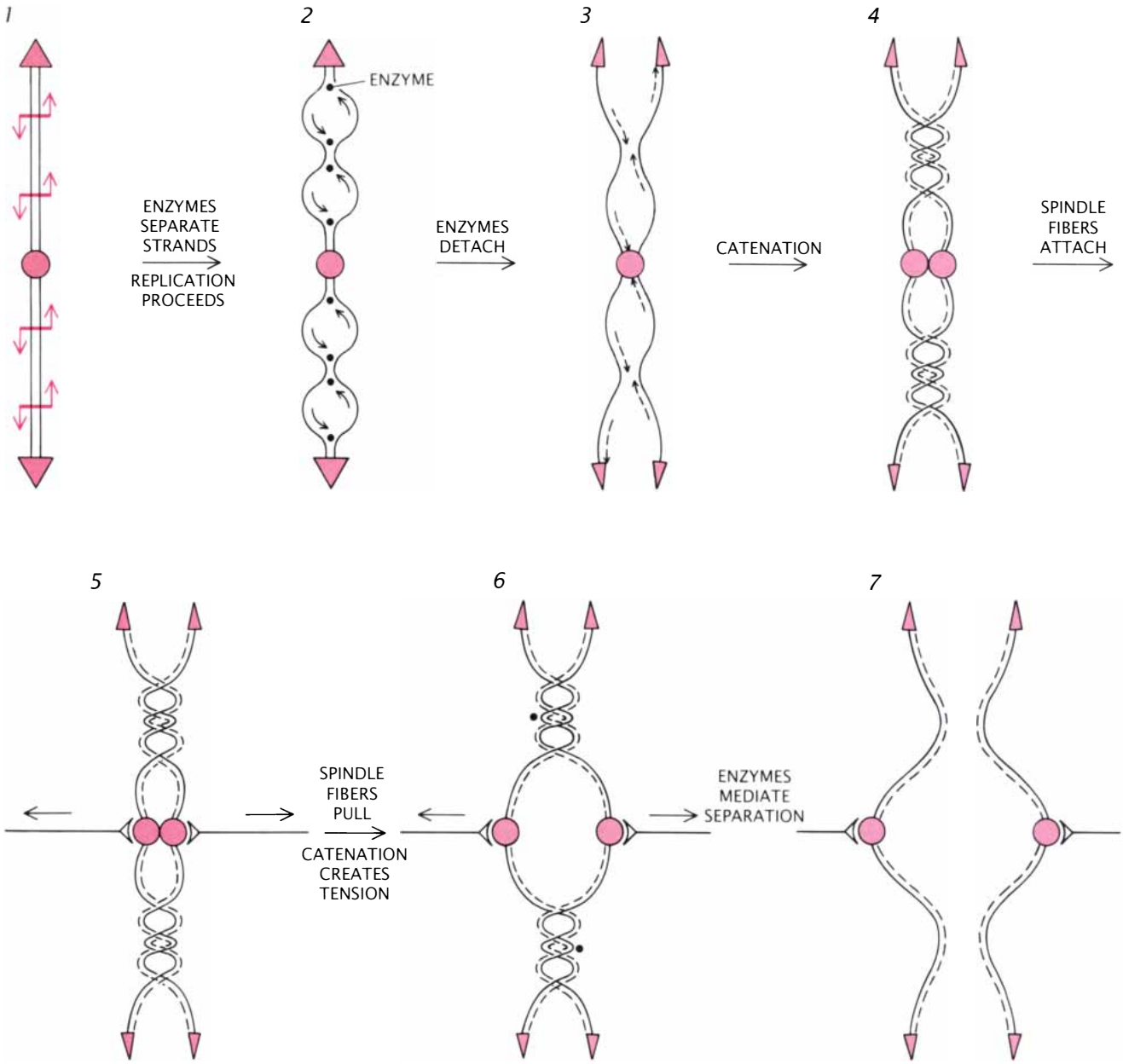
HUGHES
AIRCRAFT COMPANY

Subsidiary of GM Hughes Electronics

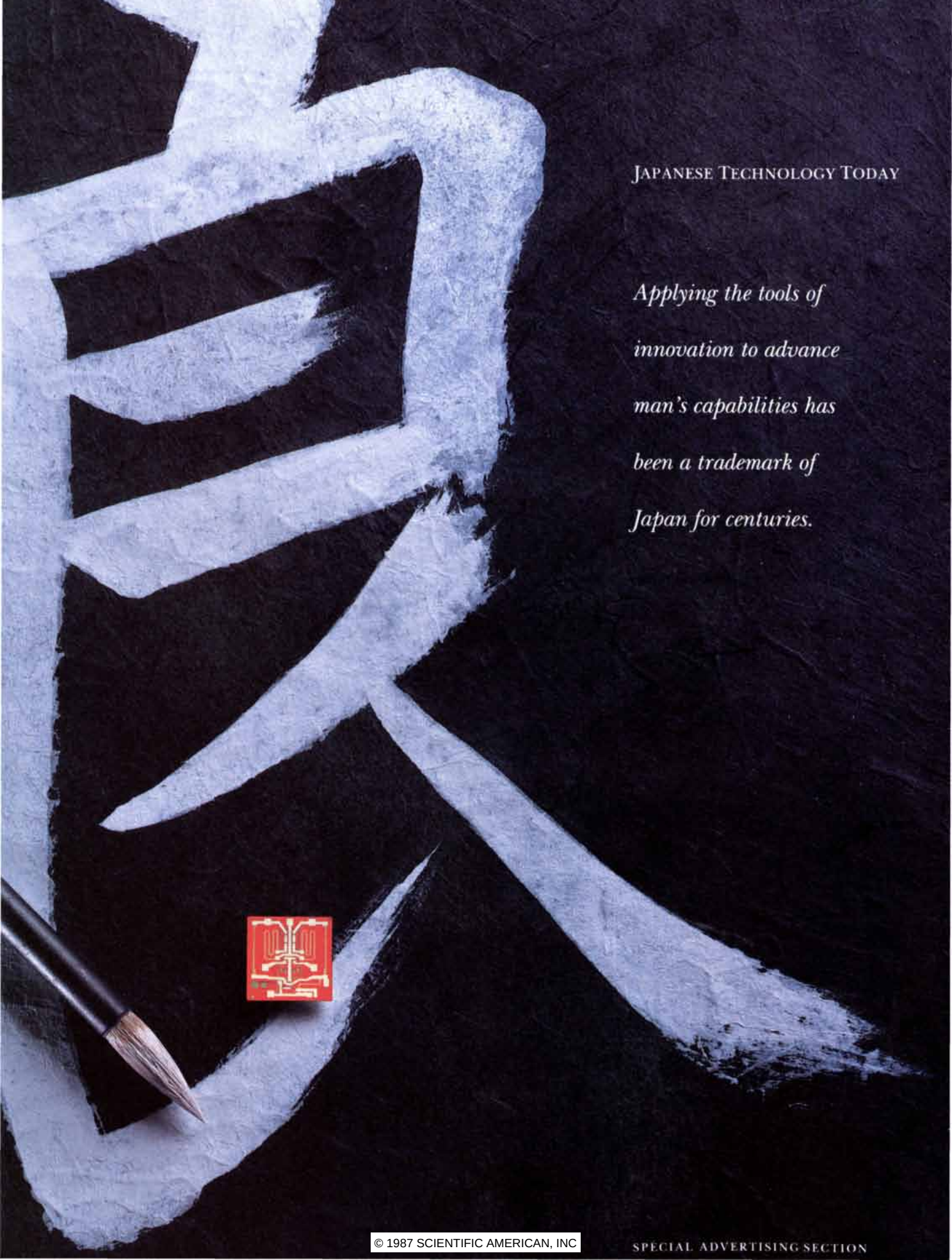
transiently arrested in mitosis. The cells detect a malfunction in segregation and delay the completion of mitosis, waiting for the problem to correct itself. Meanwhile the daughter cell continues to grow.

Artificial chromosomes have been valuable assistants in our efforts to probe the factors that govern chromosome behavior during mitosis. We expect to find out even more as the techniques for constructing them improve and expand. We and other workers are using artificial chromosomes to explore meiosis, the type of cell division that gives rise to germ cells (egg and sperm) in higher organisms. Such studies could increase the understanding of genetic disorders that are linked to aberrant chromosome inheritance, such as Down syndrome.

Artificial chromosomes have already been constructed for cloning human genes in yeast; the cloning was the work of David T. Burke, Georges F. Carle and Maynard V. Olson of the Washington University School of Medicine. Their chromosomes merely serve as vehicles for preserving and multiplying human genes, which can then be transferred to other cells for experimentation. Burke and his colleagues expect artificial chromosomes to be able to carry genes exceeding 50,000 base pairs in length, thereby filling the need for a vector that can accommodate the large genes and gene clusters that play important roles in the genetics of humans and other animals. It is not unlikely that in this respect artificial chromosomes will someday provide genetic engineers with a new degree of flexibility and control.



CATENATION MODEL may explain why sister chromatids resist segregation. Catenation, or intertwining, occurs during chromosome replication (*broken line*). Cellular enzymes that unwind the replicating strands get pushed out when two sites of synthesis (*arrows*) approach each other (*1–3*), leaving the chromatids twisted together (*4*). When spindle fibers attach (*5*) and start pulling, this catenation creates tension (*6*). Other enzymes transiently cut one of the chromatids so that the two are able to separate (*7*).



JAPANESE TECHNOLOGY TODAY

Applying the tools of innovation to advance man's capabilities has been a trademark of Japan for centuries.

"Accords have traditionally been some of the nicest cars to sit in.

They are also satisfying on the move. The engine seems to hum sweetly. Cruising is hushed. The steering, deliciously accurate. And the shifter carves to perfection.

The Honda Accord is world-class comfortable."

–Car and Driver

It's enough to make us blush.

The Accord LX HONDA

Japanese Innovation: A Personal View
By Akio Etori

Mutual understanding and a spirit of concession are necessary to settle international matters. There are many difficulties which Japan must consider. One problem: the international inequality of systems and practices in export and import regulations, customs tariffs and patent systems. If there are differences between countries in these matters, friction is unavoidable.

At present, there is a race among countries in the development of superconductors. The number of dissertations concerning new research results, and the number of patent applications are increasing. There is, however, a difference between the patent systems of the United States and Japan; this impairs the desired equality of international competition. For example, the period of time that elapses from filing a patent application to granting the patent is short in the USA, long in Japan; and the patenting of ideas is approved in one country but not the other.

Japan's traditional international trade policy has been one of "small profits, quick returns." Carried to excess, this policy could result in insupportable pressure on foreign manufacturers of comparable products, with consequent economic and social disruption. For Japan to assure responsible membership for itself in the international community, particularly in light of the emergence of groups of developing and newly industrialized nations, such excess must be avoided. Japan's alternative is development and sale of goods that can be produced in smaller quantities and sold at a larger profit, products of high quality and a high level of production sophistication.

Reaching toward innovation, with its increased focus on individual attainment, has not been easy for the Japanese; as Dr. Motokazu Uchida, a senior manager at Hitachi, has said, "We tend to believe that 'a tall tree catches much wind'." Japan is a monoracial country with strong national unity. The education system has enhanced this unity, and many Japanese share similar perceptions and life-styles. The employment system based on seniority and lifetime employment has strengthened the sense of solidarity of Japanese workers.

For future markets and future industrial vitality, however, strong individual creativity is indispensable. Japan began the process of scientific research much more recently than Western countries, so science and technology are still solidifying as perceived cultural assets. The process is continuing. ◆

Automotive Technology: Four-wheel Steering

Until recently, Japanese automotive technology led in process and quality assurance areas while tending to follow Europe in adopting such features as twin cam engines, full-time four-wheel drive, and antilock braking. But in 1987, a new technology has distinguished Japan's auto industry: four-wheel steering, a completely new approach to improving the performance and safety of cars. Nobohiko Kawamoto, president of Honda Research and Development and director of Honda Motor Company, explains: "Our basic philosophy is to make all four tires do some work in every occasion; that is why we offer four-wheel, antiskid braking and four-wheel drive on some models."

With front-drive cars, front wheels and tires normally do a disproportionate amount of work—in carrying load, in transmitting torque to the road, in steering and in braking. Since all Honda cars are front-wheel drive, Honda had a strong incentive to make the rear wheels do more work. When Honda started research into four-wheel steering ten years ago, it hoped to reduce the turning circle, thus increasing the utility of Honda cars. "If the rear wheels are turned parallel to the front ones when steering at high speed, stability is much improved," Kawamoto states. He adds, "In hard cornering, the fact that all four wheels are steered results in the cornering forces being more uniformly distributed between the front and rear tires."

Applying the tools of innovation to advance man's capabilities has been a trademark of Japan for centuries.

Today's Japanese technology continues this tradition. Intensive research in trench cells in chip technology is just one field of exploration in the following report on technological contributions from Japan.

Cover Photo: Karl Steinbrenner

The Authors

The 1987 edition of "Japanese Technology Today" was written by John Hartley and Akio Etori.

Mr. Hartley, trained as an engineer, is an industrial journalist. His most recent book, *Fighting the Recession in Manufacture,* was published in 1986. Mr. Etori is an award-winning and respected writer on science and technology in Japan, having covered these areas for over 20 years.

Conceptually, the process works as follows: When a conventionally steered car is turned at high speed by a crosswind, by a bump in the road or by driver action, the car changes direction, that is, it assumes a yaw angle relative to the direction of the road. The driver must strive to pull it back onto its original line. With four-wheel steering the car can move sideways, with its body still parallel to the direction of the road. Result: little driver effort needed to bring the car back on line.

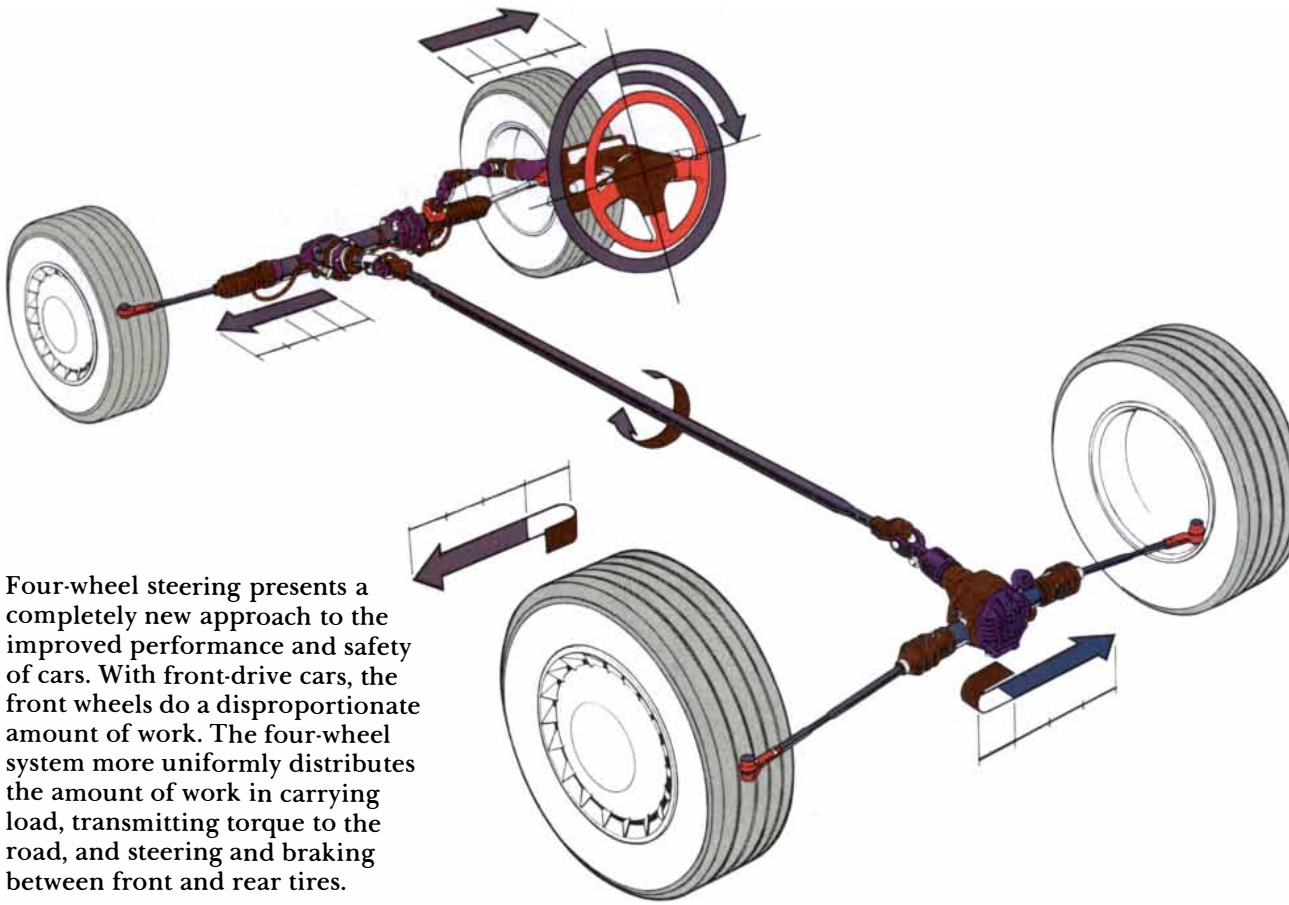
Four-wheel steering does not mean the rear wheels turn as much as those in front. They turn through small angles—up to about five degrees in the Honda system—and the front wheels do most of the work. The rear wheels comprise a neat mechanism: Kawamoto explains that the new Honda Prelude is designed so that when the steering wheel is first turned, the rear wheels turn in the same direction. But as the steering wheel is turned more in the same direction, the rear wheels turn back to the opposite direction.

This odd-seeming arrangement can work very well. When turning a sharp corner, the rear wheels first move the back of a car out a bit, increasing the angle of attack. Later, when the rear wheels straighten and then turn opposite, they allow a tighter turn radius.

Mazda has introduced patented electro-hydraulic four-wheel steering on its new 626 models. In Mazda's system the rear wheels turn opposite to the front wheels at speeds up to 25MPH, and parallel to them at higher speeds. A complex electronic control system is located at the rear of the car. It actuates a stepping motor and gear mechanism to cause steering motion to push a hydraulic valve one way at slower speeds, and the other way at speeds over 25MPH. Ryoichi Hirose, Mazda engineer responsible for the system, admits its complexity but says it was necessary to achieve required results. "We wanted active control," he says, with the main aim being improvement in safety for inexperienced drivers at high speeds. "The inexperienced driver is nervous at high speed, and overreacts in an emergency," Hirose states. "With a conventional car he may go off the road, but four-wheel steering makes the car safer."

"Japanese automakers, with access to advanced electronic technologies, are at the forefront in implementing electronics for various automobile controls. As we successively incorporate electronics in automobiles...we have to develop an "Expert System", an artificial intelligence service tool that facilitates detection of problems."

MR. KIYOSHI MATSUMOTO, EXECUTIVE VICE PRESIDENT, TOYOTA MOTOR CORP.



Four-wheel steering presents a completely new approach to the improved performance and safety of cars. With front-drive cars, the front wheels do a disproportionate amount of work. The four-wheel system more uniformly distributes the amount of work in carrying load, transmitting torque to the road, and steering and braking between front and rear tires.

Why JAL?

"I like organizations that pay attention to details. I feel at home."

Akio Morita
Chairman of the Board &
Chief Executive Officer
Sony Corporation
Tokyo, Japan

"I respect an operation when it's managed by true professionals."

Jeffrey A. Barks
Associate Dean for Master's
and Bachelor's Programs
Sloan School of Management
M.I.T.

In a world of increasingly unpredictable service standards, Japan Air Lines remains a reference of professionalism. That's one of the reasons why more people from all over the world choose JAL to fly to the Orient than any other airline. On your next trip to the Orient, fly the airline that's run the way you'd run an airline. For reservations, call your travel agent or Japan Air Lines directly at 1-800-JAL-FONE.

JAL
JAPAN AIR LINES
We give you more of the Orient.

Tokyo Osaka Seoul Hong Kong Beijing Shanghai Manila Singapore Bangkok and Beyond

Honda's Kawamoto says his firm is also "aiming at the ordinary driver, with four-wheel steering as a safety feature. But we put it on the Prelude first; that is a better handling, stable car. At this stage, it is not clear whether the mechanical or electronically controlled system is better. However, the mechanical system is simple and reliable, which is what we felt best on the first one."

Kawamoto adds, "We are studying all these systems, and think that on a mid-engine car, or one with a very long wheelbase, an electronic system might be better. In any case, we need to train people to get used to this four-wheel-steering system before we move on to other ones. I am always telling my people not to sell four-wheel steering as fashion. It is a serious technical development and so we want to go quietly, and study how it operates in the real world."

Four-wheel steering has not received universal acclaim. Some European engineers dismiss it as too costly, and some experts in Japan have reservations. Shotaro Kobayashi, editor of *Car Graphic* magazine in Japan, is recognized as a top technical writer on motoring. "It is a mixed blessing," says Kobayashi. "It is all right up to a certain point. I think the Honda system is better, because it is simpler." Kobayashi recognizes the advantages to inexperienced drivers.

Four-wheel steering is a technological innovation unique to Japan.

Innovation: Japan's Technological Trademark

The Japanese economy depends on the manufacture and sale of innovative consumer goods and industrial equipment, products which both contain, and are produced by, high-technology components and processors. The key word is 'innovative' as distinct from 'imitative'; in the 1990s it will be Japanese goods, not Western ones, which increasingly will be copied by low-price competitors.

Japanese industry is preparing for the turn-of-the-century economy by investing its funds—and the minds of its best people—in science and technology application. Despite difficult economic conditions, the current rapid rate of rise in that investment is likely to continue. Taken together with Japanese cultural strengths which integrate consensus with planning and quality/productivity concerns, Japanese technology has become a potent world force, and should continue to play a role in the manufacturing and development planning of Western industry. It makes sense that Western industry recognize the beneficial opportunities that technological industrial cooperation can reveal.

Senior Japanese executives and officials have realized that worldwide relationships are essential considerations for the future of Japan. A nation which was self-contained for centuries has made significant changes and succeeded in adapting production capabilities to market needs. The next step: successful development and commercialization of technologies which can anticipate those needs.

In February, *The New York Times* reported that Japanese investigators commissioned by The Japan Foreign Trade Council, Inc., and headed by Masahira Sakamoto, chief economist of the International Trade Institute, found "many advantages remaining in American industry." In *The Times*, the investigators concluded, "American industry has kept its technological superiority in most advanced fields. Especially in the basic and development stage, there is a huge accumulation of innovative ideas." To a large extent Japan looks toward participation in the process which gives rise to such industrial strength, and recognizes that contribution is part of the cost of participation. Japan's focus is Western, not strictly American in this area, although trans-Pacific activity promises the largest economic involvement. A "symmetrical access" project, under joint management of The Japan Society for the Promotion of Science and the private National Academy of Science in the U.S., is now under way. By improving the flow of technical information and cooperation, the project aims to redirect attention from markets to projects and resources.

Historically, Japan's prowess has been in product development, not in basic research; fundamental con-

Japan must also export its technologies. Whereas manufacturing exports may invite unemployment in importing countries, technology exports would be conducive to their industries. So we should export technologies and be paid their worth...exporting our expertise and thereby helping our trading partners."

DR. HISAO OKA, SENIOR MANAGING DIRECTOR, MITSUBISHI ELECTRIC CORP.

Throughout "Japanese Technology Today"/1987, top executives of Japanese industry speak of the function and role of research and innovation in Japan. The following most represents their judgment.

"...it's about time Japanese companies set about forming global strategies, irrespective of the stronger yen or trade friction. In addition to linking up with foreign firms who offer outstanding technology, Canon will gradually shift its research and development operations overseas."

CANON INC. PRESIDENT RYUZABURO KAKU, QUOTED IN *JAPAN ECONOMIC JOURNAL*, APRIL 1987

THE TWO MOST DYNAMIC FORCES IN PHOTOGRAPHY

Your flair, your talent, your creativity and the technological sophistication of the Professional Maxxum 9000. An unsurpassable combination. But if the potential unleashed by this camera's ground-breaking continuous autofocus is readily apparent, consider the opportunities unlocked by its many other special features. By selecting Maxxum's center-weighted metering or shifting to spot metering with highlight or shadow biasing, you can use light and color like a fine painter to add drama and sensitivity. With Automatic Multi-Program Selection to put you in the program that matches your focal length, even while you zoom, you need never be distracted from the subject before you. And with Professional Maxxum's 1/4000 sec. shutter speed and 1/250 sec. flash sync you can freeze the subject instantaneously. Other important considerations are options like auto bracketing, 5fps motordrive and a 100 exposure filmback, and the 26 lenses, 4 dedicated flash units, and 2 data backs that comprise the most complete autofocus system anywhere. With the force of Minolta technology in support of your vision, there's virtually nothing you can't do.

Motordrive and AF85mm lens optional. Be certain the valuable Minolta USA 2-yr. camera/5 yr. lens ltd. warranty is in your package. For more information see your Minolta dealeror write Minolta Corp., 101 Williams Dr., Ramsey, N.J. 07446. In Canada: Minolta Canada, Inc. Ontario.

YOU AND THE MIND OF MINOLTA

MAXXUM®
AUTOFOCUS SYSTEM

ONLY FROM THE MIND OF MINOLTA

MINOLTA

cepts and inventions from Europe and the USA have been converted into saleable products. Sony's much-travelled chairman, Akio Morita, has said, "There are many kinds of creativity—in marketing, product planning and production. Sony developed the 'Walkman', but it was not the result of creativity, just a development of existing technology." There are thousands of examples of Japanese excellence in product-development innovation, including many in the realm of consumer-electronics video. Sony's recently announced Super-VHS and ED Beta video systems are steps along an existing technology path, as are such new products as 8mm video and digital audio tape.

Japanese industrial research, of which there is a prodigious amount, still focuses largely on practical product applications. Dr. Osamu Horigami, chief research scientist at Toshiba's Energy Science and Technology Laboratory, led a team which developed wires that superconduct electricity at the temperature of liquid nitrogen, a commodity Japan can produce domestically. "We are not so interested in the maximum temperature we can reach," Dr. Horigami points out. "We are more interested in industrial applications," as indeed are many of his colleagues. At Sumitomo Electric Industry Ltd., one of the world's largest wire/cable makers, dozens of researchers and an R&D budget well into the hundreds of millions of yen have been committed to superconductivity. Should other organizations like IBM succeed in establishing patent superiority, then Sumitomo is prepared to cross-license.

"Whether Japanese have creative abilities is often debated uselessly," states Hajime Karatsu, technical adviser to Matsushita Electrical Industrial Company and professor at Tokai University, in a recent issue of *Business Tokyo Magazine*. "Ideas alone do not constitute technology. What is most important in technology is the development of methods for using it, of finding uses for various products, and ways new technology can be used. Japan does this well. The Japanese economy imports virtually all of its natural resources, processes them and makes its living from the added value."

Exporters can no longer hope to succeed by concentrating solely on making and supplying quality products. Contributions have to be made by putting to use the ability which we have cultivated in the past....in manufacturing goods we have developed an array of related technologies. By making use of that know-how, we contribute to the world of science...."

DR. MOTOKAZU UCHIDA, GENERAL MANAGER, RESEARCH & DEVELOPMENT PROMOTION CENTER, HITACHI LTD.

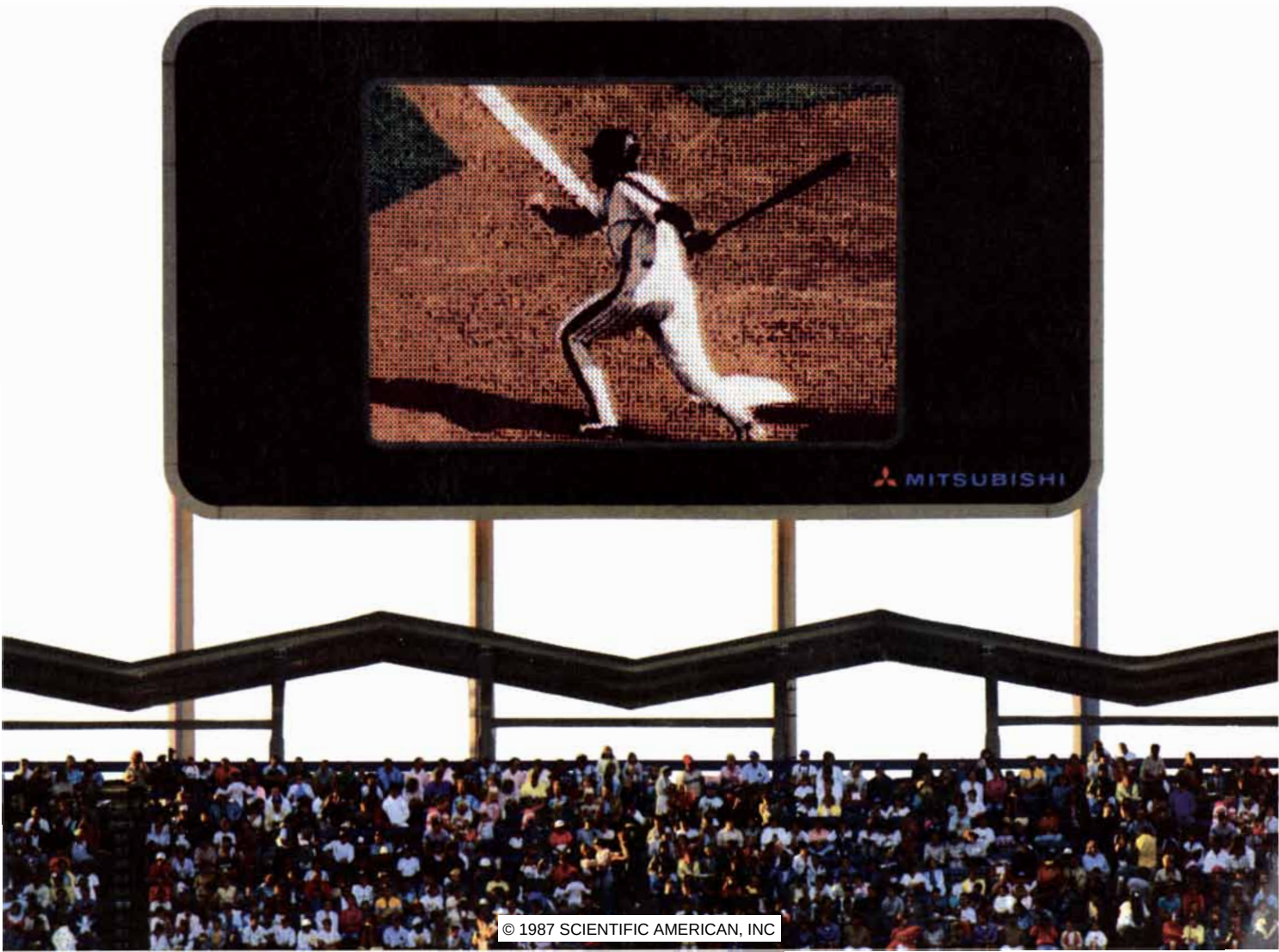
To get to the leading edge of the industrial application of new technology, any organization—even one which represents broad government sponsorship—must take in at least a portion of the basic research. The U.S. Department of Energy has quickly raised its funding of research for applications of super-conductivity. Simultaneously, Japan's Ministry of International Trade and Industry (MITI), which helps guide Japanese worldwide industrial and commercial capability, is setting up a facility in Nagoya to sponsor similar research. Thus, Japan is moving toward initiation and innovation in the crucial scientific and technical areas; indeed, researchers at the University of Tokyo under Professor Shoji Tanaka were among the first to confirm reports of progress published in Autumn 1986. One specific area of interest and experience in Japanese superconductivity research is that of fast surface travel by magnetically levitated express trains.

As any Westerner who has ridden the 'shinkansen' bullet-train among Japanese cities can attest, train travel is taken seriously there. However, the cost of high speed has itself been high, giving the new research an extra incentive for progress. In the words of Hisashi Kado, a researcher at the state-owned Electro-Technical Laboratory, "We need to develop our own original technology; more and more, the developing nations are catching us up." Kado is leading a project in which a DC Squid (acronym for 'superconductive quantum interference device') is being used to measure the magnetic field generated by activities in the brain. The Squid was developed to measure minute magnetic fields.

Extending technological development in such fashion has the effect of driving new research, as source technologies become distant from new products and processes. Japan has already established leading tech or competitive applications in a wide breadth of areas. Apart from consumer electronics, Japanese industry has excelled in production of semiconductors and microprocessors, as well as advanced materials, particularly ceramics. Automated process control, together with applications of Western quality control theory, has given the Japanese a quality "edge" in industries from automobile manufacture to ski-making. The matter of imitation vs. innovation becomes moot as there becomes an increasing scarcity of teachers to follow.

The process of technical innovation in Japan must be regarded in context. The benefits of a consensual approach to business planning do not apply as well to individualistic undertakings like invention; in Japan, four decades of economic growth have been made possible more by the 'salaryman' and by bureaucratic planning than by the entrepreneur. Japanese value their similarities more than their differences, and thus produce fewer iconoclasts and "rugged individuals" than Western cultures do. From a corporate/economic perspective, transcendence of such cultural limitations is occasioned from the organizational boardroom. Top corporate officers mandate new ventures and research, and allocate funds and housing for them. Staff for the new enterprises are drawn largely from the organization, and pursue their tasks with very high levels of enthusiasm and diligence. Thus, benefit is drawn from the collective consensus approach, which is then applied to what in the West would be regarded as high-risk, venture-capitalistic undertakings. There is certainly no shortage of either Japanese talent or

While everybody else has been trying to make their screens larger, we've made ours smaller.



Mitsubishi small
60 inches

Seven years ago at Mitsubishi, we invented Diamond Vision.® And used it to build a television screen 60 feet long by 40 feet high. With a picture so bright and clear, it can show replays in vivid detail to stadium crowds.

Outdoors.

In broad daylight.

By comparison, developing a television picture of any size for home use is a relatively small task.

Which helps explain how in 1986, we were the only television manufacturer with the technology to create a 35-inch conventional tube television. With a picture as sharp and clear as a 26-inch screen, but 86% larger.

The world's first 35-inch, direct-view television. Its 86% larger than a 26-inch set, and can deliver a stunning 560 lines of resolution.

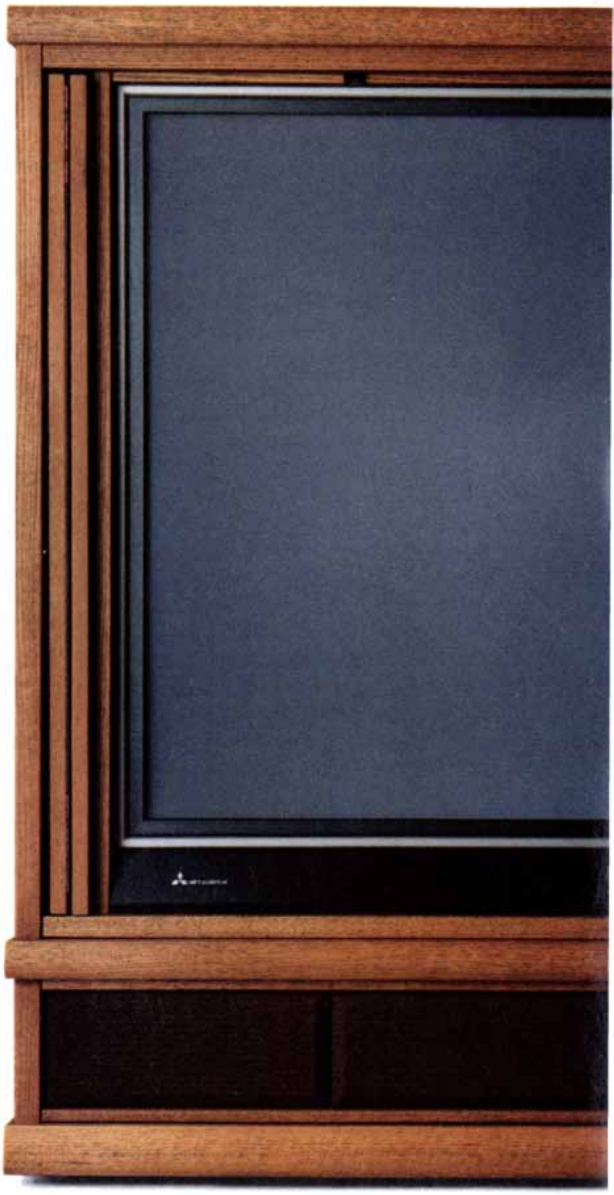
And why in 1987, we're the only television manufacturer to improve it.

Introducing Diamond Vision II.™

Traditionally, a CRT's brightness has been limited by an inhibiting layer that forms inside the cathode. By using Scandium Oxide, Mitsubishi engineers have eliminated this problem and developed the world's first truly high-current cathode.

Rather than simply using the resultant brightness for a brighter picture, we've opted to use more darkly-tinted glass on the screen for dramatically better contrast.

In addition, our new Dynamic Beam forming actively refocuses the beam at every position on the screen, enabling us to minimize distortion. So our Diamond Vision II™ offers 15% sharper focus at the center of the picture and 40% improvement at the edges.

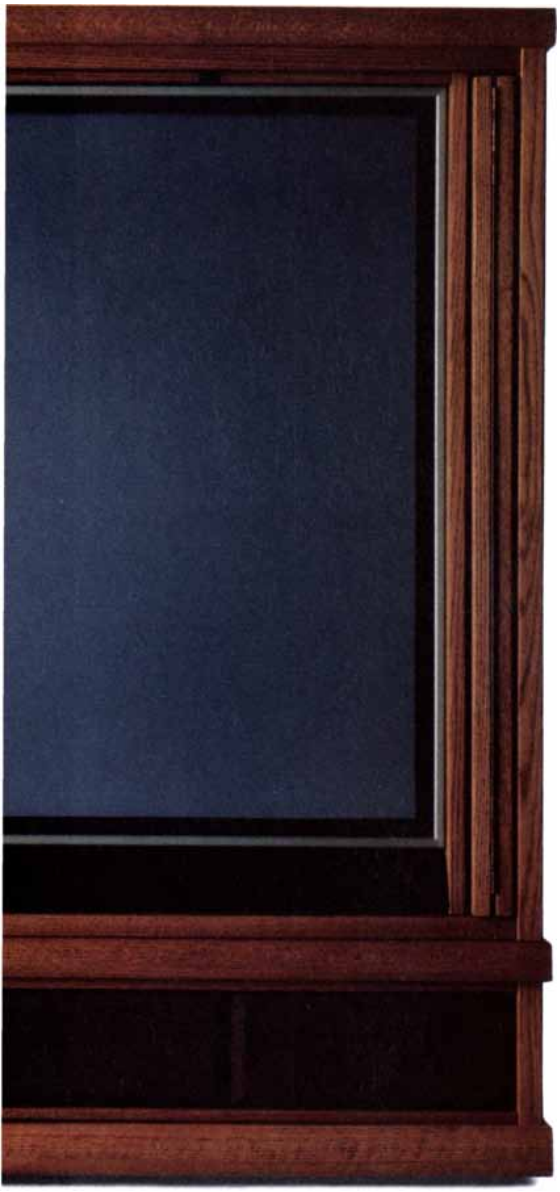


Our new 60-inch big screen television includes our Diamond Vision II™ wireless remote that includes VCR controls (for select Mitsubishi VCR's)

Even bigger news.

Nowhere are the benefits of our new Diamond Vision II™ more apparent than in our new 60-inch big screen TV. The largest picture

creen televisions.
nd under.



technology, four-speaker stereo sound and a simple, straightforward
And like all our Big Screen TV's it's completely Super VHS-ready

we've ever offered for home use. Diamond Vision II™ is also standard on our 31, 35, 40, 45 and 50-inch models. Each with an image so incredibly crisp, you can see every detail.

Our screens deliver a wide off-axis viewing, for a bright, undistorted picture from just about anywhere in the room. And when you're finished viewing, you can just close the doors and the TV becomes a discreetly handsome piece of furniture.

One that you can easily move into your home with the help of the Mitsubishi Three Diamond Card.™ You can use the Three Diamond Card to finance your new television or for that matter, any of our terrific audio/video products. Just ask your dealer for an application. Many times your credit can be approved while you wait.

Before long, you'll be enjoying your new Mitsubishi big screen at home.

Unlike a conventional set, big screen TV turns your ordinary televised sporting event into an incredible spectacle.

And you can enjoy movies with the kind of panoramic scope that small screen TV just can't offer.

Mitsubishi televisions. Available only at authorized Mitsubishi dealers and selected stadiums nationwide.

45-inch. Traditional style with genuine cherry veneer wood.

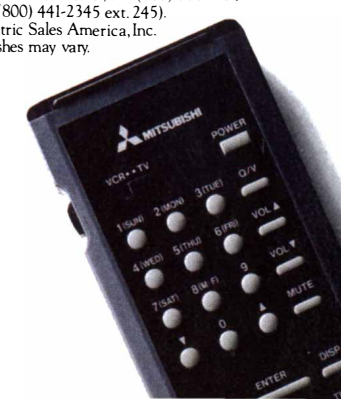
50-inch. Contemporary style with genuine oak veneer wood.

40-inch direct-view. Transitional style with genuine pecan veneer wood.

MITSUBISHI®

For the name of your authorized dealer, call (800) 556-1234 ext. 245 (In California (800) 441-2345 ext. 245).

Cabinet colors and finishes may vary.



"...Basic research is necessary, but not to quickly resolve problems related to trade surplus or economic frictions....we should try to develop an originality, by international standards, fostering our individual ideas and, based on them, devise new things."

MR. ICHIRO YOSHIYAMA, SENIOR EXECUTIVE DIRECTOR, MINOLTA CAMERA CO., INC.

brainpower, as performance of Japanese who study and/or work in Western university and corporate environments can attest.

Taizo Ueda is managing director of the Tokyo-based Honda Foundation, which holds symposia and awards a Honda Prize to scientists making major contributions to eco-technology, which is technology to help improve the natural environment. "The basic philosophy of our founder, Soichiro Honda (also founder of Honda Motor Car Company)," Ueda says, "is that basic harmony should exist between technology and ecology. The Foundation holds creativity in high esteem as well," he adds. "Honda doesn't like to imitate." In 1986 a Honda Prize was awarded to Tohoku University professor Junichi Nishizawa, who had achieved some prior reputation as an inventor of computer and communications components, and had written extensively on eco-technology.

Honda Prize awards are not restricted to Japanese eco-technologists. Though the Foundation is separate from the founding Company, the Company itself, like many other Japanese corporations, is reaching well beyond Japan's natural boundaries to contribute to, and benefit from, scientific research. Honda has funded the establishment of a research center at Cranfield Institute of Technology in Bedfordshire, England. "The Center is intended to promote mutual understanding," says Ueda, "since students and professors from any country will be able to work there. Cranfield has close liaison with universities in Aachen, Compiegne and Stockholm, and from next year will offer exchange programs with both Toyohashi Institute of Technology and Nagaoka Institute of Technology here in Japan." The new Honda program offers opportunities for Japanese researchers to increase their experience in Western research methods and environments.

Strong Japanese voices seek a steadier increase in science and basic research sourced or sponsored in Japan. Takashi Nagayama is general manager of The Science and Technology Foundation, and attained his position after being managing director of Matsushita Electric. Matsushita, the giant Osaka-based electronics firm, grew from the entrepreneurial work of Konosuke Matsushita, who subsequently established and funded the Foundation. "The Japan Prize was established, and is awarded each year by the Foundation, to recognize scientific contributions to society," Nagayama states.

"The grant system does not yet work as well in Japan as it does in the West," Nagayama suggests, "and this impacts on our universities. They receive some grants from government for research, but if they also receive funds from industry they obtain less from government. So students and professors regard research as a steady path, evenly founded along well-proven routes."

In 1983, the Japanese government established a plan to increase the amount of joint research carried out between industry and government. Called 'The Cooperative Research Project', it encouraged universities to work more with industry. In 1986, thirty-eight national universities and eleven technical colleges took part in joint projects involving 188 private firms and aggregating some $12 million in funding.

As do Western companies, Japanese firms face strong management pressure to develop techniques and products which can show an immediate financial return. They concentrate on practical designs, on making complex products easier to use, and on practical production engineering. Research is structured to that end. The powerful Ministry of International Trade and Industry (MITI) guides and supports such endeavors by coordinating companies and their research in the development of IBM-competitive computers in the late 1970s, of very-large-systems integration (VLSI) technologies, and of a "fifth-generation" computer project. Dependence upon centrally managed research is a distinctive feature of Japanese industry. Such MITI-management assists by focusing resources, but can also hamper by limiting areas of exploration.

Longer-range planning within the MITI structure includes basic technology projects like three-dimensional integrated circuits, ceramics, plastics, synthetic metals and biotechnology. One program seeks to develop large-scale robots for electric-power-station, undersea and routine maintenance applications. Others involve superconductivity and lasers. These projects follow a similiar pattern, with one or two government laboratories joining forces with private firms to develop specific products or processes. Preset, target objectives discourage 'off-the-wall' idea-researching. Moreover, since firms compete in all arenas save the one involving the joint research, there is a tendency for individual companies to withold their best work and workers, while drawing maximum benefit from each project.

U.S. advanced-technology manufacturers have learned from experience that it is intercorporate competitiveness, not government regulation, which inhibits the joint development of shared technologies. Being licensed to act cooperatively across a limited economic range does not alone assure cooperation. Kazuhiro Fuchi, director of the "fifth generation computer" project at the Institute for New Generation Computer Technology (ICOT), suspects some major Japanese computer companies are reluctant to send their most talented people to ICOT. Meanwhile, he has heard complaints from top executives at the same companies, who seek a greater return on their research invest-

Landscape of the mind

The very place from which the Nasca Indians envisioned the perfect patterns they carved on the desert's face, is the same source from which ideas spring at Epson. The mind's eye. Here Epson's highly trained engineers transform imaginative ideas into today's and tomorrow's innovations.

We were the first to develop a printer small and fast enough to work in a calculator. Expanding on our precision technology, we developed the highest quality and performance printers for PCs; establishing Epson as the world's leading printer manufacturer.

Believing in what's presumed impossible creates a basic unrest at Epson. One which impels us to create startlingly innovative products. Like a flat-screen, hand-held, high-resolution LCD TV; human-engineered ultra-friendly IBM® PC compatible computers, and a desktop performance computer small enough to fit into a briefcase. Just by looking beyond what is, to what might be.

And our imaginative spirit is never lost on earthly features like cost, performance and easy operation.

Just like the Nasca Indians, we think the best place to view things is from above; here, on the landscape of the mind, where the imagination takes flight.

IBM is a registered trademark of International Business Machines Corporation.

EPSON. Where anything is possible.

EPSON®

SEIKO EPSON CORPORATION: Suwa-shi, Japan Tel: (0266) 52-3131 Telex: 3362435 EPSON AMERICA, INC.: Los Angeles, U.S.A. Tel: (213) 373-9511/9522 Telex: 182412 EPSON DEUTSCHLAND GmbH: Düsseldorf, F.R. Germany Tel: (0211) 56030 Telex: 8584786 EPSON (U.K.) LTD.: London, U.K. Tel: (01) 902-8892 Telex: 8814169 EPSON FRANCE S.A.: Paris, France Tel: (1) 47.57.31.33 Telex: 610657 EPSON AUSTRALIA PTY., LTD.: Sydney, Australia Tel: (02) 452-5222 Telex: 75052 EPSON ELECTRONICS (SINGAPORE) PTE., LTD.: Singapore Tel: 5330477 Telex: 39536 EPSON ELECTRONICS TRADING LTD.: Hong Kong Tel: 5-8314600 Telex: 65542 EPSON ELECTRONICS TRADING LTD. (TAIWAN BRANCH): Taipei, Taiwan Tel: (02) 7160855 Telex: 24444 EPSON SEGI S.p.A.: Milan, Italy Tel: (2) 670-9136 Telex: 315132 EPSON-STI, S.A.: Barcelona, Spain Tel: (3) 239-7707/8 Telex: 50129 EPSON CANADA LTD.: Willowdale, Canada Tel: (416) 495-9955 Telex: 6966633 EPSON LATINOAMERICA S.A.: Caracas, Venezuela Tel: (02) 34·35·90 Telex: 27660 EPSON DO BRASIL INDUSTRIA E COMERCIO LTDA.: São Paulo, Brazil Tel: (011) 210-9599 Telex: 38572 EPSON CHILE S.A.: Santiago, Chile Tel: 395917 Telex: 341615 EPSON ARGENTINA S.A.: Buenos Aires, Argentina Tel: 392-2620 Telex: 24353

ments even though quick commercial payouts were not a prime objective of the program.

At present, research in Japanese companies is designed to produce practical products, and is successful by nearly any measure in doing so. At NEC, the large electronics and telecommunications corporation, Michiyuki Uenohara is executive vice president. "Thirty percent of NEC's research and development activity is conducted on behalf of our manufacturing divisions, which are responsible for detail design and engineering and production," Uenohara points out. "We transfer engineers occasionally from the relevant division to the central laboratory, to draw them more into the theoretical nature of the research work.

"That way we minimize the 'NIH', or 'not-invented-here' syndrome, which is occasionally the response of the operating divisions to lab-developed product and process designs," he concludes. NEC's application of management skills to labwork thus adds an element of innovation and originality, and over the years has produced a fair share of firsts. "In 1959, NEC demonstrated the first all-transistor computer in the world," says Uenohara, adding that in computer memories, "NEC was in the race from the beginning." The company has also developed a data-flow computer currently being used to process image data relayed by weather satellites.

Japanese industry is adapting to the new global economic environment, in which originality and innovation are essential to create and retain market leadership. The leading companies in the manufacture of high-modulus carbon fiber are Toray Industries and Toho Rayon. Both companies were dependent on mass-produced textiles just a few years ago. Giants Nippon Steel and Kawasaki Steel have diversified via licensing and joint-venture and are involved in silicon and ceramics technologies. Nippon Steel estimates that steel products will aggregate less than half of 1995 sales. Kawasaki steel has formed joint ventures with LSI Logic of the USA to produce silicon wafers, and with Pechiney of France to make hard-ferrite magnets.

The Human Frontiers Science Program has heavy support at the highest levels of both the Japanese bureaucracy and national government. A product of MITI's Agency for Industrial Science and Technology, the Program is a conscious attempt to conceive and execute a pure-research undertaking which will reflect favorably on Japan's science-technology research community. The Program is intended as well to create platforms for more functional research and development in the next generation. According to Katsuhiko Umehara, deputy director of the Program, Prime Minister Yasuhiro Nakasone sought to have Japanese researchers across several disciplines and ministries working together. Foreign reaction has been favorable and has influenced the cooperative spirit of those involved. "We want to have international workshops and give research grants," Umehara says, "and to promote the exchange of scientists between countries. As Prime Minister Nakasone says, 'we have to contribute to balance our basic-research obligations.'"

One objective of the Human Frontiers Science Program is to learn how the human perceives and recognizes things: how his memory works, how he learns, what intelligence he inherits, and how motor and behavioral control functions. Much of the work will involve research into biotechnology, an area where Japanese science is rapidly gaining strength. Antigen-antibody interaction, morphogenesis, and molecular recognition and response are included, and it is recommended that technological development be undertaken to enable the reading and analysis of the DNA nucleotide sequence of human genes. The latter is work already well into the planning stage in the USA, and a cooperative effort under the MITI Program could prove most productive.

Microprocessors

Japanese industry has shown significant success in the design and production of microprocessors (MPUs). The market within Japan is as competitive as it is internationally, with NEC tending toward market leadership but facing increasing pressure from Hitachi and Fujitsu. The latter have set up a joint project to develop a device based on the TRON open architecture. The TRON (The Realtime System Nucleus) project embraces MPUs, computers, networks and interfaces in a new type of system designed by Ken Sakumara, associate professor in the Department of Information Science at Tokyo University. Its salient feature is its projected ability to handle tasks on a simultaneous "realtime" basis, with no delay between entry and execution of an instruction.

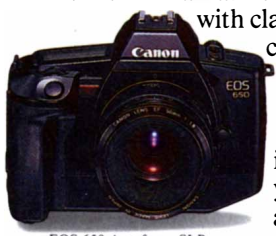
"We must let foreign businesses know that we are interested in...international affiliations in the area of research.... The possibilities for this exist in a wide range of fields. What is important, in essence, is to be able to speak a common language in research."

MR. TOSHIYA INOUE, SENIOR MANAGING DIRECTOR, VICTOR COMPANY OF JAPAN LTD.

Hitachi and Fujitsu, working on an original Hitachi design, have been joined by Mitsubishi Electric which had been developing a TRON MPU. Hajime Yasuda, manager of Strategic Product Planning for 32bit MPUs at Hitachi, explained the TRON rationale: "We decided on a proprietary architecture, and realized one company alone could not create a new architecture; we are trying to avoid duplication of effort." Toshiyuki Hani, marketing manager for MPUs at Fujitsu, claimed Fujitsu had spent two and a half years developing a 32bit MPU, but "owing to the need for standardization, we decided it was too time-consuming for us to go it alone." The companies, 32bit MPU is aimed mainly at the engineering work station market, with a version being developed for process-control applications. Based on 1.3micron design

We measure our vision by what *you* see.

As Canon celebrates its 50th anniversary, two trends, we think, are remarkably clear. In every field of human endeavor the importance of image information is growing by leaps and bounds. At the same time, we work in an atmosphere so charged with claims and counter-claims that what we say about corporate goals isn't nearly so impressive as what you see us do about them.



EOS 650 Autofocus SLR

Technological vision.

On our way to becoming one of the world's premier camera makers, our aim has been to make it simple for anyone to capture a beautiful image on film. Witness Canon EOS, the autofocus camera whose revolutionary lens-integral motor design not only captured the Japan Camera Grand Prix and the European Camera 1987-88 awards. Because of its unique autofocus system, the European Camera judges also gave it their first ever Innovation Award as well.

Commitment to image quality also drives the design of all our wide range of products for personal, medical, office and industrial use; from our personal typewriters and copiers to our X-ray equipment and ultrahigh precision steppers for semiconductor production. For another outstanding example see our new digital Color Laser Copier. The first to offer both its superb color reproduction and its advanced color, size and format editing capabilities — at a price within the reach of small, entrepreneurial businesses.


Color Laser Copier

Cooperative vision.

But impressive as they are, these innovations in image information technology are to us only part of a much bigger picture. To us, the technology we help to create is a resource we want the world to share. That's why, for example, when we first conceived our now world-famous Laser Beam Printer, we didn't keep its secrets to ourselves. By making its key technologies available to partners worldwide, we helped to create a whole new industry: desktop publishing. Now we're working with a leading air courier company on advanced G4 facsimiles to realize the potential of the telephone companies' new Integrated Service Digital Networks.


LBP-8II
Laser Beam Printer


Canon FAX 14003 Mixed Mode Terminal

Productive vision.

We recognized long ago that one-way trade is no trade at all. As early as 1973, we built a plant to produce our copiers in Giessen, West Germany. That investment was followed by others in Taiwan, the U.S., Korea and France.

Last year we concluded technical exchange agreements involving two plants in the People's Republic of China. This year we started operations at a new US$25 million business machine production facility in Newport News, Virginia. We also imported more than US$45 million worth of semiconductors for use in our products. As we celebrate our first half century growing together with partners all over the globe, we are proud that one in every three of our 35,000 employees contributes to local prosperity in his native land outside Japan.

Because, like you, when we hear people talk about vision, we like to see the results.

Canon

CANON INC
Tokyo, Japan

rules and CMOS technology, the TRON MPU will include 700,000 elements and will follow the latest design layout, with a memory management unit and a cache memory on-chip. It will be able to handle 32bit data internally, with an extensive set of instructions.

In terms of performance, the latest members of the NEC Company's V series is the most powerful Japanese design. The NEC V60/70 models were designed back from NEC's V80 chip, which will be based on 1.0micron rules and has been designed to operate at 10 million instructions per second. It is due in 1990.

The TRON Project

The TRON project involves development of an architecture for an MPU, as well as various operating systems, and a specification for the keyboard, disk drives, input/output control and other features of a computer. Additionally, the project aims to develop MPUs for specific applications from the basic architecture, including ITRON (industrial TRON), a realtime operating system for process control; BTRON (business TRON) for work stations; MTRON (macro TRON) to control networks; CTRON for databases and mainframes. Altogether TRON is a new concept in compatibility, embracing far more than just another computer-and-keyboard. Ken Sakamura of Tokyo University, leader of the TRON project, takes a very wide view.

"Our aim is to produce a computer anyone can use, and I think of computers as being far wider in scope than mainframes, minis and PCs," Sakamura enthuses. "There are, of course, computers in home appliances like TVs and VCRs, each with its own function. People can use a TV with videotex and modem adaptor to reserve space in a hotel, for example, or manage their finances on a home computer. TRON will be one machine which will do it all."

The first aim of the TRON project is development of interfaces for any devices with similar software. The microprocessor can assure that all computers sharing an instruction set will understand the same instructions, and as Sakamura points out, "We have produced the standard instruction set."

Second, the TRON MPU must be able to address sufficient memory. Sakamura foresees that the 4gigabytes available to 32bit work stations will become limiting in the late 1990s. "Because of the reduction of storage costs, more than 4Gbytes will soon be installed in some systems, and eventually the 32bit chip will become obsolete. The TRON chip can go to 64bits, giving an addressable memory capacity of 1,800Tbytes," he projects.

Sakamura's group has specified and tested four interface layers, to permit connection among TRON computers. Adopted by manufacturers, the interfaces would also permit communication with any intelligent device—a VCR, a home appliance, even a telephone/terminal.

Realtime multitasking computing is an inherent objective. "One should be able to work through several windows at once," he said. "While working, for example, one should be able to scan a scene with a portable camera and have the image transferred to memory and screen simultaneously. While existing computers may be able to handle simultaneous transactions, data is moved in and out of memory as focus moves from process to process, and procedures tend to be slow."

There is little imaginative limit in Sakamura's group's objectives. TRON external memory is to be based on optical recording via a device the size of a portable CD player and with a capacity of up to 600Mbytes. A portable TRON a half-inch thick, complete with interfaces, is planned.

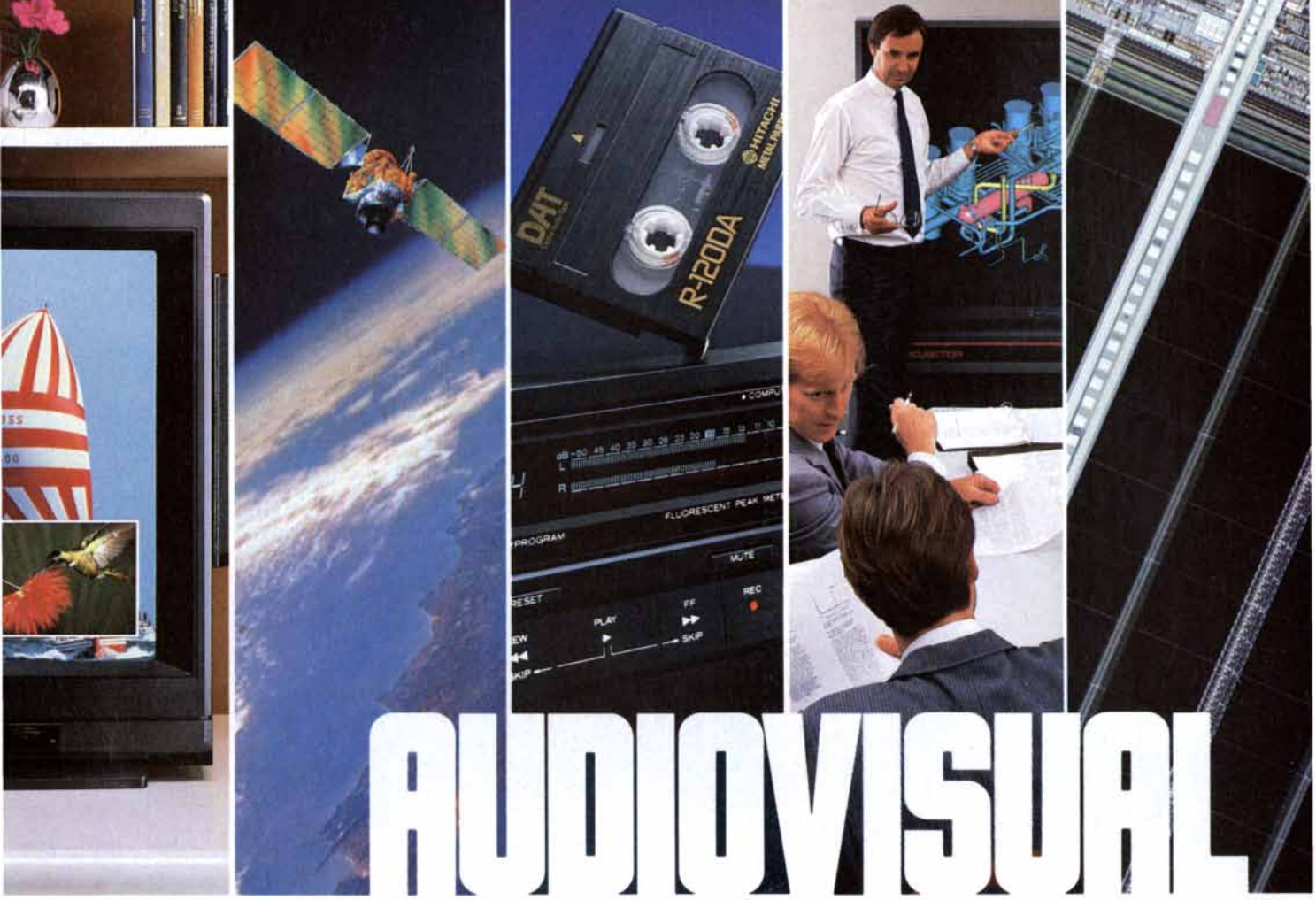
Western observers may appreciate the innovation represented in the TRON team's approach to a peripheral as prosaic as a keyboard. The team's new design is shaped to the contour of the hand. Two sets of keys are arranged in arcs opposite an operator's hands, with two smaller sets in smaller arcs opposite the thumbs position. Sakamura intends that keyboards be available in varied sizes: "Our keyboards are like shoes," he said. Keyboard space has been reserved for an electronic pen pad. Sakamura recognizes the importance of keyboard design to potential markets in Asia and elsewhere, markets in which there is no tradition of typewriter use. He seeks ultimately to design a keyboard adaptable to most languages.

"The way people think reflects their culture and language," Sakamura says, "and Arabs and Orientals are at a disadvantage in that most computer software and operating systems are based on English. The character sets for the languages of different nations can be coded and stored in the computer. The combination of keyboard and multilingual capability is what we are trying to improve in machine relationships."

To date 62 companies have joined in the TRON project, including not just Japanese companies but Western interests like AT&T, Data General, Hewlett-Packard, Honeywell and Texas Instruments.

"We are trying to orient our research staff into two different directions. One group would be committed to R&D of areas closely related to customers' needs. Members of that group would approach customers directly. The other group would step deeper into such fields as are nearer to pure science, such as bio-electronics. The ecology of man tells us that we have an efficient communications function; for example, a signal is transmitted via a sensor to the brain where it is processed. We should learn more from functions of man and study how they can be applied to communications."

DR. HARUHIKO TSUCHIYA, SENIOR MANAGER/DEVELOPMENT PLANNING, ENGINEERING STRATEGY PLANNING HEADQUARTERS, NIPPON TELEPHONE AND TELEGRAPH CO.



Hitachi's wide-ranging audiovisual technologies include Picture-in-Picture for simultaneous viewing of more than one program on a single TV screen, DAT, a high-density projection display, and frame memory used in IDTV.

AV should be more than faithful sight and sound reproduction. Ideally, it also creates drama and ambience.

Digital technology (the conversion of conventional signals into computerized zeros and ones) has led to a remarkable proliferation of audiovisual uses — in TV, for example, for more diversified and sophisticated programming and information services accessible through computer connections or videotex terminals. And this is only the beginning.

Hitachi's scientists and engineers are using digital applications such as frame memory to develop Improved Definition TV. IDTV will greatly improve picture quality without changing current broadcasting standards by doubling the density of scanning lines and increasing vertical resolution 1.5 times. This same Hitachi technology has resulted in the Digital Audio Tape recorder, which is capable of superior recording and reproduction.

Hitachi's original screen technology has led to high-density big screen projection TV, using screens up to 110 inches. It is contributing to a wholly new technology, High Definition TV. HDTV is capable of photographic quality resolution and will soon enable satellite services to transmit wide screen images that give the viewers the feeling of actually being there.

We link technology to human needs, and believe that our special knowledge will lead to numerous easy-to-use systems and products with highly advanced functions.
Our goal in audiovisual — and in medicine, energy and transportation as well — is to create and put into practice products and systems that will improve the quality of life the world around.

HITACHI

Hitachi, Ltd. Tokyo, Japan

TOYOTA SUPRA

NERVE CENTER.
Analog instruments monitor every rev. 8-way adjustable driver's Sport Seat and optional glove-soft leather interior add even more luxury to performance.

IT DRAWS A CROWD, AND LEAVES IT BEHIND.

POWER STRUCTURE.
Good-bye crowd. Turbo-charged 3.0-liter twin cam 24-valve engine with inter-cooler generates 230 hp at 5600 rpm.

It's an imposing presence. An eye-magnet. But the 1988 Toyota Supra Turbo is not simply for show. The low-slung aerodynamics separate it from the crowd, even at standstill. But the gap widens quickly when its electrifying turbocharged potency is released: 0 to 60 in a mere 6.7 seconds.* And control is precise: race-bred double wishbone fully independent suspension; ventilated 4-wheel power disc brakes; optional Anti-Lock Braking System (A.B.S.). Toyota Supra—#1 in its segment in customer satisfaction.** The crowd can look but they can't touch it.

Get More From Life... Buckle Up!

*United States Auto Club Certified Performance Figures.
**J.D. Power & Associates 1987 Customer Satisfaction with Product Quality and Dealer Service (for 1986 vehicles).

TOYOTA QUALITY
WHO COULD ASK FOR ANYTHING MORE!

The companion volume to the acclaimed PBS Series

Voyage to the heart of scientific inquiry with two of the world's great teachers

The Ring of Truth takes us to the place where ideas begin, where the tools are made, where the work of science is done. "That world," write the Morrisons, "is full of surprises."

Throughout, the quarry is not *what* we have learned about the universe but rather *how* we go in search of answers. Using the stuff of every day—doughnuts and mousetraps, fireworks and bicycles, a piece of cake, a bolt of lightning, a big yellow van that measures the earth and a guest appearance by Julia Child—the Morrisons use precept and example to bring out the deepest conclusions about science. *The Ring of Truth* is an ingenious blend of pictures and text from a pair of master teachers at their very best.

Illustrated with more than 300 photographs, two-thirds of them in color.

"In the experience of science, it is easy enough to describe the great discoveries. What is much more difficult is to convey a real understanding of how these discoveries were made. The Morrisons get their hands dirty, actually *do* experiments on the foundations of science, and sweep the reader up into an **exhilarating adventure.** *The Ring of Truth* is just what the title promises—an introduction to how we know."

—Carl Sagan

THE RING OF TRUTH

An Inquiry Into How We Know What We Know

by PHILIP and PHYLIS MORRISON

8" x 10." $24.95, now at your bookstore. For mail orders, send coupon below or call *toll-free* (credit cards only): **1-800-638-6460**

RANDOM HOUSE

RANDOM HOUSE MAIL SERVICE
Dept SA 31-2
201 East 50th Street, New York 10022

Please send me THE RING OF TRUTH (55663-1). I enclose $24.95 plus applicable sales tax. (Publisher pays postage and handling.) And, if not satisfied, I may return the book postpaid within 10 days for a full refund.

☐ Enclosed is my check/money order.
Or charge my ☐ American Express ☐ VISA ☐ MasterCard

_____
Exp. Date _____
Signature _____
Name _____
Address _____
City _____ State _____ Zip _____

Price applies to U.S. and territories only. In Canada, available through bookstores only. 002 03 SA 31-2

Synchrotron Radiation

A threadlike beam of electrons wiggling through a gauntlet of magnets in a storage ring produces the world's brightest ultraviolet light and X rays. Such radiation opens the way to new science and technology

by Herman Winick

Twenty years ago there were perhaps 50 people in the world who exploited a peculiar type of electromagnetic radiation that is generated in circular electron accelerators. The radiation is intense and highly polarized, and it contains a broad continuum of wavelengths. Yet it was then regarded largely as a nuisance, since it represented an energy "leak" that made it difficult and costly to push electrons to ever higher energies. Its few users usually had to carry out their investigations in "parasitic" programs that took advantage of whatever radiation happened to be generated during particle-physics experiments.

Today this type of radiation, which is known as synchrotron radiation, is widely regarded as a unique scientific and technological resource. It spans wavelengths from the infrared to the gamma-ray regions of the electromagnetic spectrum and is by far the most powerful continuum radiation in those spectral regions occupied by vacuum-ultraviolet (VUV) light (so named because it propagates over significant distances only in a vacuum) and X rays. As an added bonus, the radiation is also pulsed, collimated and partially coherent. The fact that all these properties are found in radiation from one source makes synchrotron radiation a remarkably versatile analytical tool for basic and applied research in physics, chemistry, biology and their numerous subfields.

Yet the usefulness of synchrotron radiation is not limited to the laboratory. The possibility of selecting certain X-ray wavelengths from a continuum source allows radiographs to be made that are "tuned" to certain elements—a fact that may be fruitfully applied in the imaging of blood vessels, for example. Furthermore, the intense X-radiation produced at synchrotron-radiation facilities can be applied to etch integrated-circuit elements that are smaller than the wavelength of visible light.

In view of the many current and potential applications of the radiation, it is not surprising that circular electron accelerators, such as synchrotrons (machines that briefly accelerate electron beams to high energy) and storage rings (machines that keep a high-energy electron beam circulating for hours), are operated as fully or partially dedicated synchrotron-radiation sources, serving more than 5,000 investigators worldwide. Indeed, the synchrotron-radiation "parasites" have in many cases consumed the host facility. Yet increasing numbers of scientists and engineers from universities, Government laboratories and industry need greater access to synchrotron radiation of ever higher intensity to pursue their work.

The demand is being met at existing facilities by opening more beam lines, which channel the radiation from a storage ring to experimental stations, and by installing magnetic devices known as wigglers and undulators, which boost the intensity of the radiation to exceedingly high values. In addition new sources of synchrotron radiation are being built throughout the world, ranging from superconducting devices that could fit into a large living room to storage rings one kilometer in circumference that will serve thousands of users.

Although synchrotron-radiation facilities can be quite large, the experimental stations at the ends of the beam lines closely resemble the setups found in small laboratories. Thus the large facilities can serve many "small science" research groups. A typical such group consists of from three to five workers, who can perform a significant measurement in a week or so with end-station equipment costing less than $100,000.

Principle of Generation

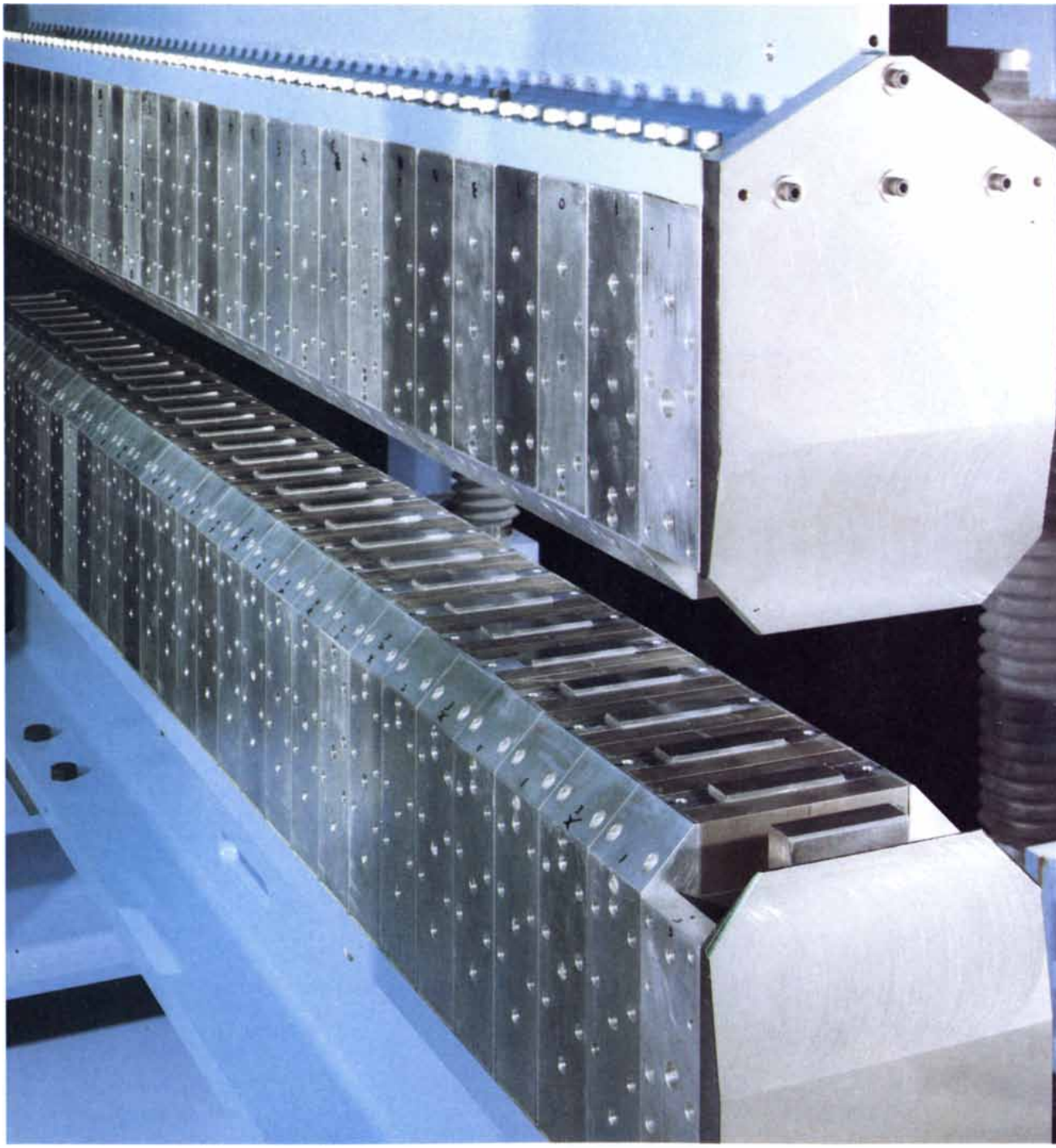
Whenever a charged particle is accelerated (or decelerated), it gives off electromagnetic radiation. Radio waves, for example, are generated when currents of electrons in an antenna oscillate at frequencies corresponding to that particular part of the electromagnetic spectrum. Because a change in direction of motion is an acceleration, a circulating current of electrons also radiates. At low speeds (well below that of light) the electrons emit very weak, low-frequency radiation in almost all directions. But if the electrons are traveling at relativistic speeds (nearly as fast as light), the radiation increases spectacularly in intensity, directionality and frequency.

Because synchrotrons and storage rings are machines that force relativistic electrons into circular paths by means of powerful magnets, they are sources of such electromagnetic radiation. (The radiation was first observed in 1947 in an electron synchrotron of the General Electric Company; hence its name.) A synchrotron brings low-energy electrons to some high energy, at which point they are made to strike a target. It can perform this operation between 50 and 60 times per second. Electrons can be injected into a storage ring at the ring's operating energy or at some lower energy. In the latter case the ring functions briefly as a synchrotron to raise the energy of the beam before it is made to circulate for hours at constant energy.

Both synchrotrons and storage rings consist of an evacuated ring-shaped chamber around which two basic types of magnets have been arranged: focusing and bending mag-

nets. A focusing magnet's multiple poles set up a nonuniform magnetic field that acts as a lens, confining the electrons in a tight beam as they travel around the chamber. A bending magnet typically establishes a uniform magnetic field that is at right angles to the direction in which the electrons are moving. Such a field bends the path of the electrons. The actual path traced by an electron beam is therefore usually not a perfect circle but rather a series of circular arcs, one at each bending magnet, connected by straight segments. Synchrotron radiation produced by electrons curving in the bending magnets emerges into tangential beam lines [*see illustration on next page*].

The ring-shaped vacuum chamber also threads through one or more cavities in which electromagnetic fields that oscillate at radio frequencies are established. These fields replenish the energy lost by electrons as they give off synchrotron radiation and can raise the energy of the electron beam if necessary. They also divide the circulating beam into electron "bunches," which are typ-



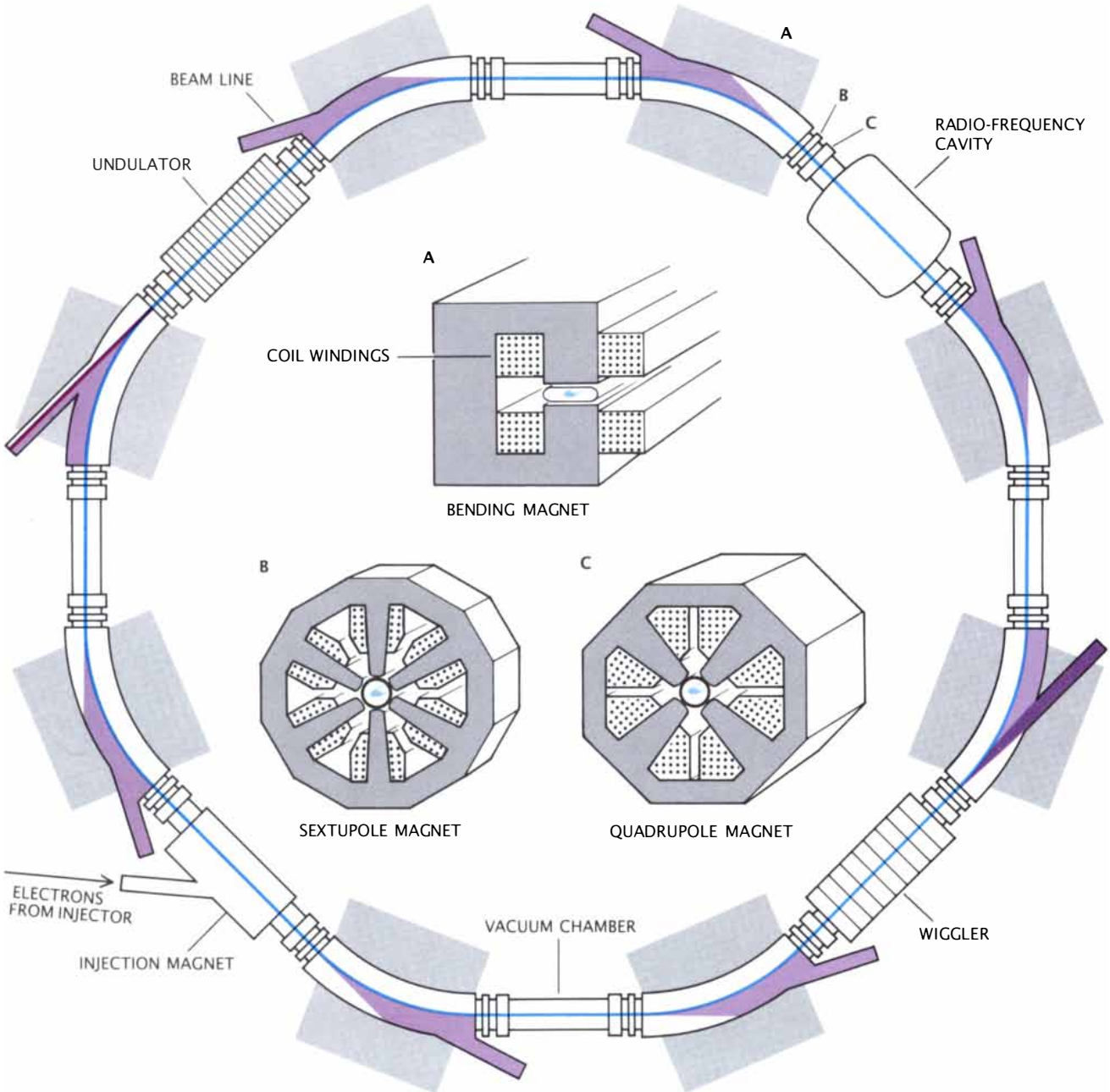
WIGGLER induces electrons traveling through the evacuated tube it straddles to emit intense ultraviolet light and X rays. The device, which was designed and built at the Lawrence Berkeley Laboratory, consists of two rows of 30 permanent-magnet poles. It was developed for a group composed of many institutions, including national laboratories (led by the Lawrence Livermore National Laboratory), various branches of the University of California system and the Stanford Synchrotron Radiation Laboratory.

ically a few centimeters in length.

Because storage rings provide a much stabler electron beam as well as a more constant radiation intensity and spectrum than synchrotrons, storage rings have become the predominant source of synchrotron radiation. The switch from synchrotrons to storage rings as radiation sources began in 1968, when Ednor M. Rowe and his colleagues adapted the Tantalus storage ring at the University of Wisconsin for this purpose.

Properties of the Radiation

Whether it is generated by a synchrotron or by a storage ring, synchrotron radiation has several characteristic properties, some of which were pointed out as early as 1898 by A. Liénard. Virtually all the radiation emitted at a given instant by a single electron as it curves in the field of a bending magnet is confined to a forward cone with a very small opening angle. This natural emission angle is numerically about equal to the ratio of the energy represented by the electron's rest mass (given by Ein-



STORAGE RING dedicated to the production of synchrotron radiation is structured around a ring-shaped vacuum chamber through which a beam of electrons (*blue*) circulates. An oscillating electromagnetic field established in a radio-frequency cavity provides energy to maintain the particles at relativistic speeds (nearly as fast as light) after they are injected into the ring from an external accelerator (not shown). Quadrupole and sextupole focusing magnets confine the electrons in a tight beam by means of fields set up by four and six poles respectively, arranged radially around the vacuum chamber. Bending magnets force the electron beam to curve, causing the particles to emit synchrotron radiation (*purple*). The ring may also include other magnetic devices, known as wigglers and undulators, that substantially increase the "brightness" of the radiation—a measure of its concentration. Pipes called beam lines channel the radiation from the various magnetic devices to experimental stations.

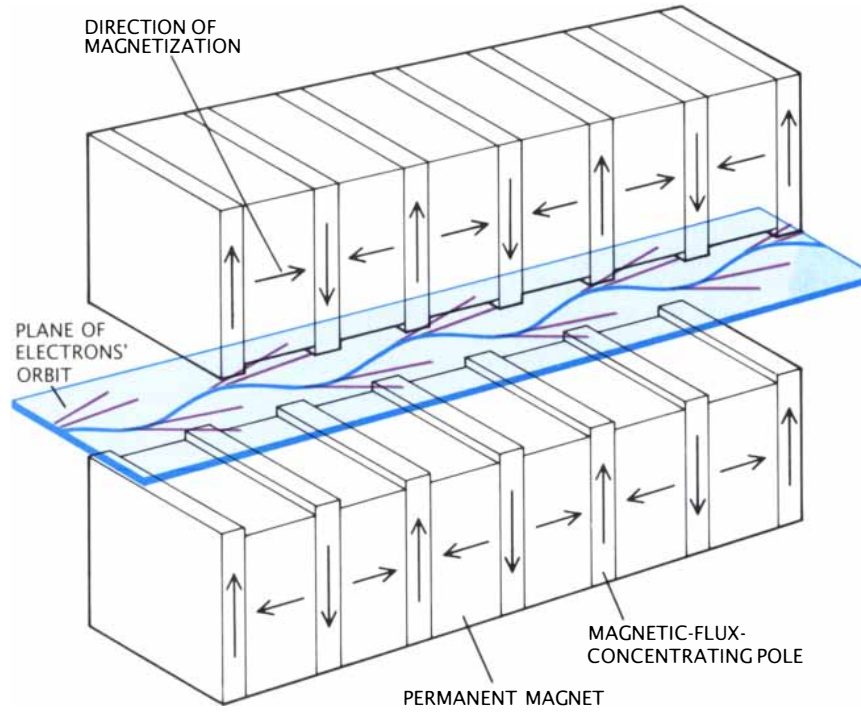
stein's famous equation $E = mc^2$) to the electron's total energy, which includes both its rest mass and its kinetic energy. An electron's energy is generally measured in units of electron volts (one electron volt is the energy gained by an electron in a vacuum as it is accelerated by a potential difference of one volt). Since the electron's rest mass represents an energy of .5 MeV (million electron volts), if an electron is given a total energy of, say, 5 GeV (billion electron volts), the natural angle over which synchrotron radiation is emitted is only .0001 radian, or about .006 degree. This can be regarded as a beam of nearly parallel rays—geometrically quite similar to the narrow beams of lasers.

As the electron traces a circular arc in a bending magnet, the axis of the narrow cone of radiation moves with the electron. Although the vertical emission angle remains small, the radiation sweeps a broad horizontal swath typically several degrees wide. Mirrors and other optical elements are used to collect as much of the horizontal swath as possible in order to deliver a large number of photons down a beam line.

Synchrotron radiation has other unique properties that conventional sources of VUV light (such as gas-discharge lamps) and X rays (such as electron-impact tubes) do not offer. For example, synchrotron radiation is highly polarized; that is, the oscillating electric field associated with the radiation vibrates primarily in the plane of the electron's orbit. In addition, because the electron beam is bunched, the radiation is emitted in pulses. These pulses can be as short as a few tenths of a billionth of a second, and, depending on the size of the ring and how many electron bunches are circulating, the rate at which the pulses are repeated can vary between about 100,000 and 500 million times per second.

There are also marked differences between the spectra of the VUV light and X-radiation produced from storage rings and those produced from conventional sources. Whereas synchrotron radiation is extremely powerful over a broad continuum of wavelengths, gas-discharge lamps and X-ray tubes emit radiation primarily in the form of a few discrete wavelengths superposed on a very weak background continuum.

Normally experimenters are interested only in a narrow bandwidth, or range of wavelengths, from the broad spectrum provided by synchrotron radiation. For a given bandwidth, synchrotron radiation from bending magnets can supply at least 1,000 times more photons per second than conventional VUV and X-ray sources. The number of photons emitted per second within a given bandwidth is called the flux of the radiation. The source characteristic of greatest importance, however, is often its brightness (sometimes called brilliance or radiance): the flux divided by the angle through which the radiation is emitted and the cross-sectional area of the source. Brightness is more or less a measure of the radiation's "concentration." Because synchrotron radiation has a large flux, is naturally emitted through a small angle and originates in an electron beam with a very small cross-sectional area (often less than a square millimeter), it has exceedingly high brightness. Indeed, synchrotron radiation from bending magnets is about a million times brighter than other VUV and X-ray sources.



INSERTION DEVICE, such as a wiggler or undulator, produces synchrotron radiation that is considerably brighter than radiation from a bending magnet. The device causes electrons to follow a sinusoidal path instead of a curved one by establishing a series of magnetic fields that alternate in polarity and that are perpendicular to the electrons' direction of travel. A wiggler enhances the brightness of the radiation produced by a given electron beam by a factor roughly equal to twice the number of full oscillations the beam undergoes. The deflections of the beam are smaller in an undulator than in a wiggler, and the radiation's brightness can in theory be increased by a factor about equal to the square of the number of oscillations, but only at discrete photon energies.

Wigglers and Undulators

The flux and brightness of synchrotron radiation can be increased even more if electrons are made to jiggle while coursing through a storage ring. This is the role of magnetic devices called wigglers and undulators, which are inserted between a storage ring's bending magnets. Both devices consist of two rows of magnets that establish many fields of alternating polarity perpendicular to the electron beam. The alternating magnetic fields cause the electron beam to move from side to side or up and down as it travels through the device, giving individual electrons a sinusoidal trajectory. As each electron curves, it emits radiation for much the same reason that it emits radiation while tracing a circular arc through a bending magnet.

The difference between wigglers and undulators lies in the amount of angular deflection produced by the magnetic poles. In a wiggler the angle is large compared with the natural emission angle of synchrotron radiation. This results in a continuous spectrum similar to what is produced in a bending magnet having the same magnetic field strength, although the radiation has greater flux and brightness. (The stronger the

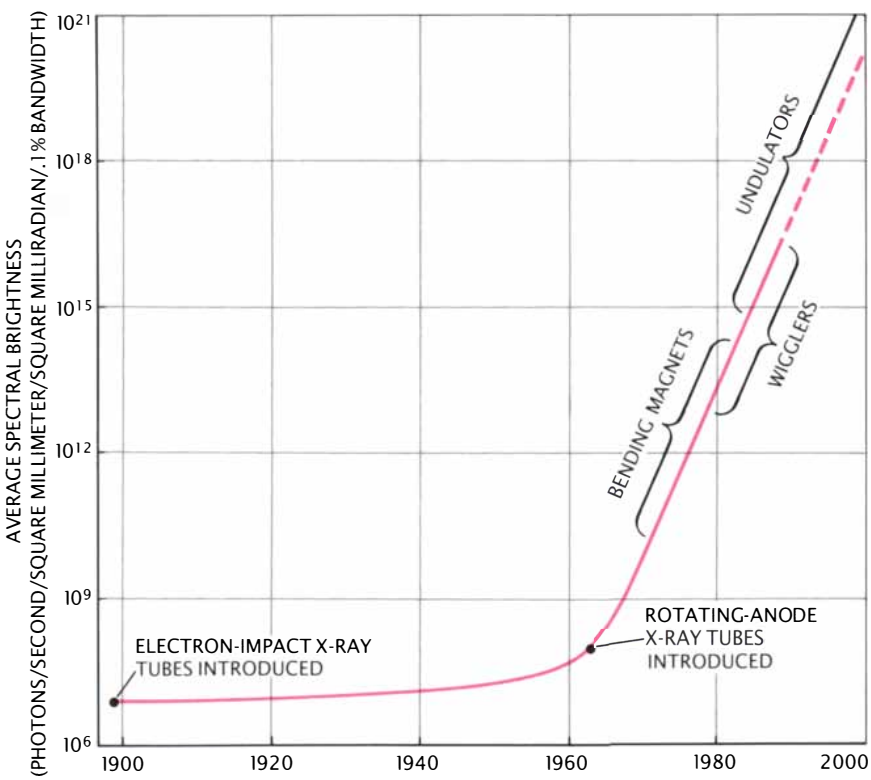
magnetic field of a bending magnet or a wiggler is, the higher the photon energies are to which the spectrum extends.) The flux and brightness of the radiation that emanates from a wiggler exceeds that of a single, large bending magnet of similar magnetic field strength by a factor approximately equal to the number of poles, or twice the number of full oscillations an electron beam undergoes.

In an undulator the angular deflection caused by each set of poles is smaller than or comparable to the natural emission angle. In this case the electromagnetic waves radiated by each electron can mutually reinforce or suppress one another to enhance the emission of certain wavelengths at certain emission angles, resulting in extremely high brightness at the enhanced wavelengths. The radiation emanating from an undulator also has a degree of coherence that is proportional to the radiation's brightness. Coherence is a measure of the ability to form patterns of illumination that result from the interference of electromagnetic waves originating at different points or arriving at different times.

For an ideal electron beam (one that does not tend to spread while it maintains a vanishingly small cross-sectional area) the brightness of the radiation produced by an undulator is greater than that produced by a bending magnet by a factor approximately equal to the square of the number of oscillations. Undulators that cause the electron beam to undergo 60 oscillations have already been built, and ones that cause several hundred oscillations are planned.

The first wiggler to serve as a radiation source for experiments was installed in 1978 in the 4-GeV SPEAR storage ring used by the Stanford Synchrotron Radiation Laboratory (SSRL). The device had six electromagnetic poles that caused three full oscillations of the electron beam and could sustain a peak magnetic field more than twice as strong as that of the bending magnets in the ring.

Although undulators were studied in linear accelerators at Stanford University during the 1950's and in synchrotrons in the U.S.S.R. during the 1970's, the devices were built chiefly to verify principles of their operation and produced primarily visible light. The development of an undulator capable of generating VUV light and X-radiation was hampered before 1979, because it was believed such an undulator would have to incorporate complex and costly superconducting coils in its electromagnetic poles so as to avoid overheating of the coils.



BRIGHTNESS OF X RAYS has increased by many orders of magnitude since the advent of synchrotron-radiation sources. Undulators in storage rings are the brightest source.

This situation changed following the seminal work of Klaus Halbach of the Lawrence Berkeley Laboratory (LBL). Halbach showed that, in place of electromagnets, strong permanent magnets made from rare-earth elements and cobalt could be arranged in simple arrays to construct alternating-field devices. Such devices have the property that the field strength in the gap between the poles remains constant as all linear dimensions are reduced in scale. Hence short-pole permanent-magnet insertion devices can be made (allowing more electron-beam oscillations in a given length), whereas devices based on electromagnets quickly run into coil-heating problems as they are scaled down. Halbach's work implied that permanent-magnet undulators suitable for use as VUV and X-ray sources in storage rings could be made rather simply and economically.

The first undulator intended as an X-ray source for experiments was installed in the SPEAR storage ring in late 1980, the result of a joint effort of the LBL and the SSRL. It delivered X-radiation having photon energies of about 1 KeV (1,000 electron volts) that was more than 100 times brighter than the radiation supplied by the ring's bending magnets. In its first application as a radiation source in 1981, a pinhole was employed to select the proper bandwidth for the experiment, taking advantage of the natural correlation between wavelength and emission angle of undulator radiation. The undulator delivered four orders of magnitude more photons than had been available earlier from a bending-magnet beam line. A similar device was built at about the same time in Novosibirsk in the U.S.S.R. by Gennady Kulipanov and Niccolai Vinokurov.

The demonstration gave the SSRL and other laboratories a strong impetus to develop insertion devices, namely wigglers and undulators, for their rings. At present a total of about 20 insertion-device beam lines are in use as radiation sources around the world, and more are under construction. Major efforts are under way to add insertion-device beam lines to the largest facilities fully dedicated to the production of synchrotron radiation, such as the Photon Factory in Japan and the National Synchrotron Light Source at Brookhaven.

Scientific Applications

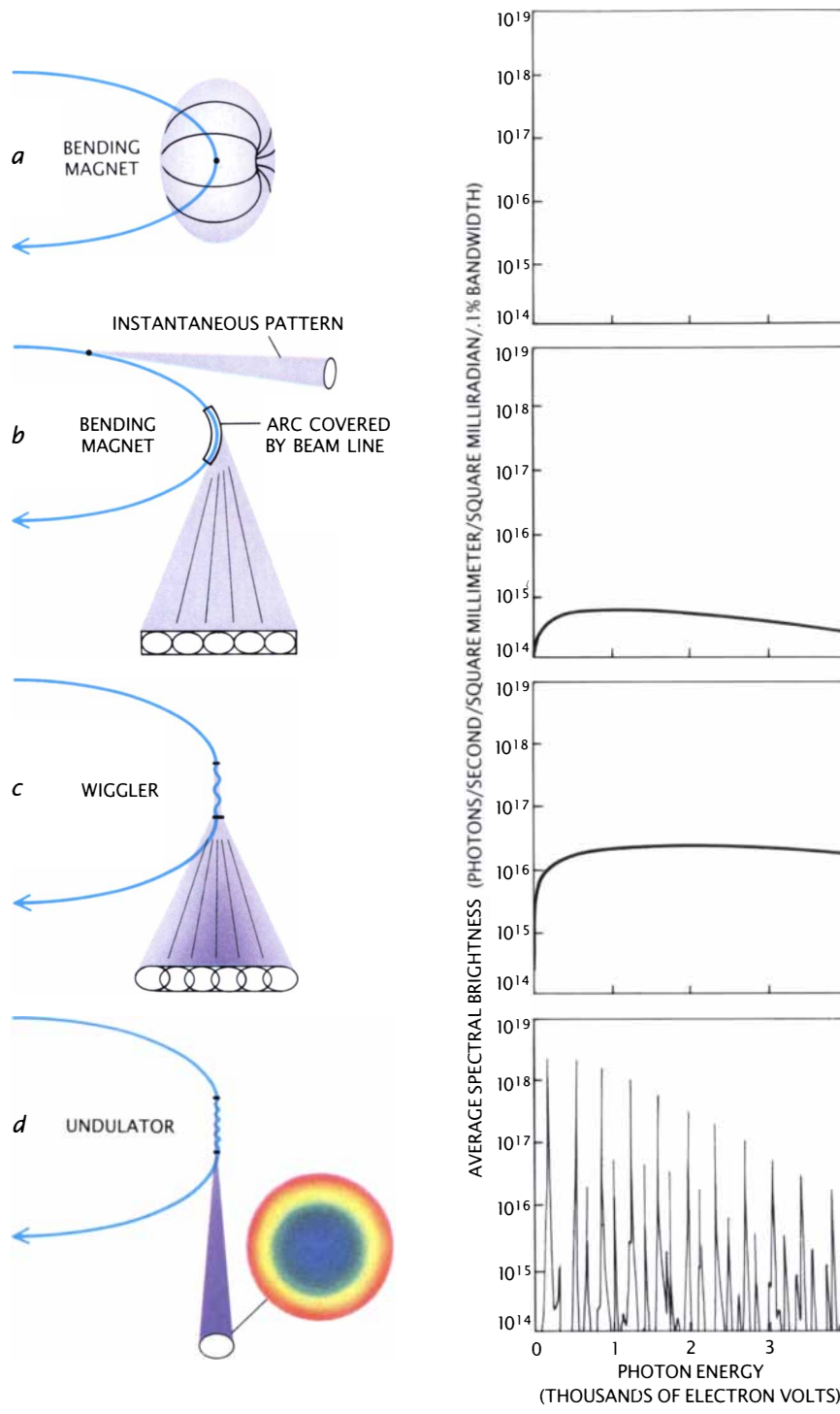
A broad range of scientific applications will stand to benefit from these

improvements in synchrotron-radiation sources. Wavelengths within the VUV and X-ray ranges spanned by synchrotron radiation have the same dimensions as characteristic structures in atoms and molecules, making the radiation an excellent probe of the structure of matter. Moreover, synchrotron radiation is "tunable": a very narrow bandwidth can be selected from the intense, broadband source by means of a monochromator. In the VUV domain a monochromator is a diffraction grating: a slab of polished material on which many closely spaced grooves have been scratched or etched.

In the X-ray domain gratings cannot be employed, because it is not possible to make the grooves close enough to one another to match the shorter wavelengths of the radiation. Instead planes of atoms in a large monolithic crystal are employed. A particular wavelength is reflected by the electrons of the atoms in a process known as Bragg scattering. In this way beams of X rays can be made with energies defined to about one part in 10,000 and in some cases 100,000. Recently investigators succeeded in defining the energy of X-ray beams down to a phenomenal one part in a trillion by scattering synchrotron radiation off the nuclei of atoms in a crystalline lattice (so-called nuclear Bragg scattering). The fact that synchrotron radiation can be tuned so precisely significantly extends the range of phenomena that can be studied using these X-ray beams.

The tunability of synchrotron radiation makes it possible to exploit the fact that each element exhibits a sharp increase in absorption at certain wavelengths, called absorption edges. The absorption edges arise from the fact that an electron in a given energy level can absorb a photon (and escape from an atom) only when the photon's energy exceeds a certain well-defined value known as the electron's binding energy.

In a technique known as extended X-ray absorption fine structure (EXAFS) much detailed information about the local environment around atoms of a particular element in a complex material is extracted from measurements of X-ray absorption made in a continuous scan of the photon energy fromjust below an absorption edge of the element to several hundred volts above the edge [see "The Analysis of Materials by X-Ray Absorption," by Edward A. Stern; SCIENTIFIC AMERICAN, April, 1976]. A



EMISSION PATTERN AND ENERGY SPECTRUM of photons generated in a storage ring depend on the speed of the electrons and the device through which they travel. If the electrons were made to pass through a bending magnet at speeds much less than the speed of light (*a*), they would emit radiation in almost all directions, but the radiation would be so diffuse and of such low energy that it would not even appear on the graph. Electrons traveling through a bending magnet at relativistic speeds (*b*) produce much more intense radiation, including photons at higher energies, emitted in a narrow cone tangent to their curved path. Because the axis of the cone coincides with the electrons' direction of travel, it traces a swath as the electrons circulate. Relativistic electrons in a wiggler (*c*) produce a similar but much brighter beam that results from the superposition of radiation arising from each electron-beam oscillation. The beam emanating from an undulator (*d*) is much more tightly focused and therefore brighter. Owing to interference effects, photons are concentrated according to their energy (indicated by different colors) at specific angles from the axis of the beam, making it possible to employ a pinhole or an annular slit to select radiation of a particular energy.

New Cavalier RS coupe. You may never fall for another car's lines.

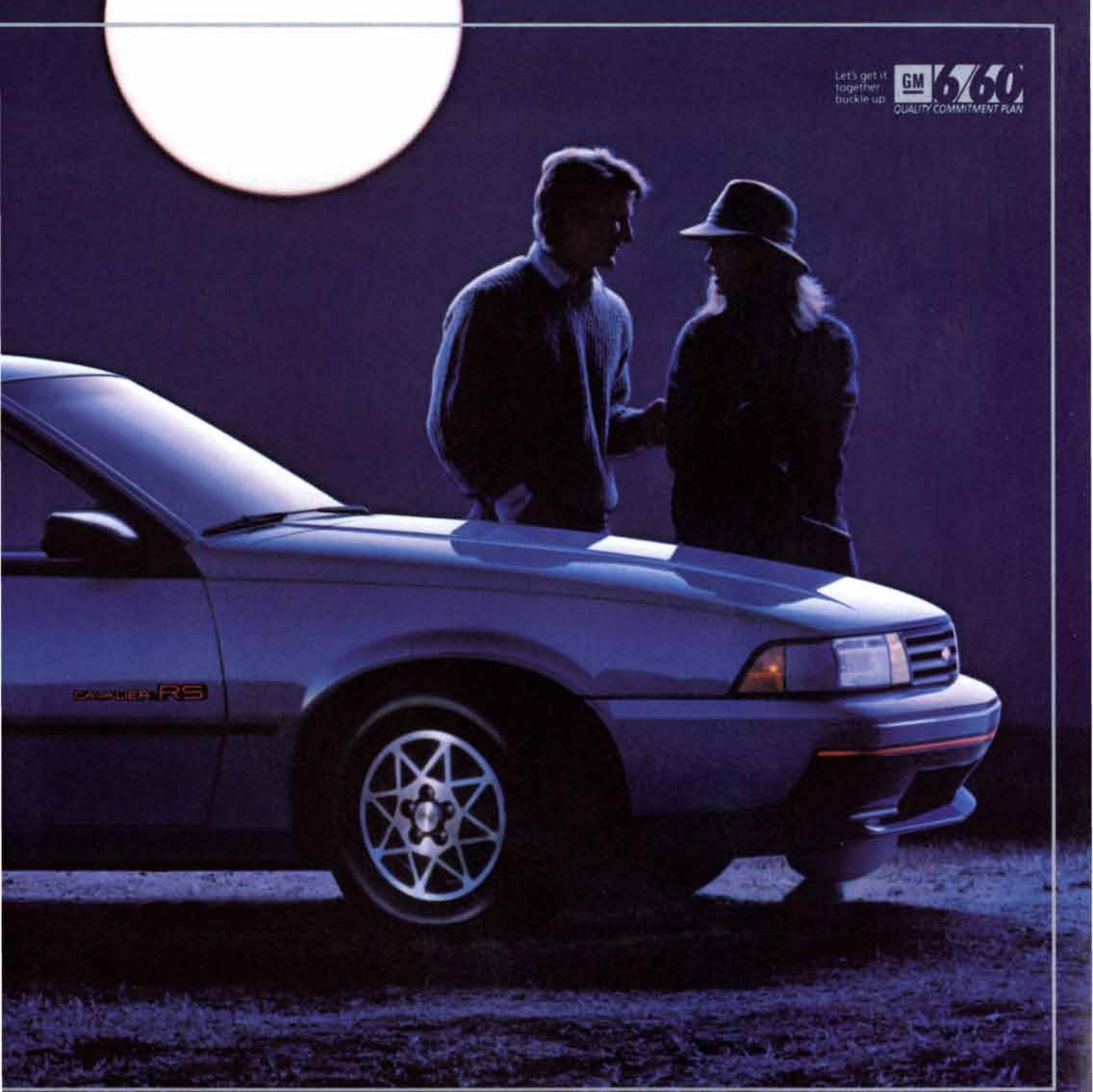
No matter what promises others make, this car's got the lines that will move you: the 1988 Cavalier RS. The silhouette turning heads in the fast lane.

Aerostyled and uplifted, inside and out, from halogen headlamps to shapely rear end, it romances the road on a front-wheel drivetrain and available sport suspension. Then cozies up its 5-passenger interior with ergonomically designed seating. Rich cloth covering. Cut-pile carpeting. And a standard AM/FM stereo.*

All of which helped make the Chevy Cavalier one of America's most popular car lines.

So before you're taken by another, check out the Cavalier RS. Once you set your sights on these lines, few others will lure you away.

*May be deleted for credit.



- Electronic Fuel Injected 2.0 Liter engine.
- Rack-and-pinion power steering.
- Optional aluminum wheels.
- Performance-tuned shock valving.
- 6-year/60,000-mile powertrain warranty.
  See your Chevrolet dealer for terms and conditions of this limited warranty.

THE *Heartbeat* OF AMERICA

TODAY'S CHEVROLET

synchrotron-radiation source makes it possible to do such a photon-energy scan in a few minutes (in some cases in a few seconds); with conventional X-ray generators days or weeks were necessary. Hence by employing synchrotron radiation instead of conventional X rays, investigators can measure many more samples, study more dilute species and follow structural changes during chemical reactions as they occur. In short, synchrotron radiation has transformed the technique of EXAFS from an interesting curiosity to a broadly applicable tool of analytical chemistry.

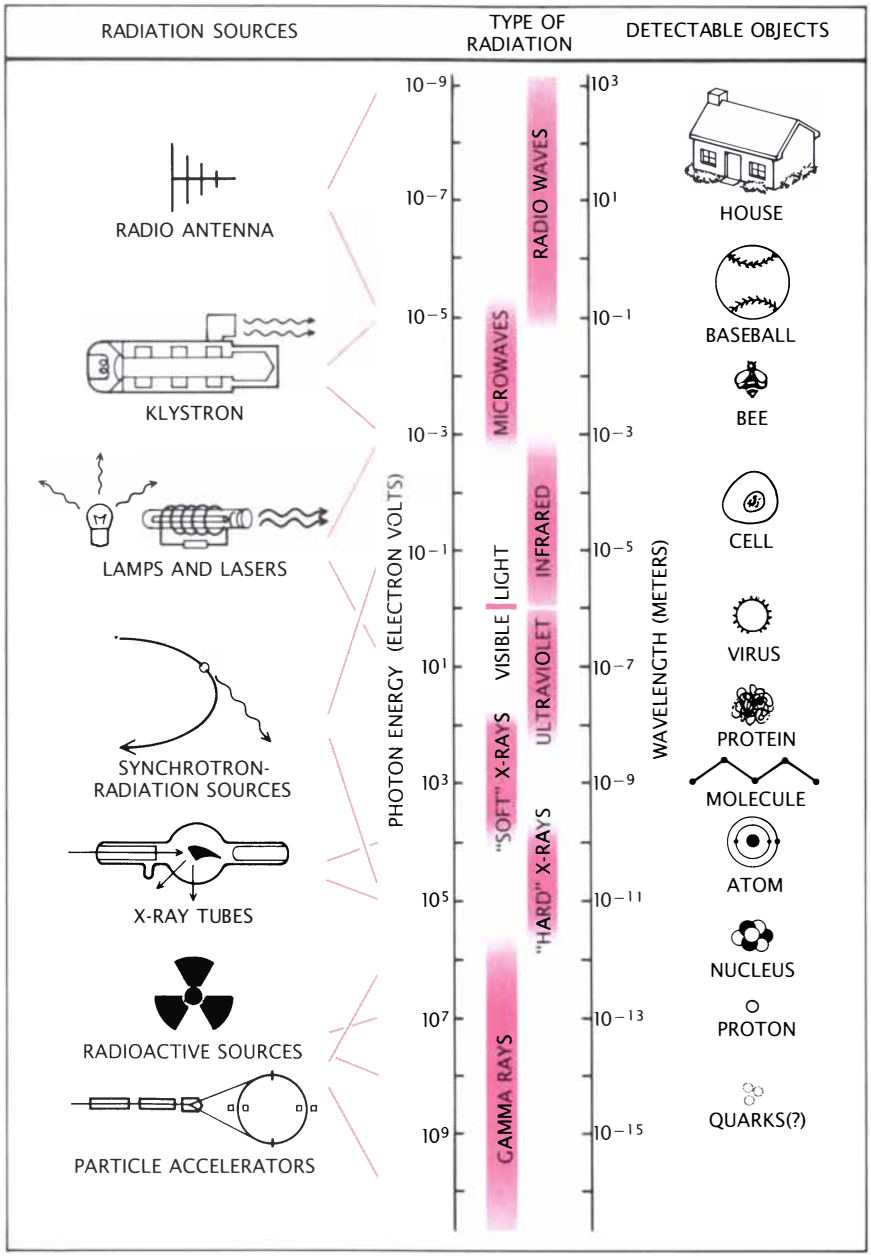
The tunability, brightness, collimation and beam stability of synchrotron radiation also recently made it possible to develop X-ray microtomography: a new microscopic-imaging technique that produces views of cross-sectional slices through a small sample. The technique yields images with a resolution approaching a micrometer (a millionth of a meter)—1,000 times better than conventional tomographic scans. By tuning the X-ray energies to bracket particular absorption edges, images can be processed to produce element-specific structural maps.

The diffraction, or scattering, of X rays rather than their absorption can serve as a probe of the magnetic properties of materials and phase transitions in two-dimensional systems such as surface layers and liquid-crystal films. In both these applications the effects of interest are very weak, and only extremely high-intensity radiation yields a reasonable number of scattered X rays. From the angles and energies at which X rays are scattered considerable information can be obtained. These and other X-ray techniques have opened up entirely new possibilities for learning about atomic arrangements in many condensed-matter systems.

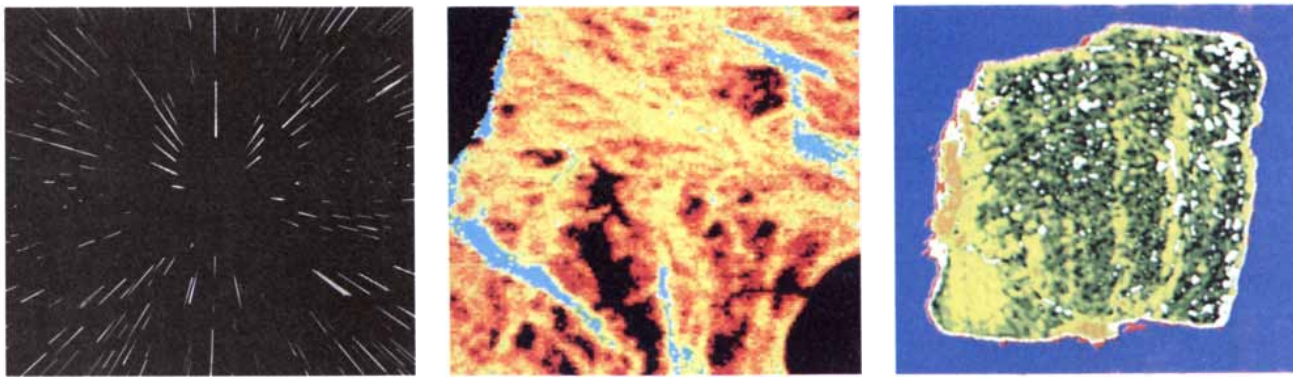
Other properties characteristic of synchrotron radiation can be exploited as well. For example, the radiation's polarization makes it possible to determine the orientation of molecules adsorbed on surfaces, since the probability that an electron in a molecule will absorb an X-ray photon depends on the orientation of the molecule with respect to the direction of polarization of the photon. The information gleaned from these types of measurements is of great importance, since such surface phenomena underlie the corrosion of materials and the catalytic processes on which the refining of oil depends.

Similarly, the pulsed nature of synchrotron radiation can be turned to advantage in many experiments, for instance those designed to determine the lifetime of excited states of atoms or molecules. A sharp, well-defined pulse of synchrotron radiation turns on a timer at the same instant that it strikes a sample of material, exciting atoms or molecules into a higher energy level. The timer is then switched off when a detector registers the arrival of a particle or photon emitted by the excited atom or molecule as it decays to a lower energy level.

Most intriguing perhaps are the potential applications made possible by the coherent radiation provided by undulators. Although it is not as coherent as visible light from some lasers, undulator radiation has much higher levels of coherence than was previously available in the VUV and X-ray regions of the spectrum. As more beam lines and more powerful



SPECTRUM OF ELECTROMAGNETIC RADIATION represents the range of photon wavelengths or energies that are most scientifically useful. The spectrum can be covered only by a multiplicity of radiation sources. Radiation employed to detect an object must have a wavelength that is equal to or less than the object's dimensions. Synchrotron-radiation sources provide exceedingly bright radiation in the ultraviolet and X-ray regions, which span the wavelengths suited to study atomic and molecular structures.



IMAGES made with synchrotron radiation display its versatility. X rays produced by storage rings can be scattered from a protein crystal to make a diffraction pattern (*left*) that is recorded on film. Such patterns yield information about the crystal's structure. (The streaks are an artifact of the crystal's rotation during the exposure.) Because each element's absorption of X rays increases sharply at characteristic photon energies, known as absorption edges, the "tunable" X-radiation that can be produced at synchrotron-radiation sources allows element-specific images to be made. A scanning X-ray micrograph of a piece of human skull (*middle*) illustrates the calcium distribution in the bone. The sensitivity of X rays to specific elements can also be put to use in X-ray microtomography, a technique that enables investigators to reconstruct a two-dimensional slice of a sample from numerous X-ray projections. Such an image reveals the internal microstructure of a small lump of coal in cross section (*right*).

undulators are installed in storage rings, applications based on this coherence, such as X-ray holographic imaging, will no doubt be advanced.

Technological Applications

The unique characteristics of synchrotron radiation open up new capabilities for medicine and industry as well. The capability to tune the radiation to an element's absorption edges, for example, can be fruitfully applied in medical diagnostic radiography. Currently angiograms, or images of blood vessels, are obtained by injecting an iodine-containing solution into the circulation and exposing the patient to conventionally generated X rays. Because these X rays are not tuned to the absorption edge of iodine and are also relatively weak, a rather high concentration of iodine is required. Moreover, in the case of angiograms of the coronary arteries the exposure must be no longer than a few milliseconds to prevent blurring due to heart motion. The required iodine concentration is achieved in coronary angiography by injecting the solution through a catheter that has been worked down a major artery until it is in the vicinity of the coronary arteries. This procedure produces radiograms of excellent quality, but it is invasive and associated with significant risks.

The sensitivity to contrast agents such as iodine can be enhanced by two or more orders of magnitude by noting the difference between two exposures made with monochromatic X-ray beams, one beam with an energy just above the absorption edge and the other with an energy just below it [*see illustration on next page*]. Such beams can be readily produced using synchrotron radiation. Taking the difference between the two images enhances the contrast between the blood vessels carrying the iodine and the other tissues of the body, since the nonvascular tissue is in effect "subtracted out."

This dual-energy subtraction technique may make it possible to image the coronary arteries in less time and with a much smaller iodine concentration than is needed for conventional angiography. The lower concentration may eliminate the need for an interarterial catheter, thereby lowering the risk considerably.

The collimated and intense X-radiation that can be generated in storage rings has yet another application, potentially of great significance to the electronics industry: the manufacture of high-density integrated circuits. Integrated-circuit chips are currently manufactured by applying optical lithographic techniques to etch successive layers of circuitry into a wafer of a semiconductor material. The pattern of the circuitry is defined by coating the wafer with a light-sensitive substance called photoresist and illuminating the coated wafer with visible or ultraviolet light through a "mask," or a stencil imprinted with the pattern. Depending on the type of photoresist, either the exposed or the unexposed photoresist is easily washed away and the remaining sections of photoresist outline the pattern.

The minimum size, and hence the maximum number, of elements in a circuit is set by diffraction effects that blur the pattern exposed through the mask. The extent of these effects depends on the wavelength of radiation illuminating the mask. If visible or ultraviolet light is employed, only features larger than about 500 nanometers can be replicated accurately. Shorter-wavelength light ("soft" X rays with a wavelength of between one nanometer and two nanometers seem optimal) would make it possible to replicate faithfully features on the order of 50 nanometers or less. In a related application, called X-ray micromachining, nonelectronic structures with very large height-to-width ratios can be carved out of the surface of materials if the photoresist is illuminated with higher-energy X rays. This lithographic technique can be applied to make microscopic filters and nozzles.

In order for X-ray lithography to be economical as a mass-production technique, however, exposure times of about a second are required. It appears that only synchrotron-radiation sources can supply X-radiation at the necessary intensities to make such exposure times feasible. The possibility that X-ray lithography will be the technique of choice to mass-produce circuit elements smaller than those produced by visible- or ultraviolet-light lithography has led groups in Japan, West Germany, Great Britain and the U.S. to design special storage rings for the purpose. To achieve the most compact production facility, some storage-ring designs incorporate superconducting bending magnets.

In spite of the myriad applications

of synchrotron radiation, none of the currently operating storage rings is capable of achieving the full potential from insertion devices, in particular undulators. This has led to an intense worldwide activity to design and construct the next generation of storage rings specifically designed to accommodate wigglers and undulators. The first of these, the .8-GeV Super-ACO ring in Orsay, France, began operation early in 1987. Several other rings have recently been authorized for construction. These include the 1.3-GeV Taiwan Light Source, the 6-GeV European Synchrotron Radiation Facility in Grenoble, the 1.5-GeV Advanced Light Source at Berkeley, a 1.5-GeV ring in Trieste and a 2.5-GeV ring in Moscow. Other groups are in the process of developing proposals for such rings, including groups at the Argonne National Laboratory and at laboratories in Brazil, West Germany and the U.S.S.R.

These new storage rings will have many straight sections to accommodate wigglers and undulators, and they will have electron-beam properties optimized to extend significantly the performance of these insertion devices. In particular the stored electron beams in these rings will have a small diameter and will tend not to dilate—a combination of properties termed low emittance.

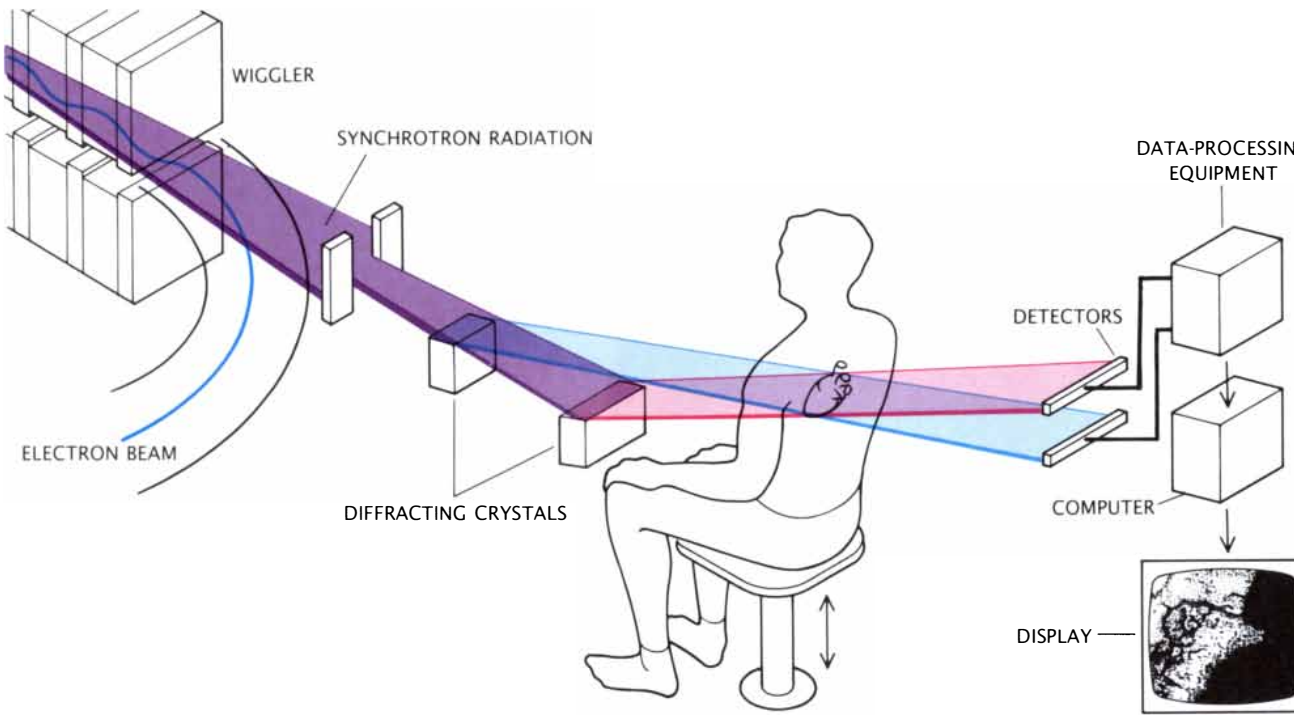
As was recognized in 1976 by Ken Green and Rena Chasman of the Brookhaven National Laboratory, the lower the emittance of an electron beam is, the higher will be the brightness of the radiation beams it produces. Most existing rings have emittances of 100 nanometer-radians or more. Some rings under construction have design values between five and 10 nanometer-radians, which will allow undulators in them to have a brightness about 100 times higher than undulators in existing rings.

Yet the maximum practical levels of brightness and coherent power for undulator radiation are reached only at the so-called diffraction limit, in which the emittance of the electron beam is numerically about equal to one-tenth of the wavelength of the radiation. Any lower emittance would not improve the radiation owing to fundamental diffraction effects. Thus a diffraction-limited electron beam capable of producing .1 nanometer (12.4 KeV) X rays would have an emittance of about .01 nanometer-radian.

Improving Ring Performance

Achieving such exceedingly low emittances clearly poses a daunting challenge to designers of storage rings. The emittance of the beam in an electron storage ring is determined by an equilibrium established between processes that cause individual electrons to oscillate transversely to the beam and processes that damp these oscillations, which are known as betatron oscillations. Betatron oscillations are excited by emissions of photons that are at a slight angle to the direction of motion of an electron, since the emissions cause the electron to recoil in a direction transverse to its motion. Radio-frequency cavities damp the betatron oscillations as they restore the energy lost to synchrotron radiation, because the cavities impart momentum to electrons only in the beam's direction of motion.

The resultant electron-beam emittance depends on the specific arrangement of bending and focusing



CORONARY ANGIOGRAPHY (imaging of the coronary arteries) is more sensitive and potentially less risky when synchrotron radiation is used instead of conventional X-ray sources. An iodine-containing liquid has been injected into the patient's blood. Special diffracting crystals convert synchrotron radiation into beams with photon energies just above and just below an absorption edge of iodine. The two beams converge at the patient's heart and then diverge to strike separate detectors. The higher-energy beam is absorbed nearly as much as the lower-energy beam by soft tissue and bone, but the iodine absorbs much more of the higher-energy beam. Taking the difference between the images generated by the two detectors in effect "subtracts out" most of the patient's body tissues and bone, leaving an image in which the coronary arteries and other blood vessels are enhanced. Conventional angiography, because it is less sensitive, requires that the iodine be infused directly into the arteries.

magnets, which make up what is called the magnet lattice of the ring. The magnet lattice must also provide adequate "dynamic aperture": room for electrons to deviate in stable betatron oscillations from the path designed for the electron beam. Electrons freshly injected into the ring undergo such deviations; they are also induced when individual electrons scatter from residual gas molecules in the vacuum system. The oscillating electrons eventually accumulate at the core of the stored beam as a result of the combined damping effect of synchrotron radiation and radio-frequency cavities, nevertheless the magnet lattice must provide sufficient dynamic aperture to keep the beam from losing too many electrons until the damping takes place. A large dynamic aperture also eases tolerances on the alignment and field quality of the magnets and facilitates operation of the ring.

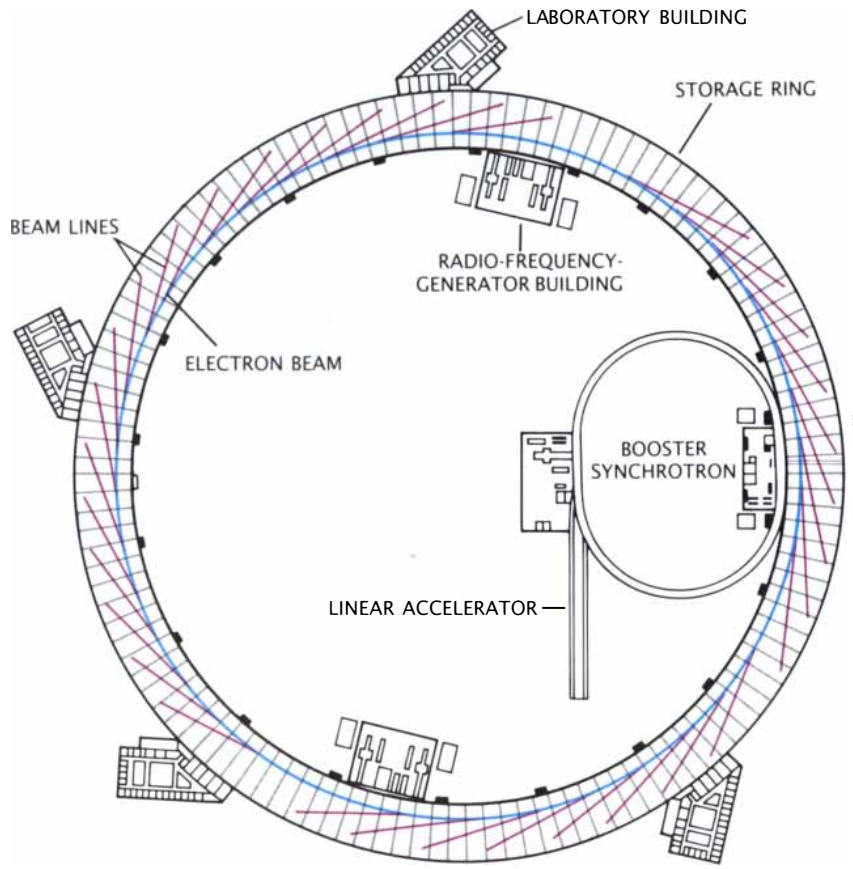
Low emittance can be generally achieved by focusing the beam with many strong quadrupole and sextupole magnets. Unfortunately the nonlinear character of sextupole fields, which are needed to correct aberrations in the focusing effects of the quadrupole fields, results in a limiting of the range of stable particle oscillation—that is, in a small dynamic aperture. As a result particular care must be taken in the design of the magnet lattice so that a large enough dynamic aperture is preserved as the emittance is reduced.

Another possible way to achieve very low emittance is to place tens or even hundreds of meters of specially designed wigglers in the ring. This approach may be a good way to obtain extremely low emittance, because it should not reduce the dynamic aperture. Helmut Wiedemann of Stanford University has calculated that if 200 meters of damping wigglers are employed, the emittance of the PEP storage ring at Stanford could be reduced by about a factor of 10 when the ring is operated at 6 GeV, reaching a respectable value of about .5 nanometer-radian.

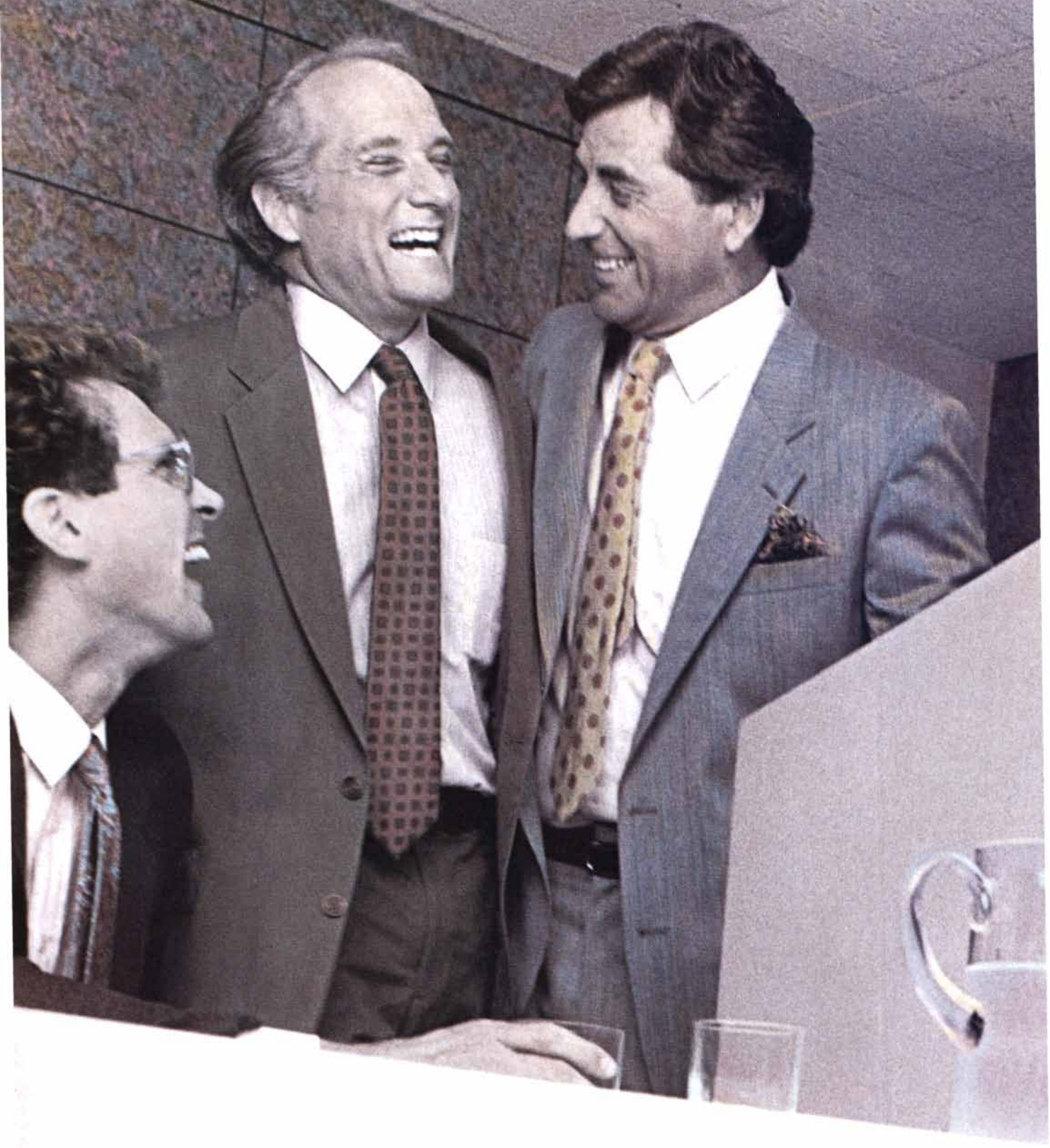
There are many other technical challenges involved in constructing synchrotron-radiation sources optimized for insertion devices. For example, positive ions produced from interactions of the electron beam with the residual gas in the vacuum chamber—even at operating pressures of a few trillionths of atmospheric pressure—are trapped by the negatively charged electron beam, causing increased scattering and uncontrolled focusing effects. The best solution to the ion-trapping problem in storage rings appears to be to store not electrons but positrons (the antimatter counterparts of electrons). Positive ions are repelled by the positron beam, which also produces synchrotron radiation.

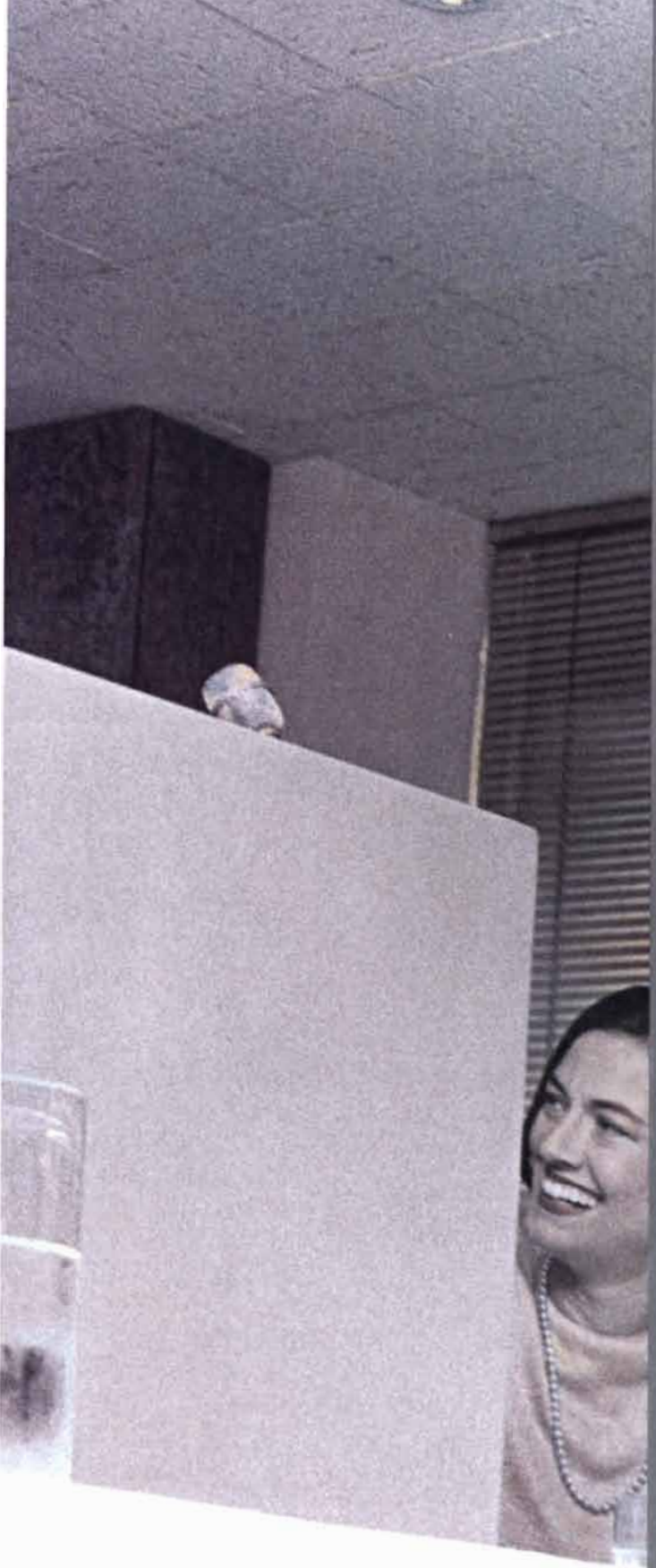
The higher radiated flux and power density in new storage rings also bring new problems. Insertion devices currently in operation on multi-GeV rings produce a total radiated power of several kilowatts, and beam-line components are exposed to a power density of up to 10 kilowatts per square centimeter. This is about the power density found in a welding arc! The even higher power density that future synchrotron-radiation sources will yield therefore requires further improvements in the design of beam-line components, particularly optical components that must maintain fine precision under the intense thermal stress. The higher flux also necessitates the development of high-speed photon detectors and more powerful data-processing systems. In short, virtually all the components of existing beam lines and experimental stations must be upgraded to match the capabilities of the next-generation sources now being designed and built.

Nevertheless, the effort entailed will be worthwhile. As these new rings are brought into operation, a large community of workers in basic and applied research will be able to tackle many known phenomena that are inscrutable to investigators employing current sources of radiation. Even more exciting is the expectation that these new sources will open the way to important new science. Indeed, the past two decades have shown that with each order-of-magnitude improvement in flux, brightness, coherence or some other property of synchrotron radiation, novel applications have come—in many cases unimagined until the source was actually available.



LARGE SYNCHROTRON-RADIATION FACILITY, such as the one depicted here, can reach a kilometer in circumference and serve thousands of investigators carrying out small-scale experiments at the ends of many beam lines. The storage ring, which circulates electrons or positrons with energies between 6 and 8 GeV (billion electron volts), is designed to produce extremely bright radiation with photon energies well above 10 KeV (thousand electron volts). One such facility is being built in Grenoble, France, as a collaborative European project; others are being proposed in Japan and the U.S.

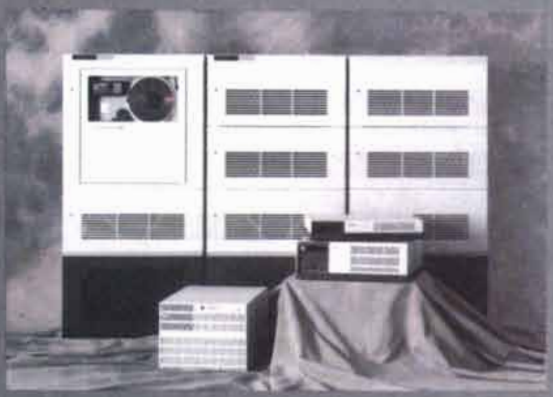




"They say this place won't be the same without me,

but I'd rather talk about you. And what we've accomplished this past year. 'The year of the 3B Computers,' some of you call it. Last June, Marketing and Sales and Accounting were like separate companies. Cut off by separate systems. Separate software. Now Fran's bean-counters can share what they know with us. Ditto for Kelly's people. And what we do helps them. And if you don't believe it's working, ask Fran. All her numbers are up. Way up. Which is a relief to me, because I know some of you wondered what an old sales guy like me knew about computers in the first place. Frankly, not much. But I know AT&T.

AT&T comes through."



The AT&T 3B Computer Family

Now no department need be an island. The same 3B Computer that knits up to 300 individuals into an efficient departmental workgroup can break down the barriers *between* departments.

Reason: AT&T's unique ability to put computers in touch. Every 3B Computer connects with every other 3B. And to the worlds of IBM,* DEC** and many others.

Result: Departmental users are no longer prisoners of their own system. Information can be shared across all departments, and used wherever it can help your business succeed.

AT&T 3B Computers and a world of applications are available right now. Learn more. Call 1 800 247-1212.

From equipment to networking, from computers to communications, AT&T is the right choice.

*IBM is a registered trademark of International Business Machines Corp.
**DEC is a trademark of Digital Equipment Corp.

AT&T
The right choice.

The Ancestry of the Giant Panda

Is the panda a bear? Is it a raccoon? Or does it actually belong in a family of its own? Molecular analysis provides new insights into this long-standing genealogical problem

by Stephen J. O'Brien

The shy and gentle panda, *Ailuropoda melanoleuca,* leads a nomadic life high in alpine bamboo forests on the edge of the Tibetan plateau in western China. With its huge eyespots, Mickey Mouse ears and playful, juvenile appearance, it is one of the most widely recognized and cherished of all animal species. Yet from an evolutionary point of view it has also been one of the most enigmatic. For almost 120 years biologists have disagreed about the correct taxonomic placement of this species, alternatively placing it with bears in the family Ursidae, with raccoons in the family Procyonidae or in its own family, the Ailuropodidae. My studies and those of my colleagues at the National Cancer Institute and the National Zoological Park in Washington suggest that the riddle of the panda's ancestry has now been solved.

The controversy began in the year 1869 when the giant panda was first described by Père Armand David, a French missionary and naturalist based in China. David recognized the panda as a species new to science and named it *Ursus melanoleucus* (for black-and-white bear). The following year Alphonse Milne-Edwards, a colleague of David's (and later director of the Museum of Natural History in Paris), examined skeletal material sent to him by David and concluded that the giant panda was more closely related to the lesser or red panda, *Ailurus fulgens* (also from China), than it was to a bear. Because the ancestry of the red panda had already been established (it was considered a member of the raccoon family), Milne-Edwards believed the bearlike appearance of the giant panda was a holdover from the past, a reflection of its descent from an ancestor common to both bears and raccoons. He renamed it *Ailuropoda melanoleuca.*

Since that time more than 40 treatises on the subject of the giant panda's ancestry have been published, but in spite of these labors no clear-cut consensus emerged from them. As recently as 1964 D. Dwight Davis, curator of mammals at the Field Museum of Natural History in Chicago, published an extensive account of the anatomy of *Ailuropoda* and concluded on the basis of a comparative study of 50 organ systems that the panda was a member of the bear family. Soon thereafter Desmond Morris, curator at the Zoological Society of London, and R. F. Ewer of the University of Ghana independently con-

LING-LING and her male partner Hsing-Hsing at the National Zoological Park in Washington (*right*), like all giant pandas, have a slow, ambling gait that reflects the long

cluded on behavioral and morphological grounds that the giant panda belonged in the raccoon family. More recently ethologists John F. Eisenberg of the University of Florida and George B. Schaller of the New York Zoological Society recommended that the giant panda be placed in its own family, the Ailuropodidae. That view was endorsed in 1986 by a team of Chinese scientists including several from the Beijing Zoo and Beijing University, who published a lengthy anatomical account of the panda.

Such problems reflect the difficulties encountered by systematists who must identify species and determine their relationships to other organisms. One of the key decisions a systematist must make is whether a character is homologous to a similar character in another species or whether it is merely analogous. If a trait is homologous, it has arisen as the result of common descent; that is, all the species possessing it have descended from a common ancestor. All members of the class Mammalia, for example, have body hair, young that are nourished by the mammary glands of females, and four-chambered hearts. All members of the cat family have retractable claws, four toes on their hind feet and contractile pupils. The degree of evolutionary relationship is determined by counting the number of homologous characteristics shared by different species; the more closely related two species are, the greater the number of traits they have in common is.

One difficulty in identifying a homologous character is that its genetic basis is rarely understood. A small morphological change may result from extensive reorganization of multiple genes and thus represent a quantum jump in evolution at the molecular level. Conversely, an apparently major structural change (such as the appearance of an opposable thumb in the panda) may result from small genetic changes.

If a trait possessed by two or more species is analogous, it means those species owe their similarity to convergence, a process by which unrelated organisms evolve similar structures in response to the same environment. The wings of birds and insects are an example of convergence; they perform identical functions but have arisen by independent means. Because convergent characters do not share a common descent, they may confound the deductive process and mislead the naturalist. That is the case with the giant panda.

The giant panda looks like a bear but has many traits that are not at all bearlike. Like the red panda, it is largely herbivorous, existing primarily on a diet of bamboo; it has a massive head, large, flattened teeth and well-developed jaws and jaw muscles that provide the grinding power

hours they spend upright feeding on bamboo. Red pandas (*left*) are raccoons, not bears; like the giant panda, however, they feed primarily on bamboo and have some anatomical features, such as flat grinding teeth with multiple cusps, in common with it.

GIANT PANDAS are indigenous to China, where they live in regions (*color*) of alpine forest on the edge of the Tibetan plateau. Habitat loss, attributable to human population growth, deforestation and the cyclical dying off of bamboo forests, threatens the panda with extinction. Fewer than 1,000 pandas exist in these regions of China today.

necessary for its high-fiber vegetarian diet. One of the giant panda's unique features is that it is the only species outside the apes known to have a functional opposable thumb. The two types of thumb are strictly analogous; the panda thumb, unlike the primate thumb (which is a modified fifth digit), is an extension of the radial sesamoid wristbone and functions as a sixth digit, enabling the panda to strip the leaves from bamboo stalks with considerable dexterity. These characteristics and others have combined to produce an ambling creature specialized for sitting upright on its hindquarters for long periods of time (as much as 10 to 12 hours per day) eating bamboo in a peaceful predator-free environment.

In addition the giant panda is an atypical bear for reasons not directly related to its diet. Alpine bears generally hibernate, whereas the giant panda does not. (Bamboo is thought to provide insufficient fuel for hibernation.) Another characteristic of the panda that has confused taxonomists is its vocal sounds: the animal does not growl or roar but bleats like a sheep. On a cellular level, biologists have been perplexed by the fact that in both number and morphology its chromosomes, of which there are 21 pairs, more closely resemble those of the red panda (22 pairs) than they do those of bears (37 pairs).

My colleagues at the National Cancer Institute and I became interested in the question of the giant panda's ancestry in 1983, when we were asked to determine the paternity of a cub born to Ling-Ling, one of two pandas at the National Zoo. The question had arisen as the result of rather bizarre circumstances. After several years of cohabitation Ling-Ling and her male partner Hsing-Hsing had failed to mate, and a decision was made to artificially inseminate Ling-Ling with sperm from Chia-Chia, a panda in captivity at the London Zoo. To everyone's amazement, 24 hours before the scheduled insemination procedure Ling-Ling mated with Hsing-Hsing for the first time. In spite of this, zoo officials decided to proceed with the artificial insemination and Ling-Ling received sperm from Chia-Chia in addition to the sperm she had already received from Hsing-Hsing. This maximized the chances of successful fertilization but left the paternity of the cub in question.

Although the panda cub died within a few hours of birth, paternity was readily established (as it can be for human beings) by a method called gel electrophoresis. Different proteins were isolated from the cub's tissue samples and placed on a gel matrix, where they separated when exposed to an electric field. By comparing the cub's patterns of separation (identical proteins travel the same distance on the gel) with those of both Hsing-Hsing and Chia-Chia, we were able to determine that the father was indeed Hsing-Hsing.

We then asked ourselves whether we could employ similar genetic techniques to obtain data relevant to the ancestry of the giant panda. If the panda was more similar on a molecular level to raccoons than to bears, or vice versa, this would certainly provide compelling evidence favoring one genealogy over the other. According to the rules of homology, we could expect the proteins of the giant panda (and the DNA sequences that encode them) to be most like those of its closest relatives. We set out to compare homologous proteins from the giant panda, the red panda, the raccoon and several bears.

To make sense of our results one needs to understand the "molecular clock" hypothesis. This concept, originally articulated in 1962 by Emile Zuckerkandl and Linus Pauling, then at the California Institute of Technology, is simple and powerful. It is based on the premise that the genetic material of reproductively isolated populations diverges continuously over the course of time. The populations diverge because random mutations arise in their DNA (the hereditary material) and are passed on to subsequent generations. These mutations may occur in regions of DNA that code for particular amino acids (the building blocks of proteins), or they may affect noncoding regions of DNA that have no apparent effect on the composition of the organism. In either case, because they accumulate with time, the extent to which two species differ—either in a sequence of DNA nucleotides or in the amino acid sequence of a particular protein—is a good indication of relationship. On the basis of this information, geneticists can estimate the approximate point in time

when two species diverged from a common ancestor.

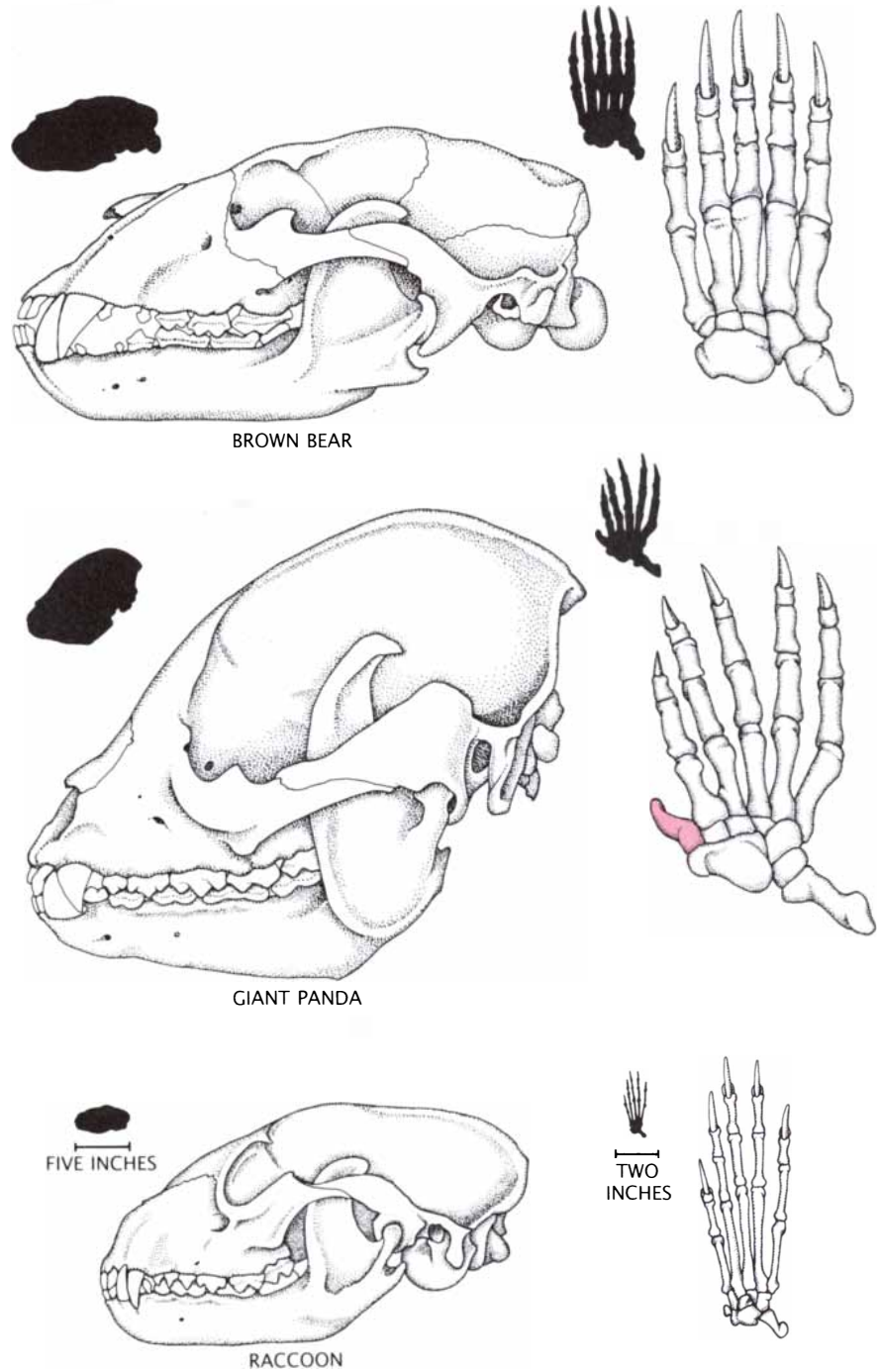
The emergence of the molecular-clock hypothesis laid the foundation for an entirely new field of biology known as molecular evolution. In spite of certain reservations (the rate at which mutations accumulate varies from molecule to molecule and may not be the same for all species), most evolutionary biologists acknowledge that genes do mutate in a steady, time-dependent manner. For this reason molecular data can be a valuable adjunct to more traditional methods of taxonomic study.

We initiated our research on the ancestry of the giant panda by collecting blood and skin samples from six species in the family Ursidae (the brown bear, the polar bear, the Asiatic black bear, the American black bear, the sun bear and the sloth bear) and from three species in the family Procyonidae (the red panda, the raccoon and the kinkajou).

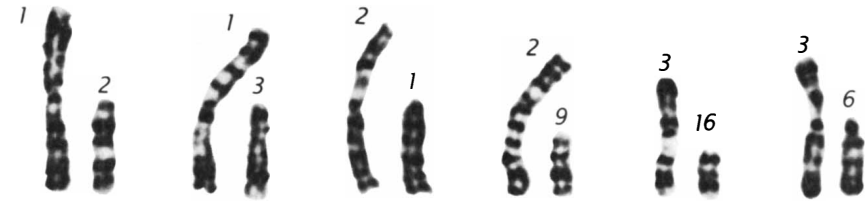
We began our analysis with a technique known as DNA hybridization. Fibroblast cells (chosen because they divide readily in culture) from a selected species are grown in a medium containing radioactive nucleotides. As the fibroblasts divide, they incorporate the radioactive nucleotides into their newly synthesized DNA. Single strands of this radioactive DNA from one species are then mixed in the laboratory with single strands of nonradioactive DNA from a second species. The two types are allowed to form double-strand hybrid DNA molecules; the extent of hybridization is assessed by measuring the radioactivity of the newly formed DNA. The sequence homology of DNA from the two species is indicated by the temperature at which the hybrid strands unwind when subjected to heat. This so-called melting temperature is inversely proportional to the differences in the composition of the two DNA strands that formed the hybrid DNA molecule.

A matrix of melting temperatures was generated by Raoul E. Benveniste, one of my colleagues at the National Cancer Institute, from DNA-hybridization studies on the giant panda and its relatives. We used these data and the principles of the molecular-clock hypothesis to construct a phylogenetic diagram (a classification that expresses evolutionary relationships between organisms). According to our data, the procyonids were the first group to split from an ancestor common to both the bear and the raccoon families; soon afterward the red panda diverged from the main lineage of the raccoon family. More important, our data indicate that the giant panda is more closely related to bears than to raccoons.

Although we were fairly confident of our results, we needed to carry our analysis several steps further. One reason was that we wanted to test our results for consistency; could we obtain the same phylogeny with other methods and molecules? We turned to gel electrophoresis, one of the techniques we had used to determine the paternity of Ling-Ling's cub. We wanted to measure genetic similarity by comparing mutational differences in genes that code for isozymes: dif-



ANATOMY of the giant panda differs considerably from that of its relatives, the bears and raccoons. The distinctive skull, for example, represents an adaptation to herbivory: the enlarged cranium provides extra grinding power for macerating bamboo. The opposable thumb, formed by an extension of the wristbone (*shown in color*), is an adaptation enabling the panda to strip leaves from stalks of bamboo. Many such specializations have confused taxonomists studying the evolutionary relationships of the panda.



DISTINCTIVE CHROMOSOMES of the giant panda, *Ailuropoda melanoleuca,* which were once considered a strong argument against including pandas and bears in the same family, have now been shown to be the result of simple chromosome fusion. Detailed analysis of banding patterns reveals that the panda's chromosomes are derived from an ancestral bear (now extinct), whose chromosomes apparently fused end to end; the centromere acts as a central connecting point. Representative chromosomes (*1–3*) from the giant panda (*in each pair at left*) are compared with their homologues from the brown bear *Ursus arctos* (*in each pair at right*). In the first frame they are paired with one homologue; in the next frame they are inverted 180 degrees and paired with their second homologue. By matching the banding patterns it can be determined, for example, that one half of chromosome 1 in the giant panda is homologous with chromosome 2 in the brown bear and that the other half is homologous with bear chromosome 3.

ferent versions of homologous enzymes. The relative mobilities of some 50 homologous enzymes isolated from nine different species were compared and the results were used to estimate the "genetic distance" (the average number of mutational differences per gene) between representative individuals from the various species.

We found (after compensating for the limited resolution of electrophoresis and the relative rates of amino acid substitution) that the genetic-distance values between the species do vary in proportion to the length of time those species have been reproductively isolated. We were able to construct a phylogenetic diagram based on our isozyme genetic-distance values that was structurally equivalent to the tree produced by DNA-hybridization data. The independent agreement of the DNA and isozyme trees increased our confidence in both the phylogenetic tree we had constructed and the fidelity of the molecular clock.

Next we turned our attention to a molecular technique called two-dimensional electrophoresis. Again, one measures protein migration on a gel, but the method makes it possible to compare large numbers of proteins (more than 300) at the same time. As before, one can estimate the genetic distance between various species based on the degree to which their protein profiles (determined by protein mobility on a gel) differ from one another. I asked David Goldman of the National Institute of Mental Health to examine fibroblast proteins from the giant panda and its relatives. His results were resounding in their corroboration of our previous findings.

As a final confirmation of the reliability of our phylogenetic diagram, we compared our results with those obtained by Vincent M. Sarich of the University of California at Berkeley, who in 1973 had carried out a series of immunological tests on the panda and its relatives. These tests estimate the evolutionary distance between species by comparing the extent to which a serum protein (such as albumin) from one species will react with antibodies raised against that protein from a different species. The fewer the antibody molecules that bind, the more distant the relationship between the species. Although Sarich's results were equivocal with respect to the red panda, they agreed nicely with what my colleagues and I had found out about the giant panda: that it is more closely related to bears than it is to raccoons.

Once we had established the consistency of our data, we needed to set the time on our molecular clock by calibrating our phylogenetic tree with data from the geologic record. This is necessary in order to determine the chronology of branching points on the tree. Although we now knew which species were most closely related to one another, we did not know how long ago in the earth's history they had diverged and split into separate lineages.

One way to calibrate an evolutionary tree based on molecular data is to correlate the rate of molecular evolution in that group (in this case nine species in the order Carnivora) with a second, unrelated group, such as the primates, whose protein molecules appear to be advancing at the same rate as the carnivores and whose fossil history is well documented. If we could show that two species of primates had the same molecular-distance values as two species of bears, we could assume that both groups diverged at about the same time. Put a different way, if we could show that the albumin molecule is as different in the black bear and the sun bear as it is in the gorilla and the chimpanzee, we would know that these groups diverged at approximately the same point in geologic time.

We repeated the same set of molecular analyses we had carried out on the panda on several species of primates. Because our data matched those obtained by other geneticists studying molecular evolution in primates, we could have confidence in the accuracy of our results. Based on an estimated date of African ape–human divergence at 35 million years ago, we were able to conclude that between 30 and 50 million years ago the ancestors of the modern ursids and procyonids split into two separate lineages.

Within 10 million years of that event the procyonid group diverged into Old World procyonids (those that inhabit Europe, Africa and Asia), represented today by the red panda, and New World procyonids (those that inhabit North and South America), represented today by the raccoon, the coati, the olingo and the kinkajou. At about the same time that the gibbons split from the great apes (from 18 to 25 million years ago), the ancestors of the giant panda diverged from the main ursid line. This event came nearly 20 million years after the initial divergence of the ursid and procyonid lines. The brown, black, sloth and sun bears diverged from one another approximately six to 10 million years ago.

Although our molecular data appeared finally to settle the question of the panda's ancestry, they failed to address one last piece of the puzzle: the bizarre chromosomes seen in the giant panda. Bears in the genus *Ursus* have 74 single-armed chromosomes with the centromere at one end. The giant panda, on the other hand, has only 42 chromosomes and most of them have two arms. To some biologists these differences were proof alone that the bear and the panda

could not possibly belong to the same family.

By now we were suspicious of this conclusion, and so I asked my colleague William G. Nash of the National Cancer Institute to reexamine the chromosomes of the giant panda. Nash had recently developed special techniques for examining the distinctive light-and-dark banding patterns on chromosomes by culturing the cells in the presence of a type of retrovirus that transforms normal, slowly dividing cells into rapidly dividing cells. He examined the chromosomes just prior to the metaphase stage of cell division, when they can be stained to display exquisitely detailed banding patterns.

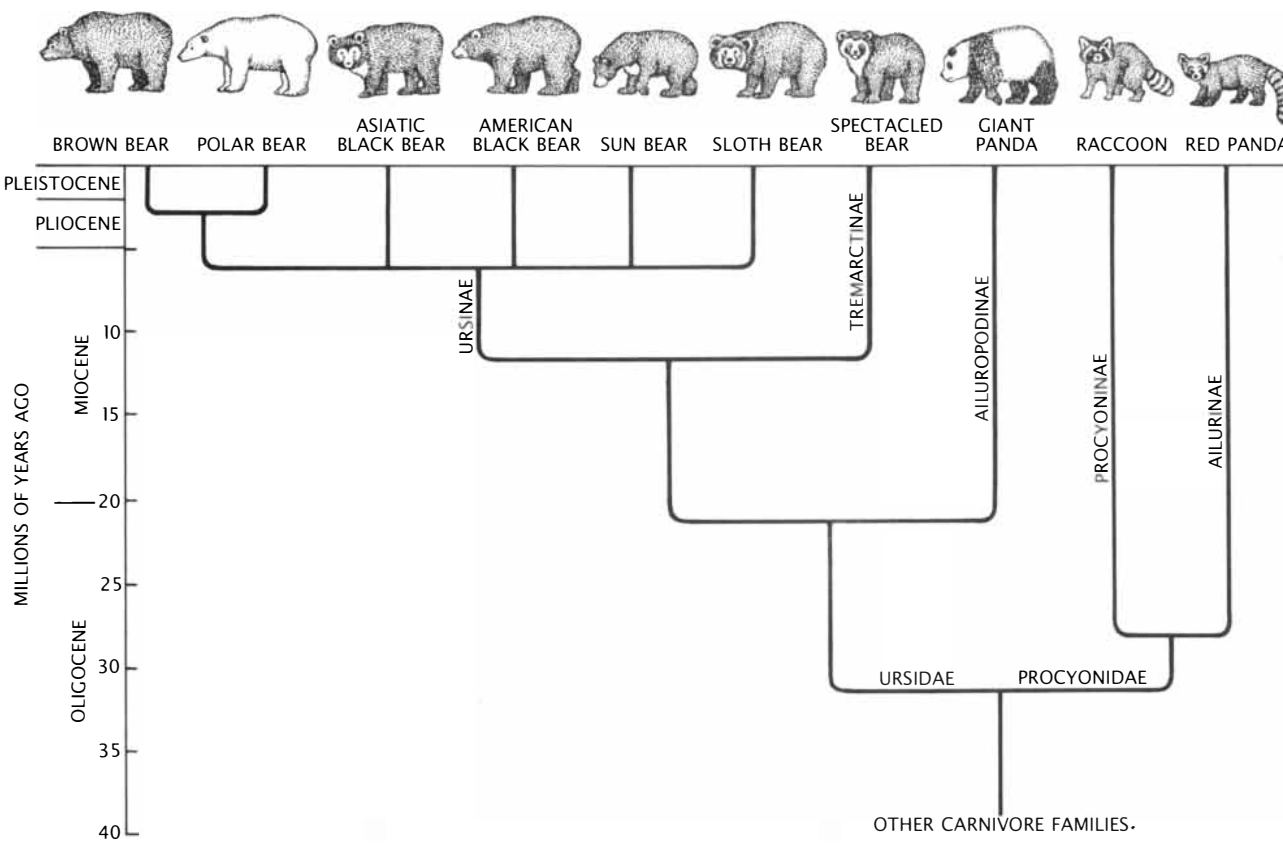
On comparing chromosomes from the giant panda with chromosomes from the six *Ursus* bears, Nash made a startling discovery: the banding patterns of the six bear species were virtually identical with those observed on just the arms of the giant panda's chromosomes. We concluded that giant panda chromosomes were nothing more than pairs of bear chromosomes attached head to head with common centromeres. What looked like a major reorganization of chromosomes was in fact a simple fusion of chromosomes in the line leading to the giant panda.

On the basis of our results we were able to make a final taxonomic recommendation, namely that the giant panda belongs with the bears in the family Ursidae and the red panda belongs with the raccoon in the Procyonidae. Since both the giant panda and the red panda diverged from ancestors early in their respective family's history, we believe separate subfamily status is justified for both.

In addition we believe our data cast new light on rates of evolution. The giant panda has been cited as a striking example of rapid (or "punctuated") evolution because its divergence from bears was thought to have been a relatively recent event, one that took place less than two million years ago. Our results, however, suggest that the panda actually diverged from its bear ancestors as long as 15 to 25 million years ago. If these older dates are correct, they are more consistent with gradualism: evolution that takes place as the result of small, gradual changes.

While carrying out this study, I became impressed by the power of molecular data to resolve evolutionary relationships. Nevertheless, it is important to mention that such data are limited in at least two ways. First, they cannot gauge elapsed time in the strict sense and therefore must be calibrated by the fossil record. Second, because most mutations are random, they say little about the forces of natural selection responsible for morphological change.

We nonetheless believe molecular analysis offers an objective approach to taxonomic problems and provides additional evidence that can be combined with traditional anatomical or behavioral data to achieve more accurate understanding of relationships between species. In the case of the giant panda, the molecular clock enabled us to finally settle the question of this animal's true ancestry and thus resolve a problem that had perplexed taxonomists for more than a century.



PHYLOGENETIC TREE based on data obtained from modern molecular methods places the giant panda in the Ursidae, or bear family. Within that family the panda is distinctive enough to warrant placement in its own subfamily, the Ailuropodinae. The red panda is retained in the Procyonidae, or raccoon family, where it too is granted separate subfamily status. The molecular clock suggests that the raccoon and bear families diverged from a common carnivorous ancestor some 35 to 40 million years ago.

Demons, Engines and the Second Law

Since 1871 physicists have been trying to resolve the conundrum of Maxwell's demon: a creature that seems to violate the second law of thermodynamics. An answer comes from the theory of computing

by Charles H. Bennett

One manifestation of the second law of thermodynamics is that such devices as refrigerators, which create inequalities of temperature, require energy in order to operate. Conversely, an existing inequality of temperature can be exploited to do useful work—for example by a steam engine, which exploits the temperature difference between its hot boiler and its cold condenser. Yet in 1871 the Scottish physicist James Clerk Maxwell suggested, in his *Theory of Heat*, that a creature small enough to see and handle individual molecules might be exempt from this law. It might be able to create and sustain differences in temperature without doing any work:

"...if we conceive a being whose faculties are so sharpened that he can follow every molecule in its course, such a being, whose attributes are still as essentially finite as our own, would be able to do what is at present impossible to us. For we have seen that the molecules in a vessel full of air at uniform temperature are moving with velocities by no means uniform.... Now let us suppose that such a vessel is divided into two portions, A and B, by a division in which there is a small hole, and that a being, who can see the individual molecules, opens and closes this hole, so as to allow only the swifter molecules to pass from A to B, and only the slower ones to pass from B to A. He will thus, without expenditure of work, raise the temperature of B and lower that of A, in contradiction to the second law of thermodynamics."

The "being" soon came to be called Maxwell's demon, because of its far-reaching subversive effects on the natural order of things. Chief among these effects would be to abolish the need for energy sources such as oil, uranium and sunlight. Machines of all kinds could be operated without batteries, fuel tanks or power cords. For example, the demon would enable one to run a steam engine continuously without fuel, by keeping the engine's boiler perpetually hot and its condenser perpetually cold.

To protect the second law, physicists have proposed various reasons the demon cannot function as Maxwell described. Surprisingly, nearly all these proposals have been flawed. Often flaws arose because workers had been misled by advances in other fields of physics; many of them thought (incorrectly, as it turns out) that various limitations imposed by quantum theory invalidated Maxwell's demon.

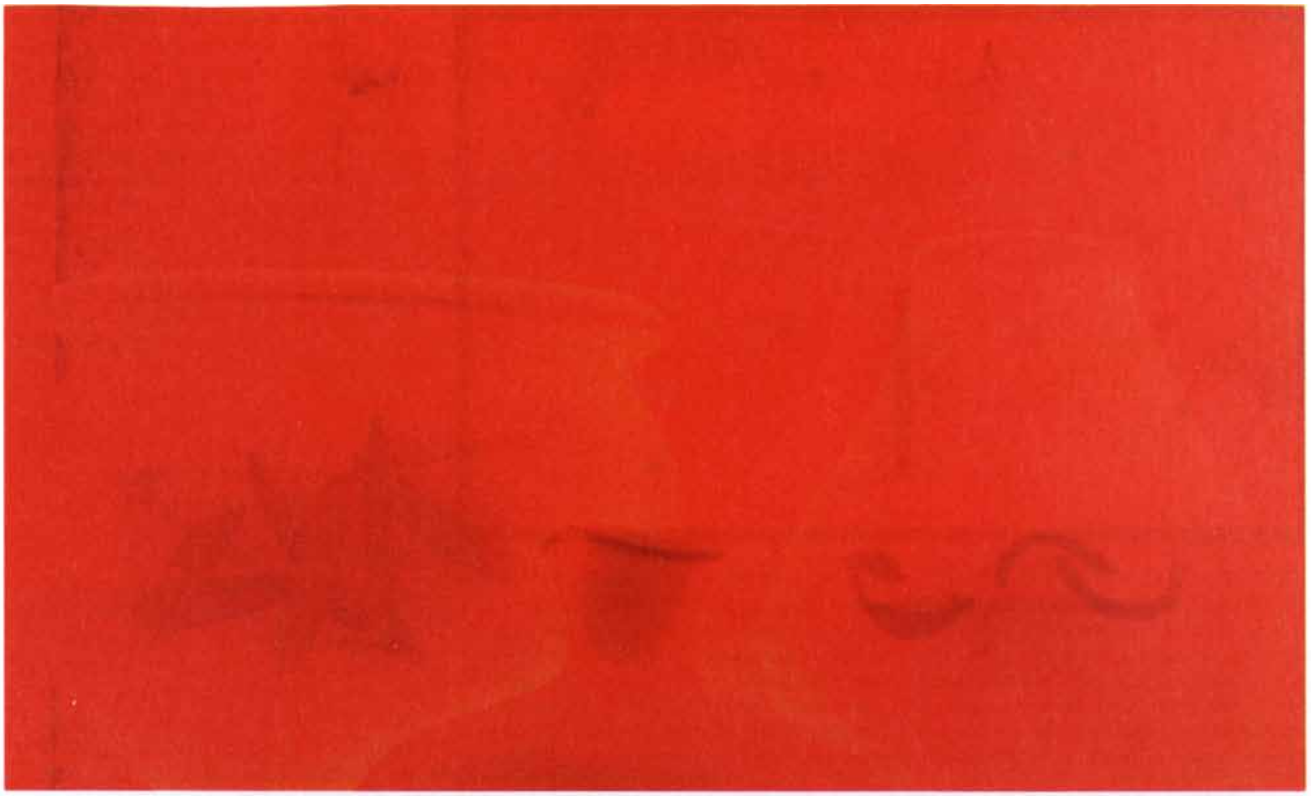
The correct answer—the real reason Maxwell's demon cannot violate the second law—has been uncovered only recently. It is the unexpected result of a very different line of research: research on the energy requirements of computers.

Since Maxwell's day numerous versions of the demon have been proposed. One of the simplest creates a pressure difference (rather than a temperature difference) by allowing all molecules, fast or slow, to pass from B to A but preventing them from passing from A to B. Eventually most of the molecules will be concentrated in A and a partial vacuum will be created in B. This demon is if anything more plausible than Maxwell's original demon, since it would not need to be able to see or think. It is not immediately evident why such a demon—a one-way valve for molecules—could not be realized as some simple inanimate device, for instance a miniature spring-loaded trapdoor.

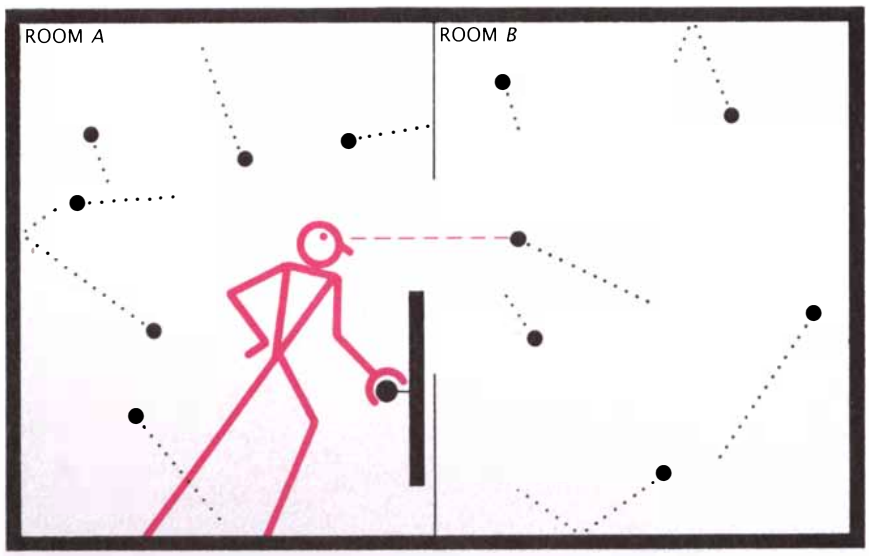
Like Maxwell's original demon, the "pressure demon" could be a source of limitless power for machines. For example, pneumatic drills of the kind used to cut holes in streets generally run on compressed air from a tank kept full by a gasoline-powered compressor. A one-way valve for air molecules could substitute for the compressor, effortlessly collecting air from the surroundings into the high-pressure tank.

One might think such an arrangement would violate the law of conservation of energy (otherwise known as the first law of thermodynamics), but it would not. The energy for cutting concrete could be taken from heat in the air collected by the one-way valve; the air's temperature would drop as it passed through the machinery. There is nothing in the first law to prevent an engine from supplying all its energy needs from the ambient heat of its environment, or even from the waste heat of its own friction and exhaust. It is the second law that prohibits such engines.

To analyze the demon's actions closely, then, one must understand some of the subtleties of the second law. The second law was originally expressed as a restriction on the possible transformations of heat and work, but it is now seen as being fundamentally a statement about the increase of disorder in the universe. According to the second law, the entropy, or disorder, of the universe as a whole cannot be made to decrease. This means that only two kinds of events are possible: events during which the entropy of the universe increases and events during which it remains constant. The former are



UNIFORM GLOW in a hot furnace (*top*) demonstrates one consequence of the second law of thermodynamics: it is impossible to distinguish objects in a vessel at uniform temperature without an external light source hotter than the vessel's ambient temperature. In a vessel at uniform temperature objects glow in such a way that exactly the same intensity and color of light come from the surface of every object (even objects that have different reflectances and colors). The reason is that if any object appeared darker than its surroundings, it would absorb energy at the expense of its neighbors. As a result it would become hotter and its neighbors would become cooler. According to the second law, however, objects that are initially at the same temperature cannot spontaneously come to have different temperatures. (In this photograph some contrast is visible because the temperature inside the furnace is not exactly uniform.) By an external light source, intrinsic differences in reflectance are visible (*bottom*).



MAXWELL'S DEMON, described in 1871 by James Clerk Maxwell, seems able to violate the second law of thermodynamics. The demon controls a sliding door that blocks a hole in a wall between rooms containing gas at equal temperatures and pressures. It observes molecules approaching the hole and opens and closes the door to allow fast-moving molecules to pass from room A to room B but not vice versa. Slow-moving molecules, conversely, are allowed to pass only from B to A. As the demon sorts, B heats up and A cools. According to the second law, a certain amount of work is required to create a temperature difference, but the work of sliding a door can be made negligibly small.

known as irreversible processes because to undo them would violate the second law; the latter are called reversible processes. One can decrease the entropy of a given system by doing work on it, but in doing the work one would increase the entropy of another system (or that of the first system's environment) by an equal or greater amount.

A classic irreversible process, and one that helps in defining the concept of entropy a little more precisely, is called free expansion. Suppose a chamber filled with gas is separated by a partition from a vacuum chamber of the same size. If a small hole is made in the partition, gas will escape (that is, it will expand freely) into the formerly empty chamber until both chambers are filled equally.

The reason the molecules spread out to fill both chambers is mathematical rather than physical, if such a distinction can be made. The numbers of molecules on the two sides of the partition tend to equalize not because the molecules repel one another and move as far apart as possible, but rather because their many collisions with the walls of the container and with one another tend to distribute them randomly throughout the available space, until about half of them are on one side of the partition and about half are on the other side.

Since the spreading of the molecules is due to chance rather than to repulsion, there is a chance that all the molecules might return simultaneously to the chamber from which they came. If there are n molecules, however, the probability of all of them returning to their original chamber is the same as the probability of tossing n coins and having them all come up "heads": $1/2^n$. Thus for any sizable number of molecules (and there are about 300,000,000,000,000,000,000,000 molecules in a gram of hydrogen) the free expansion is an effectively irreversible process: a process whose spontaneous undoing, although possible, is so unlikely that one can say with confidence it will never be observed.

The disordered state—the state in which the gas has spread into both chambers rather than residing compactly in a single chamber—is more probable than the ordered state. That is, there are more configurations of molecules in which the molecules occupy both chambers, just as, when 100 coins are tossed, there are more ways to achieve a total of 50 heads and 50 tails than there are to achieve 100 heads and no tails. In saying that the entropy of the universe tends to increase, the second law is simply noting that the universe tends to fall into more probable states as time passes.

Can this concept be quantified? In other words, can one say how much the disorder of the gas has increased after it has spread out to fill both chambers? Consider a single molecule in the gas. A molecule that can roam throughout both chambers has twice as many possible positions as a molecule confined to a single chamber: there are twice as many ways for a molecule to occupy the two-chamber apparatus. If there are two molecules in the two-chamber apparatus, each molecule has twice as many possible positions as it would have in a single chamber, and so the system as a whole has 2×2, or four, times as many possible configurations. If there are three molecules, the system has $2 \times 2 \times 2$, or eight, times as many possible configurations.

In general, if there are n molecules in the gas, the gas can fill two chambers in 2^n times more ways than it can fill a single chamber. The gas in the two-chamber apparatus is said to have 2^n times as many "accessible states" as the gas in a single chamber. In the same way, the number of accessible states in most systems depends exponentially on the number of molecules.

The entropy of a system is therefore defined as the logarithm of the number of accessible states. In the example of the two-chamber gas apparatus, a 2^n-fold increase in the number of accessible states is an increase in entropy of n bits, or binary units. (The base of the logarithm—and hence the size of a unit of entropy—is arbitrary; it is conventional to choose base 2 and binary units.) The logarithmic scale has the advantage of making the entropy of a sample of matter, like its energy or mass, roughly proportional to the number of molecules in the sample. One can draw an analogy to a computer memory: an n-bit memory, other things being equal, has size, weight and cost that are roughly proportional to n, although the number of distinct states possible in the memory is 2^n.

The earliest statements of the second law did not mention randomness or disorder; they concerned heat, work and temperature. How can these concepts be related to our quantitative definition of entropy?

The molecules in any sample of matter are always in motion. The speed and direction of each molecule are random, but the average speed of the molecules is proportional to the square root of the sample's temperature (as measured from absolute

zero). As the temperature of a sample is raised (and the average speed increases) the velocities of individual molecules come to be distributed over a greater range than they are when the average speed is low.

When the average speed is high, then, every molecule in the sample has a greater range of velocities available to it, just as a molecule in a two-chamber gas apparatus has a greater range of positions available to it than a molecule in a single-chamber apparatus has. There are thus more accessible states at high temperatures than there are at low temperatures. The motion is more disordered at high temperatures, because it is harder to predict what the velocity of any molecule will be.

Disorder of molecular motion and disorder of molecular positions must both be counted in determining the entropy of a system. The entropy of a gas can be increased either by allowing the gas to occupy a greater volume or by increasing its temperature so that its molecular motion becomes more disorderly.

Any flow of heat therefore carries entropy with it. To be precise, it turns out that a heat flow carries an amount of entropy proportional to the quantity of heat flowing divided by the temperature at which the flow takes place. Hence a flow from a hot body to a cold body raises the entropy of the cold body more than it lowers the entropy of the hot one: the same amount of heat leaves the hot body as enters the cold body, but in figuring the entropy decrease of the hot body one divides by a high temperature, whereas in figuring the entropy increase of the cold body one divides by a low temperature. A heat flow from a hot to a cold body thus raises the entropy of the universe.

Our more precise definition of entropy gives us a better understanding of why Maxwell's demon seems to violate the second law. By its sorting action the demon is causing heat to flow from room A to room B, even after room B has become warmer than room A. The demon is therefore lowering the entropy of room A by a greater amount than it is raising the entropy of room B. The demon therefore decreases the entropy of the universe as a whole—a thermodynamic impossibility.

In his description of the demon Maxwell made it clear he believed in the validity of the second law. He suggested that perhaps human beings are unable to violate the second law (by doing what the demon can do) simply because they lack the demon's ability to see and handle individual molecules. This is not a completely satisfying exorcism of the demon, because it leaves open the question of whether a being able to see and handle individual molecules, if such a being did exist, could violate the second law.

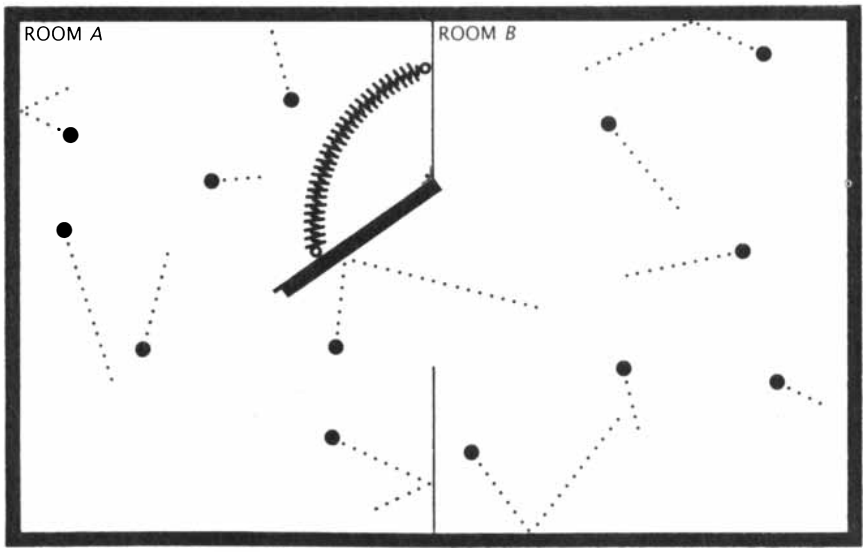
One way to uncover the reasons Maxwell's demon cannot work is to analyze and refute various simple, inanimate devices that might function as demons, such as the miniature spring-loaded trapdoor mentioned above, which acts as a one-way valve for molecules.

Imagine that the door opens to the left. If the demon works as it is supposed to, then every time a molecule from the room on the right strikes the door, the door swings open and the molecule passes into the room on the left. When a molecule from the left strikes the door, however, the door slams shut, trapping the molecule. Eventually all the molecules are trapped on the left, and the demon has compressed the gas (reducing its entropy) without doing any work.

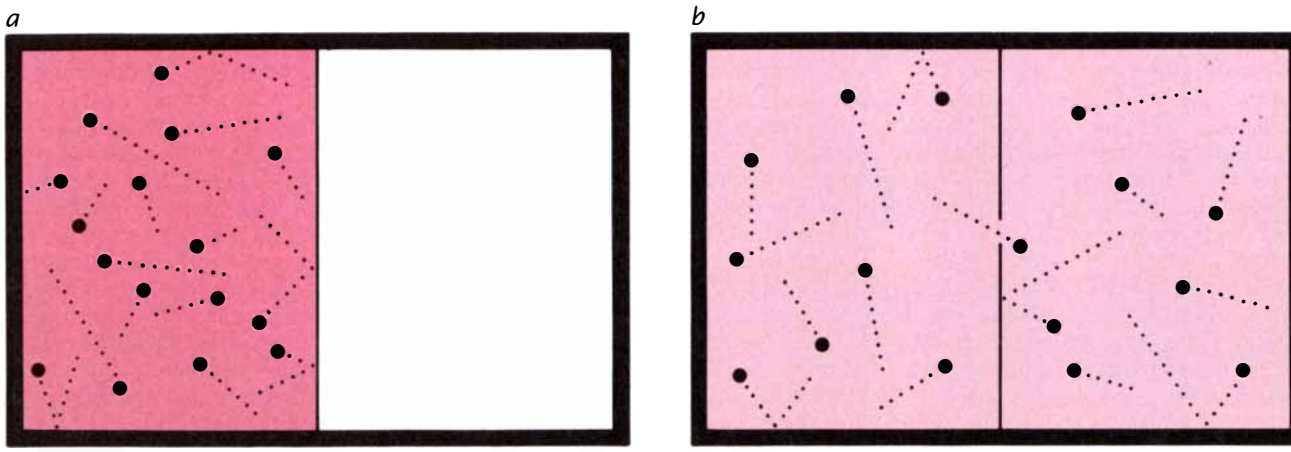
How is the trapdoor demon flawed? First of all, the spring holding the door shut must be rather weak. The work of opening the door against the spring's force must be comparable to the average kinetic energy of the gas molecules. In 1912 Marian Smoluchowski pointed out that because the door is repeatedly being struck by molecules, it will eventually acquire its own kinetic energy of random motion (that is, heat energy). The door's energy of random motion will be about the same as that of the molecules striking it, and so the door will jiggle on its hinges and swing open and shut (remember that the door is very small), alternately bouncing against its jamb and swinging open against the force of the spring.

When the door is open, it obviously cannot function as a one-way valve, since molecules can pass freely in both directions. One might still hope that the door would act as an inefficient demon, trapping at least a small excess of gas on the left, but it cannot do even that. Any tendency the door has to act as a one-way valve, opening to let a molecule go from the right to the left, is exactly counteracted by its tendency to do the reverse: to slam shut against a molecule that has wandered in front of it, actively pushing the molecule from the room on the left to the one on the right (aided by the force of the spring).

The two processes—a molecule pushing its way past the door from right to left, and the door pushing a molecule from left to right—are me-



TRAPDOOR "DEMON" is a form of Maxwell's demon designed to operate automatically and to create an inequality of pressure, not of temperature. A spring-loaded trapdoor blocks a hole between two rooms initially containing gas at equal temperatures and pressures. The door swings open in only one direction in order to admit molecules from room B into room A but not vice versa. Eventually, one might think, molecules will accumulate in A at the expense of B, creating an inequality of pressure. Actually the inequality does not build up. The trapdoor, heated by collisions with molecules, jiggles open and closed randomly because of thermal energy. When it is open, it is not a one-way valve, and as it closes it may push a molecule from A into B. The latter process takes place as often as its inverse, in which a molecule from B pushes past the door into A.



FREE EXPANSION of a gas is a thermodynamically irreversible process: one in which the entropy, or disorder, of the universe increases. A gas is initially confined in one chamber of a two-chamber apparatus (*a*). The barrier between the chambers is pierced, and molecules leak from one chamber into the other until approximately the same number of molecules are in both (*b*).

chanical reverses of each other: a motion picture of one, shown backward, would look like the other. In an environment at a constant temperature and pressure both processes would take place equally often, and the ability of the trapdoor to act as a one-way valve would be exactly zero. It cannot work as a demon.

In environments where the pressure is not equal on both sides of the door, of course, such devices do function. Large-scale versions, built with macroscopic doors and springs, can be seen on ventilator fans designed to blow stale air out of restaurants without admitting gusts of outside air when the fan is off. Microscopic versions would function in much the same way, allowing molecules to pass if there were excess pressure on one side but shutting off the flow if there were excess pressure on the other. The devices would not violate the second law, because they could only allow pressures to equalize; they could never form regions of excess pressure.

Even though a simple mechanical demon cannot work, perhaps an intelligent one can. Indeed, some time after Maxwell had described the demon, many investigators came to believe intelligence was the critical property that enabled the demon to operate. For example, in a 1914 paper Smoluchowski wrote: "As far as we know today, there is no automatic permanently effective perpetual-motion machine, in spite of the molecular fluctuations, but such a device might, perhaps, function regularly if it were appropriately operated by intelligent beings."

The physicist Leo Szilard attempted a quantitative analysis of this question in a paper published in 1929, "On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings." Although the title seems to imply an intelligent demon could violate the second law, the body of the article is devoted to refuting this notion and to arguing that no being, intelligent or not, can do so. Szilard thought the observation, or measurement, the demon must make (for example, to see which side a molecule is coming from) cannot be done without also doing enough work to cause an increase in entropy sufficient to prevent a violation of the second law.

Szilard considered a demon that differed in several ways from Maxwell's; his demon has since come to be called Szilard's engine. (The engine I shall describe here differs slightly from Szilard's original one.) The engine's main component is a cylinder in which there is a single molecule in random thermal motion. Each end of the cylinder is blocked by a piston, and a thin, movable partition can be inserted in the middle of the cylinder to trap the molecule in one half of the cylinder or the other [*see illustration on opposite page*]. The engine is also equipped with devices for finding which half of the apparatus the molecule is in and a memory for recording that information.

The engine's cycle consists of six steps. In the first step the partition is inserted, trapping the molecule on one side or the other. Szilard argued that the work necessary to insert the partition can in principle be made negligibly small.

In the next step the engine determines which half of the apparatus the molecule has been trapped in. The engine's memory device has three possible states: a blank state to signify that no measurement has been made, an *L* to signify that the molecule has been observed in the left half of the apparatus, and an *R* to signify that the molecule has been observed in the right half. When the measurement is made, the memory switches from the blank state to one of the other two.

The third step, which might be called a compression stroke, depends on the knowledge gained during the preceding step. The piston on the side that does not contain the molecule is pushed in until it touches the partition. Unlike the compression stroke of an internal-combustion engine, this compression stroke requires no work, because the piston is "compressing" empty space; the molecule, which is trapped on the other side of the partition, cannot resist the piston's movement.

Then, in the fourth step, the partition is removed, allowing the molecule to collide with the piston that has just been advanced. The molecule's collisions exert a pressure on the face of the piston.

In the fifth step, which might be called the power stroke, the pressure of the molecule drives the piston backward to its original position, doing work on it. The energy the molecule gives to the piston is replaced by heat conducted through the cylinder walls from the environment, and so the molecule continues moving at the same average speed. The effect of the power stroke is therefore to convert heat from the surroundings into mechanical work done on the piston.

In the sixth step the engine erases its memory, returning it to the blank state. The engine now has exactly the same configuration it had at the

beginning of the cycle, and the cycle can be repeated.

Overall, the six steps appear to have converted heat from the surroundings into work while returning the gas and the engine to the same state they were in at the start. If no other change has occurred during the cycle of operation, the entropy of the universe as a whole has been lowered. In principle the cycle can be repeated as often as the experimenter wants, leading to an arbitrarily large violation of the second law.

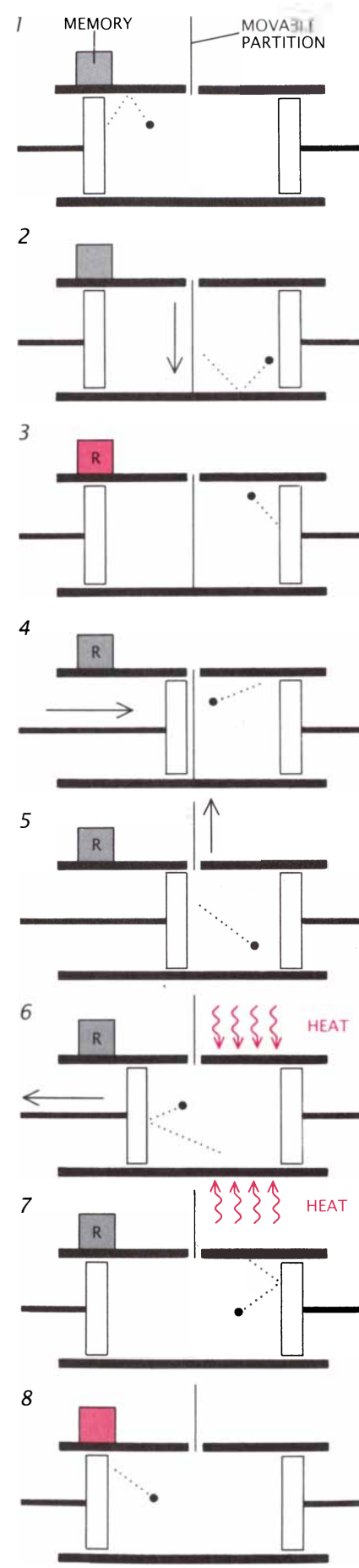
Szilard's way out of this predicament was to postulate that the act of measurement, in which the molecule's position is determined, brings about an increase in entropy sufficient to compensate for the decrease in entropy brought about during the power stroke. Szilard was somewhat vague about the nature and location of the increase in entropy, but in the years after he published his paper a number of physicists, notably Leon Brillouin (the author, in 1956, of the widely read book *Science and Information Theory*) and Denis Gabor (best known as the inventor of holography), tried to substantiate the postulated irreversibility of measurement. In particular they tried to determine what the cost should be, in terms of energy and entropy, of observing a molecule by aiming light at it and observing the reflections.

In their work, Brillouin and Gabor drew on a theory that had been developed since Maxwell's time: the quantum theory of radiation. According to the classical wave theory of light (to which Maxwell made fundamental contributions), the energy of a light ray can be made arbitrarily small. According to the quantum theory, however, light consists of energy packets called photons. The energy of a photon depends on its wavelength, or color, and it is impossible to detect less than one photon of light. Brillouin argued that for a molecule to be observed it must scatter at least one photon of a probe beam, and that when the photon's energy is dissipated into heat, the dissipation must produce an entropy increase at least as great as the entropy decrease Szilard's engine could achieve as the result of information gained about the scattering molecule.

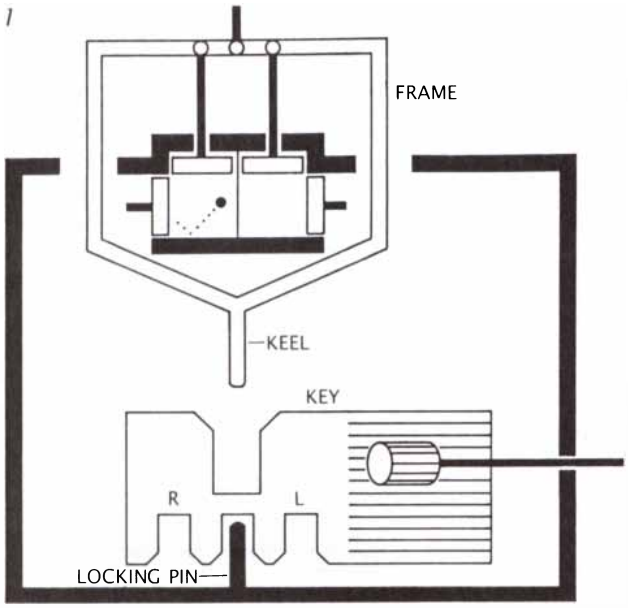
Why not simply use a probe beam of photons that have very low energies? The scheme will not work, because of another, more complicated, consequence of the quantum theory. According to the quantum theory of radiation, any vessel whose walls and interior are all at a single constant temperature becomes filled with a "gas" of photons: a bath of radiation. The wavelengths of the photons depend on the temperature of the vessel. Such a photon gas constitutes the uniform red or orange glow inside a hot furnace. (At room temperature the photons are mostly in the infrared part of the spectrum and are therefore invisible.)

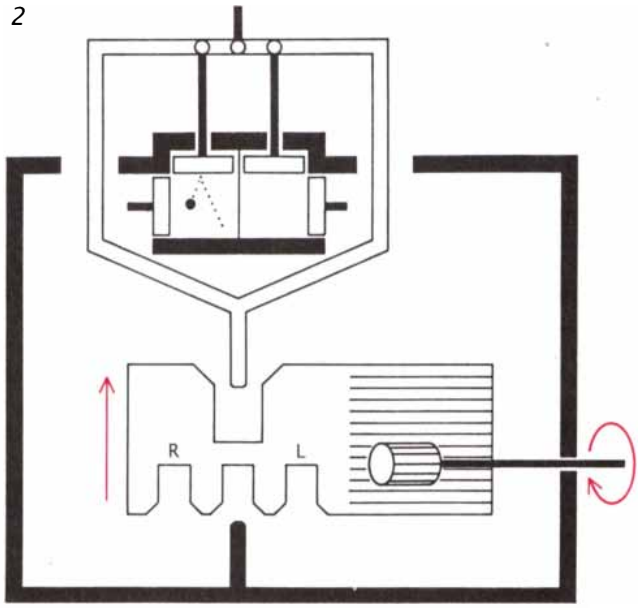
The photon gas might seem at first to be a handy source of light by which the demon could observe gas molecules (thereby saving itself the entropy cost of a flashlight). One of the surprising consequences of the second law, however (a consequence discovered by Gustav Robert Kirchhoff in 1859), is that it is impossible to see anything in a vessel at uniform temperature by the light of the vessel's own glow. If one looks into a kiln in which pots are being fired, for example, one will see a uniform orange glow almost devoid of contrast, even though the pots in the kiln may have very different colors, brightnesses and surface textures.

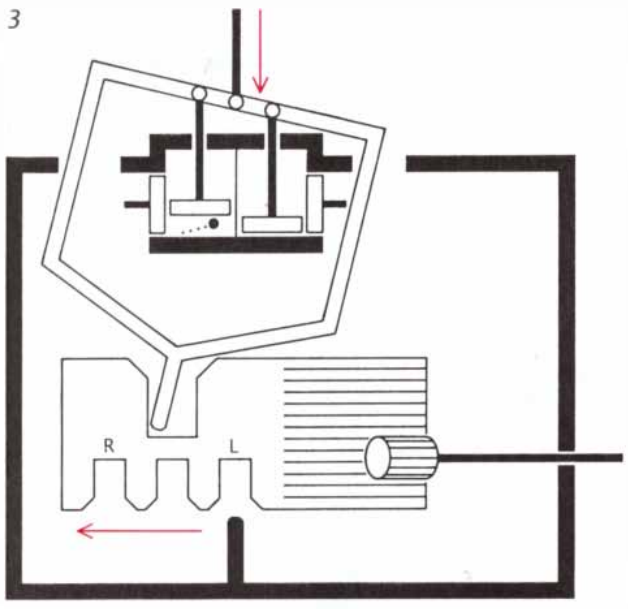
The objects in the hot kiln look as if they are all the same color and brightness, but they are not, as one can verify by shining a bright light on them from outside the kiln. The reason the objects nearly disappear by the light of the kiln must therefore

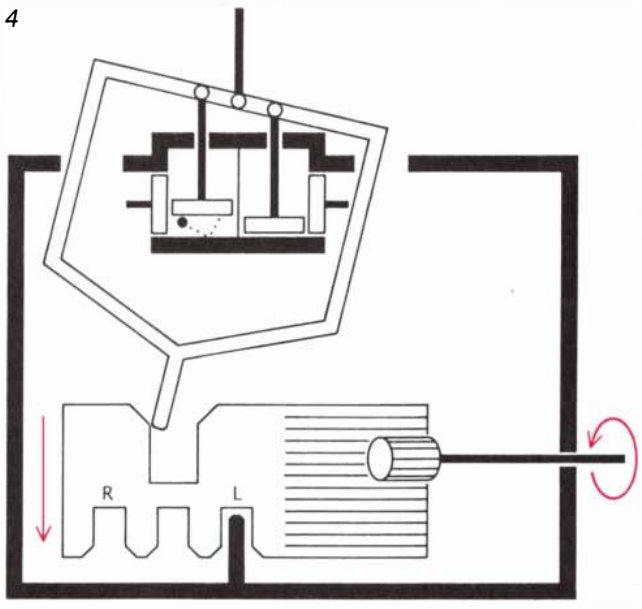


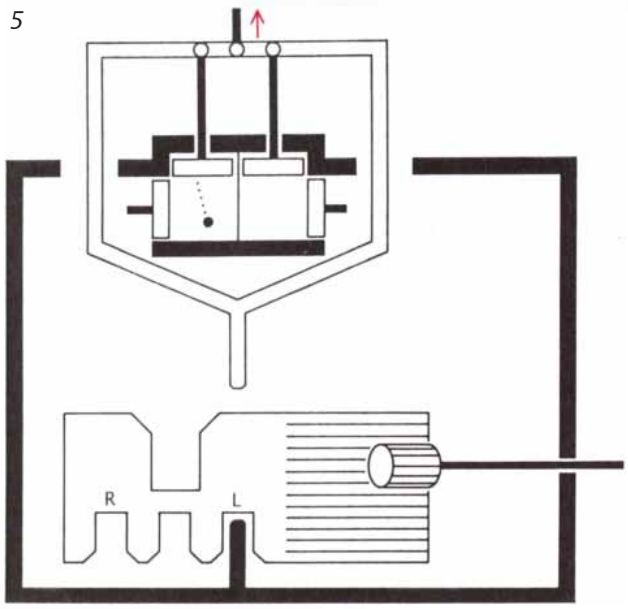
SZILARD ENGINE, modeled after a machine described in 1929 by Leo Szilard, seems to convert heat from its surroundings into work, contrary to the second law. The engine (*1*) is a cylinder that is blocked off at both ends by pistons; it is equipped with a movable partition and devices for observing the cylinder's contents and recording the results of observations. The cylinder contains a single molecule. At the start of the engine's cycle (*2*) the partition is lowered, trapping the molecule in one half of the cylinder. The observational devices determine and record which half contains the molecule (*3*), and the piston from the other half is pushed in until it touches the partition (*4*). Moving the piston requires no work, since it compresses empty space. Then the partition is withdrawn (*5*) and the molecule strikes the piston, pushing it backward (*6*). (The one-molecule gas "expands" against the piston.) Energy lost by the molecule as it works against the piston is replaced by heat from the environment. When the piston has returned to its original position (*7*), the memory is erased (*8*) and the cycle can begin again.











MEASUREMENT APPARATUS, designed by the author to fit the Szilard engine, determines which half of the cylinder the molecule is trapped in without doing appreciable work. A slightly modified Szilard engine sits near the top of the apparatus (*1*) within a boat-shaped frame; a second pair of pistons has replaced part of the cylinder wall. Below the frame is a key, whose position on a locking pin indicates the state of the machine's memory. At the start of the measurement the memory is in a neutral state, and the partition has been lowered so that the molecule is trapped in one side of the apparatus. To begin the measurement (*2*) the key is moved up so that it disengages from the locking pin and engages a "keel" at the bottom of the frame. Then the frame is pressed down (*3*). The piston in the half of the cylinder containing no molecule is able to descend completely, but the piston in the other half cannot, because of the pressure of the molecule. As a result the frame tilts and the keel pushes the key to one side. The key, in its new position, is moved down to engage the locking pin (*4*), and the frame is allowed to move back up (*5*), undoing any work that was done in compressing the molecule when the frame was pressed down. The key's position indicates which half of the cylinder the molecule is in, but the work required for the operation can be made negligible. To reverse the operation one would do the steps in reverse order.

be that dark (that is, nonreflective) objects glow proportionately more brightly than light (reflective) objects, so that the total light intensity leaving any object (reflected and emitted light combined) is the same.

To see why this strange leveling of intensity must take place, suppose it did not occur and think about the consequences for the second law. Suppose two objects, say a vase and a pot, are placed close together in a kiln at uniform temperature. If the intensity of light leaving the vase toward the pot were greater than that leaving the pot toward the vase, energy would flow from the vase to the pot. The pot would become warmer and the vase would become cooler.

Thus, without the expenditure of work, two regions that were once at a uniform temperature would come to different temperatures, just as if a Maxwell's demon had been sitting between them, and the second law would be violated. Therefore if the second law is to be valid, objects in a vessel at uniform temperature cannot have different surface intensities.

In order to see the objects in a furnace, then, one must shine light in from an external source, such as a flashlight that has a filament hotter than the furnace's temperature. In daily life such light sources—the sun, for example—make it possible for us to see objects in vessels that are uniformly at room temperature.

Brillouin, Gabor and others, armed with an understanding of the photon gas, argued that Maxwell's demon cannot observe the molecules it sorts without some kind of light source. Therefore, they said, the demon cannot violate the second law. Every time it observes a molecule the demon must dissipate the energy of at least one photon; the energy of that photon must be greater than a minimum energy determined by the temperature of the gas in which the demon sits. Such arguments, although they are not completely rigorous, seemed to substantiate Szilard's belief that acquiring a given amount of information entails producing a corresponding amount of entropy.

The next major progress toward banishing the demon was a side effect of research by Rolf Landauer of IBM on the thermodynamics of data processing. Certain data-processing operations, such as the copying of data from one device into another, are analogous to measurements, in that one device acquires information about the state of the other. Hence it

was generally believed in the 1950's that data-processing operations were intrinsically irreversible (in the thermodynamic sense of the word), just as Szilard had argued that measurement in general is irreversible. It was thought that any kind of data operation required the generation and removal of at least one bit's worth of heat for every bit of data to be processed. (This is an extremely small quantity of heat: roughly one ten-billionth of the heat actually generated by existing electronic circuits.)

In about 1960 Landauer analyzed the question more thoroughly. He found that some data operations are indeed thermodynamically costly but others, including, under certain conditions, copying data from one device to another, are free of any fundamental thermodynamic limit [see "The Fundamental Physical Limits of Computation," by Charles H. Bennett and Rolf Landauer; SCIENTIFIC AMERICAN, July, 1985].

Landauer's proof begins with the premise that distinct logical states of a computer must be represented by distinct physical states of the computer's hardware. For example, every possible state of the computer's memory must be represented by a distinct physical configuration (that is, a distinct set of currents, voltages, fields and so forth).

Suppose a memory register of n bits is cleared; in other words, suppose the value in each location is set at zero, regardless of the previous value. Before the operation the register as a whole could have been in any of 2^n states. After the operation the register can be in only one state. The operation has therefore compressed many logical states into one, much as a piston might compress a gas.

By Landauer's premise, in order to compress a computer's logical state one must also compress its physical state: one must lower the entropy of its hardware. According to the second law, this decrease in the entropy of the computer's hardware cannot be accomplished without a compensating increase in the entropy of the computer's environment. Hence one cannot clear a memory register without generating heat and adding to the entropy of the environment. Clearing a memory is a thermodynamically irreversible operation.

Landauer identified several other operations that are thermodynamically irreversible. What all these operations have in common is that they discard information about the computer's past state. In Landauer's phrase, such operations are "logically irreversible."

The connection of these ideas to the problem of the measurement, implicit in Landauer's work and in the reversible models of computation developed during the 1970's by Edward Fredkin of M.I.T., myself and others, became explicit in 1982, when I proposed that they provide the correct explanation of Maxwell's demon. Consider the operating cycle of Szilard's engine. The last step, in which the engine's memory is reset to a blank state, is logically irreversible, because it compresses two states of the machine's memory ("The molecule is on the left" and "The molecule is on the right") into one ("The molecule's position has not yet been measured"). Thus the engine cannot reset its memory without adding at least one bit of entropy to the environment. This converts all the work that had been gained in the power stroke back into heat.

What about the measurement step? Is it thermodynamically costly as well? In that case the engine would add to the entropy of the universe twice: once in measuring the molecule's position and again in resetting its memory after the power stroke. Actually the measurement does not have to be thermodynamically costly. There are ways to observe molecules other than by bouncing light off them. To prove this point I have designed a reversible measuring device, which measures and records the position of the molecule without undergoing any thermodynamically irreversible steps.

We have, then, found the reason the demon cannot violate the second law: in order to observe a molecule, it must first forget the results of previous observations. Forgetting results, or discarding information, is thermodynamically costly.

If the demon had a very large memory, of course, it could simply remember the results of all its measurements. There would then be no logically irreversible step and the engine would convert one bit's worth of heat into work in each cycle. The trouble is that the cycle would not then be a true cycle: every time around, the engine's memory, initially blank, would acquire another random bit. The correct thermodynamic interpretation of this situation would be to say the engine increases the entropy of its memory in order to decrease the entropy of its environment.

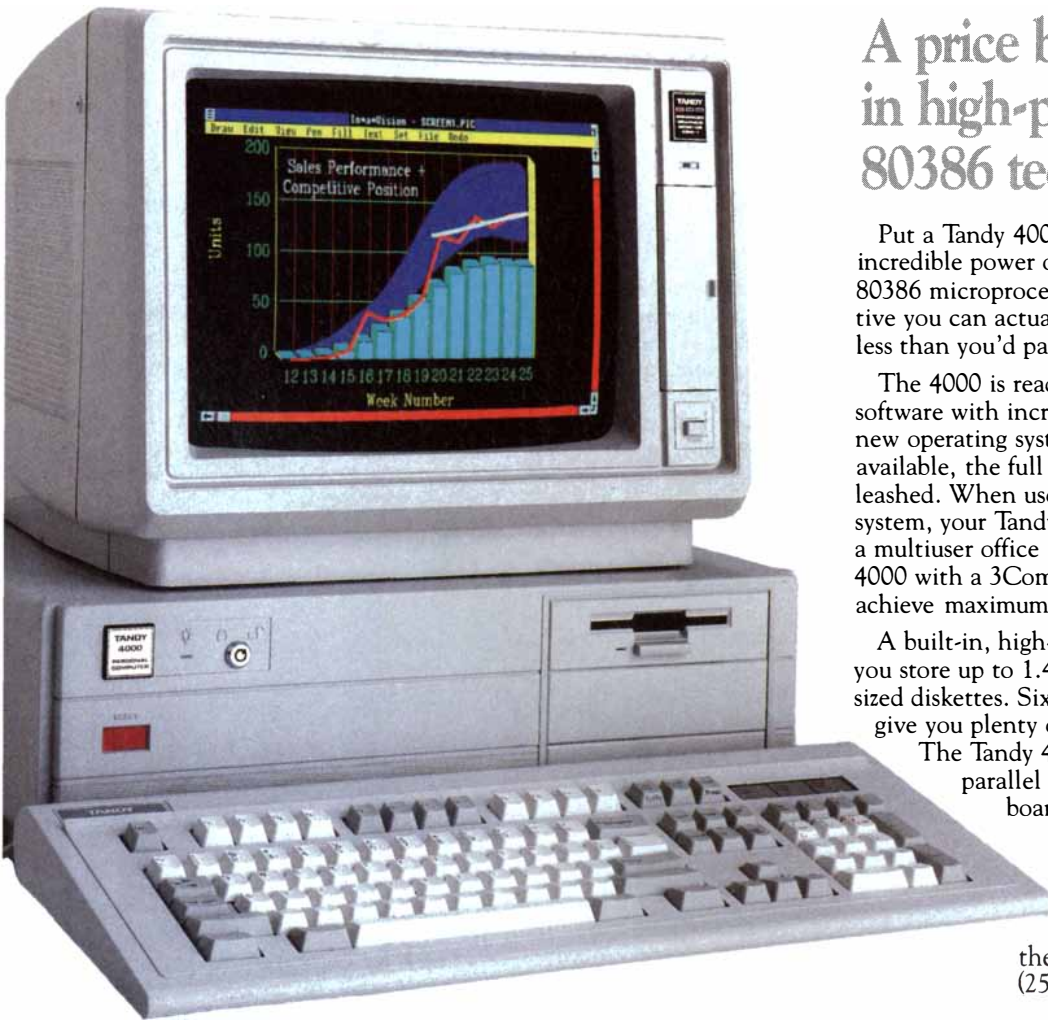
Attributing the gain in entropy to the resetting step rather than to the measurement step may seem to be a mere bookkeeping formality, since any complete cycle of Szilard's engine must include both steps, but considerable confusion can be avoided if one draws a clear distinction between the acquisition of new information and the destruction of old information. The confusion may or may not have existed in Szilard's mind. In most of his paper he refers to measurement as the irreversible step, but at one point he makes an accounting of entropy changes during the cycle and finds, without explicitly commenting on it, that the increase in entropy takes place during the resetting of the memory.

If subsequent workers had pursued this aspect of Szilard's paper, they would have come to our present understanding of Maxwell's demon. Their failure to do so is an irony in the history of science: the advancement of one branch of physics (the quantum theory of radiation) apparently delayed progress in another branch (thermodynamics). One aspect of quantum mechanics that reinforced the idea that a fundamental thermodynamic price must be paid for acquiring information is the uncertainty principle, which holds that certain sets of measurements cannot be carried out with more than a certain degree of precision. Although the uncertainty principle sounds similar to Szilard's hypothesis that measurements have an irreducible entropy cost, in fact it is fundamentally different. Szilard's hypothesis concerns the thermodynamic cost of measurements, whereas the uncertainty principle concerns the possibility of their being made at all, whatever their thermodynamic cost.

Another source of confusion is that we do not generally think of information as a liability. We pay to have newspapers delivered, not taken away. Intuitively, the demon's record of past actions seems to be a valuable (or at worst a useless) commodity. But for the demon "yesterday's newspaper" (the result of a previous measurement) takes up valuable space, and the cost of clearing that space neutralizes the benefit the demon derived from the newspaper when it was fresh. Perhaps the increasing awareness of environmental pollution and the information explosion brought on by computers have made the idea that information can have a negative value seem more natural now than it would have seemed earlier in this century.

Tandy Computers: Because there is no better value.™

The New Tandy® 4000



A price breakthrough in high-performance 80386 technology.

Put a Tandy 4000 on your desk and unleash the incredible power of the 32-bit, 16-megaherz 80386 microprocessor. The 4000 is so cost effective you can actually configure a 386 system for less than you'd pay for a competitor's 286 model.

The 4000 is ready to run current PC and AT® software with incredible new speed. And when new operating systems such as OS/2™ become available, the full potential of 80386 can be unleashed. When used with the XENIX® operating system, your Tandy 4000 can become the heart of a multiuser office system. Or configure the Tandy 4000 with a 3Com® workgroup file server to achieve maximum productivity.

A built-in, high-capacity 3½″ disk drive lets you store up to 1.4 million characters on pocket-sized diskettes. Six AT slots and two XT™ slots give you plenty of room for system expansion. The Tandy 4000 also comes with serial and parallel ports, a 101-key enhanced keyboard, and a keylock with chassis-lock mechanism for system protection.

Come into a Radio Shack Computer Center and see the Tandy 4000—only $2599. (25-5000)

Send me a new 1988 computer catalog.

Mail To: Radio Shack
Dept. 88-A-909
300 One Tandy Center
Fort Worth, TX 76102

Name _____
Company _____
Address _____
City _____ State _____
ZIP _____
Phone _____

Radio Shack®
COMPUTER CENTERS
A DIVISION OF TANDY CORPORATION

Price applies at Radio Shack Computer Centers and participating stores and dealers. Monitor, display adapter and operating systems sold separately. AT/Reg. TM and XT/TM IBM. XENIX/Reg. TM and OS/2/ TM Microsoft Corp. 3Com/Reg. TM 3Com Corp.

Early Farming in Northwestern Europe

Agriculture spread north and west from eastern Europe 8,000 years ago. Recent work based on distribution maps of Stone Age sites is showing how the "invasion" took place and what happened afterward

by John M. Howell

How did agriculture come to northern Europe? The general outlines of that event have been known for some time. Not long after 6000 B.C. members of the Linear Pottery Culture began moving north and west from eastern Europe, bringing with them a way of life based on cultivation. Yet the detailed picture of this "invasion" has changed dramatically in the last few years. As recently as the 1970's it was generally held that the Linear Pottery Culture spread in a more or less uniform pattern, diffusing up the major rivers in linear belts. Now, however, it is known that the pattern was neither linear nor uniform. On the contrary, the Linear Pottery sites were grouped in "settlement cells": dense clusters of villages that formed the basic units of settlement. Some of the settlement cells lasted until not long before the time of Christ.

One of the main tools in this reevaluation has been the distribution map. Once perceived as little more than a means of locating sites, distribution maps have come to be seen as rich in potential. Such maps not only were instrumental in demonstrating the existence of settlement cells but also have been used to show striking changes later in the Neolithic (the period associated with the introduction of agriculture). In the early Neolithic—the Linear Pottery phase—sites tended to be found in the bottom of river valleys. Thereafter the sites moved up onto the edge of the plateau above the valley, then out on the plateau itself. These changes were accompanied by significant technical and social developments, which are coming into focus for the first time through use of the once humble distribution map.

The Linear Pottery Culture takes its name from the incised and impressed decoration that appears on its ceramic artifacts. Although there is some regional variation in that pottery, the most striking feature of Linear Pottery sites is their cultural uniformity, a uniformity that applies both to the artifacts and to the sites themselves. Linear Pottery sites consist of hamlets or small villages, occasionally surrounded by a stockade. Each village is made up of wood-framed longhouses measuring, say, five meters wide and 30 meters long. Many had three subdivisions, which may have been for people, grain and livestock respectively. That the beginnings of social differentiation may have been present is indicated by the existence at some sites of one house that is even longer than the others (although built in the same style).

The uniformity of the Linear Pottery villages is one sign that agriculture was introduced into Europe by immigrants and not merely by the spread of new ideas. The uniform material culture associated with the villages includes an array of unsophisticated flintwork, small axes and nonmonumental burials. The burials are rarely in cemeteries; often they are found in ditches flanking the longhouses from which clay for the walls was taken. The remains of animals and plants indicate an agricultural way of life; sheep and cattle were the main domesticated animals. In areas where organic remains are well preserved, a mixture of cereals, vegetables and pulses (plants such as the lentil) is frequently found.

From about 5500 B.C. the Linear Pottery groups had begun moving north and west from their origins in the Carpathian Mountains, and by 4500 B.C. they had reached the Paris basin. The original, unrigorous examination suggested that the sites were to be found in a linear arc around the Danube, and the name Danube Culture was initially given to them. It was observed that in choosing their sites the early farmers had

LATE NEOLITHIC TOMB near Saumur in the Loire valley of France was a communal burying place that may have held 200 or more burials. The tomb is of the type

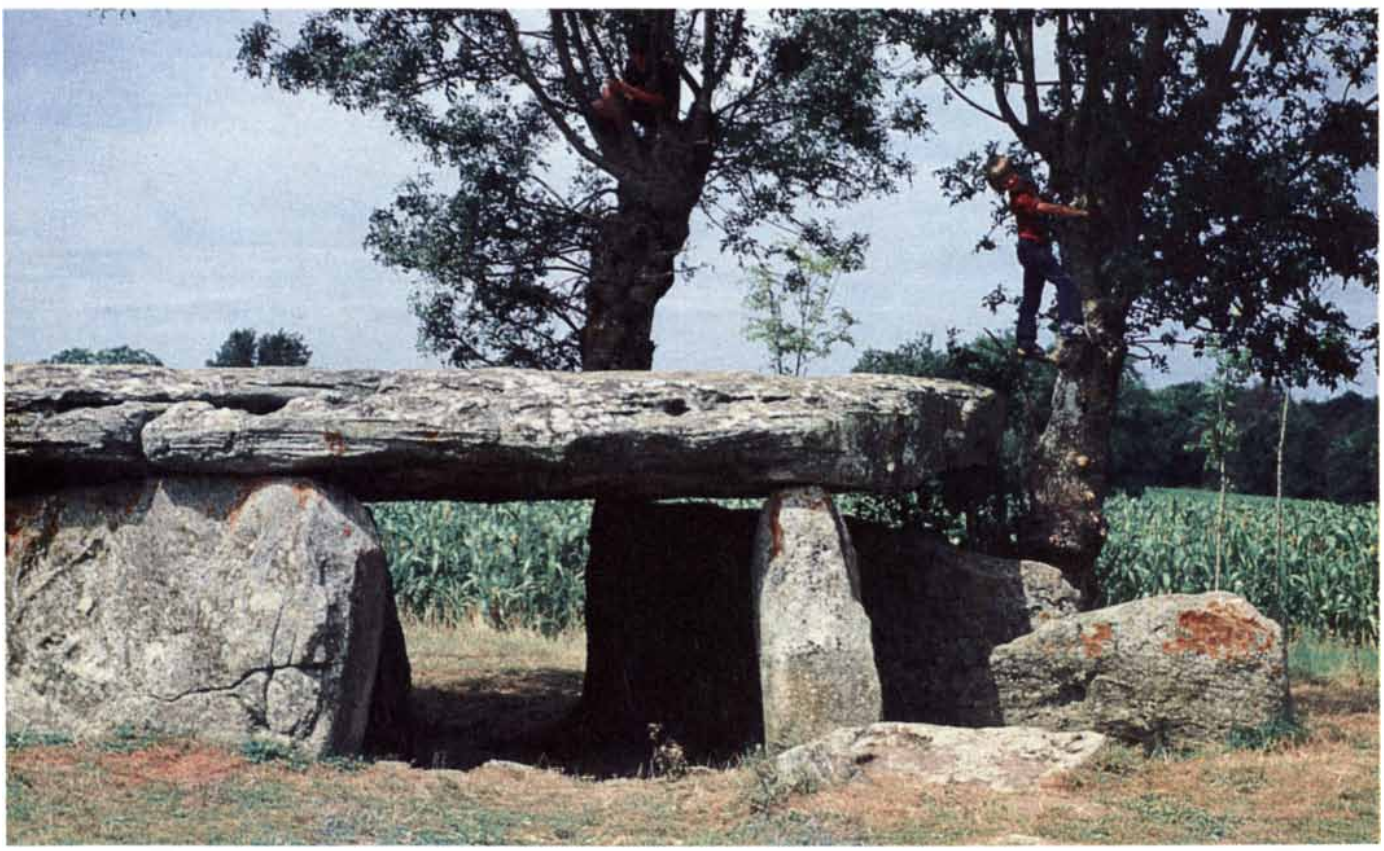
shown a preference for the light, glacial soils known as loess. Following early models of agricultural development, the villagers were thought to have practiced swiddening, a form of agriculture in which the fields are exploited for a season or until they are exhausted or overgrown with weeds; then they are abandoned. (Swiddening is frequently carried out by slash-and-burn methods.)

Recent work has shown that much of that picture is wrong. It began to be changed in the early 1970's, largely through the work of Janusz Kruk of the Institute of Sciences in Cracow, who studied early Neolithic settlement patterns near that city. Kruk's model was based on the integration of economic, social and ecological information with distribution maps, and it has become the cornerstone of subsequent interpretations. Much of the subsequent interpretation (including my own) stemmed from the University of Oxford and consisted of extending the geographic approach to poorly studied areas elsewhere in Europe. In my case the area of interest was in France, and I focused on the region drained by three major rivers: the Seine, the Oise and the Marne.

It should not be thought that all the new work came from Oxford. Indeed, the central concept of the settlement cell was developed not by English investigators but by German ones, including Bernhard Sielmann, then at the Frankfurt Museum. One reason the idea was formulated in Germany is that the settlements near the Rhine are densely clustered in the valley bottoms in a way that could give rise to the notion of a cell, or unit, rather than a linear pattern. Following Sielmann, archaeologists recognized similar clusters in the Dutch Limburg (along the Maas and its tributaries) and on the Aldenhoven plateau near Cologne.

Between the Limburg and the Loire, where my work was done, the picture is not as clear. In that region the glacial loess favored by the Linear Pottery farmers is not found in large expanses in the broad valley bottom. Instead the walls of deep valleys drop steeply as much as 30 meters from the promontories above. The valleys are too narrow for large clusters of villages, and so settlement cells are not immediately evident. Nevertheless, they are there. A careful examination I carried out on the basis of excavations, field surveys, museum collections and aerial photographs revealed areas of dense settlement contrasting with the absence of sites in other areas. The small settlement cells may take up only one or two segments of a narrow river valley, but they are analogues of the more obvious units found in Germany and Holland.

Whether the sites are clustered in tight settlement cells or are in the more apparently linear French arrangement, they are always found in the valley bottom. In the valley the river's meanders carve out "islands" of rich alluvial soil, and on such islands the farmers built their villages. The islands are above the level of regular flooding but are low enough so that they would have been been flooded periodically. These alluvial parcels range in area from 125 to more than 500 hectares (a hectare



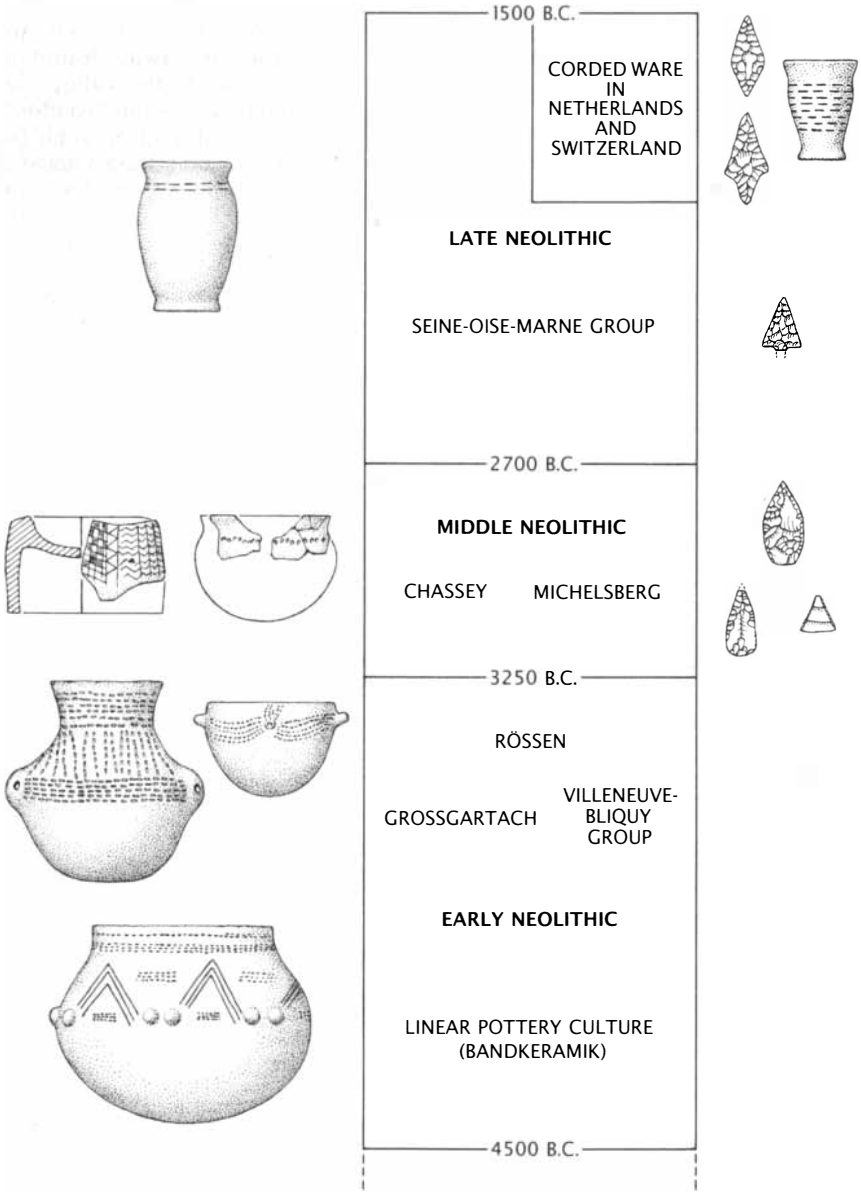
called *allée couverte* (covered passage). The side walls stand about 1.5 meters high. The stones were probably originally covered by a mound of earth. Work done by the author shows that in parts of France during the late Neolithic period communal tombs provided a geographic focus for scattered family houses. That finding is part of a recent reevaluation of spatial patterns in Neolithic culture based on distribution maps of Neolithic sites. The photograph was made by Ian Kinnes of the British Museum.

is about 2.5 acres), and their size suggests how many people they could have supported. Estimates of the amount of land required to feed each person by prehistoric wheat production vary from half a hectare to a hectare. Therefore the largest islands might have supported 500 villagers, or perhaps even more.

The ecology of the alluvial islands has been reconstructed by analyzing soil samples from pits and postholes of the Linear Pottery villages. Such work indicates that the valley bottoms were covered with a rich, dark soil similar to the "black earth" of eastern Europe. At the time of occupation the soil supported abundant meadows. In those meadows the Linear Pottery villagers did not practice swiddening, as had been envisioned in the earlier model. Rather, as work done in the mid-1970's showed, they practiced a form of horticulture in which plots were cultivated for many years. The fixed-plot system gave rise to considerable stability: it has been shown that some villages were occupied for 500 years or more.

The Pan-European culture based on this simple agricultural system lasted until about 4000 B.C. Soon after that date ecological and social forces began to alter the cultural landscape dramatically. Part of the alteration was due to the fact that within the Linear Pottery area subregions had begun to develop distinctive cultural styles. For example, the form of pottery known as Limburg ware is found from the Dutch Limburg down to Picardy in France across the lowland arc extending to the west of the Rhine. The entire lowland arc should probably be considered a unit-settlement phenomenon, which endured long enough for a unique cultural style to develop. The same thing undoubtedly happened elsewhere, and as the early Neolithic waned its cultural uniformity was replaced by a variety of styles.

By the beginning of the middle Neolithic in about 3500 B.C. even more profound developments were under way, as is shown by changes in the type, location and distribution of sites. Unlike Linear Pottery villages, middle Neolithic sites can be found outside the ecological niche of the valley bottom. Along with the change in location came a sharp increase in the proportion of sites defended by ditches and heavy palisades built of tree trunks. Apparently society had entered a phase of warfare. It had also become more dependent on cattle, which in the middle Neolithic account for more than 50 percent of animal remains, much more than in the earlier period.



NEOLITHIC PERIOD IN EUROPE spanned more than 3,000 years. The Neolithic is the period associated with the beginning of agriculture. The bearers of agriculture to northwestern Europe were members of the Linear Pottery culture. During early Neolithic times the Linear Pottery culture pervaded Europe. In the middle Neolithic, cultural uniformity gave way to regional styles such as Grossgartach, Chassey and Michelsberg. Flint arrowheads proliferated, partly because society had become more warlike.

What sense can one make out of these changes? First, it should be noted that some of their roots were already present in late Linear Pottery culture. Excavations at a site called Langweiler 3 on the Aldenhoven plateau and at Menneville in the Aisne valley revealed massive, fortified enclosures surrounded by ditches. Resemblances in construction suggest that the two enclosures are contemporary, and ceramic artifacts found with them are of a late Linear Pottery style known as Grossgartach. Similar enclosures have been found in the valley of the Seine at Barbuise-Courtavent and in the valley of the Aisne at Berry-au-Bac. Most ditched enclosures, however, date from between 3500 and 2700 B.C., in the mid-Neolithic. It would seem that forces present at the end of the early Neolithic were somehow accelerated, ultimately giving rise to communities that look very different from the earlier ones.

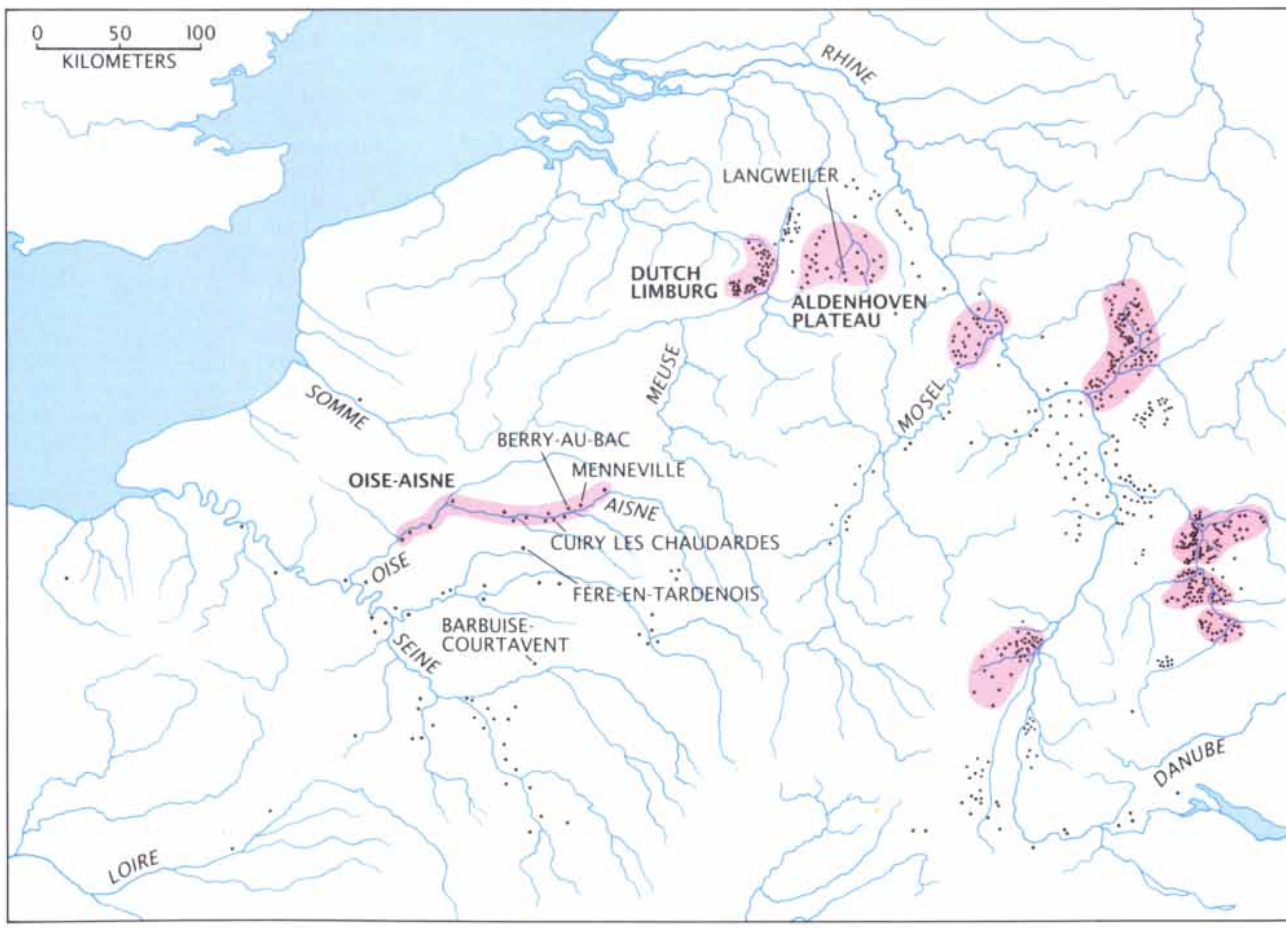
In trying to understand what those forces might have been, prehistori-

ans of northwestern Europe were initially thrown back on the concept of cultural influences. The pottery associated with the fortified middle Neolithic enclosures in France and Germany is generally of two types, which are known as Chassey and Michelsberg. Chassey culture covered central and southern France and Michelsberg covered central Germany. Much effort has gone into understanding the relations between them, based on the assumption that Chassey culture had affinities with the Mediterranean region and Michelsberg had affinities with central Europe. Understanding the influences of one on the other, it was thought, might help to unravel the social developments of the middle Neolithic.

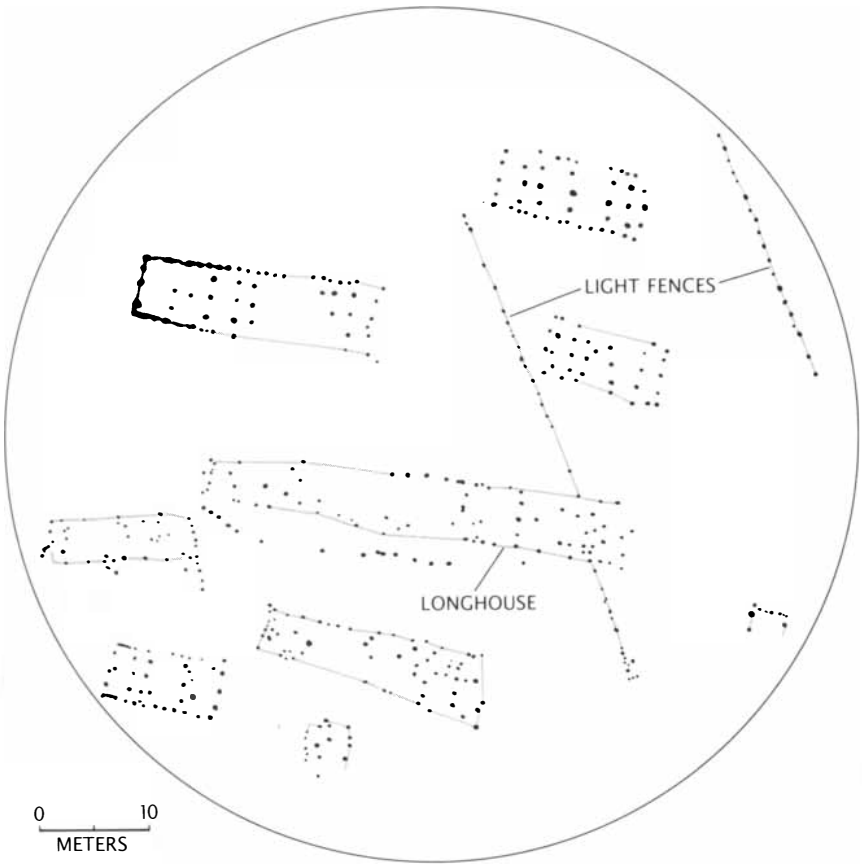
Yet such arguments were largely misplaced, in part because they lacked a firm geographic basis. In the mountainous terrain of Switzerland analogous "central European" and "Mediterranean" groups occupied different valleys. In the flat, open terrain of the Paris basin, however, cultural boundaries were not as distinct. There was undoubtedly much contact between groups with slightly different cultures. Indeed, the interaction was so intense that the styles intermingled: north of the Loire the main difference between them is that so-called Chassey pottery is decorated, whereas Michelsberg is plain. In other words, north of the Loire the tradition is firmly central European but with a number of outside influences.

Recently more fruitful work has concentrated on the aspects of social development that are common to both Chassey and Michelsberg traditions. In that work heavy emphasis has been placed on examination of spatial patterns, in particular the expansion out of the valley bottoms into new landscape zones. In the valley of the Somme it is clear that late Linear Pottery implements (for example stone axes) begin to be found outside the alluvial islands. A more detailed look at the valleys of the Aisne and the Oise shows the alluvial islands were by no means abandoned; enclosures are still found in the valley bottom. A new group of enclosures is found, however, in previously unoccupied territory on the edge of the plateau, particularly on the points of readily defensible promontories. Most of these new sites overlook the sites of early Neolithic villages on the valley floor.

Along with changes in location went a shift in the distribution pattern. Early Neolithic sites seem to be dotted more or less randomly over the alluvial islands. Middle Neolithic sites, on the other hand, tend to be grouped within their larger clusters. The form of the grouping varies with the locale. In the valley of the Seine, for instance, enclosures often are paired on the cliffs, facing each



LINEAR POTTERY SITES are grouped in "settlement cells," whose existence was first recognized in the 1970's. Linear Pottery groups moved north and west from eastern Europe along the main rivers. It was once thought the sites were randomly distributed in river valleys. The picture changed with the advent of the concept of settlement cells. The cells, shown in red, are groups of linked sites that were settled at roughly the same time. On the Aldenhoven plateau near Cologne and the Dutch Limburg the cells are dense clusters. Along the Oise and the Aisne rivers in France the valleys are steeper, and the cells are more linear.



LINEAR POTTERY VILLAGE near the hamlet of Cuiry les Chaudardes in the Aisne valley was excavated beginning in the 1970's by workers from the University of Paris. The main structures of the village were longhouses. The longhouses had timber frames that were filled in with mud and roofed with reeds, grass or rushes. The presence of one house that is longer than the others (*center*) may indicate that the village had a chief. Among the houses were light fences made of saplings, probably for enclosing animals. The circular excavation may cover less than a fourth of the Linear Pottery village.

other across the valley. Were they enemies? Allies? (Given the vagaries of the archaeological record, it is also possible that they were not precisely contemporary.) In the valley of the Aisne, Linear Pottery sites had been randomly distributed statistically on the lower terrace near the river. Middle Neolithic sites there, however, are found in statistically nonrandom clusters; within clusters the sites are about two kilometers apart.

Although some economic changes did accompany the spatial shifts, the overall basis of the economy did not change. The new sites on the edge of the plateau above the river are positioned to take advantage of the source of the many small streams that begin at the spring line. Such a pattern shows that mid-Neolithic communities, like their early Neolithic forebears, lacked heavy plows (they probably had light ones) along with irrigation and were dependent on high groundwater for their farming. What does seem to be new is the greater reliance on cattle, which require fairly large areas of grazing land. The search for grazing land may have been one of the factors motivating the upward movement out of the valley bottom.

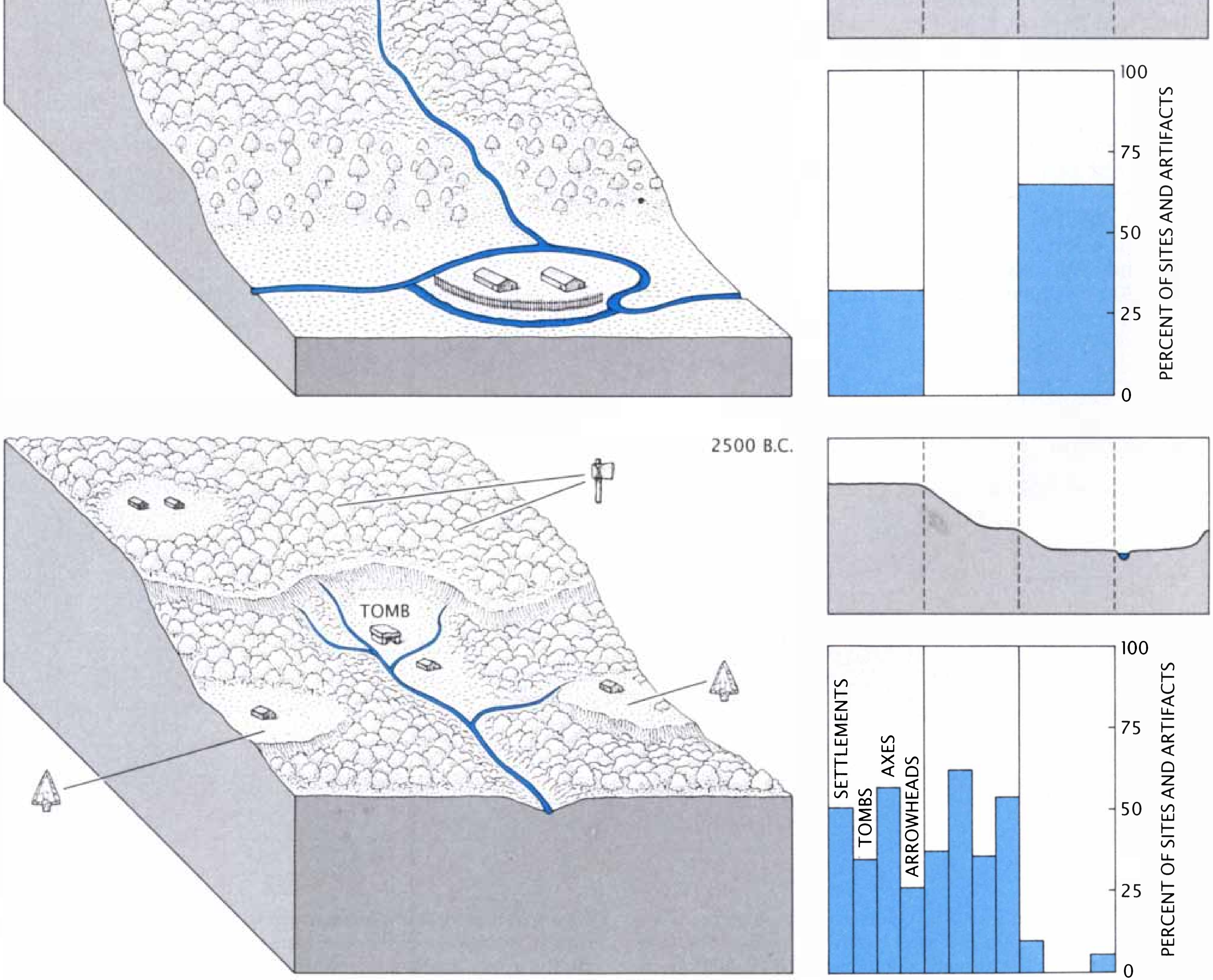
In its turn the search for pastureland may have given rise to a pioneering spirit that led to long-distance contact between groups. It is clear that the middle Neolithic was a time of increasing long-distance exchange. In addition to the search for pasturage, the pioneering spirit may have been encouraged by a need for scarce resources, such as workable flint or specific volcanic rocks for making axes. In Britain at this time such rocks were traded over distances of 600 kilometers or more. The trading of flint and stone in France has not been as carefully analyzed. Yet intensive flint mining did take place in the Oise, and it may have provided the basis for long-distance exchanges. The Oise is an area where fortified mid-Neolithic enclosures are densely concentrated, and those enclosures may have had a key role in redistributing raw materials and prestige objects (such as particularly highly decorated pottery) through trade.

Can a model be formulated to account for some of the striking changes seen in the middle Neolithic? Yes. The underlying stimulus for change may have been an interaction of climatic decline and natural-resource pressure. By the end of the early Neolithic much of the best alluvial soil had been claimed by Linear Pottery immigrants. In effect the frontier had been closed, and expansion no longer provided an easy outlet for population pressure. At the same time the technology needed for exploiting virgin land on the plateau (such as the heavy plow) was lacking. In such a situation deterioration of the climate might well have led to warfare. Hostilities would have resulted in the centralization of sites in defensible positions and in the construction of fortifications.

Various pieces of evidence suggest that this model captures at least some of the truth. The evidence for a more warlike society comes not only from the massive fortifications but also from the diversity of arrowhead types that proliferated in the middle Neolithic. Indeed at some sites, such as Crickley Hill in the English Cotswolds, arrowheads found in the rubble of a burned palisade offer unmistakable proof of hostilities.

Of crucial significance for such a model, of course, is the evidence for a deteriorating climate. Some data do suggest that in the mid-Neolithic the climate became colder and wetter,

SPATIAL DISTRIBUTION OF SITES changed sharply during the Stone Age. In the early Neolithic (*top*), Linear Pottery villages were on rich "islands" of alluvial soil carved by river meanders. By the mid-Neolithic (*middle*) some sites had moved up to promontories above the valley; many sites were heavily fortified. In the late Neolithic (*bottom*) sites became less centralized and moved far back on the plateau between rivers. Different landscape zones were exploited for specialized activities, as is indicated by the distribution of artifacts such as axes (for clearing forests) and arrowheads (for hunting).



which would quite likely have reduced agricultural yields. Pollen diagrams (counts of pollen grains from contemporary soil samples that yield a distribution of plant species) from the Aisne region suggest that beginning in late Linear Pottery times the proportion of water plants increased. By the late Neolithic much of the alluvial zone in the Aisne appears to have been transformed into a marsh. Work on oxygen-isotope ratios in marine fossils from the south of France (ratios sensitive to temperature) confirms a cold spell beginning in about 3500 B.C. In the mountainous zones of Europe the tree line descended at about the same time, indicating a colder climate.

Social pressure resulting from a scarcity of arable land is more difficult to document than changes in climate, because it leaves few physical traces. Nevertheless, it seems probable that in the context of a worsening climate middle Neolithic communities might have experienced land pressure. After all, by the beginning of the middle Neolithic many alluvial islands had been occupied for from 500 to 1,000 years. After a millennium of farming, the available land might have been exhausted, leading to sharp competition for fertile fields.

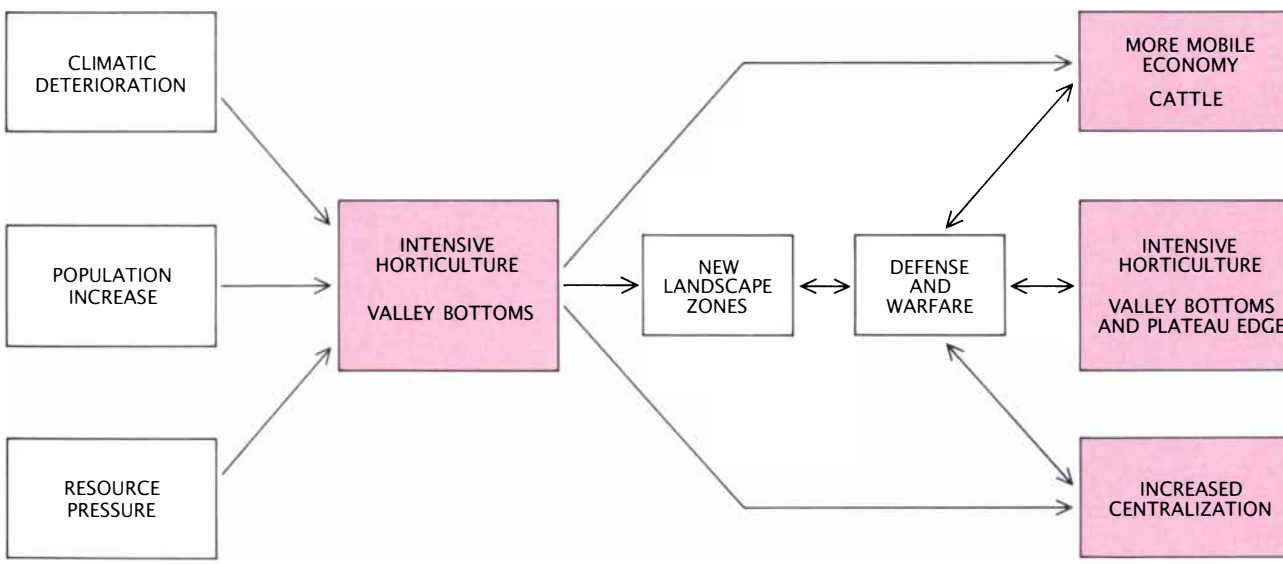
Additional evidence will no doubt be required before the model proposed above can be rigorously confirmed or rejected. Whatever the outcome of that process, it is clear that by the end of the middle Neolithic in about 2700 B.C. further changes had begun. In France the beginning of the late Neolithic is marked by the emergence of a new cultural grouping that takes its name from the Seine, Oise and Marne. Typical sites of the Seine-Oise-Marne culture are strikingly different from those of mid-Neolithic communities. Huge, fortified enclosures were no longer built, and in their place one finds small, scattered remains, each the site of a dwelling for perhaps only one family group.

Although they are not centralized the way middle Neolithic sites are, the late Neolithic dwellings did not exist in isolation. Evidence from the upper reaches of the Ourcq River near the town of Fère-en-Tardenois shows that the houses were grouped in what might be called an expanded village: a cluster of widely separated structures with a burial place as their common geographic focus. Sometimes the burial is marked by a megalithic monument, but more often it is simply a communal pit containing as many as 200 or more bodies.

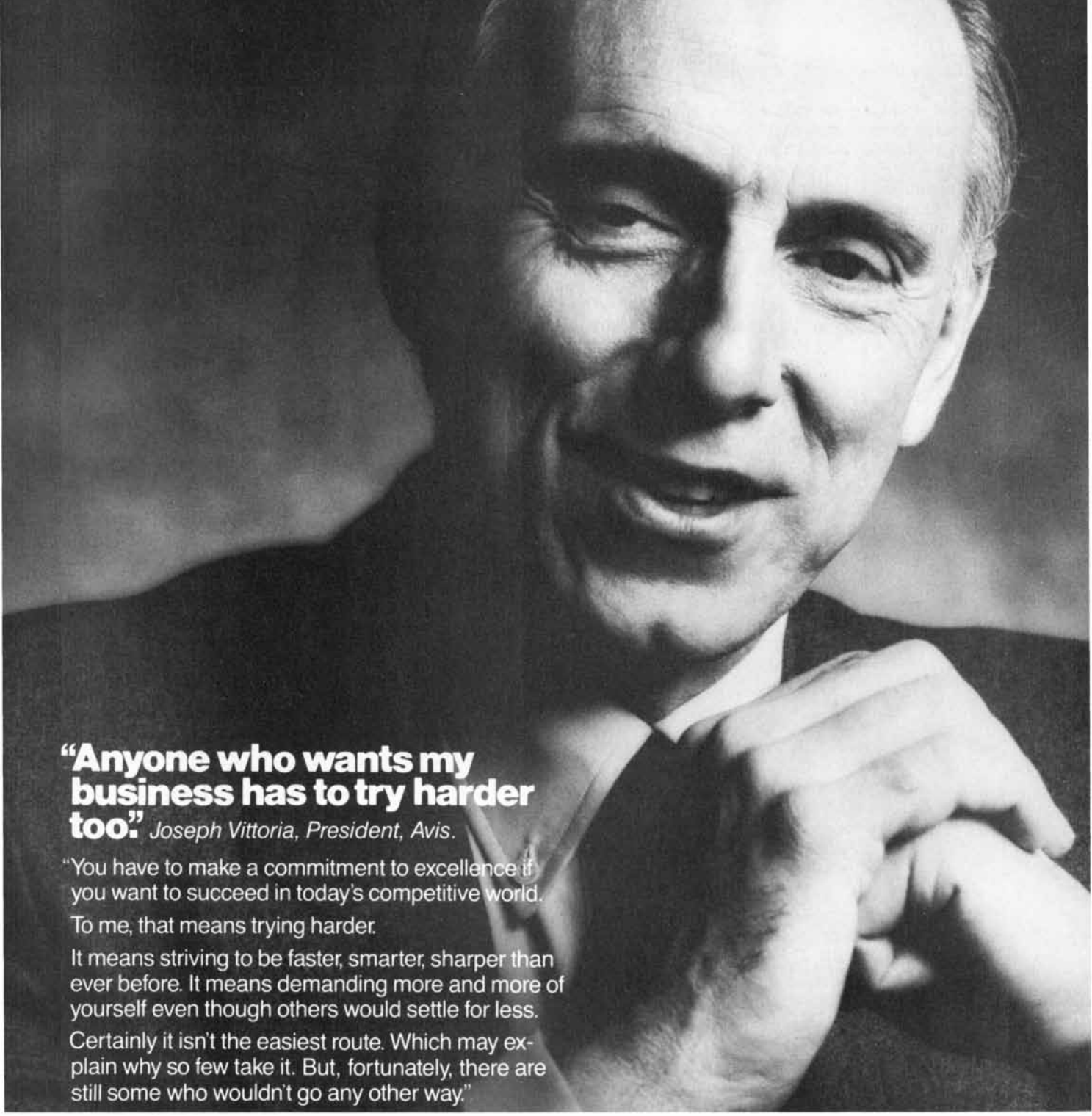
A striking feature of the late Neolithic clusters is that they have expanded far beyond the promontories and onto the plateau. I constructed a simple site typology of late Neolithic sites in northern France and located them in ecological zones. Four types of finds were plotted: houses, burial sites, axes and arrowheads. Houses and axes were typically found on the plateau between the rivers. Burial sites tended to be on hillsides covered with the thin, chalky soil called rendzina; arrowheads were likely to be found on hillsides covered with sandy deposits, which were probably rich hunting grounds.

Clearly, then, in late Neolithic culture different activities took place in their own landscape zones. Even more significant, agriculture (associated with the house sites) had moved out on the plateau, implying an economic shift. As I have described, the soil near the spring line, which provided the basis of middle Neolithic farming, was sufficiently moist and fertile to be worked without heavy plows. On the plateau, which is watered chiefly by rain, the heavy plow is needed to turn the soil and bring up moist earth from below. Indeed, the appearance of Seine-Oise-Marne culture is roughly simultaneous with the spread of the heavy plow and means of pulling it. Equipped with those implements, individual families could farm the previously unworkable soil of the plateau and achieve unprecedented economic independence.

The introduction of the plow appears to have been only one aspect of a general economic transformation that took place during the late Neolithic. Charcoal layers at sites cleared of forest during late Neolithic times suggest that fixed-plot methods had given way to a slash-and-burn mode. That conclusion is reinforced by the fact that most of the Seine-Oise-Marne dwellings are small and were occupied briefly compared with the length of time that the burial sites were used. There is also evidence that the pig (an animal well suited to newly cleared forest edges) had tak-



AUTHOR'S MODEL for the social changes that took place during the mid-Neolithic period is based partly on climatic deterioration. A wetter, colder climate, combined with population increases, may have led to pressure on resources such as land for farming. Resource pressure in turn may have yielded warfare, with the accompanying fortification and centralization of sites.



At Lufthansa, we're always working harder.
It's evident in our outstanding on-time arrival record, our meticulously maintained fleet, and our friendly, courteous service.

Perhaps that's why, in 161 cities worldwide, people prefer flying Lufthansa.

People expect the world of us.

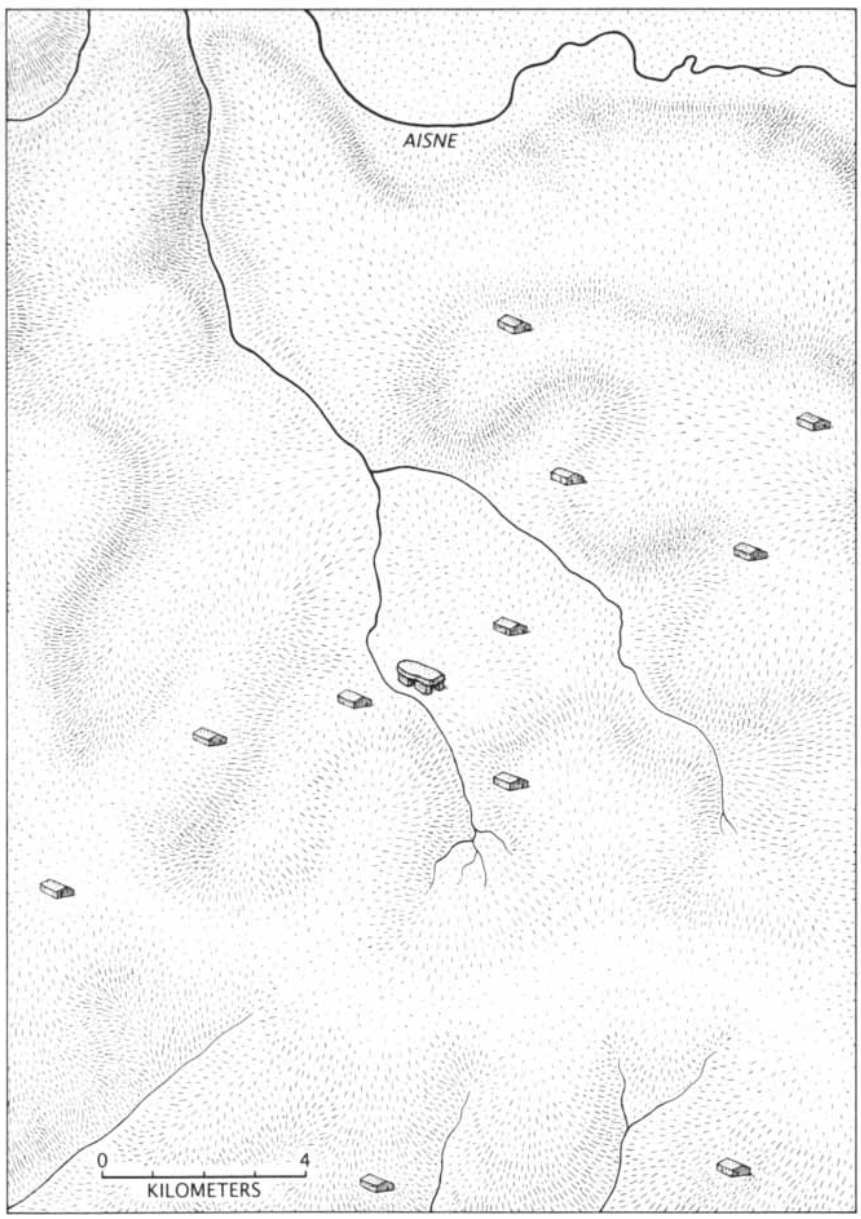
Lufthansa
German Airlines

Lufthansa is a participant in United's Mileage Plus er program. See your Travel Agent for details.

en on a greatly increased economic significance.

Recent thinking in archaeology has tended to explain the linked appearance of plow, traction and new domesticates as a "Secondary Products Revolution," which in certain parts of Europe added milk and wool to the inventory of agricultural products. Indeed, in some regions the development of wool as a cash crop, accompanied by the introduction of metallurgy, heralded the end of the late Neolithic period and the beginning of the Bronze Age. One branch of Seine-Oise-Marne culture (the Horgen group in Switzerland) underwent just such a transition from about the middle of the third millennium B.C. onward.

In the heartland of the Seine-Oise-Marne culture, however, the very success of the expansion onto the plateau seems to have inhibited further social development. In the Paris basin the amount of land accessible to plow-based agriculture was quite large in relation to the population. Unlike some other regions, including Switzerland, where the limited availability of land led to renewed demographic pressure, in much of France simple outward expansion made it possible to maintain social and economic structures that elsewhere were forced out of existence by Bronze Age innovations.



LATE NEOLITHIC DWELLINGS centered on a funerary monument are near the village of Vaux, which is south of the Aisne valley. In its general construction the tomb resembles the one shown in the illustration on pages 118 and 119. It was probably a communal burial site, although the direct evidence was removed by looting in the 19th century. The Neolithic sites scattered around the tomb were dwellings for individual families.

Thus it seems that resource pressure—in particular demographic pressure on arable land—was one of the driving forces of social and cultural change during much of the Neolithic period in northwestern Europe. Population pressure appears to have been one of the key factors in causing the remarkable centralization and fortification of settlements that appeared during the mid-Neolithic. In addition, resource pressure had a role in determining that in areas such as Switzerland and Britain, where arable land was limited, the Stone Age ended sooner than it did in the Seine-Oise-Marne region, where the plateaus offered ample room for continued expansion.

The crucial significance of arable land for social structure is perhaps not surprising in an early agricultural society. Yet for the Neolithic period that significance could not have been understood without increased scholarly emphasis on spatial distribution, both on the scale of Europe and the smaller scale of local areas. Indeed, the distribution map is one of the tools that has survived a notable recent transition in archaeology and emerged with increased scholarly respect.

One often hears the term "New Archaeology" applied to approaches in archaeology that were developed in the 1960's. In contrast to the "Old Archaeology," which was largely empirical in its bent and shied away from overarching theories, the New Archaeology draws heavily on theories developed in anthropology, linguistics and geography.

The uses of the distribution map reflect these fundamental changes within the discipline of archaeology. Instead of simply being an instrument of location, such maps have become the starting point for elaborating theories to explain underlying events. In the case of the Neolithic period in northwestern Europe, that is just what has happened. Beginning with such spatial concepts as settlement cells, centralization and scattering, and ecological zones, it has been possible to explain some of the underlying forces that determined what form early agriculture took on the plains of northern Europe.

Great Products at Great Prices from a Great Company...

ELEK-TEK

DATA STORAGE

i·MEGA BERNOULLI BOX
For PC/XT/AT and most compatibles

10 meg (single) **$950**
20 meg (single) **1275**
20 meg (10 + 10) **1450**
40 meg (20 + 20) **1650**
Bernoulli Plus (80+20+20) **2999**
interface priced separately
Bernoulli Cartridges
IOMEGA 10 meg—3 pak **155**
IOMEGA 20 meg—3 pak **239**
Dysan 10 meg—3 pak **140**
BB HEAD CLEANING KIT 59

HARD DRIVE CARDS

Plus Development 20/40 MB . CALL
Western Digital 20 MB **$399**
Western Digital 30 MB **550**

MULTI FUNCTION BOARDS

AST
Advantage 128 . . . . . **$229**
I/O Mini II **79**
Six pak + with 384K **169**
Six pak premium with 512K **229**
Rampage PC **229**
Rampage 286 . . . . . **279**

GENERIC Multi-function Board
(clock & calendar, 2 ser., 1 par.)
0K **89**
384K **159**

INTEL ABOVEBOARD
All products . . . . . **CALL**

ORCHID TECHNOLOGY
All products . . . . . **CALL**

GENERIC COMPATIBLES OF MOST OF THE ABOVE . . . . . CALL

HARD DRIVES

SEAGATE
Internal Hard Disk Subsystems
20 Megabyte $319
Each system includes:
- Seagate Drive
- Controller
- Documentation
- 2 Cables
- 1 yr limited warranty
- For use in IBM PC, XT, AT&T, most compatibles

30 MB RLL ½ ht. kit 349
40 MB Drive full ht. for AT 599
80 MB Drive full ht. for AT 999

MINI SCRIBE
30 MB ½ height **369**
40 MB full height **675**

ARCHIVE XL
Tape Backup

- 40 MB Capacity
- QIC—40 Tape and File Structure Format
- Uses Standard floppy interface

Mfr. Sugg. Ret. $419 **$329**

MODEMS

Hayes
Model 1200

HAYES *Internal*
(including software)
1200B 1200 Baud **$279**
2400B 2400 Baud **429**

HAYES *External*
(no software included)
1200 1200 Baud **289**
2400 2400 Baud **439**

OTHER MODEMS
Hayes compatible
Internal *(including software)*
Practical Peripherals 1200B **84**
Everex 1200 . . . . . **94**
Everex 2400 . . . . . **189**
External *(no software included)*
U S Robotics 1200 baud **199**
U S Robotics 2400 baud **359**
Practical Peripherals 1200 baud . **109**
Practical Peripherals 2400 baud . **179**
Novation Parrot 1200 . . . . . **92**
Everex 2400 . . . . . **239**

CHIPS

64K RAM (150 or 200 NS) **$1.35**
128K RAM **2.20**
Samsung 256K RAM 150 NS **2.60**
Major 256K 100/120/150 NS **3.30**
256K Static RAM **7.00**
NEC V-20 . . . . . **9.00**
Intel 8087 math chip (5MH) **110**
Intel 80287 math chip (6MH) . . 175

HEWLETT PACKARD
LaserJet
Toner Cartridges
HP92285A for Original LaserJet
HP92295A for LaserJet Series II
Mfr. Sugg. Ret. $115 **$80**

PRINTERS

EPSON
FX 86e

Epson's FX-86e is the industry's standard DOT MATRIX PRINTER. Features include 8K buffer, NLQ, 40/200 cps.

OTHER EPSON PRINTERS
FX 86e (New)200 CPS . . . . . **CALL**
FX 286e wide carriage (New) . . **CALL**
LX 800 180 CPS . . . . . **CALL**
EX 800 300 CPS . . . . . **CALL**
EX 1000 300 CPS . . . . . **CALL**
LQ 800 NLQ (Parallel) . . . . . **CALL**
LQ 1000 NLQ (Parallel) . . . . . **CALL**
LQ 2500 NLQ 324 CPS . . . . . **CALL**
CR 420i NLQ 420 CPS . . . . . **CALL**

SAVE OVER 80%
EPSON SQ-2000
wide carriage Ink jet printer
Mfr. Sugg. Ret. $2395 **$399**

PRINT BUFFERS

PRACTICAL PERIPHERALS
64K parallel or serial **$129**
256K parallel or serial **159**

HANZON
64K universal **250**
64K memory expansion **56**

QUADRAM
Microfazers . . . . . **CALL**

MONITORS

BEST BUY
Samsung 1464W 13" RGB **$275**
Samsung 1252G 12" Amber TTL . **99**
Tatung 1360 RGB **339**
Sanyo 9212 12" Amber Composite **84**

NEC
Multisync **549**

AMDEK
310A (TTL) Amber **129**
410A, 410G, 410W . . . . . **CALL**
600S RGB color **339**
722C **450**

PRINCETON GRAPHICS
All Models . . . . . **CALL**

SONY MONITORS
All Models . . . . . **CALL**

Huge Discounts on
TOSHIBA NEC OKIDATA XEROX / Diablo

DICONIX 8213
Portable Printer
Perfect for Laptop Computers
Call for Price

SPECIAL PURCHASE!
Diablo D-625 Daisywheel 25CPS
Elek-Tek Price $399*
*Price includes IBM parallel interface

TRIPPLITE & PERMA POWER POWER PROTECTION

Surge protector strips
PermaPower 6 outlet **$28**
Isobar4-6 4 outlet **42**
Isobar8-15 8 outlet **50**
Line conditioners/stabilizers
TL1200 4 out 1200W **135**
TL1800 6 out 1800W **199**
Command consoles
TLCCI 612 . . . . . **79**
PermaRS562 Power Commander **65**
Standby power systems
TLBC450 450 watt system **360**
TLBC750 750 watt system **535**
TLBC1200 1000 watt system . . . **850**
TLBC2000 . . . . . **1299**
SAFE400VA 400 volt system . . . **350**
SAFE1200VA 1200 volt system . . **899**

PUSH BUTTON SWITCH BOXES

448010 2 pos. 25-pin AB **$29**
448052 2 pos. 36-pin AB **32**

ROTARY SWITCH BOXES

448028 3 pos. 25-pin ABC **$29**
448060 3 pos. 36-pin ABC **33**

VIDEO BOARDS

Hercules Graphics PLUS **$179**
Generic Hercules compatible **80**
Hercules color **155**
Generic color compatible **80**
Hercules InColor Card **329**
Vega Deluxe **289**
Everex EGA 480 . . . . . **150**
Generic EGA ½ card **139**
Quadram Prosync **369**
Paradise EGA 480 . . . . . **175**
Paradise EGA 350 . . . . . **139**
Paradise Hi res **90**
Paradise Modular Graphics **99**

HUGE SAVINGS ON DISKETTES *Call for Quantity pricing for 10 boxes or more*

(*Packed in flip n' file) (**unformatted)	Dysan	maxell	3M	ELEK-TEK	SONY	
3½" SSDD	14.00	13.00	13.00*	—	13.00	12.00
3½" DSDD	18.00	20.00	20.00	—	20.00	16.00
3½" DSDDHD	—	—	49.00	—	49.00	—
5¼" SSDD	8.00	8.00	8.50	—	—	7.00
5¼" DSDD	10.00	10.00	10.00	9.00	—	8.00
5¼" DSDD96TPI	12.00	20.00	20.00	—	—	10.00
5¼" DSDDHD (for IBM AT)	12.50	20.00	20.00	12.50	—	17.00
8" SSDD**	19.00	20.00	20.00	—	—	—
8" DSDD**	23.00	23.00	23.00	—	—	—

3M DATA CARTRIDGES

3M	1	10	25
DC100A	$13.50	$13.00	$12.00
DC300A	17.50	17.00	16.00
DC300XL/P	21.50	21.00	20.00
DC600A	23.00	22.50	21.50
DC1000	14.50	14.00	13.50
DC1000 alpha**	16.50	16.00	15.00
DC1000 delta*	16.50	16.00	15.00
DC2000	19.00	18.50	18.00
DC2000 gamma ***	22.50	22.00	21.00
DC2000 alpha ***	22.50	22.00	21.00

* formatted 10 MB Cartridges
** formatted 20 MB Cartridges
*** formatted 40 MB Cartridges

ELEK-TEK

6557 North Lincoln Avenue, Chicago, Illinois 60645 312-677-7660
Call Toll Free 800-621-1269—Except IL, AK • Canadian Toll Free 800-458-9133

Corp. Accts. Invited. Min. Ord. $15.00. Visa or MasterCard by Mail or Phone. Mail Cashier's Check, Mon. Ord., Personal Check Add $4.00 1st item. (AK, HI, P.R., Canada add $10.00 first item) $1.00 ea. add'l shpg. & handl. Shipments to IL address add 7% tax. Prices subj. to change. WRITE for free catalog. RETURN POLICY: All sales are final except defectives which will be replaced with identical merchandise only within various time *rames as permitted by manufacturer's warranty policy. ALL ELEK-TEK MERCHANDISE IS BRAND NEW, FIRST QUALITY AND COMPLETE. Delivery subject to availability. DUNS #09-718-0517

Modeling Tidal Power

Computer simulations show that a tidal power dam in the Bay of Fundy would raise tide levels as far away as Boston. The models can now be used to help assess the environmental and economic cost of tidal power

by David A. Greenberg

The highest tides in the world are found in the upper Bay of Fundy, which lies between New Brunswick and Nova Scotia. The difference between low water and high water there is typically 12 meters, and for extreme tides the range can be close to 17 meters. As the water approaches its peak level it often rises at the rate of more than a foot every 10 minutes, and in some places unwary beachcombers must race to reach dry land ahead of the surging tide. The water of the St. John River pours upstream over the Reversing Falls, and in the rivers at the head of the bay one can see tidal bores: the shallow bottom slows the leading edge of the advancing tide, so that a moving cascade seems to propagate up the river.

For centuries the tides have ruled the lives of people along the Fundy coast. At the same time there were always those who dreamed of harnessing the tide for their own benefit. As early as 1609 settlers built mills driven by tidal power. In this century the dream is to exploit the tides to generate electricity. The most effective way to extract power from the tides would be to build a dam across part of the bay. At high tide water would fill the reservoir behind the dam; after the tide had receded enough the water would be released through turbines, much as it is in an ordinary hydroelectric dam. The total flow at each tide through the Bay of Fundy could theoretically generate 400 million kilowatt-hours of energy, which would be equivalent to the output of 250 large nuclear power plants. Realizing just a small part of that potential would make a useful contribution to the energy demands of eastern North America.

The governments of Nova Scotia and of Canada have already built a test plant, on an inlet of the bay at Annapolis Royal, that has a generating capacity of 20,000 kilowatts, and they have seriously examined three proposed sites for a commercial tidal power dam in the Bay of Fundy. The largest of the proposed projects, in Minas Basin, would release through its turbines 55,000 cubic meters of water per second, or about seven times the flow of the St. Lawrence River at Montreal. It would be capable of generating some 5,000 megawatts of electric power. The projects would have been competitive economically a few years ago, when oil prices reached their highest level; they remain an attractive option for the near future because, unlike power from nuclear or fossil fuels, tidal power is a clean, reliable and—on the face of it—safe source of renewable energy.

Yet any engineering project of this size is bound to have an impact on the environment. In order to study what some of these effects might be, my colleagues and I have developed a computerized model. It suggests that a tidal power dam would cause a small increase in tide levels over a large area extending down the Gulf of Maine to Boston and Cape Cod, an increase that could have significant environmental consequences. We cannot yet predict the exact environmental impact, but we think our model is now accurate enough to play an essential role in assessing the true costs of tidal power.

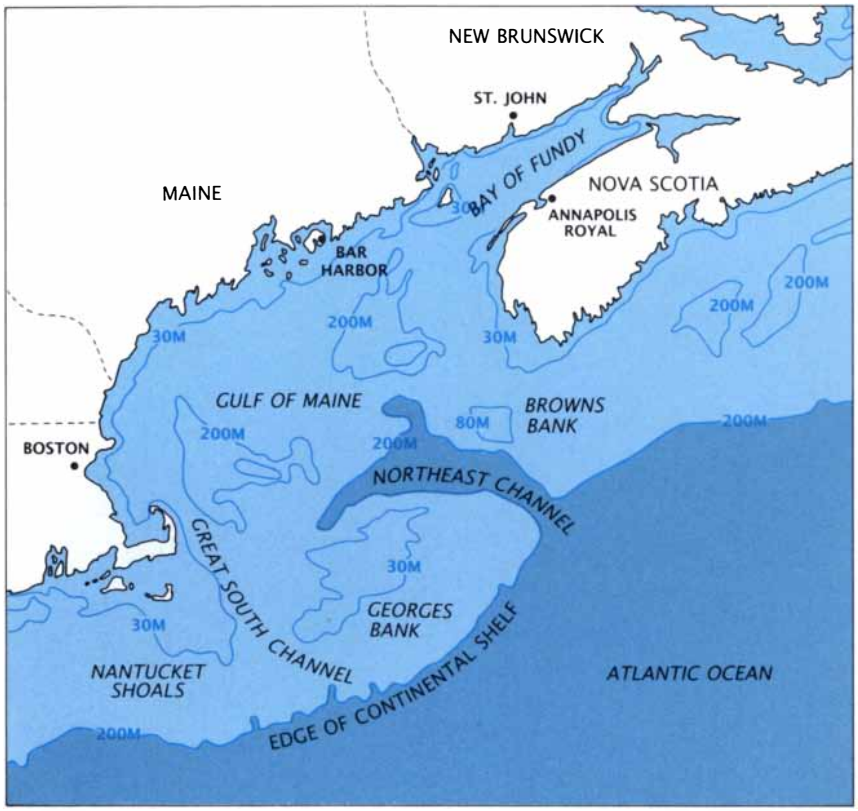
The scientific explanation for the high tides in the Bay of Fundy has been the subject of much speculation for decades. The gravitational forces of the sun, earth and moon generate the tides in the deep sea, which are in turn the primary driving force for the tides in the shallower waters of the continental shelf. But on the continental shelf itself local factors such as the shape of the shoreline, the depth of the water and the width of the shelf affect the tide even more than the astronomical forces. The exceptionally high tides in the Bay of Fundy are thought to result from the physical proportions of the bay, which give rise to a phenomenon called resonance.

Resonance can occur in any system in which there is a natural frequency of oscillation. When the system is driven by an external force at the same frequency as its natural frequency, there is a sudden and drastic increase in the amplitude of the oscillations. For example, a crystal wine goblet, by virtue of its shape and substance, rings at a certain tone, or frequency, when it is struck. A singer can produce the same tone, generating sound waves that reinforce the natural vibrations of the goblet and can cause it to shiver so violently that it breaks.

A closer analogy to what takes place in the Bay of Fundy is the oscillation of water in a bathtub. If a standard bathtub is filled to a depth of 20 centimeters, a wave of water sloshes back and forth about once every two seconds. Now, if one were to help the wave by giving it a push every time it went by, the wave would quickly increase in size and threaten to overflow the sides of the tub. One might think that pushing more often would have an even greater effect, but in fact it would cause only light splashing because the pushing would interfere with the natural motion of the wave. The period of the wave (the length of time it takes to complete a full oscillation) is determined by the depth of the water and the length of the tub.

For some 50 years it was thought the depth and length of the Bay of Fundy were such that a deep-water wave would move into the bay and

DRAMATIC CHANGE is a daily fact of life for the inhabitants of the small fishing port of Halls Harbour, Nova Scotia. The tide amplitudes here are average by the standards of the Bay of Fundy, and yet they govern the work patterns of the fishermen. Boats can sail only at high tide and must be docked carefully so that they come to rest safely on the exposed harbor bed at low tide.



GULF OF MAINE and Bay of Fundy must both be included in a study of the tides in the bay and the effects of building a power dam. The highest tides are found in Minas Basin and the swiftest tidal currents off Cape Split. Shepody Bay, Cumberland Basin and Minas Basin are proposed sites for tidal dams. There is a test plant at Annapolis Royal.

be reflected out again with a period very close to the period of the lunar tide. The tidal period is 12.4 hours; the bay's "first resonant mode" was estimated to be somewhere between 10.5 and 12 hours. In 1968, however, Desiraju B. Rao, then at the marine-sciences branch of the Department of Energy, Mines and Resources, calculated on the basis of an early tidal model that the resonant period of the bay is actually shorter: about nine hours. That is close enough for some amplification but not close enough for true resonance. The resonance thesis is nonetheless serviceable—if the model considers not only the Bay of Fundy but also the Gulf of Maine beyond it. The bay and gulf combined turn out to form one tidal system, which has a resonant period of about 13.3 hours.

The resonance thesis predicts that if the length of the bay were to be changed by constructing a tidal power dam at its head, the bay's resonant period would change; because the bay is near resonance with the tide cycle, even slight changes in its length could cause significant changes in tidal amplification. In order to model the tides and predict changes due to tidal power barriers, we need to include several significant factors other than the length of the bay. In the upper bay, for example, the exposure of shallow areas at low tide and their submergence at high tide can shift the land-water boundary by several kilometers. Off Cape Split, between Minas Channel and Minas Basin (the area with the largest tides), there is a narrows that severely constricts the flow of the tide. Currents of more than four meters per second are found there in the flood tide and the ebb tide, leading to unusually high dissipation of tidal energy—which also must be computed in our model.

Our model, like any mathematical model, is a tool for predicting. A deterministic view of nature proposes that if one knew all the physical laws and the state of the entire universe at a given instant, one could predict the future. Unfortunately it is not possible to solve all the relevant equations of motion for even a small system. With the advent of computers, scientists have found one way around this intractable problem by constructing mathematical models that are simplified approximations of the real world. The models themselves are often extremely complex, with a large number of parts, each of which calls for huge amounts of computation. Yet they can yield important results in a reasonable length of time. The challenge in designing a good model is to break down the problem to the proper scale and to choose the physical laws that are most relevant to the problem. There is always a tradeoff between the accuracy afforded by a more complex model and the faster computation time of a simpler one.

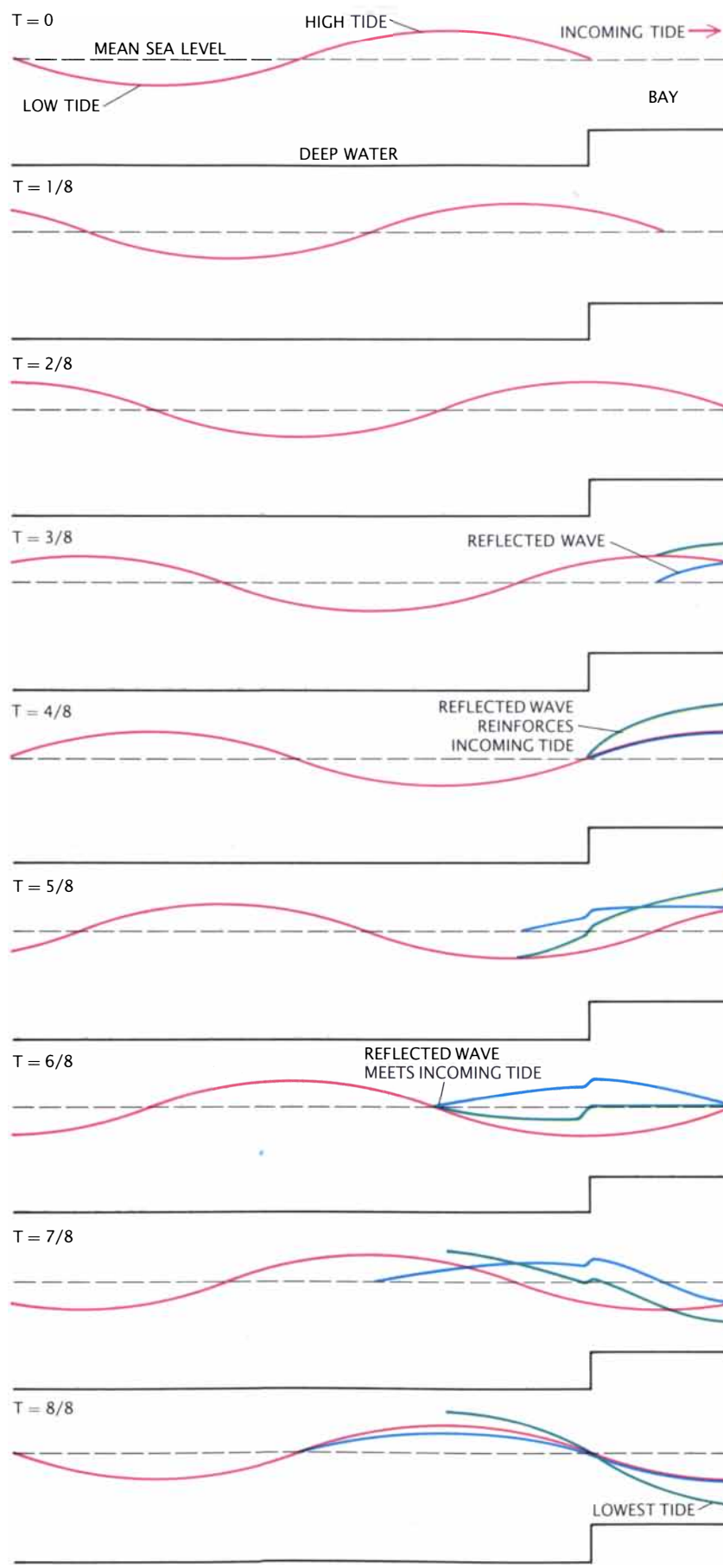
In modeling the Bay of Fundy and Gulf of Maine we are concerned not with the motions of individual molecules but with the tide-driven movement of large volumes of water. We therefore divided the area into a grid in which each grid cell represents a square column of water extending from the sea floor to the surface. Each cell obeys the physical laws that govern the motion of fluids in the sea. (One can consider the currents to be the same from top to bottom.) Then, if we know the state of the tidal currents and water levels at a certain time, we can apply our model to calculate the currents and elevations at some future time. Moreover, we can predict what would happen if the physical area were changed, for example by a dam.

The law of conservation of mass dictates that the water level of a grid cell goes up if there is more water coming in across the sides of the cell than is going out. If the water level is higher in one cell than in an adjacent cell, the water is accelerated toward the one with the lower level. Because of the earth's rotation, any flow between the cells leads to an acceleration in a perpendicular direction (to the right in the Northern Hemisphere). At the same time the flow is opposed by a frictional force that is proportional to the square of the current's speed. In certain places, such as a narrow, fast channel or a sharp bend, we must also account for the acceleration of the water as the current changes speed and direction in a short distance. In most tidal calculations this effect is small and can be ignored, but it is not the case for the upper Bay of Fundy.

The size of the grid cells is what determines the resolution of the model and the amount of computation

needed. The smallest island, bay or subsurface ridge that can be represented in the model is one that is about the size of a single cell. The shortest barrier or spit would be one whose length is that of the side of a cell. The maximum time step that can be used is the time it takes for a shallow-water wave to traverse one grid cell. Since many computations are made for each cell, high resolution (many cells for a small area) would require more computation than lower resolution. For example, to improve the resolution of a model by a factor of two one would have to square the number of grid cells and use twice the number of time steps (because the wave crosses the smaller cells in half the time), which would require an eightfold increase in the computational effort. (That is to say, the computational effort is inversely proportional to the cube of the length of the side of a grid cell.)

In order to reduce the computation time needed to run our model of the Bay of Fundy and Gulf of Maine, we developed a progressive mesh-refinement scheme that yields very fine resolution (1.7 kilometers) in the upper Bay of Fundy, with progressively larger grid cells (2.3 and seven kilometers), and coarsest resolution (21 kilometers) in the Gulf of Maine. Extensive testing has indicated that the accuracy of the model's predictions is not affected by coarse resolution in the Gulf of Maine but is noticeably affected by resolution at the head of the bay. The highest resolution is required in Minas Basin, around Cape Split and around intertidal areas, where the extreme differ-



DEEP-SEA TIDE is amplified by resonant interaction with an appropriately shaped bay. The tide (*red line*) enters the bay at $t = 0$. At $t = 2/8$ the tide reaches the head of the bay and is reflected back out (*blue line*). The incoming and outgoing waves combine to produce the observed water level (*green line*). At $t = 4/8$ the waves reinforce to give the peak tide amplitude. At the end of one cycle the combined troughs result in the lowest tide; at the edge of the continental shelf the outgoing crest meets the next tide coming in. In the real world the deep sea is 10 times deeper than the continental-shelf waters. When the reflected wave leaves the bay, its amplitude falls by an amount inversely proportional to the change in depth of the water. (The amplitude did not increase when the tide entered the bay because some water bounced off the edge of the shelf and was sent back to the deep sea.)

No other system of keeping up can compare with SCIENTIFIC AMERICAN *Medicine.*

YOUR SYSTEM: a time-consuming, futile struggle to keep up with the information explosion.

The classic texts are convenient references—but the information they contain is obsolete before publication.

Like many physicians, you probably rely on the texts you first used in medical school. But even using the most recent editions, you find material that no longer reflects current clinical thinking—and they lack the latest information on such topics as herpes, oncogenes, AIDS, and photon imaging.

Reading stacks of journals alerts you to recent developments—but can't give you quick answers on patient management.

Struggling through the hundreds of journal pages published each month—even on only the really significant advances in the field—is arduous and memory-taxing. And it's a task that costs physicians valuable time—their most precious resource.

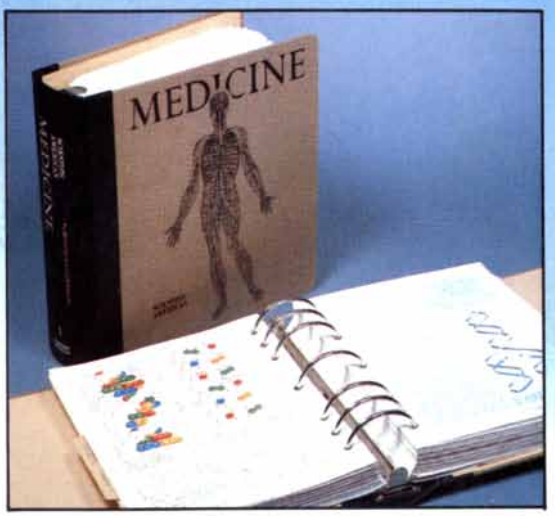
Review courses cover clinical advances—but, months later, do you recall the details of a new procedure or unfamiliar drug?

Seminars can also be costly and make you lose valuable time away from your practice—expenses that may amount to several thousand dollars. And, the speaker's skill often determines how much you learn.

SCIENTIFIC AMERICAN

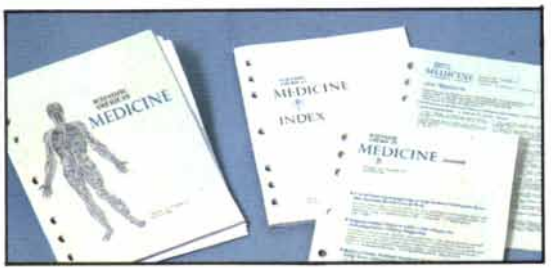
MEDI

OUR SYSTEM: a rewarding, efficient way to keep yourself up-to-date—and save hundreds of hours of your time for patient care.



A comprehensive, 2,300-page text in two loose-leaf volumes, incorporating the latest advances in medical practice as of the month you subscribe.

This superbly designed, heavily illustrated resource, called "the best written of all [the internal medicine] books" by *JAMA* (251:807, 1984), provides a practical, comprehensive description of patient care in 15 subspecialties. And, because the text is updated each month, the clinical recommendations reflect all the current findings. A practice-oriented index and bibliography of recent articles further enhance the efficiency of the text.



Each month, six to nine replacement chapters to update your text *plus* new references, a four- to five-page news bulletin, and a completely new index.

You'd have to read hundreds of journal pages each month—and memorize the contents—to get the same information SCIENTIFIC AMERICAN *Medicine* contains. With our updated text, you read only the information you really need. Our authors, largely from Harvard and Stanford, sort through the literature and monitor developments, incorporating the significant advances into our chapters.

At no additional cost, a 32-credit CME program, to save you valuable patient-care time and the expense of attending review courses.

Earn 32 Category 1 or prescribed credits per year with our convenient self-study patient management problems; each simulates a real-life clinical situation. Choose either the complimentary printed version or, at a modest extra charge, the floppy-disk version for your IBM® PC/PC Jr., Macintosh™, or Apple® IIe or II+ (or compatible).

Your subscription includes:
the two-volume, 2,300-page loose-leaf text
and, each month, six to nine replacement chapters,
a newsletter, new references,
and a completely revised index.

Or call toll-free: 1-800-345-8112
(in Penn. 1-800-662-2444)

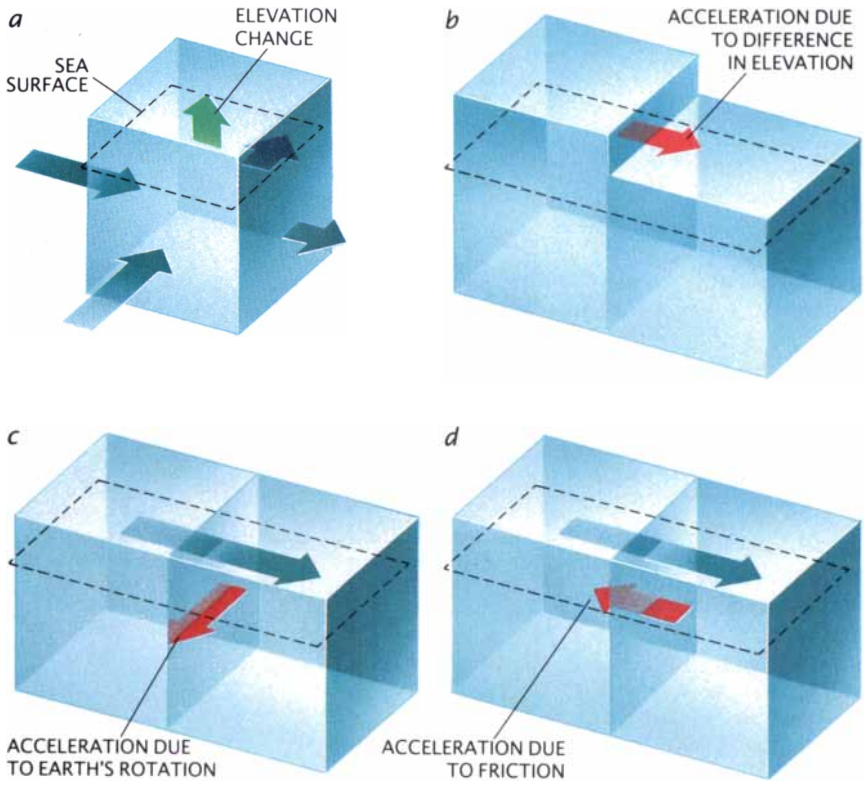
CINE

SCIENTIFIC AMERICAN MEDICINE
415 MADISON AVENUE, NEW YORK, N.Y. 10017

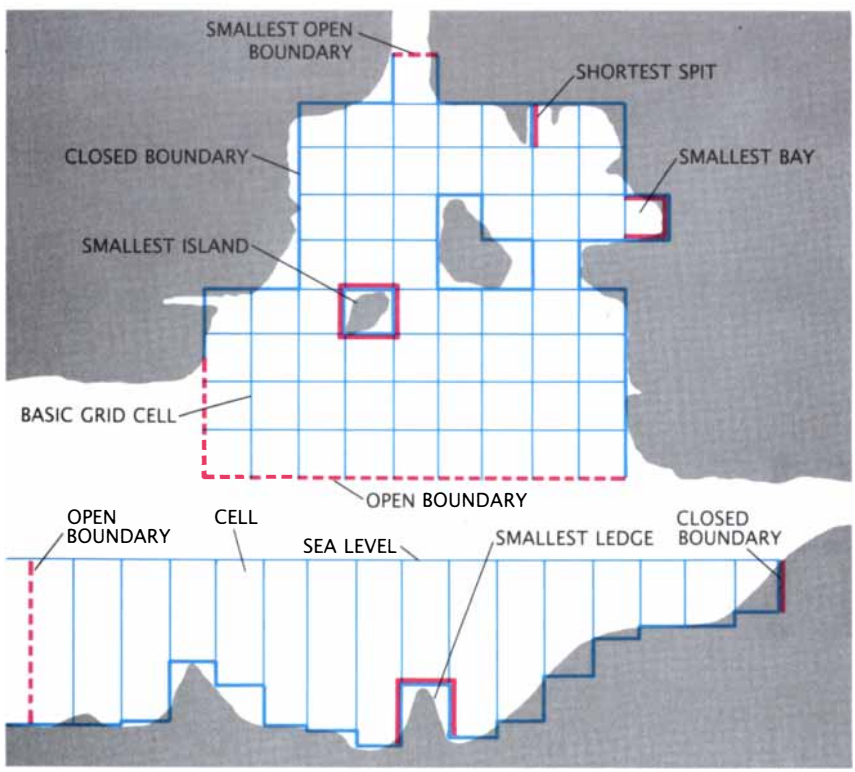
☐ **Yes,** I'd like to try the SCIENTIFIC AMERICAN *Medicine* system. Please enter my subscription at a first-year price of US$245*.
☐ Also enroll me in the CME program. I'd prefer:
☐ the printed version at **no additional charge.**
☐ the floppy-disk version at an additional $US97.*
Computer type or compatible:
☐ IBM® PC/PC Jr. (128K RAM, DOS 2.0/2.1)
☐ Macintosh™ (512K RAM)
☐ Apple® IIe (64K RAM)
Apple® II+ (64K RAM) 80-col. card by: ☐ Apple® ☐ Videx®
☐ Enroll me in the floppy-disk CME **only** at US$152.*
(Note model above.)
☐ Check enclosed* ☐ Bill me
Signature _____
☐ VISA ☐ MasterCard Exp. Date _____
Account No. _____
Name _____
Specialty _____
Address _____
City _____ State _____ Zip Code _____

*Add sales tax for Calif., Ill., Mich., or N.Y. Allow 8 weeks for delivery. Add US$10 for shipping to Canada. IBM is a registered trademark of International Business Machines Corporation. Apple is a registered trademark of Apple Computer, Inc. Videx is a registered trademark of Videx, Inc. Macintosh™ is a trademark of McIntosh Laboratory, Inc. and is used by Apple Computer, Inc. with express permission.

7M



CONSERVATION OF MASS dictates that if more water flows into a grid cell than flows out, the surface rises (*a*). A difference in surface level between cells results in acceleration of water from the high to the low side (*b*). The force of the earth's rotation pushes the water in a direction perpendicular to flow (*c*). Bottom friction opposes the flow (*d*).



DIMENSIONS of a single grid cell determine the size of the smallest island, bay, spit of land, open boundary or subsurface ledge that can be resolved by the tidal model.

ence between the low- and high-water lines changes land boundaries and currents dramatically. Without good resolution in these areas it would be difficult to reproduce the existing tidal regime accurately and our predictions would be on shaky ground.

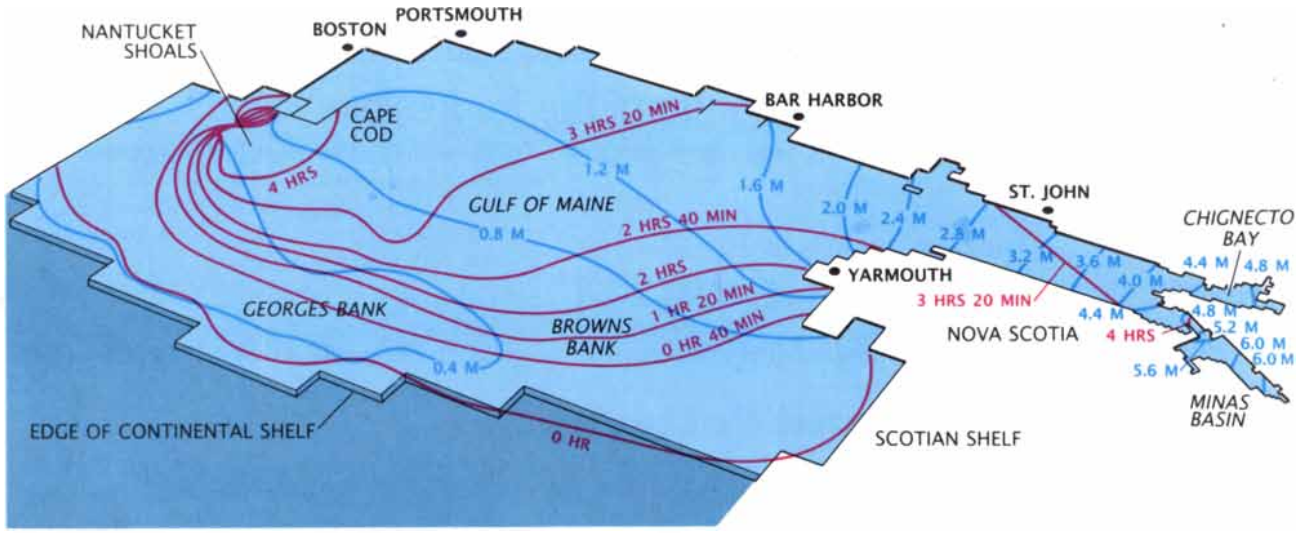
In order to complete the specification of the model, we need to provide information about what takes place at its boundaries. The shoreline is a "closed boundary": no water can flow across the side of a grid cell that represents the shore. Large areas at the head of the Bay of Fundy become dry at low tide, and so cells are removed from the computations as the water level falls and are added when the level rises to represent a moving land boundary.

At the grid cells on the open-sea boundaries, for example at the edge of the continental shelf, we simulate the movement of water into and out of the model. The deep-sea tides drive the water over the continental shelf into the Gulf of Maine, providing the conditions that set the model in motion. The conditions can be specified either as the tidal elevation or the tidal flow (the elevation is easier to measure than the flow).

In applying our model to see how a dam would affect tides we must ensure that the open-sea boundary is far enough away so that tides there would not be significantly changed by introducing the dam. The edge of the continental shelf is one logical boundary, since the shallow shelf area has only marginal influence on deep-sea tides. The remaining sea boundaries in this study are one shelf width away from the Gulf of Maine, far enough so that effects from the upper bay would be negligible. (The remainder of the model is bounded by land, which is simulated by the closed-boundary condition in which no current can pass through a cell.)

The choice of the sea boundaries introduces the greatest uncertainty in the model, because actually no part of the ocean is far enough away to experience absolutely no effects from the dam. Christopher Garrett and his co-workers at Dalhousie University and I used a more analytical model of the entire North Atlantic to approximate the effects on far-off areas outside the Gulf of Maine. Using these figures, we determined the limits of uncertainty in the tidal model and concluded that its predictions of the change in tide elevations should be accurate to within 25 percent.

Before using our model to make



TIDAL MODEL accurately reproduced the observed height and timing of the lunar tide. The blue lines show the predicted height of the high tide above the mean sea level. The red lines time the progress of the tide from its arrival at the edge of the continental shelf. The bay-and-gulf system would resonate if its length corresponded to a quarter of a full wave cycle (*see illustration on page 128C*). The tide takes approximately from 3.3 to four hours to travel from near the edge of the continental shelf to Minas Basin—close enough to a quarter of the 13.3-hour period of oscillation of the bay and gulf to account for the large tide amplitudes.

predictions we had to calibrate it to ensure that it reproduced the existing tidal regime accurately. Deep-sea tide gauges provided measurements of the tides at the open-sea boundaries of the area covered by the model. Using these data, we ran the model and obtained predictions for tidal elevations and currents, which were then compared with actual observations at more than 70 sites. Successive adjustments and refinements brought the model to its present form, which agrees well with the observed tides.

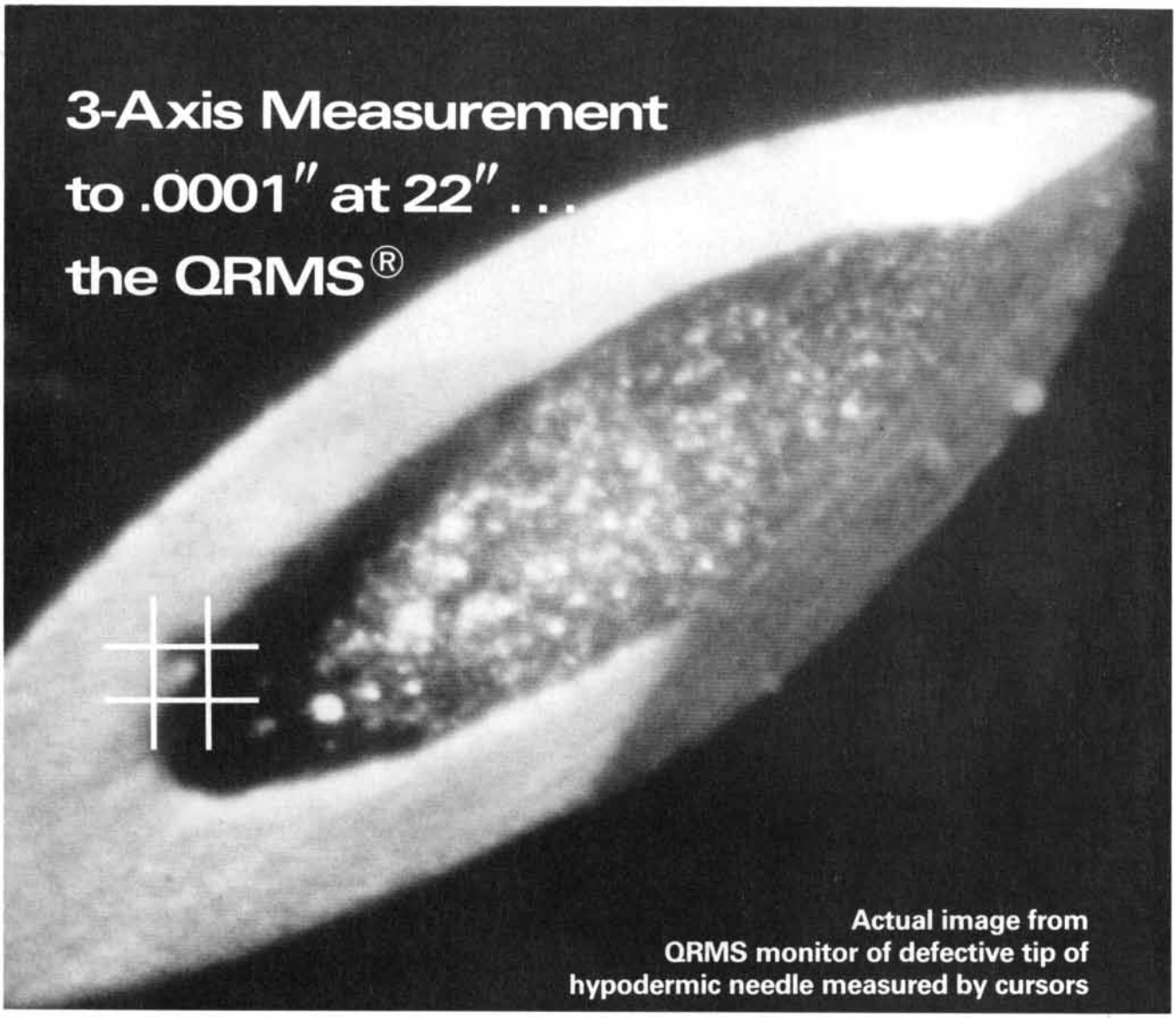
A principal objective of the calibration process was to determine the proper coefficients of friction. These account approximately for bottom roughness and nontidal background currents, both of which create frictional forces that diminish the height of the tidal wave and delay its arrival time. The frictional characteristics were checked by comparison with the long-term variation in tides. The semidiurnal lunar tide varies in the deep sea by plus or minus 3.7 percent over 18.6 years owing to variations in the declination of the moon's orbit. Observations indicate that in the Bay of Fundy and Gulf of Maine friction reduces the modulation to plus or minus 2.4 percent. The model's response to a 3.7 percent increase in input tide at the open-sea boundary was very close to the observed value of 2.4 percent. With the model accurately reproducing the effects of friction, we could go on to predict changes in tidal regime with greater confidence.

Before presenting our predictions I need to describe the proposed tidal power plants whose effect we want to determine. The largest scheme proposed for the upper Bay of Fundy would have a maximum capacity of about 5,000 megawatts and an average actual production of perhaps 1,500 megawatts—much more than could be absorbed by the electrical grids in Atlantic Canada. The output could, however, be integrated into the large northeastern U.S. grid, amounting to an immediate saving of some 23 million barrels of oil every year.

Existing and proposed tidal power developments in Canada, France, South Korea and the U.S.S.R. involve dams that trap the high-tide water in a reservoir and then release it through turbines. Special turbines are needed because the head (the difference in water level across the dam) available for tidal power generation is low compared with that of a large hydroelectric plant. The turbines would be large machines, with diameters of seven meters or more, that turn relatively slowly (at about 60 revolutions per minute).

For the project proposed for the upper Bay of Fundy, 100 or more such turbines would be installed in a dam spanning eight kilometers. The reservoir behind the dam would be filled through sluices by the rising tide. There would be a short wait after high tide for the tide outside the barrier to fall enough to create the minimum head required for power generation, and then the turbines would be engaged for about seven hours—until the drop in the reservoir reduced the head to an unusable level, when the turbines would be closed. As the tide again rose above the reservoir level, the sluices would open and the cycle would begin again. This one-way mode of generation, employed at the present test plant in Annapolis Royal, is called ebb generation. Two-way generation is also possible and would yield more power, but the greater complexity of design would make the power more expensive.

To simulate the presence of a tidal power barrier, we needed a special set of rules for the flow of water through the grid cells at the barrier. For example, the difference in water level across the barrier determines whether sluices or turbines are operating and affects the direction of the flow of water. The velocity calculation at the dam cannot conserve momentum, because energy is being removed from the water. Instead there is a hydraulic relation in which the flow through the barrier is proportional to the square root of the head across the barrier. The number of sluices and turbines being simulated

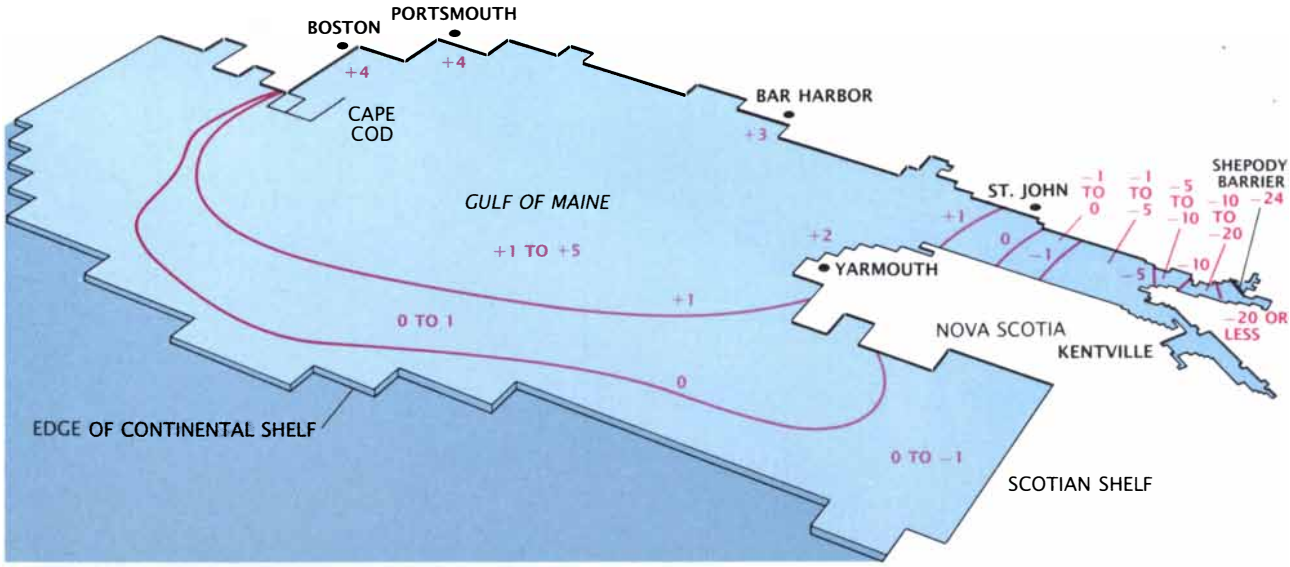


Questar Corporation announces its Remote Measurement System — the QRMS — an optical AT-based, keyboard-controlled and menu-driven gauging, alignment and imaging system. Non-contact digital imaging calculates linear and angular measurement in 3 axes to .0001″ and does it with on-line monitoring in a high-speed production environment without disturbing the dynamic process. It has full image and data archival capabilities, will generate gray-scale profiles and 3-D maps, calibrates without reference, can include numeric and statistical processing of data, and is portable. The QRMS, which incorporates the award-winning Questar long-range microscope, is a powerful tool for research and production. You no longer need to stop the process, remove the part, inspect the part, replace the part. You simply

MEASURE IT . . . CONTROL IT . . . DOCUMENT IT

QUESTAR

P.O. Box 59, Dept. 3 (215) 862-5277



SHEPODY BAY barrier would diminish tide amplitudes at the head of the Bay of Fundy while increasing it in the Gulf of Maine. (The map gives the change in centimeters.) The barrier would shorten the oscillation periods of both the bay (nine hours) and the combined bay and gulf (13.3 hours). The bay's period would move further from the 12.4-hour lunar period, and so tides there would fall. Conversely, the period of the bay and gulf together would move closer to the lunar period, causing an increase.

along the side of each cell also determines the rate at which water flows in and out of the cell.

I shall now describe the results of simulations carried out for two of the proposed power plants being considered most seriously. The first simulation was of a small barrier across Shepody Bay. This dam would have an installed generation capacity of 1,600 megawatts. At the barrier itself, the model predicts, the amplitude of the tidal sea-level elevation would decrease by 24 centimeters. (Amplitude is measured from the mean water level and is equal to half the tidal range; in this case high water would be 24 centimeters lower and low water would be 24 centimeters higher.) There would be a decrease throughout the head of the bay, with the effect diminishing and turning into an increase in tidal amplitude in the central part of the bay near St. John. The coast of the Gulf of Maine from Bar Harbor to Cape Cod would experience an increase in tidal amplitude ranging from three to four centimeters.

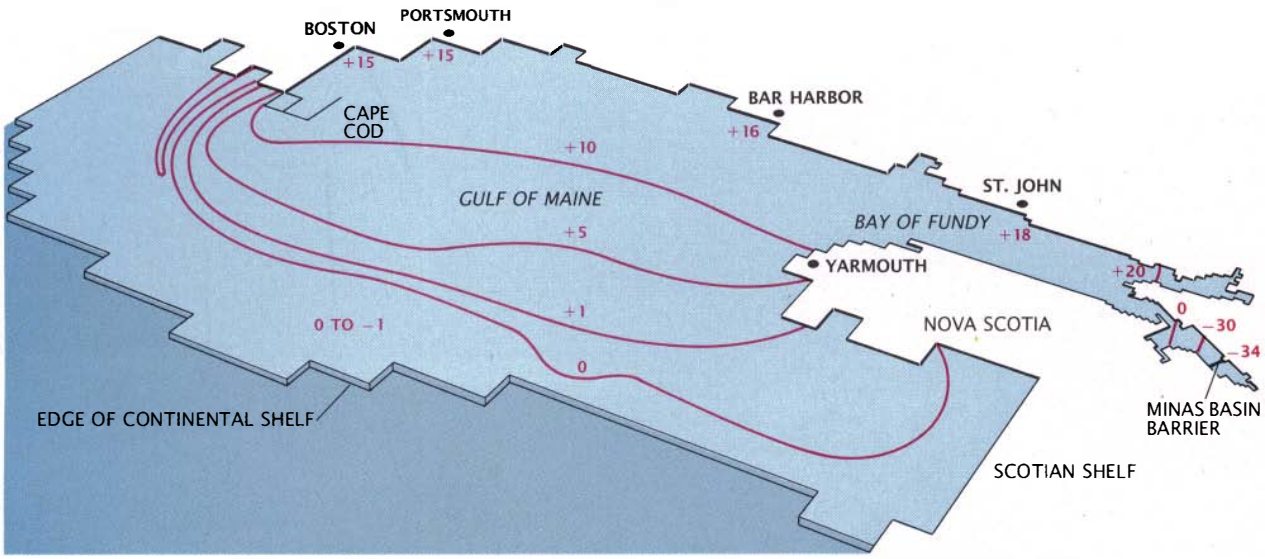
The second simulation was of the largest project now being considered, a barrier eight kilometers long across part of Minas Basin with an installed capacity of about 5,000 megawatts. The model predicts that the installation would lead to a 34-centimeter decrease in tidal amplitude at the barrier, with the effect dropping off rapidly as one moves out toward the bay. Throughout the greater part of the bay and the Gulf of Maine there would be an increase in tidal amplitude ranging from more than 20 centimeters in the upper part of Chignecto Bay to 15 centimeters in Massachusetts Bay. (In both simulations these effects fall off toward the model's open-sea boundaries.)

The nature of the changes can be explained qualitatively in terms of the resonance thesis that I described above. Putting in a barrier shortens the Bay of Fundy, moving the bay's nine-hour period further away from the tidal period of 12.4 hours and causing a decrease in the tidal amplification from the mouth to the head of the bay. In contrast, the reduction of the 13.3-hour resonant period of the entire bay-and-gulf system by about 10 minutes brings it closer to the tidal period and thereby causes increased tides.

In the case of the Shepody Bay project the decreased amplification in the Bay of Fundy dominates at the head of the bay, whereas the increased amplification for the bay-and-gulf dominates in the Gulf of Maine. In the central part of the Bay of Fundy the two effects cancel. The Minas Basin simulation, on the other hand, is complicated by the effects of friction. Normally a large amount of energy is dissipated by friction in the narrows off Cape Split. Because the dam decreases flow through this channel, the tide moves in more slowly and less energy is dissipated, so that just inside the cape the water would carry more energy and cause the tides to rise higher. The combination of reduced dissipation and changing resonance pattern gives rise to increased tidal amplitudes everywhere except near the barrier.

The changes in the tides that are predicted by our model would have effects all along the coast of the Bay of Fundy and the Gulf of Maine. Would these effects be significant? The changes appear to be minor compared with natural variations in local sea levels. At Boston, for example, the high tide can vary from less than one meter to more than two meters above mean sea level in the course of the year. Local winds commonly whip up surface waves with an amplitude of one meter or more and large storm systems have created surges of as much as two meters. Measurement of the earth's crustal movement indicates that the Boston area is sinking (with respect to mean sea level) at a rate of from 15 to 30 centimeters per century. The greenhouse effect could, according to the best predictions, cause a general rise in sea level of one meter over the next 100 years.

Compared with these variations, one might think that the maximum 15-centimeter increase in tidal amplitude we calculate for Massachusetts Bay would not be noticeable, but in areas that are already sensitive to sea-level variation such an increase



MINAS BASIN barrier would have a markedly different effect. Tide amplitudes would fall near the dam and rise by 10 to 20 centimeters in the Gulf of Maine. Unlike the Shepody Bay dam, it would not cause a drop in the area around Cape Split, even though it would cause a similar change in resonance characteristics. Instead the barrier would slow the current, causing less energy to be dissipated in the narrows off the cape. The water retains more energy and rises, offsetting the resonance effects.

could indeed be significant. For the sea to rise to flood levels there must usually be some combination of high tide, large waves and a storm surge. An additional small increase in the level of high-tide might precipitate floods under conditions that otherwise would not cause problems. Hence one might expect an increase in the number and severity of floods, and areas previously thought safe could find themselves within new floodplain limits. Many wharfs and shore structures have been built close to the present high-tide level; some of the older ones are already strained by the rising sea level, and they would be subjected to further stress if there were even a slight increase in high tide.

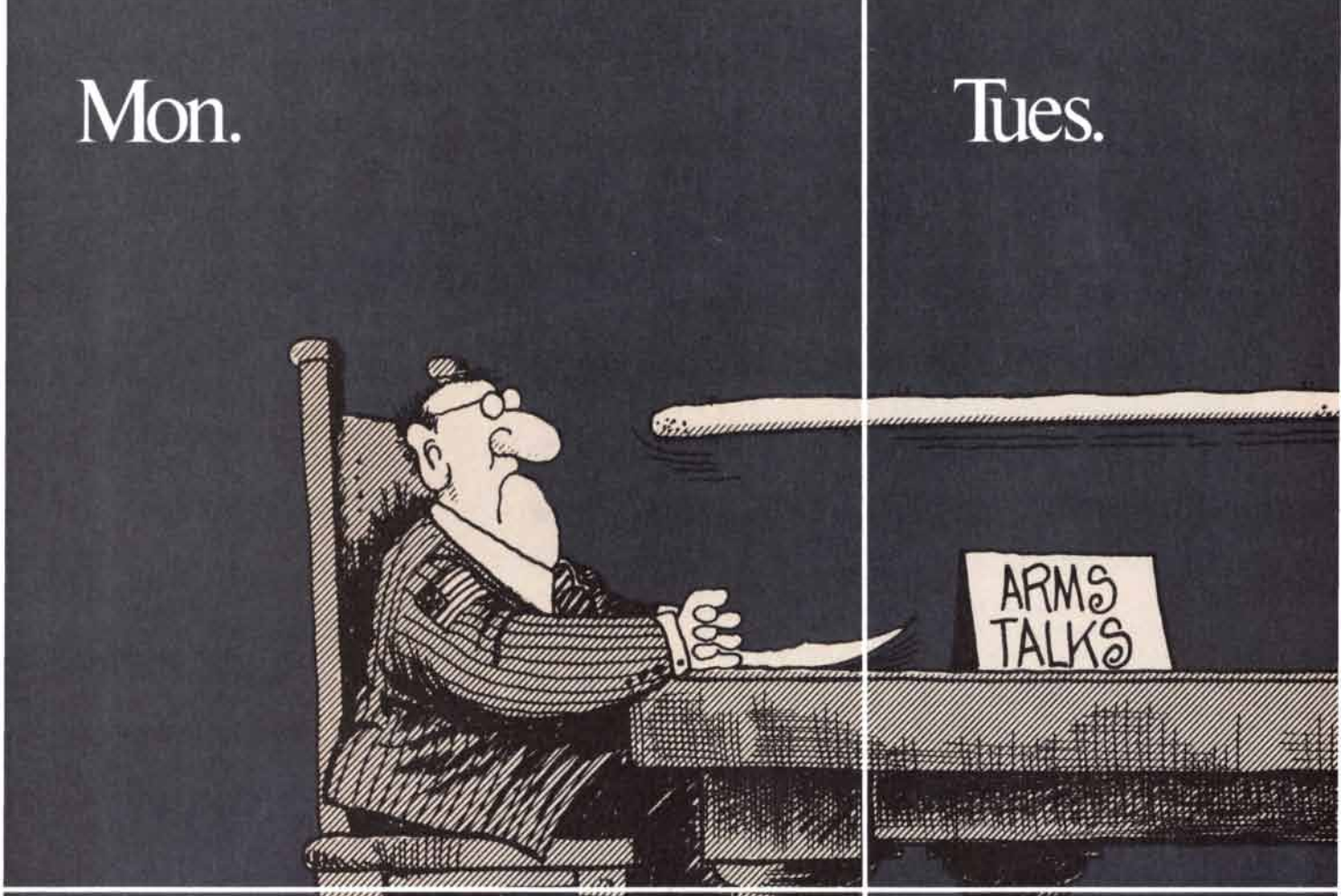
Moreover, tidal currents would be altered by the changes in tide levels. These currents play a major role in mixing the water in the Bay of Fundy and the Gulf of Maine. In summer the surface water is warmer and less salty than the bottom water, and therefore lighter. Strong tidal currents churn the water, cooling the surface layers and warming the bottom. A change in the surface temperature could affect the recreational use of some areas and even modify local weather characteristics, such as the formation of fog. Swifter tidal currents could also worsen shore erosion along the coast, which is already a major problem in certain areas along the Bay of Fundy and the Gulf of Maine. In addition, longer intervals of stronger tidal currents would hamper marine navigation in many areas where it is already complicated.

A tidal power dam could also have widespread ecological consequences. Intertidal zones (those that are alternately flooded and exposed by the tide) support complex ecosystems. Many species are tied to particular parts of the zones and to specific bottom material. An ecosystem may or may not be able to adapt to changes in the zone's tidal limits, depending on whether or not the change in tide levels alters its bottom characteristics. If, for example, the higher-tide region were to be extended past a sandy bottom to a higher rocky area, an existing salt marsh might be transformed into a rock-and-weed ecological system. Worse yet, if people were to build dikes against the flooding brought on by higher tides, the marsh might be destroyed altogether.

The waters around and behind the tidal power dam would be most severely affected. Such a barrier would cross the migration path of some fish species. It is not yet known whether the migrating fish could be redirected to fish passages to avoid the turbines. The tidal currents in the upper Bay of Fundy are strong enough to mix the water completely from top to bottom. The surface water is kept cool throughout the summer. Slower tidal currents in the head pond could allow the water to stratify, with a warmer layer on top and a cooler layer deeper in the basin. The stratification would change the ecosystem's dynamics, since the warmer surface layers would probably carry less sediment and would support organisms different from those found in the present cool and muddy water environment.

In a world that is struggling to cope with acid rain, the greenhouse effect, political instability among oil-producing nations and the disposal of nuclear waste, tidal power offers an attractive source of renewable, nonpolluting energy. The total harnessable tidal power energy in the world can supply only a fraction of global energy demands, but as conventional power sources are exhausted, or as their economic and environmental costs become too high, tidal power can make a contribution. There are bound to be environmental impacts of tidal power—as there are of any energy source currently available. The effects of tidal power may in the end prove to be less costly than those of other sources.

Our work has shown that tidal models are now accurate enough to predict the effects of power dams on the behavior of tides. It is up to environmental engineers and policy makers to capitalize on this tool by making an effort to consider the full ramifications of tidal power installations and their effects and deciding whether the price justifies making a commitment to tidal power.



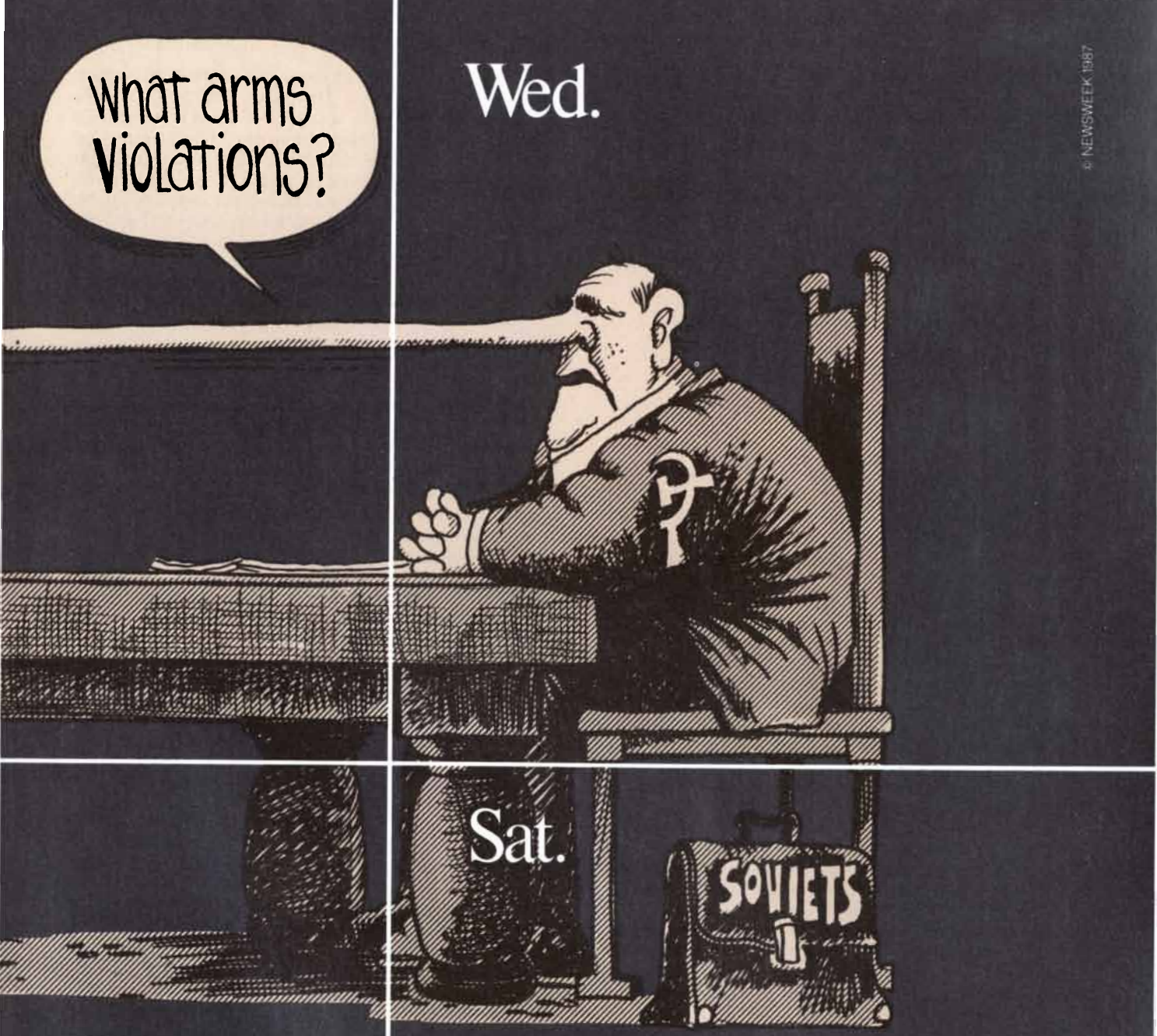
The week a lighter perspective gave you a clearer view.

Every week there's a page in Newsweek that takes a lighter look at some heavy subjects.

It's a page called Perspectives. A gathering of cartoons and comments on the world around us and on some of the people and events that we might otherwise take too seriously.

Sometimes witty, often biting, always revealing, Perspectives brings our readers statements overheard in hallways or on the airwaves, material spotted in newspapers across the country and from other countries, lunacies and wisdom in speeches and interviews that—quite simply—help keep things in perspective.

It's this variety of opinion and viewpoint that makes Newsweek different from



other newsweeklies. Like our economics specialist, Robert J. Samuelson. And our consumer affairs expert, Jane Bryant Quinn. Or columnists Meg Greenfield and George Will. They're another reason Newsweek readers know that every week in Newsweek they'll find the news and more.

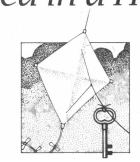
They'll find a sense of perspective that has won Newsweek more awards than any other news magazine. And that keeps so many more readers coming back for more.

Newsweek

Why it happened. What it means.

THE AMATEUR SCIENTIST

Fluid interfaces, including fractal flows, can be studied in a Hele-Shaw cell



by Jearl Walker

Monitoring the flow of a fluid in three dimensions is difficult because the motion can be quite complex and one's view of the flow's interior is often obscured. Just before the turn of the century the English engineer Henry S. Hele-Shaw devised a way to simplify the flow by confining it to a thin layer. A Hele-Shaw cell consists of two parallel transparent plates separated by a narrow gap. Spacers between the plates keep the fluid in the cell. With this apparatus one can view and photograph a fluid that is either trapped in the cell or made to flow through it from end to end.

For half a century the device was merely a convenient way of observing fluid flow. Then, in about 1956, Sir Geoffrey Taylor, a British physicist renowned for his work in the physics of fluids, recognized that the cell could serve to model a flow problem important to the oil industry. In some oil fields oil is trapped in a porous material and hence cannot be pumped out directly, as a pool of oil can. One solution is to pump water through the material and so drive the oil to form a pool, from which it can be recovered. Taylor saw the problem on a grander scale. How do two fluids behave when one of them (such as water) with a low viscosity drives another (such as oil) that has a greater viscosity? In his analysis Taylor found that the interface between the two fluids can become unstable, generating wavelike shapes.

When the unstable nature of the interface was verified, flows within a Hele-Shaw cell drew fresh attention. Checking various fluids, investigators discovered that the unstable interfaces give rise to many surprising shapes. Research done in the past two years has shown that under certain circumstances an interface develops a fractal appearance. The term fractal refers to a type of symmetry in which a pattern found at one magnification is present at other magnifications as well; such a pattern is said to be self-similar. Although the patterns seen in Hele-Shaw cells are fractal only for a limited range of magnifications, they have generated much interest in the study of unstable fluid interfaces.

I set out to reproduce some of the research by constructing a simple Hele-Shaw cell. Although the cell is often made from plate glass, I chose acrylic plates to make it easier to drill holes in the walls of the cell. The square plates, measuring 15 inches on a side and three-quarters of an inch in thickness, were bought from a local plastic supply house for about $40. Thinner (and less expensive) plates would serve, but a thickness of less than half an inch would probably allow the plates to buckle when they are mounted in the cell.

In some experiments I wanted to inject a fluid into the center of the cell. To accommodate the needle of a syringe I drilled a hole through the acrylic plate that was to be the top of the cell. The hole was an eighth of an inch in diameter. To facilitate the injection I also bent the tip of one of the needles so that it would slip slightly under the top plate.

I experimented with several types of spacers. The best spacers proved to be the narrow, thin strips of foam rubber that are sold for weatherproofing windows. I chose strips that were a quarter of an inch wide. After cutting the strips to fit the perimeter of the bottom plate, I stuck them in place with their sticky side down.

At one end of the cell I cut several narrow openings in the foam-rubber strip so that fluid could escape if the cell was full when additional fluid had to be added during an experiment. The plates were held together with six or more small C-clamps. To prepare for an experiment I removed the clamps and the top plate and then poured a liquid onto the bottom plate. If I wanted the entire cell to be full during the experiment, I poured enough liquid so that about half of the bottom plate was covered. When I added the top plate and squeezed the plates together with the clamps, the liquid would fill the cell, with any extra leaking through the holes in the spacer.

To avoid bubbles I usually poured the liquid onto the half of the bottom plate nearest to me and then laid the top plate down carefully, beginning at my side of the cell. Bubbles were nonetheless sometimes still present when the clamps were tightened. To remove them I stuffed a paper towel into the central hole and stood the cell on end. After the bubbles had migrated to the top of the cell, most of them popped when they reached the openings in the spacer.

When the clamps were fully tightened, the plates were separated by about a millimeter. The spacing is crucial to some of the experiments because the size of the patterns developed by the interface depends on it. Some of the patterns also depend on the left-to-right width of the cell: they may not appear if the cell is too narrow. This can present a problem in homemade cells, because you want the cell to be wide enough to promote the patterns but not so wide that the plates buckle.

To photograph the patterns I positioned a flood lamp to shine upward between two tables of equal height. Across the gap between the tables I taped a sheet of white drawing paper to diffuse the light and to protect the lamp from any spilled liquid. After filling a cell I placed it above the paper so that it rested, on its clamps, on the tables.

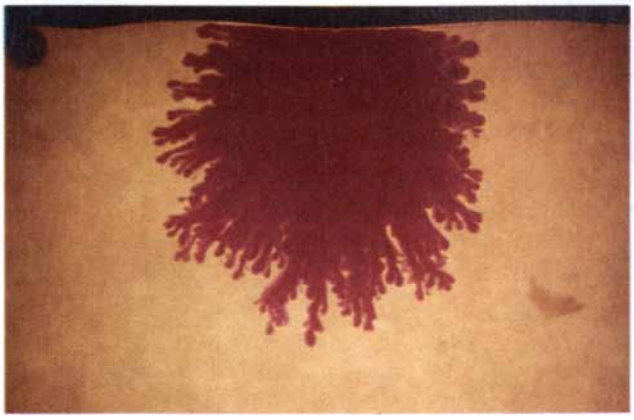
I first investigated experiments that were reported in 1958 by Taylor and Philip G. Saffman of the University of Cambridge. Imagine a Hele-Shaw cell filled on one side with a fluid of low viscosity and on the other side with a fluid of high viscosity; there is a straight interface between them. If the low-viscosity fluid is forced against the high-viscosity one so that the interface moves slowly, the interface remains straight. If the speed exceeds a critical value, however, the interface grows unstable, breaking up into wavelike shapes that become more pronounced with

Chocolate syrup draining in a Hele-Shaw cell

A pattern in glycerin with a loosened clamp

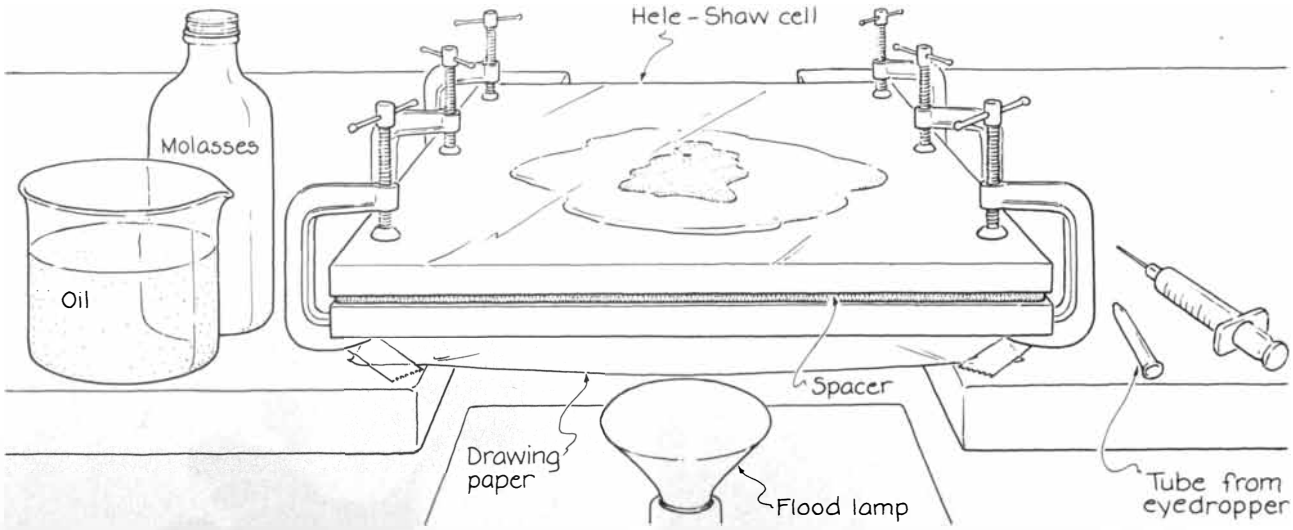
Air blown into dark corn syrup

Dyed water injected into condensed milk



Cellulose at a high concentration

Dyed water injected into gelatin



An arrangement for viewing a Hele-Shaw cell

time. The interface soon resembles a glove: fingers of the low-viscosity fluid extend into the high-viscosity fluid. The instability and the growth of fingers stem from the difference between the viscosities of the two fluids. Interfacial tension (or surface tension, if one of the fluids is air) along the interface opposes the instability, because it attempts to reduce the surface area of the interface. Still, if the interface moves fast enough, the fingers form and grow.

Saffman demonstrated the appearance of fingers with a vertical cell in which air overlay glycerin. The air served as the low-viscosity fluid. When additional air was forced into the top of the cell, glycerin was forced out of the bottom. The speed of the air-glycerin interface was high enough so that the interface developed waves, which grew into fingers of air stretching down into the glycerin. Surprisingly, one finger soon dominated the pattern, with the other fingers quickly becoming frozen in length. The width of the dominant finger eventually was half the width of the cell.

I could replicate Saffman's experiment only crudely. To do so I filled part of my Hele-Shaw cell with corn oil and then stood it on end to drain the oil to the bottom. When the oil had settled, I laid the cell down on the table and partially opened the middle clamp on the side nearest me. The resulting sudden expansion of the space between the plates forced the oil to flow toward me, breaking up the oil-air interface into a series of short fingers. I also produced fingers of air in another experiment. With the oil placed in the cell so that it surrounded the central hole, I inserted the tip of an eye-dropper tube into the hole and drew oil rapidly out of the cell into the tube. When the perimeter of the pool of oil raced toward the hole, fingers of air formed.

Earlier this year Tony Maxworthy of the University of Southern California reported a related experiment. He partially filled a cell with silicone oil, rotating the cell to wet the entire interior. Then he stood the cell on end. When the oil had drained to the bottom and the oil-air interface was straight, he turned the cell over so that the oil was above the air. In some trials the final orientation of the cell was vertical; in others the cell was at various angles to the horizontal.

According to Taylor's analysis, when the more viscous fluid is driven into the less viscous one, the interface is unstable only if it moves at a speed lower than a certain critical value. For each angle Maxworthy tried, the interface should move fast enough to be stable, and yet in each trial fingers of air reached up into the descending oil. Apparently when the cell was initially rotated into position for a trial, the interface began moving at a low speed, allowing the fingers to form. Once the speed exceeded the critical value, the stabilization of the interface tended to promote the growth of a few of the fingers while arresting the growth of the others. Eventually one finger dominated.

When the cell was tilted from the horizontal by only a few degrees, the spacing between the fingers was wide and the dominant finger developed only slowly. When the cell was vertical, the spacing between fingers was narrower and the dominant finger grew rapidly; in addition the fingers split into smaller fingers, a process known as bifurcation. The splitting is due to the high speed of the interface when the cell is vertical and to the narrowness of the gap between the plates.

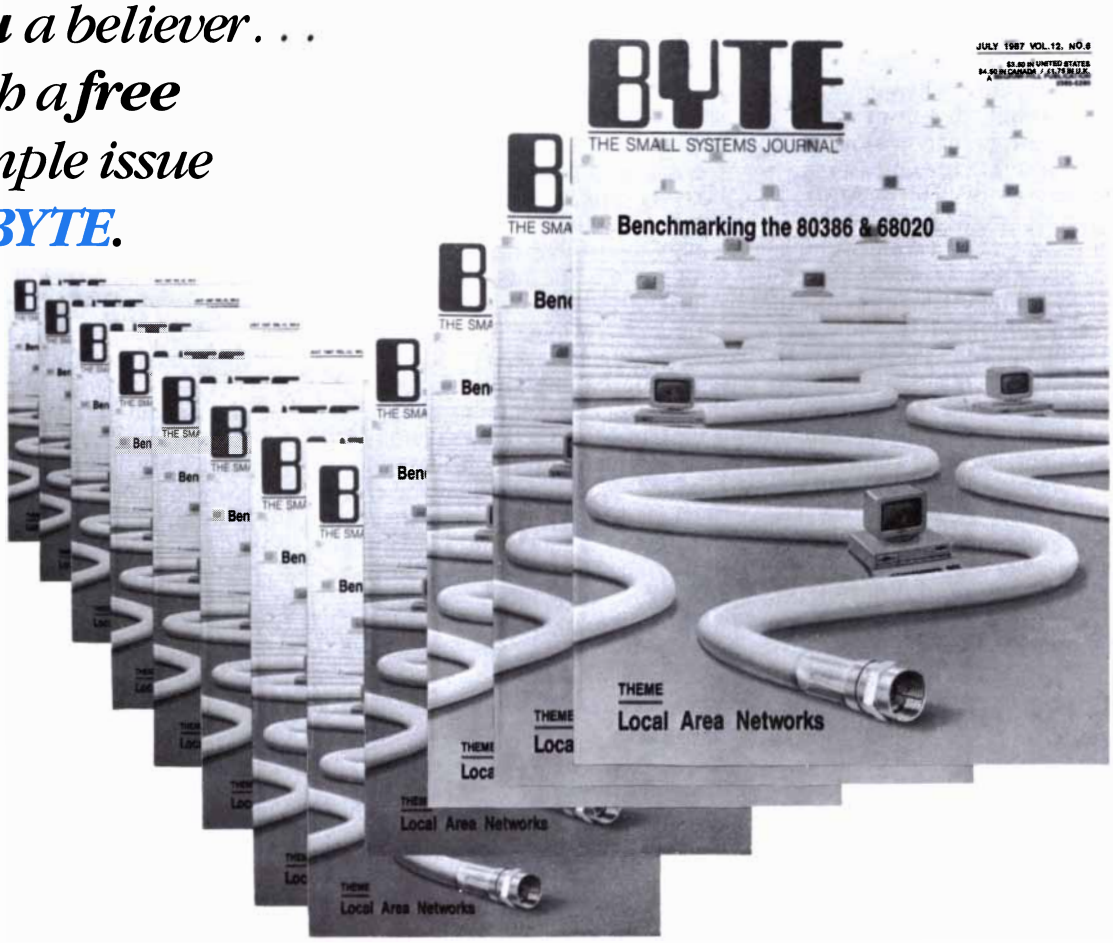
The experiment is easy to do with a homemade cell. After placing the cell on end to drain corn oil to the bottom of the cell (the central hole was plugged), I inverted it. Waves and then fingers formed quickly. When the fingers of air began to bifurcate, I laid the cell flat and photographed the fingers.

I repeated the experiment with other liquids. One liquid was Slime, a gel sold in toy stores. Slime is a member of a peculiar class of liquids that is designated non-Newtonian. The viscosity of such liquids changes when the liquid is sheared or stressed; in Slime's case the viscosity increases. To the Slime I added an equal volume of water; I stirred the mixture and partially filled the cell with it. When I inverted the cell, the interface quickly developed fingers of air that were interspaced with narrow fingers of the mixture. The head of each liquid finger was a blob wider than the rest of the finger. Thick molasses yielded similar liquid fingers, except that the blobs were more pronounced and the fingers tended to pinch off after a while, releasing the blobs.

In another series of experiments I injected dyed water into various liquids in the cell. The liquids included glycerin, corn oil and molasses. The dye was methyl violet, which I mixed

SEEING IS BELIEVING!

And we want to make ***you*** *a believer . . .*
with a ***free***
sample issue
of ***BYTE****.*



BYTE is a magazine that sells itself. But don't just take our word for it. Let us send you a free get-acquainted issue. Just one look at BYTE and you'll see why BYTE is recognized and respected everywhere as the number one small-systems journal.

BYTE provides you with how-to information on innovative applications you can quickly put to use on your micro. And every issue brings you straight-forward appraisals of the latest developments in hardware and software products.

You'll be able to evaluate the latest microprocessors, memory and storage equipment, modems, power components and other peripherals. You'll learn all about the new languages, applications programs, programming tools, operating systems and other software . . . to help you increase efficiency and expand the uses and applications of your equipment/software in the office or in your home.

Here's how BYTE's no-strings, free trial offer works:

Just complete and mail the postage-paid card adjacent to this page. We'll send you a copy of BYTE's big, thick current issue, along with our bill for an introductory subscription—11 more monthly issues plus our annual IBM issue.

If you like what you read in your sample copy, send us your payment. If not, return the bill marked "CANCEL" and that ends it . . . with no questions asked. You keep the trial issue.

So, take advantage of this quick, easy, no-risk way to get acquainted with BYTE. Mail the card today. And see for yourself why BYTE is the world's leading microcomputer magazine. We want to make a believer out of you.

BYTE

with the water. I filled the syringe with the dyed water, taking care not to draw air into the instrument. Then I injected the dyed water into the central hole (using the bent needle) or through the foam spacer along the perimeter (using a straight needle).

When the dyed water is injected centrally, it spreads quickly to form a beautiful pattern with lobes that bifurcate repeatedly; when the water is injected through the side of the cell, it forms a pattern that is similar but more like a fern. Lincoln Paterson of Geomechanics in Victoria, Australia, has recently studied the structures that form when dyed water is injected into glycerin. He found that the size of the smallest features on the lobes is approximately four times the spacing between the plates of the cell. The dependence is due to the nature of the waves that develop on the interface between the water and the glycerin. The most unstable wave, which dominates the production of the lobes, has a wavelength that is about four times the spacing between the plates. Water and glycerin are miscible, but their diffusion into each other is sufficiently slow so that the patterns form and then last long enough to be photographed.

My best patterns appeared when I stood the cell on one end and injected the dyed water through the foam spacer at the other end. After the water spreads, some of it falls through the glycerin, undergoing numerous bifurcations.

Paterson also studied what happens when air is injected into glycerin from a central hole. In this case the fluids (air and glycerin) are immiscible. According to theory, an interface between immiscible fluids is unstable only if the wavelengths of the waves at the interface exceed a lower limit. Initially the air bubble is a circle whose circumference is too small to support any wave with a wavelength greater than the lower limit. Only when the bubble grows larger does the interface become unstable. Then it blossoms into a flowerlike pattern with bifurcating lobes. I was able to produce such patterns by inserting the eye-dropper tube into the central hole of my cell and blowing air through it. The resulting patterns are difficult to photograph in glycerin, which is clear; they are more pronounced in colored molasses or various dark syrups.

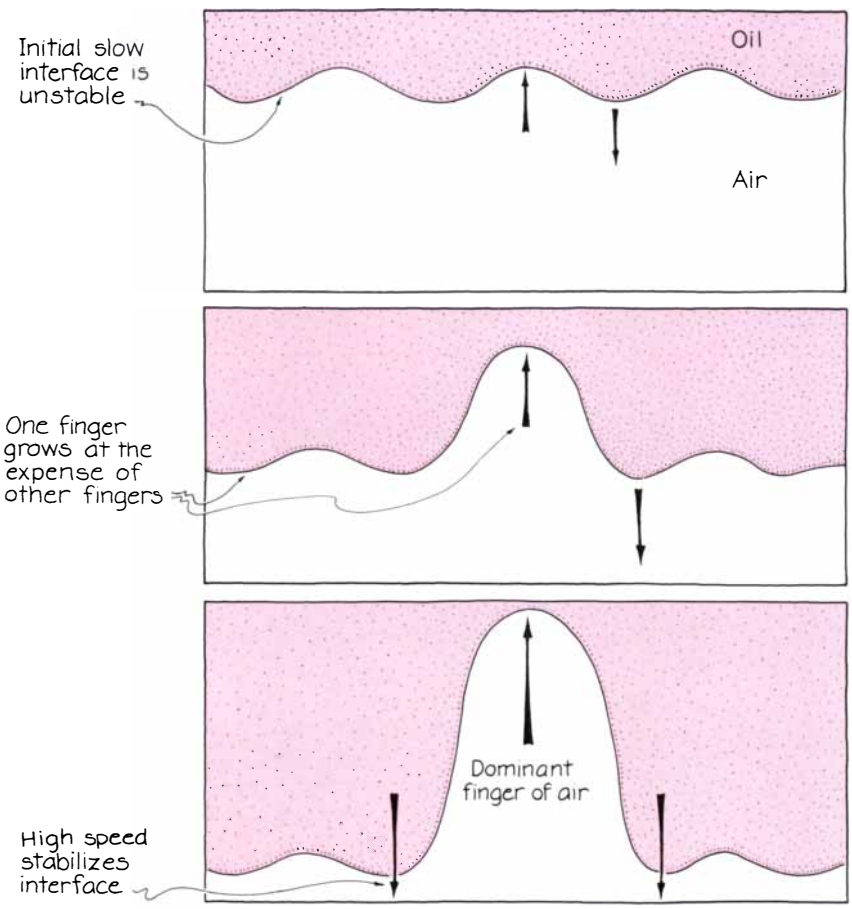
In 1985 Johann Nittmann and Gérard Daccord of Dowell Schlumberger in France and H. Eugene Stanley of Boston University reported that they had produced fractal-like fingers by injecting dyed water into an aqueous polymer solution. The fingers resembled the limbs of a bare tree, with short, narrow stems that jutted from main stems. The polymer solution was a polysaccharide with a high molecular mass. Since the solution was aqueous, there was essentially no interfacial tension between the solution and the injected water. Without that stabilizing factor the interface underwent abrupt changes in direction, giving rise to the fractal pattern.

Polysaccharides include cellulose, glycogen, chitin and starch. Attempting to duplicate the fractal fingers, I investigated a solution of water and common cornstarch. The solution is non-Newtonian, its viscosity increasing under stress. When the solution was thick, I was not able to inject the dyed water into it. I was more successful with moderate cornstarch concentrations: the dyed water created fernlike patterns that had the same thick branches and wide bifurcations I had seen in previous experiments. As I diluted the cornstarch solution more, the patterns became less interesting, until finally there was a simple ellipse. Various other starches yielded similar results.

I fared better with a mixture of water and cellulose. (The cellulose is available from biological-supply houses.) When the concentration of cellulose is moderate, the injected dyed water forms fingers with abrupt turns—something like the patterns Nittmann, Daccord and Stanley photographed but not as narrow.

Casting about my kitchen for other materials, I decided to test a common gelatin dessert (Jell-O brand). After it gelled, I spooned a small amount into the cell and gradually clamped the plates together, allowing the gelatin to ooze between them. Then I injected dyed water through one edge. After a momentary hesitation, the water burst through the gelatin, creating an elegant fractal pattern.

I have only touched on the possibilities of the research that can be carried out with a Hele-Shaw cell. You may want to try your hand with other liquids. Perhaps you will find other common substances that yield fractal patterns. Since the field is far from having been fully explored, you may find entirely new patterns.



The growth of fingers in descending oil

The September/October AUDIO UPDATE™ presents . . .

- ***CHOLESTEROL AND HEART DISEASE***

Claude Lenfant, M.D., Director, National Heart, Lung, and Blood Institute

New evidence of decrease in risk after cholesterol reduction; dietary and pharmacologic treatment of hypercholesterolemia; developments in apolipoprotein gene research.

- ***CANCER THERAPY AND MULTIDRUG RESISTANCE***

Michael Gottesman, M.D., Chief, Molecular Cell Genetics Section, Laboratory of Molecular Biology, National Cancer Institute

Characterization of the gene conferring multidrug resistance; alternative strategies to standard chemotherapy.

- ***RETROVIRUSES AND HUMAN DISEASE***

Myron Essex, D.V.M., Ph.D., Chairman, Department of Cancer Biology, Harvard University School of Public Health

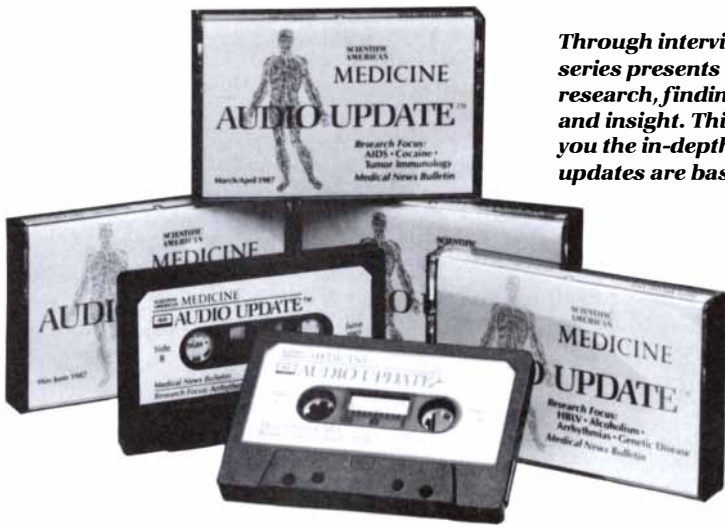
The HIVs, LAVs, HTLVs, and simian viruses: history and epidemiology; AIDS: speculation about insect vectors, cofactors, vaccines, and future developments.

- ***ANTIMICROBIAL CHEMOTHERAPY***

Harvey B. Simon, M.D., Assistant Professor of Medicine, Harvard Medical School; Infectious Disease Unit, Massachusetts General Hospital

Developments in the treatment of microbial diseases: monobactams, cephalosporins, quinolones, and other agents.

- ***INFLUENZA OUTLOOK: 1987–1988***



Through interviews with leading authorities, the Audio Update™ cassette series presents the latest developments in medicine. They'll discuss their research, findings, and conclusions, sharing with you their knowledge and insight. This unique resource goes beyond the written text to give you the in-depth background on which the SCIENTIFIC AMERICAN* Medicine *updates are based.

A unique information program . . .

Use the Audio Update™ tapes to enrich your basic knowledge. Journals, books, and seminars keep you up to date on the facts, but Audio Update™ gives you an extra benefit: the personal insights of these medical newsmakers. Their invaluable commentaries present the human implications as well as the scientific aspects of illness—the "art of medicine."

You can order single tapes or subscribe . . .

The yearly Audio Update™ series consists of six tapes. A new tape will be available every two months. Timeliness and importance will determine the content of the tapes—our goal is always to bring you the latest developments by the scientists making them.

Order your Audio Update™ tapes individually or by subscription. Each tape costs only US$9.95. Or, save money with a yearly subscription of US$50—you'll receive six bi-monthly tapes at a discount from the single-tape cost. (For prices outside the U.S. and Canada, see the coupon below.)

Send for your Audio Update™ tape today—and start listening to the voices of medicine.

SCIENTIFIC AMERICAN MEDICINE AUDIO UPDATE™

A6

415 Madison Avenue, New York, NY 10017

YES, I want to hear the latest medical news—AUDIO UPDATE™.

I'm ordering:

☐ **Single tape:** send me the audiocassette containing the September/October program—I've enclosed my check for US$9.95* (single orders require prepayment by check).

☐ **Subscription:** start my yearly subscription with the September/October program. I'll receive 6 bimonthly tapes for US$50* (a discount from the single-issue price).

☐ Check enclosed* ☐ VISA ☐ MasterCard

Exp. date _____ Account No. _____

Name _____

Address _____

_____

_____ Zip Code _____

Specialty _____

*Prices apply to U.S. and Canada. Please add sales tax for DC, IL, MA, MI, or NY. Allow two weeks for your order to be processed.

Prices outside U.S. and Canada: Single tape—US$15; Subscription—US$75. in U.S. dollars.

COMPUTER RECREATIONS

Beauty and profundity: the Mandelbrot set and a flock of its cousins called Julia

by A. K. Dewdney

The Mandelbrot set, since its appearance in these pages in August, 1985, has emerged as the newest and brightest star in the firmament of popular mathematics. It is both beautiful and profound. Indeed, the beauty of the Mandelbrot set is only a veil over its significance: the casual observer sees a miniature riot of filaments and curlicues near the set's boundary without suspecting that such patterns encode the various forms of chaos and order [*see upper illustration on page 142*].

In the 1985 foray into the Mandelbrot set much was left unsaid. The set has an important connection with stability and chaos in dynamical systems. The connection is made by way of closely related sets that are called Julia sets, after the French mathematician Gaston Julia. One Julia set corresponds to each point inside (or outside) the Mandelbrot set. Julia sets have an intrinsic fractal beauty of their own [*see lower illustration on page 142*]. Before turning to them I had better review the set named after Benoit B. Mandelbrot, a scientist at the IBM Corporation's Thomas J. Watson Research Center in Yorktown Heights, N.Y.

The Mandelbrot set inhabits the complex plane, an ordinary plane with some numbers attached to it. To be more precise, each point on the complex plane is represented by a number of the form $a + bi$. It does no harm to imagine that a and b are coordinates of the point. The a coordinate is called the real part of the number $a + bi$ and the b coordinate is called the imaginary part. The i acts partly as a marker to remind the reader which part is which. Complex numbers can be added by summing the coordinates separately. The result is a new complex number. Complex numbers are multiplied as they are in high school algebra:

$$\begin{array}{r} 3 + 7i \\ \times\ 2 - 4i \\ \hline 6 + 14i \\ \quad - 12i - 28i^2 \\ \hline 6 + 2i - 28i^2 \end{array}$$

In order for the result to be a complex number, the term $28i^2$ must be reduced by using the most important property of imaginary numbers, namely $i^2 = -1$. Thus $6 + 2i - 28i^2$ becomes $34 + 2i$. The key formula can now be presented. It opens the door to the Mandelbrot set, draws forth the Julia sets and, in a peculiar way, brings order to chaos:

$$z \leftarrow z^2 + c$$

Here z and c are complex numbers, each composed of a real and an imaginary part. The z is squared and the c is added by means of complex multiplication and addition respectively. The formula springs to life when it is iterated: computed repeatedly, with the previous value of z being used to get the next one. The resulting succession of complex numbers amounts to a strange little jig on the complex plane. At each iteration of the formula the newest complex number z lies at some distance from its predecessor. Distance is critical in computing the Mandelbrot set.

I like to think of the succession of complex numbers (points on the complex plane) produced by the formula as the wanderings of the initial point. Does it long for infinity, long to dance ever outward on the complex plane? Some complex numbers enjoy that fate. Others are forever confined within a certain area having a complicated shape. They might be called prisoners. Their prison, the area of confinement, has fractal walls.

In the foregoing description I have implied that the iterative process is repeated over and over again. But how is one to select c and the initial value of z? One answer is to make z always zero and to select different values of c. Will the prisoner escape? The experiment is repeated again and again, with c varying systematically over a portion of the complex plane. If the prisoner escapes, color c white; otherwise color c black. The prison walls take on the shape of the Mandelbrot set. If instead of coloring the escapees white one gives them a color that varies with the speed of escape, even more beautiful images emerge.

Under the rule just described, z started with the complex value 0, namely $0 + 0i$. What would happen if some other fixed starting value were adopted, say $z = 3.5 + 6i$? Would one get a different shape for the resulting set? Actually the result is always a deformed version of the Mandelbrot set. One prefers the canonical object.

Under the opposite rule, when c is fixed and z plays the role of initial point, the resulting set looks rather different from the Mandelbrot set. It—or rather its boundary—is called a Julia set. I would have liked to say "the" Julia set, but the sets are legion: for each fixed value of c one uses in the iteration formula, a new and different Julia set appears, filled in with prisoners.

My inspiration for this return visit to the Mandelbrot set lies in *The Beauty of Fractals,* a large-format book by Heinz-Otto Peitgen and Peter H. Richter of the University of Bremen [see "Bibliography," page 150]. With its stunning images in black and white and in color, it is both a mathematics textbook and a coffee-table adornment. The lore of the Mandelbrot set, its associated Julia sets and other complex systems is crystallized in theorems and occasionally spelled out in readable paragraphs.

Let me describe the effect of one theorem. A reader who writes a program to display Julia sets might notice that for some values of c the sets are obviously connected, or of a piece, but that for other values of c the sets are not connected. What makes the difference? The answer is both simple and charming: if the point c is chosen from inside the Mandelbrot set, the corresponding Julia

set is connected. If, on the other hand, c is selected from outside the Mandelbrot set, the Julia set for c is not connected.

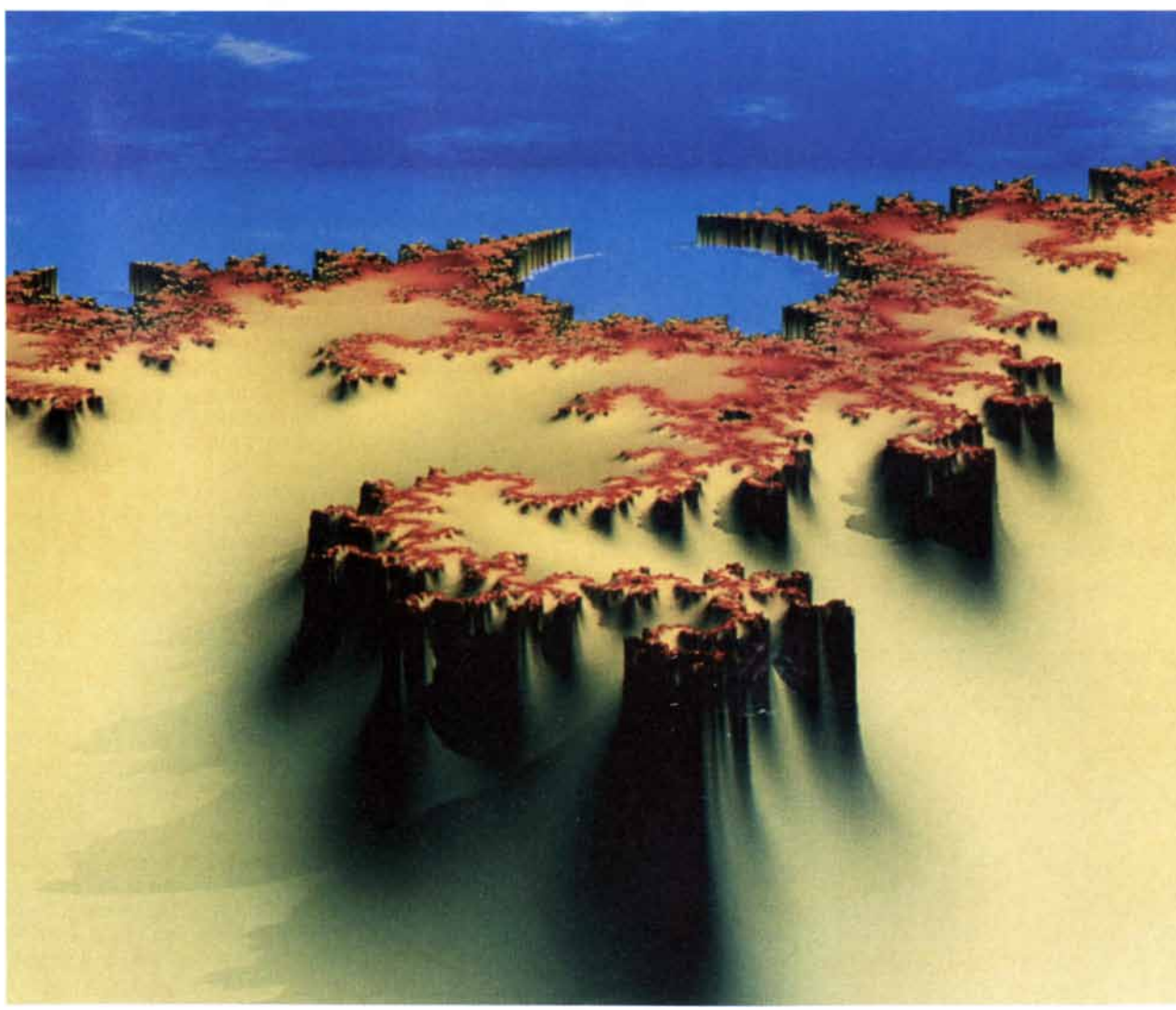
One could make a fascinating motion picture illustrating the applicable theorem. Draw a straight line L from an arbitrary point inside the Mandelbrot set to another point outside it and imagine that a point c moves slowly and steadily along L, inside the Mandelbrot set and toward its boundary. The associated Julia set becomes increasingly pinched and crinkled in appearance until, when c reaches the boundary of the Mandelbrot set, the Julia set shrinks to a fragile, dendritic skeleton enclosing no area whatever. As c passes beyond the boundary, the corresponding Julia set explodes into fractal dust.

Readers willing and able to write a program can explore the Mandelbrot set and the Julia sets by embedding certain basic algorithms in the language of their choice. The algorithms share the core iterative process, which depends heavily on a certain theorem: if the size of the so-called iterate z ever reaches 2, it is destined to leak into infinity, never to return. This distinguishes escapees from prisoners, for the most part. The algorithm allows 100 iterations for z to reach 2. Because a relatively small number of would-be escapees will not reach magnitude 2 by 100 iterations, the discriminant is not 100 percent accurate. One could, of course, allow 1,000 iterations for a slightly more accurate image, but this becomes time-consuming even on fast computers.

The magnitude of a complex number $a + bi$ is simply the square root of $a^2 + b^2$, in other words, its distance from complex 0. The core algorithm follows:

```
n ← 0
while n < 100 and mag(z) < 2
    z ← z² + c
    n ← n + 1
color the current point
```

Here an index variable n starts life at 0. Within a while-loop that controls the iteration process, n increases by 1 at each iteration. The while-loop continues to turn the crank of the basic formula as long as n has not reached 100 and the magnitude of z has not reached 2. If either condition fails, the algorithm exits from the loop. The manner of coloring the current picture point is left up to the reader. The color assigned will of course depend in some simple way on n, the slowness with which z es-



The Mandelbrot set rendered as a lake; the mountains represent the dynamics of the surrounding points

caped or failed to escape. The reader must also bear in mind that the picture point is a pair of screen coordinates that will differ from the coordinates of the complex number being plotted there.

The reader's program must contain a separate computation of the magnitude of z, represented by $mag(z)$ above. Indeed, because most programming languages have no provision for complex numbers, z must be retained in two-part form, say x (the real part) and y (the imaginary part); c must receive the same treatment, perhaps as a and b. The following would therefore be closer to a working program:

```
n ← 0
while n < 100 and x² + y² < 4
    xx ← x² − y² + a
    y ← 2xy + b
    x ← xx
    n ← n + 1
color the current point
```

Astute readers will have noticed the small trick introduced in this version of the basic process: rather than comparing the square root of the quantity $x^2 + y^2$ with 2, the algorithm compares the quantity itself with 4. The result is the same, and one can avoid continually invoking the square-root function, which can be time-consuming. The variable xx temporarily holds the newly computed x value while a new y value is being computed. The old x value is thus saved for the latter computation before it is replaced by xx.

The program I called MANDELZOOM in my first foray into the subject two years ago is now reintroduced in slightly more detailed form. It wraps the basic algorithm inside a loop that systematically varies the complex number c rather than its parts, a and b. If the picture is to be 100 by 100 pixels (grid points) in size, for example, there must be a double loop:

```
gap ← side/100
a ← acorner
for j ← 1 to 100
    a ← a + gap
    b ← bcorner
    for k ← 1 to 100
        b ← b + gap
        x ← 0
        y ← 0
        [basic algorithm]
```

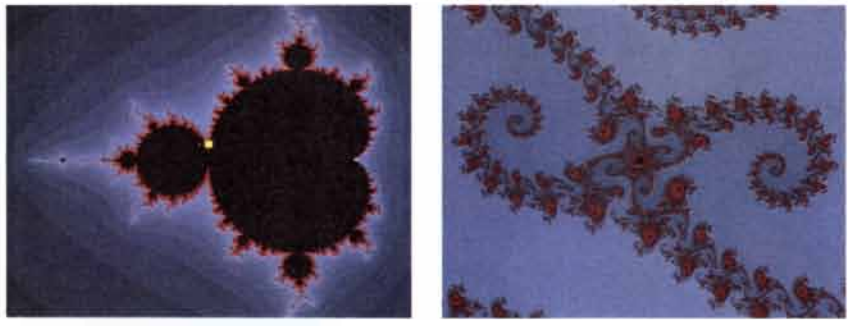
Before reaching these instructions MANDELZOOM allows the user of the program to input the complex number that sits in one corner of the square of interest. It has the coordinates *acorner* and *bcorner*, the smallest values that a and b will assume in the square. This square, specified by the user of MANDELZOOM, gives the algorithm its name. It is a window through which we can peep. It can be made very, very small, in effect zooming us into the part of the set over which the window is placed. MANDELZOOM must also request the user to input a value for *side*, the width of the picture on the complex plane. The algorithm then computes the gap between successive complex numbers c, thereby giving proper increments to a and b.

The values of the indices j and k do not enter any of the computations within the double loop, and so one is free to change the indices to some more useful form. For example, instead of j and k running from 1 to 100, they could run through 100 successive screen coordinates each. When the basic algorithm has decided what color to assign to the iterate z, the color is plotted at coordinates (j,k).

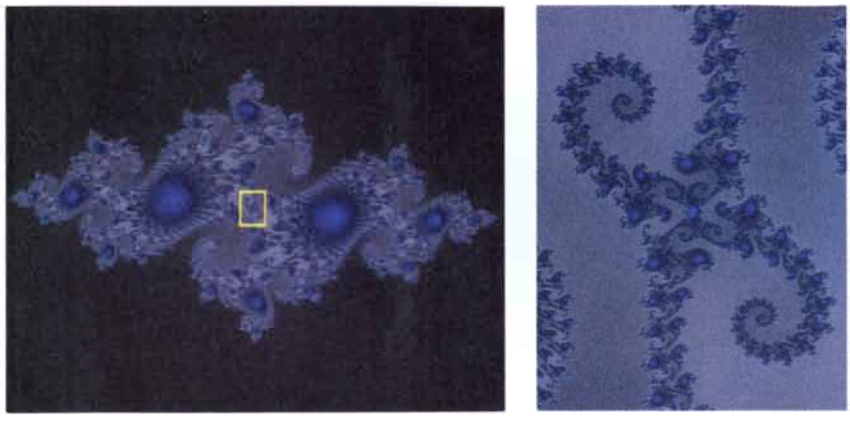
I cannot leave MANDELZOOM without mentioning a modification suggested by Peitgen. Instead of comparing the magnitude of the iterate z with 2, use 100 or even 1,000. Once the magnitude reaches 2, after all, it increases very quickly; it attains such values in just a few iterations. Still, different iterates pass such a threshold value at different speeds. The speeds themselves can be colored—and colored continuously! Red may evolve into orange, provided one has a reasonably sophisticated palette. In any event, it was by means of this technique that the color images accompanying this article were produced. Peitgen likens the speeds to values of an electrostatic field surrounding the Mandelbrot set. The "field" values are represented in the imaginary Mandelbrot landscape [*see illustration on preceding page*] as slopes of a mountain range surrounding what can only be called Lake Mandelbrot.

Although I am not enamored of the name, I feel compelled to follow earlier usage and call the program that generates images of Julia sets JULIAZOOM. Here also we can zoom in on a set to examine it at great magnifications. JULIAZOOM uses the same core algorithm as MANDELZOOM but wraps it in a somewhat different rind.

JULIAZOOM first asks the user for *xcorner*, *ycorner* and *side*. It also asks for a value of c given in terms of the



***The Mandelbrot set* (left), *a section of which* (square) *is magnified* (right)**



***Julia set* (left) *corresponding to center of picture at top right; magnified section* (right)**

variables *a* and *b*. It then employs the same double loop with some important differences:

```
gap ← side/100
x ← xcorner
for j ← 1 to 100
    x ← x + gap
    y ← ycorner
    for k ← 1 to 100
        y ← y + gap
        [basic algorithm]
```

The basic algorithm colors screen points according to the slowness with which iterates reach (or fail to reach) the magic threshold of magnitude 2. Some of the most effective graphics result from the simplest assignments. For color monitors even three colors assigned according to the following kind of scheme can be striking: assign the first color to points with a slowness value (*n*) of from 0 to 10, the second color to points with a slowness of from 11 to 20, the third color to a slowness of from 21 to 30—and then go back to the first color for the next decade, and so on. Monochrome monitors can get black-and-white (or green-and-yellow) effects by using the two colors and alternating by decades.

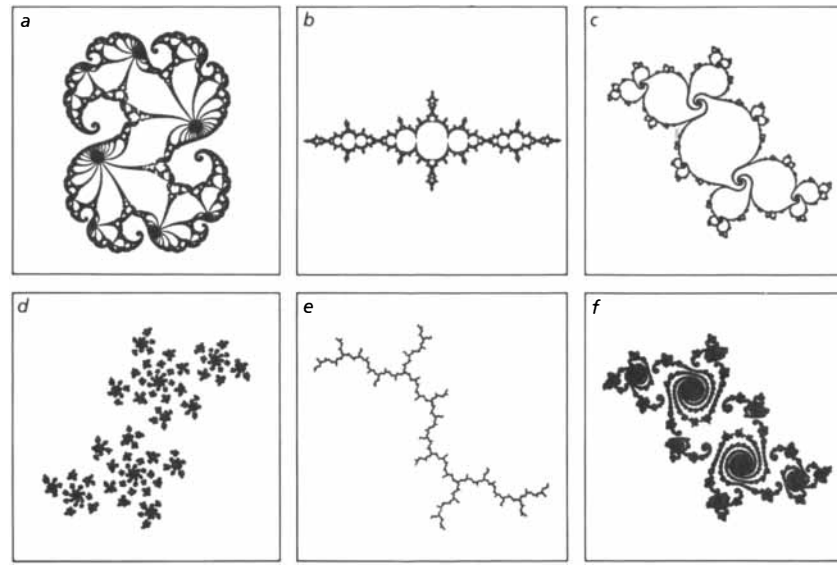
Equipped with a working version of MANDELZOOM or JULIAZOOM (perhaps both), readers will be able to explore these magnificent and meaningful fractal sets on their own. One can roam the complex plane in the vicinity of the sets or zoom in on specific parts with the computer microscope provided above. Up to the limits of resolution set by the arithmetic precision of one's machine, either set reveals amazing detail. For the guidance of voyagers into the infinitesimal world of fractals, the following coordinate ranges enclose both kinds of set on all four sides:

Julia sets: x and y from -1.8 to $+1.8$;
Mandelbrot set: x from -2.25 to $+.75$,
y from -1.8 to $+1.5$

This past July I explored the subject of chaos in dynamical systems such as pendulums and electronic circuits. A simple iteration formula involving not complex numbers but real numbers summarized the activity of such systems:

$$x \leftarrow rx(1 - x)$$

The formula is evidently quadratic: multiplied out, it contains a squared term. Depending on how the parameter value r is chosen, the formula behaves simply or strangely when it is iterated. For each r value, the iterates settle down into an orbit, a set of values that x systematically visits. At a critical value in the neighborhood of 3.5699 the iterates oscillate wildly and more or less unpredictably among a host of values. This corresponds to the situation where an underlying system, whether double pendulum or electronic circuit, becomes completely stymied in its search for stability. It wanders wildly in a manner that cannot be predicted in advance. Chaos.



Six Julia sets, some connected **(a, b, c** ***and*** **e)** ***and some not*** **(d** ***and*** **f)**

A similar phenomenon occurs with the complex iteration formula described here, $z \leftarrow z^2 + c$. For a given value of c, however, there is more than one attractive orbit, depending on how the initial value of z is chosen. If z has a relatively small initial magnitude, it will gravitate to a specific point. If the value is large, it will get larger without limit: infinity is the attractor. The specific point and infinity itself constitute two separate one-point attractive orbits for points on the complex plane. The boundary between their domains of attraction, the Julia set itself, is unbelievably crinkled and crumpled. It is also an orbit, but it is not attractive in the technical sense. Points already in the boundary hop about chaotically within it. The Julia set is not easy to compute directly, however, because a computer's numerical precision may not allow one to specify points that are exactly on the boundary to begin with; under iteration the precision degrades and the iterate wanders off into the night.

Each possible value of c results, as I have indicated above, in a new and different Julia set. In a sense the Mandelbrot set summarizes, in one fell swoop, all possible Julia sets. It describes the fate of iterates of complex 0 for all possible values of c. For some Julia sets the chaotic region is a mere treelike figure or even a symmetrical sprinkling of pepper. Readers will remember that such Julia sets correspond to c values on or beyond the boundary of the Mandelbrot set.

I met Peitgen recently at a conference in Asilomar, Calif. As we strolled along the beach between talks he described the Mandelbrot set as a kind of vast book in which each Julia set is a mere page. From the position of c in the Mandelbrot set one can predict the generic behavior of iterates in terms of the overall shape and size of the associated Julia set. There is more to it all than mere connectedness. For example, if c is chosen from the neck between the main body of the Mandelbrot set and one of its buds, the corresponding Julia set turns out to be pinched into necks and buds of its own. The analogy of the Mandelbrot set as a kind of dictionary for the Julia sets implies a fundamental distinction between the two. The Julia set is self-similar but the Mandelbrot set (even its boundary) is not. If it were self-similar, Peitgen argues, it would not have the capacity to encode its uncountable infinity of relatives called Julia.

There is much more of all this in *The Beauty of Fractals*. I am most grateful to Peitgen for providing the images illustrating this article. Not all

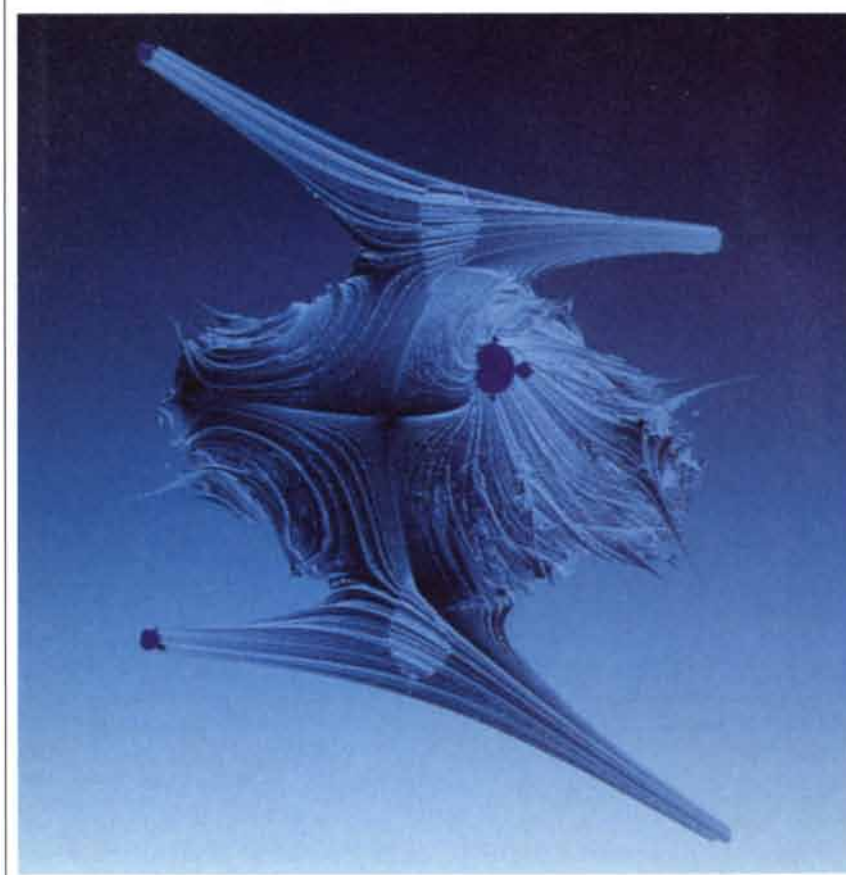
of them were produced in the ways I have described, but there are recipes in the book.

As a final note about current research by investigators of dynamic systems, I shall mention one more object that lurks in higher-dimensional space: a monstrous object resulting from cubic iterations, with z^3 rather than z^2 in the Mandelbrot formula. The object is four-dimensional, with tendrils curling off in inconceivable directions. Yet three-dimensional cross sections of it can be computed, such as the one below.

The column on algopuzzles in June continues to draw mail from around the world. Last month's reprise on the subject is continued this month as promised. Before plunging into the desert to reveal answers to the last two puzzles posed in the June column, I must correct myself on the subject of trains: Manuel Blum, a computer scientist at the University of California at Berkeley, has pointed out that the amount of work required for one train to pass another by means of the algorithm given is really proportional to n^3 (not n^2), where n is the number of cars in the train. Roughly speaking, each of n cars is moved n units of length n times. Blum has found an algorithm that does the same job with an amount of work proportional to $n^2 \times \log n$. Unfortunately there is no space here for Blum's algorithm.

The first desert-fox problem concerned a truck that could carry one 50-gallon drum of fuel at a time as well as the 10 gallons of fuel in the truck's tank. I showed last month that from a depot of two drums the truck could travel a maximum distance of 733⅓ miles before running out of fuel.

Algorithms submitted by Chester Nowogorski of Naples, Fla., and Norman Rokke of Wintersville, Ohio, indicate that a truck can use three drums of fuel to travel the grand total of 860 miles before running out of gas. Most readers attempting the feat came up short of this figure. Indeed, general formulas supplied by most readers also fell short when 3 was substituted for n, the total number of barrels employed. I cannot, therefore, vouch for accuracy in formulas such as the one sent by Lawrence Leinweber of Cleveland, Ohio. It was, however, typical of formulas



A three-dimensional cross section of the four-dimensional Mandelbrot set

submitted that yielded the greatest distances:

$$5 \sum_{i=1}^{n} \frac{100}{2i-1} - \frac{100}{2n-1} + 100$$

The Greek sigma (Σ) symbolizes summation: form n terms with successive values from 1 to n substituted for i in $100/(2i-1)$. Add all n terms together and then multiply by 5.

The second desert-fox puzzle allows a patrol car to refuel from n fuel depots placed at arbitrary points around a circular route to be followed by the car. The amount of fuel in each depot is also arbitrary, except that the total amount thus deposited is exactly enough for the patrol car to finish its route—provided it does not start at the wrong depot. Its own fuel tank will hold any amount of fuel but is initially empty. Where must the patrol car begin?

A number of readers, including Arnold V. Loveridge of Long Beach, Calif., arrived at an ingenious visualization of the problem. Take an imaginary tour with the patrol car starting at any depot and heading in any direction. Keep a graph of the fuel in the car's tank and keep driving even if the car runs out of gas. In this event the graph spills into the negative zone. With each refill at a depot the graph zigs up and then begins another long, slow decline. Ultimately the car returns to its initial point. The driver must now examine the graph and select the depot where the car's on-board gas was minimum (before refueling). That is the depot at which the car must begin.

Chaos was probed by numerous people ready and willing to program the iteration formula supplied in the August column. In fact, a number of people, including Howard Mark of Suffern, N.Y., followed my suggestion of tracking the iterative process itself by weaving a mirrored web about a parabola. Here convergence of the process could be seen as the web wove itself into a stable shape. But in the chaotic regime it filled a portion of the screen with solid white, a confusion of squares that settled to no discernible pattern. Charles A. Plantz of West Brownsville, Pa., used his microcomputer as a microscope, zeroing in on the stirrup-shaped arc of the bifurcation diagram just above one of the chaotic regions. He found not the pepper and salt I would have expected but folds and strata invading the texture of chaos.

BOOKS

The prodigy wrestles the champ, the petroleum heresy, Stone Age cinema—on location at Lascaux

by Philip Morrison

RABI: SCIENTIST AND CITIZEN, by John S. Rigden. Basic Books, Inc., Publishers ($21.95).

"Life in America is not so difficult. There any young man can become an assistant professor." That was the prickly judgment of theorist Arnold Sommerfeld, whose renowned school in Munich turned out between the wars more physics Ph.D.'s than any other in Germany. We know the remark from one who himself came back from Munich and Hamburg and Leipzig, invited to Columbia University in 1929 as lecturer on the new quantum mechanics. That man was Isidor Isaac Rabi, 31. A celebrated young visitor, Werner Heisenberg, had impressed the Columbia chairman by his big buildup for Rabi, then at work in Leipzig. Rabi's own highly original but quite unnoticed experimental thesis at Columbia was three years behind him; the new Ph.D. had gone away to return with exactly what he had sought: "People of my generation went abroad...and learned not the subject, but the taste for it, the style, the quality, the tradition. We knew the libretto, but we had to learn the music."

In Europe, Rabi had found himself side by side with other lively, eager and brilliant American advanced students, including Linus Pauling, Edward Condon, Robert Oppenheimer and H. P. Robertson. Laboratory doors securely locked from seven to seven did not fit the American style of night work; Otto Stern, the director, generous, intuitive, full of insight, even approved breaches of the rule. Those Americans all came home to become young professors who knew the score. Under such leadership the new quantum physics, nurtured by the experimental talent and vigor then already at home in America, sang loudly. (True, there was other stimulation, not least of which was the diaspora of the German physicists themselves under the Nazi scourge.) Students of every generation since have been able to catch the music of physics most sweetly here on American shores, although within the past decade or so the sounds from overseas are once more attractive.

In a brief epilogue to this book, I. I. Rabi, reflecting on a long life in science, is cautionary, although he sees a ray of hope: "Physics is now a very mature science and may be losing its attraction to young Americans of eager minds.... In the twentieth century, physics begins to lose its hard, fructifying contact with the real world.... The wonderful phenomena and discoveries have been made hard to reach.... We can rely on nobody but ourselves to get out of this intellectual slump. We must start with our schools, the training of our teachers, and the restoration of our ideals of learning for its own sake."

Rabi's rich career is easily periodized. We meet a Brownsville kid in an Orthodox Jewish family, poor immigrants from Galicia living in three rooms behind their grocery store. Soon he could read all the Yiddish books in his home. Schoolbooks were dull. A classmate brought news of the Carnegie Library branch; the boy haunted it, reading the children's books in order. He came at once to a little book on astronomy. There he read with wonder the story of the Copernican system. The child had watched the sky (once a rising full moon neatly aligned with the street "scared the hell out of me," he recalled) and had long puzzled over the mystery and power of the days of Creation. His loving parents knew one answer only to his questions about dawn, dusk, the shifting form of the moon; all was done by God's untiring hand. That day he came back from the library, "full of delight," and announced to his mother and father: "It's all very simple, who needs God?"

He was a formidable kid. His room was soon filled with wireless gear; he built a noisy spark station, all his own work except for the headphones. A magazine paid him two dollars for a piece on the design, written while he was still in grade school. The boy easily outtalked his pious family: his bar mitzvah was held not in synagogue but at home. His commentary in Yiddish to the old men who attended told "How the Electric Light Works." A brilliant if quiet student, he stood first in the State of New York on the Regents exam in history, yet he was also streetwise, skeptical and deeply proud of his culture and his gifts. He completed shop-centered Manual Training High, where one "math teacher was marvelous," to win a scholarship, which with family help took him off to distant, rural, liberal, scientific Cornell. He was graduated in chemistry; "unusually brilliant," a professor wrote in recommendation. But his mind had never been caught there. He spent three or four years doing odd jobs, reading and debating with fellow clever drifters. Still warmed by the faith and love of his puzzled family, back he went to Cornell as an ill-prepared but determined graduate student in physics.

He stayed only a year, to return to Columbia, in part to follow one Helen Newmark, a Cornell summer student who had entered Hunter College. That year, however, had brought him an epiphany: "I was there and I had the feeling I was there." Physics was parsimonious, powerful, general; it could carry him where his colorful manganese solutions never went, into the heart of the atom, "nearer to God." "Somewhere way down, I'm an Orthodox Jew.... Not religion in a secular way, but religion as inspirer of a way of looking at things.... Choosing physics means, in some way, you're not going to choose trivialities.... When you're doing good physics, you're wrestling with the Champ."

For a decade results grew in the frugal molecular-beam laboratory he created in 1931 at Columbia, each beam a tiny pure world where atoms travel one by one along alternative paths through the vacuum guided by cleverly manipulated electromagnetic fields. Hydrogen became the stuff to study; it was the simplest of substances, simple enough so that every nuance the delicate techniques

found could be analyzed. Rabi himself was impatient, useless in the day-to-day work of the laboratory, even seen as hazardous to the equipment; his part was ideas, vision, explanations and inspired strategic choice of how to spend every hour and every dollar. He was the kind of boss who was present at every success, at whatever hour it came.

The magnetic-resonance physics, innately quantum-mechanical, that arose then in Rabi's laboratory has by now spread past physics, to chemistry, medical diagnosis, atomic clocks. Presented here is a branching tree of extensions, applications and insightful results in the field compiled paper by paper. The first was published before Stern's famous split-beam work (space quantization it was called in 1924); the last reported the hydrogen radio-line studies in cosmic gas in 1958. There were about 100 of them even at that time. The Nobel prize came to I. I. Rabi in 1944; he was then fully engaged by the war, both with radar and with the nascent bomb.

Rabi never returned full time to the profound yet simple life of physicist wrestling with the angel. The beam laboratory flourished. Now he was department chairman; the anti-Semitism that had disfigured even the Ivy League had not survived the war that had turned Rabi into an international figure. He was both research director of the famous radar lab at M.I.T. in Cambridge, and a valued adviser to the Los Alamos bomb laboratory headed by Rabi's lifetime friend, Robert Oppenheimer. (The contrast and the parallel between their characters and life stories make a fascinating topic for Rabi's reflections.) With Oppenheimer he helped to originate the hopeful scheme of international control for all atomic energy. He then conceived and argued for a great Federal laboratory that would of right be opened to physicists from all campuses, and there was Brookhaven. He envisioned and proposed at UNESCO a generalization of that idea: an all-European laboratory. CERN came into being two years later at Geneva: "We have just signed the Agreement which constitutes the official birth of the project you fathered.... Mother and Child are doing well, and the Doctors send you their greetings," the European negotiators wrote him once the 11 states had signed in 1952.

He submitted along with Fermi a report to President Truman, explicitly citing moral arguments against building the H-bomb. After *Sputnik,* whose impact his committee had foreseen, he persuaded Eisenhower—they got along well—to set up the President's Science Advisory Committee. It flourished under Ike; by the time of Nixon scientists critical of the war had assumed the appearance of adversaries rather than advisers. The PSAC never recovered its initial influence. For Rabi, a man of buoyant internal harmony, there could come no tragedy like Einstein's, a pacifist drawn into the most terrible of wars, nor Oppenheimer's, a man both lauded and persecuted for the same influential service.

The tragic element of Rabi's life is a potential one, a possible future that he shares with all of us, more keenly perhaps because more responsibly. "I don't mind dying. My ancestors did that. What I do mind is the destruction of civilization." Our country matured, kept free by the oceans from external military threat. "Every time a superior weapon is developed it turns us into a continental state.... So our policy should have been all along to get every arms reduction. Combat all arms as far as we could; this was our natural advantage. But after the Second World War we just became intoxicated." The potion may be wearing off; hardware advantage seems more fleeting now, except perhaps from the windows of the Oval Office.

The readable volume is not an autobiography, although it evokes one. Its physicist-author, John S. Rigden of the University of Missouri at St. Louis, has talked so intimately with Rabi and his wife Helen that Rigden has both cited and captured—and in his turn been enfolded in—the articulate, sharp, metaphorical tone of that wise old man. Background is sometimes filled in through less perceptive sources; critical examination of the man and the high events he recounts is not absent, but in its forbearance it too is more or less autobiographical. Read the book; both the questions and the answers will stay with you.

POWER FROM THE EARTH: DEEP EARTH GAS—ENERGY FOR THE FUTURE, by Thomas Gold. J. M. Dent & Sons, Ltd., Aldine House, 33 Wellbeck Street, London W1M 8LX (£12.95), to be published in the U.S. by Harper & Row Publishers, Inc.

Everyone knows of the two classical branches of chemistry, the inorganic and the organic. Once a great gulf parted them: the difference between the living and the dead. Step by step that gulf has been bridged, until today the organic chemists think of their specialty as simply that of the myriad compounds of the cosmically rather abundant light atom, carbon. No one is surprised any longer by our synthesis of the simpler substances of life. Nor is life the only natural source of such compounds. There is more alcohol—very dilute, to be sure—in a single interstellar cloud than in all our biosphere, and entire oceans of methane are found out among the giant planets and their moons. That sounds interesting, if academic; surely it bears on such profound questions as those of the origin and spread of life in the universe. In this tightly argued, quite nonmathematical book by Thomas Gold, one of the most original of today's natural scientists, the topic also grows into an issue of breathtaking practicality, easily tied to today's headlines.

One single map here can make the case. It is a map of the Middle East, on which large and small black dots mark the 100 oil and gas fields that lie densely strewn along a narrow band lying roughly on the line of the Persian Gulf, extended 500 miles inland to southeastern Turkey. Below this area, about 2 percent of the earth's land surface, lies more than half of the recoverable oil. Oil and gas are hydrocarbons; we group them as fossil fuel, for they have long been held, with coal, to be the modified remains of buried past life. The local geologic settings vary greatly over the region, from flat deposits to folded mountains. They show a range of ages. The cap rocks that must seal every deposit of such volatile fluids are not particularly good here; big seepages of bitumen are commonplace, ancient. Those disparate reservoirs underground are brimful; this is a surplus of supply, not a matter of unusually tight seals. The oils there show a "notable homogeneity," the authorities report, and are believed to have a common origin. Who could dissent? But what is that origin? The prevailing opinion is that the enviable energy wealth of the Gulf region shows there must have been rich life there for a long time, cap rocks sealing huge masses of buried organic sediment. It is the theory that oil comes from life that leads to the conclusions; here, Gold says, the inference is inverted.

It is more natural to argue quite the other way. Hydrocarbons have no

major source within the life of the biosphere, this heresy maintains, but instead leak out almost everywhere from hundreds of miles down in the earth's mantle. Those deep sources flow generously below the Gulf region for unknown reasons; the rather uniform material has filled every overlying trap, old or young, steep or flat, deep or shallow. Local context does not much matter; the phenomenon is regional, even global. The Persian Gulf is only the greatest of such examples; actually most petroleum provinces share the quality of independence from evident local causes. Ignorance cannot by itself much support such a novel hypothesis; yet it is freeing to grasp that our rich geologic knowledge of oil seems at least neutral on the origin of the largest find. Although on the whole the literature is strongly on the side of oil from life, the Soviet experts lean the other way.

There is not much doubt that hydrocarbons were brought into the forming earth in quantity; the plausible models of planetary genesis agree, and the stuff is plentiful in the meteorites, which still bring a little in as carbonaceous chondrites. Better yet, carbon manifestly erupts from the mantle to the surface in the striking form of diamond. Unoxidized carbon has aggregated into pure crystals at mantle temperature and pressure, and rare samples have been swept upward in an explosive jet of gas within diamond pipes. A diamond pipe is a ferociously concentrated cold gas volcano; we have never seen one erupt in our time, but they appear worldwide, during all geologic periods. There is plenty of carbon down there, enough to supply the entire carbonaceous crust tens of times over, the ubiquitous limestones no less than the rare fuelstuffs. Is it all oxidized?

Light matter floats upward. Diamonds rise with explosive gas volcanoes, fluid lava pours out in the familiar volcanic eruptions, and there are even mud volcanoes, mountain-size structures of mud built up from flows of fine sediments mixed with water. Lava volcanoes carry up steam and carbon dioxide gas. Mud volcanoes pour out methane instead. Gold proposes a unification: mostly it is methane that originally outgases from the depths. When it rises in hot lava along with steam, the gas is oxidized; when it rises with cooler water, it mixes mechanically to yield mud, the methane remaining chemically unchanged. Sometimes volcanoes and even earthquakes are accompanied by gases that flame or explode when they reach the open air: more methane. (A readable collection of anecdotes renders this kind of event likely; the overall argument does not depend on the sometimes ingenuous but often persuasive evidence, qualitative at best.)

The most interesting physical story here is of the mechanics of the flow of light fluid up through rock. It is important to recognize that gas or water in the pores of rock under high pressure must be almost in pressure balance with the rock. The pores do not close up; after all, the earth is a globe because the rocks below give way to the pressure. Was the gas of Mount St. Helens released from solution near the top, or leaked from deep below in open pathways through the melt? Earthquakes can arise both from the elastic failure of overstressed rock, as the wisdom of the day favors, and from the invasion of high-pressure gases to promote rock failure under stress, cracks that grow because gas pressure has balanced the healing compressive forces of the environment.

There are more quantitative lines of discussion. The helium-nitrogen-methane ratios found in gas wells show order from this new point of view. The carbon-isotope history fits this way too. Even coal seams, holding the big fossil cycads we have all admired, may well reflect less the growth and burial of a swamp than its posthumous enrichment by agelong flows of fluid hydrocarbons, catalyzed by the carbon already present into dissociation to leave more carbon. Most coal fossils are in fact replacement fossils. The nature of the complex processes behind coalification is still open. Quantity, however, remains a stumbling block. Diamonds run less than one gram to a ton of rock in the gemmiest of diamond pipes; if the pure carbon came up from the depths in pounds to the ton as cosmic abundance puts it, the case would be easy.

Drill deep, says Gold, wherever you are! The gas is there; the economics depends on what cracks you reach. A huge meteorite-impact crater lies buried under the granites of central Sweden; its trace is called the Siljan Ring. Around the ring are oil seeps, flammable gases in wells and gravitational signs of deep-lying porosity. All of this is in the ancient granite shield, where fossil life seems of no consequence. Hydrocarbons would have to come from deep below. Backed by Swedish and American investments, a well is now being drilled some three to five miles into the Siljan Ring, the wildest of cats. Already the signs show gas in amounts beyond any local origin; whether this first well will pay depends on the rock porosity at the depth reached.

Will there be enough methane to be won in the first 10 miles or so of the upper mantle to fuel us all for centuries? No one can now say. The diversity of this ingenious theory promises more even than gas strikes: a fresh grasp of the processes of the dynamic planet. The issue is joined in this even-tempered advocacy of a geochemical heresy; is all fossil energy the delayed gift of our nuclear sun, or is some of it the reworked legacy of unknown parent stars?

THE CAVE OF LASCAUX: THE FINAL PHOTOGRAPHS, by Mario Ruspoli. Harry N. Abrams, Inc., Publishers ($45).

About 17,000 years ago the wide halls and low passages of the limestone cave of Lascaux were splendidly decorated; a few generations later the Magdalenians abandoned the place. In 1940 its hidden art was discovered by four young men of the neighborhood; in 1963, although celebrated, the cave was closed once again. Then a few years ago a party of six men, led by Mario Ruspoli, a gifted filmmaker, shot under the most restrictive of conditions a full 150 reels of 16-millimeter color film. Their aim was a comprehensive documentation for the French Ministry of Culture: the first such cinematic monograph, its title *Corpus Lascaux.*

Ruspoli became a devotee of prehistory during his years of work on Lascaux, supported by the active collaboration of the archaeologist André Leroi-Gourhan, "my guide in the world of darkness." Ruspoli died in 1986, his book of the film still in press. More than an amateur's work, the book shines in vivid image and text from the experience of a visual artist deeply engaged within those fitfully lighted passages.

That film was no easy location shoot. The crew was excluded except for the three spring weeks of lowest cave temperature each year from 1981 through 1983. They could enter the cave only two or three times a week, to shoot under "a draconian timetable" in sessions that lasted for only two hours. Light came

almost entirely from two hand-held 100-watt battery-powered quartz lanterns, plenty of photons for the Fujicolor film. Each take was filmed by two Aaton cameras simultaneously; the film was sent off each night to Paris for processing, and the rushes came back a day or two later. The final expertly edited work presents the entire cave in four two-hour silent shows, each a journey through one section of the cave and a long treatment of its decorations. It has been the basis for television films as well.

Walking along the twisted, slippery paths intent on the images caught in their moving lights, the filmmakers were like the old artists with their dim oil lamps. "It became clear to us that the cave art at Lascaux was not made to be viewed like a series of fixed paintings, but as images in movement glimpsed by the Paleolithic initiates as they walked through the dark." One still picture shows paintings in a big chamber ahead so placed as to be seen with pleasing symmetry through the natural opening before entry, the light necessarily leading the visitor. At the Great Black Cow, in what is called the Nave, it was necessary to have three people holding the cameraman up on an awkward rock ledge, too narrow to allow distance from the image surface. The painters of the past must have had a platform resting on wood posts for the long, careful work they did there; some of the image was painted 13 feet above the ledge. (The floor level is unchanged.) It is archaeologically well established that wood scaffolding was set in the cave; holes made with flint picks are frequent in the walls, some still bearing the packed clay that helped to hold the pointed wood platform beams in place. Below in the debris on the floor lie the confirming lines of fallen tool flints and wasted pigment.

Lascaux itself is no longer to be visited; there will be none but fleeting and anxious inspections for the foreseeable future. In 1963 the green disease, a diatom-algal infection, seeded from the living world above by the trampling of 1,000 visitors a day, began to overgrow the wonderful pictures. Antibiotics and formalin controlled it. But a white affliction could not be cured as easily. It was a mineral disorder, the growth of opaque white calcite crystals over the pigment surfaces, induced by the increase in temperature, humidity and carbon dioxide that entered with the life and the breath of the public. Limitations were tried; years of research were put in to seek the best preventives. There remained the responsibility of conserving a unique treasure visibly changing at a rate thousands of times higher than the rate during its solitude. The cave was permanently closed. It is kept cooled, gently breathing, its deepest levels pumped free of carbon dioxide, all well guarded against seepage of water. So far that regime seems to work. Hence the strictures against the filmmakers, who were to be the last persistent intruders.

Stranger yet, the cave of Lascaux has in a way reproduced itself. A couple of hundred yards from the entrance, in an old quarry, Lascaux II has been built as a tunnel into a huge blockhouse of concrete. The gallery walls of more than a third of the old cave have been modeled in three dimensions, using precise stereophotogrammetric data and a mechanical means of locating many points in space to guide teams of sculptors. They reproduced in a special mortar the natural relief of the cave interior accurate to a matter of inches. There are 500 tons of modeled concrete; on those shaped walls the main paintings of Lascaux have been devotedly reproduced in texture, color and form, even in technique, by the painter Monique Peytral. Since its completion in 1984, the facsimile has awed 300,000 visitors each year "as if Lascaux had just reopened."

The lamps of Lascaux are photographed here: two or three candlelike flames burn together in a shallow limestone cup on wicks of juniper twig fed by the renal fat of ox or calf. Also shown is a perforated staff, apparently a straightener for reindeer-antler weapons shaped by soaking and fire. The frequency and decoration of these specialized tools have long been a mystery; it now seems that a hunter would need to carry his personal tool, to keep the spearpoint true in changing weather.

The rich images, the symbols and signs, the enigmatic scenes, the hints of the shaman are all clearly shown and discussed in brief, cogent essays particularly by the archaeologists Brigitte and Gilles Delluc, who also supply a fine bibliography. The enigmatic Unicorn closes their chapter. Is it a chimera? A lynx? A shaman in disguise? A portrait based on oral tradition? "It is the symbol of all that we still do not know...a trial attempt and at the same time a masterstroke by the earliest Magdalenians."

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, Utah 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.

Photocopying rights are hereby granted by Scientific American, Inc., to libraries and others registered with the Copyright Clearance Center (CCC) to photocopy articles in this issue of SCIENTIFIC AMERICAN for the flat fee of $1.25 per copy of each article or any part thereof. Such clearance does not extend to the photocopying of articles for promotion or other commercial purposes. Correspondence and payment should be addressed to Copyright Clearance Center, Inc., 21 Congress Street, Salem, Mass. 01970. Specify CCC Reference Number ISSN 0036-8733/87. $1.25 + 0.00.

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

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

Address subscription correspondence to Subscription Manager, SCIENTIFIC AMERICAN, P.O. Box 953, Farmingdale, N.Y. 11737. Telephone inquiries: 1–800–227–5782. The date of the last issue on subscriptions appears at the 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

BIBLIOGRAPHY

Readers interested in further explanation of the subjects covered by the articles in this issue may find the following lists of publications helpful.

DIET AND CANCER

DIET, NUTRITION, AND CANCER. Committee on Diet, Nutrition, and Cancer, Assembly of Life Sciences. National Research Council, National Academy Press, 1982.

PATHOLOGICAL AND BIOCHEMICAL ASPECTS OF TUMOUR PROMOTION. R. Marian Hicks in *Carcinogenesis,* Vol. 4, No. 10, pages 1209–1214; October, 1983.

DIET AND CANCER—AN OVERVIEW. Walter C. Willett and Brian MacMahon in *The New England Journal of Medicine,* Vol. 310, No. 11, pages 697–703; March 15, 1984.

THE NEW EMPHASIS IN CANCER CONTROL. Peter Greenwald and Joseph W. Cullen in *Journal of the National Cancer Institute,* Vol. 74, No. 3, pages 543–551; March, 1985.

DIET, NUTRITION, AND CANCER: A CRITICAL EVALUATION. Edited by Bandaru S. Reddy and Leonard A. Cohen. CRC Press, 1986.

HELIUM-RICH SUPERNOVAS

A PECULIAR SUPERNOVA IN THE SPIRAL GALAXY NGC4618. Alexei V. Filippenko and Wallace L. W. Sargent in *Nature,* Vol. 316, No. 6027, pages 407–412; August 1, 1985.

TYPE IB SUPERNOVAE 1983N AND 1985F: OXYGEN-RICH LATE TIME SPECTRA. C. M. Gaskell, E. Cappellaro, H. L. Dinerstein, D. R. Harkness and J. C. Wheeler in *The Astrophysical Journal,* Vol. 305, No. 2, Part 2, pages L77–L80; July 15, 1986.

THE EARLY SPECTRAL PHASE OF TYPE IB SUPERNOVAE: EVIDENCE FOR HELIUM. R. P. Harkness, J. C. Wheeler et al. in *The Astrophysical Journal,* Vol. 317, No. 1, Part 1, pages 355–367; June 1, 1987.

SUPERNOVA 1987A IN THE LARGE MAGELLANIC CLOUD: THE EXPLOSION OF A $\sim 20\ M_{\odot}$ STAR WHICH HAS EXPERIENCED MASS LOSS? S. E. Woosley, P. A. Pinto, P. G. Martin and Thomas A. Weaver in *The Astrophysical Journal,* Vol. 318, No. 2, Part 1, pages 664–673; July 15, 1987.

ARTIFICIAL CHROMOSOMES

ISOLATION OF A YEAST CENTROMERE AND CONSTRUCTION OF FUNCTIONAL SMALL CIRCULAR CHROMOSOMES. Louise Clarke and John Carbon in *Nature,* Vol. 287, No. 5782, pages 504–509; October 9, 1980.

CONSTRUCTION OF ARTIFICIAL CHROMOSOMES IN YEAST. Andrew W. Murray and Jack W. Szostak in *Nature,* Vol. 305, No. 5931, pages 189–193; September 15, 1983.

THE MOLECULAR STRUCTURE OF CENTROMERES AND TELOMERES. E. H. Blackburn and J. W. Szostak in *Annual Review of Biochemistry,* Vol. 53, pages 163–194; 1984.

CHROMOSOME LENGTH CONTROLS MITOTIC CHROMOSOME SEGREGATION IN YEAST. Andrew W. Murray, Neil P. Schultes and Jack W. Szostak in *Cell,* Vol. 45, No. 4, pages 529–536; 1986.

SYNCHROTRON RADIATION

SYNCHROTRON RADIATION: TECHNIQUES AND APPLICATIONS. C. Kunz. Springer-Verlag, 1979.

SYNCHROTRON RADIATION RESEARCH. Edited by Herman Winick and S. Doniach. Plenum Press, 1980.

SYNCHROTRON RADIATION RESEARCH—AN OVERVIEW. Arthur Bienenstock and Herman Winick in *Physics Today,* Vol. 36, No. 6, pages 48–58; June, 1983.

THE ANCESTRY OF THE GIANT PANDA

THE GIANT PANDA: A MORPHOLOGICAL STUDY OF EVOLUTIONARY MECHANISMS. Dwight D. Davis in *Fieldiana: Zoology Memoirs,* Vol. 3; 1964.

ON THE TAXONOMIC AND PHYLOGENETIC POSITION OF THE GIANT PANDA *AILUROPODA MELANOLEUCA*. V. E. Thenius in *Saugetierkundliche Mitteilungen,* Vol. 44, pages 286–305; 1979.

A MOLECULAR SOLUTION TO THE RIDDLE OF THE GIANT PANDA'S PHYLOGENY. Stephen J. O'Brien, William G. Nash, David E. Wildt, Mitchell E. Bush and Raoul E. Benveniste in *Nature,* Vol. 317, No. 6033, pages 140–144; September 12, 1985.

DEMONS, ENGINES AND THE SECOND LAW

IRREVERSIBILITY AND HEAT GENERATION IN THE COMPUTING PROCESS. R. Landauer in *IBM Journal of Research and Development,* Vol. 5, No. 3, pages 183–191; July, 1961.

SCIENCE AND INFORMATION THEORY.

Leon Brillouin. Academic Press, 1962.
THE THERMODYNAMICS OF COMPUTATION—A REVIEW. Charles H. Bennett in *International Journal of Theoretical Physics,* Vol. 21, No. 12, pages 905–940; December, 1982.
ON THE DECREASE OF ENTROPY IN A THERMODYNAMIC SYSTEM BY THE INTERVENTION OF INTELLIGENT BEINGS. Leo Szilard in *Quantum Theory and Measurement,* edited by John Archibald Wheeler and Wojciech Hubert Zurek. Princeton University Press, 1983.

EARLY FARMING IN NORTHWESTERN EUROPE

SETTLEMENT AND ECONOMY IN NEOLITHIC NORTHERN FRANCE. John M. Howell. B.A.R. International Series 157, British Archaeological Reports, 1983.
ANCIENT FRANCE: NEOLITHIC SOCIETIES AND THEIR LANDSCAPES, 6000–2000 BC. Edited by Christopher Scarre. Edinburgh University Press [1983?].

MODELING TIDAL POWER

BIOLOGICAL CLOCKS OF THE TIDAL ZONE. John D. Palmer in *Scientific American,* Vol. 232, No. 2, pages 70–79; February, 1975.
REASSESSMENT OF FUNDY TIDAL POWER. Bay of Fundy Tidal Power Review Board, 1977.
MARINE FORECASTING: PREDICTABILITY AND MODELLING IN OCEAN HYDRODYNAMICS. Edited by Jacques C. J. Nihoul. Elsevier Science Publishing Company, 1979.

THE AMATEUR SCIENTIST

FINGERING WITH MISCIBLE FLUIDS IN A HELE SHAW CELL. Lincoln Paterson in *The Physics of Fluids,* Vol. 28, No. 1, pages 26–30; January, 1985.
FRACTAL GROWTH OF VISCOUS FINGERS: QUANTITATIVE CHARACTERIZATION OF A FLUID INSTABILITY PHENOMENON. Johann Nittmann, Gérard Daccord and H. Eugene Stanley in *Nature,* Vol. 314, No. 6007, pages 141–144; March 14, 1985.
THE NONLINEAR GROWTH OF A GRAVITATIONALLY UNSTABLE INTERFACE IN A HELE-SHAW CELL. T. Maxworthy in *Journal of Fluid Mechanics,* Vol. 177, pages 207–232; April, 1987.

COMPUTER RECREATIONS

THE BEAUTY OF FRACTALS. Heinz-Otto Peitgen and Peter H. Richter. Springer-Verlag, 1986.

Words.

Along with its share of four-wheel drive awards, Jeep Cherokee has won a lot of hard-core fans. But Cherokee's success is hardly a mystery. To understand it, just look at all this Jeep has going for it.

Under the hood is the most powerful optional engine in Cherokee's class: a 4 litre 6-cylinder engine that cranks out an incredible 177 horsepower and 224 foot-pounds of torque. It not only gives Cherokee greater towing capability, it makes it the fastest, most powerful vehicle of its kind.

Harnessing all those horses is a versatile 5-speed transmission. Or you can choose an optional 4-speed automatic overdrive transmission with two shift modes to suit your needs.

Cherokee also gives you the choice of two or four doors. Two- or four-wheel drive. And two shift-on-the-fly four-wheel drive systems.

Any way you go, you'll have room and comfort for five adults. And up to 71 cubic feet of cargo space with the rear seat folded down.

But Cherokee is more than the sum of its parts. Because Cherokee is all Jeep. And that's what makes it unlike any other vehicle, on or off the road.

For further information, call toll-free: 1-800-JEEP-211.

Buckle up for safety.

Music.

Only in a Jeep® Cherokee

Jeep