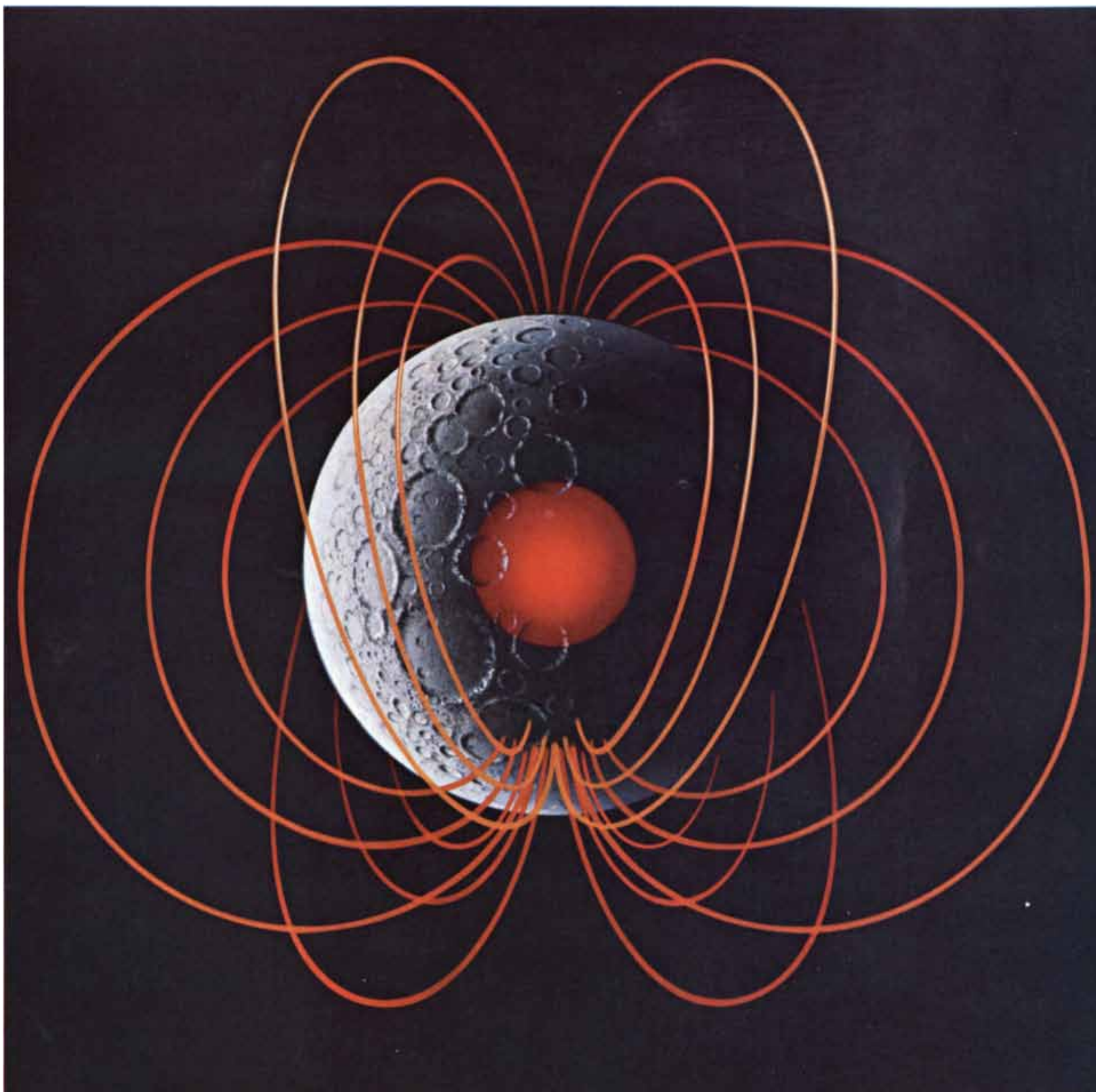
SCIENTIFIC AMERICAN

DECEMBER 1987
$3.50

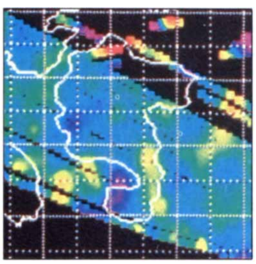
Animal cells and military tanks may crawl by similar mechanisms.
Did loops of cosmic string once seed the growth of galaxies?
Neural networks inspire an alternative approach to computation.

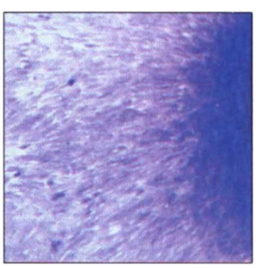


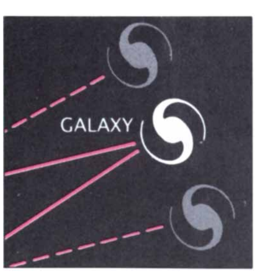
Lunar Magnetic Fields *nearly twice as strong as the earth's may have been generated billions of years ago in a core of molten iron.*

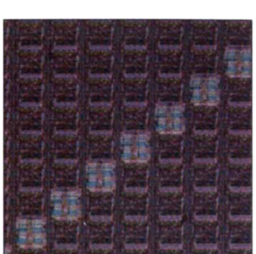
SCIENTIFIC AMERICAN

December 1987 Volume 257 Number 6









DEPARTMENTS

THE COVER shows the magnetic lines of force that may have emanated from the moon some four billion years ago (see "The Moon's Ancient Magnetism," by S. K. Runcorn, page 60). Magnetic surveys and the analysis of lunar rocks suggest that the moon may once have had a molten iron core that generated a magnetic field. The data hint that the ancient moon was repeatedly shifted with respect to its spin axis, perhaps by the impact of lunar satellites.

THE ILLUSTRATIONS

Cover painting by George V. Kelvin

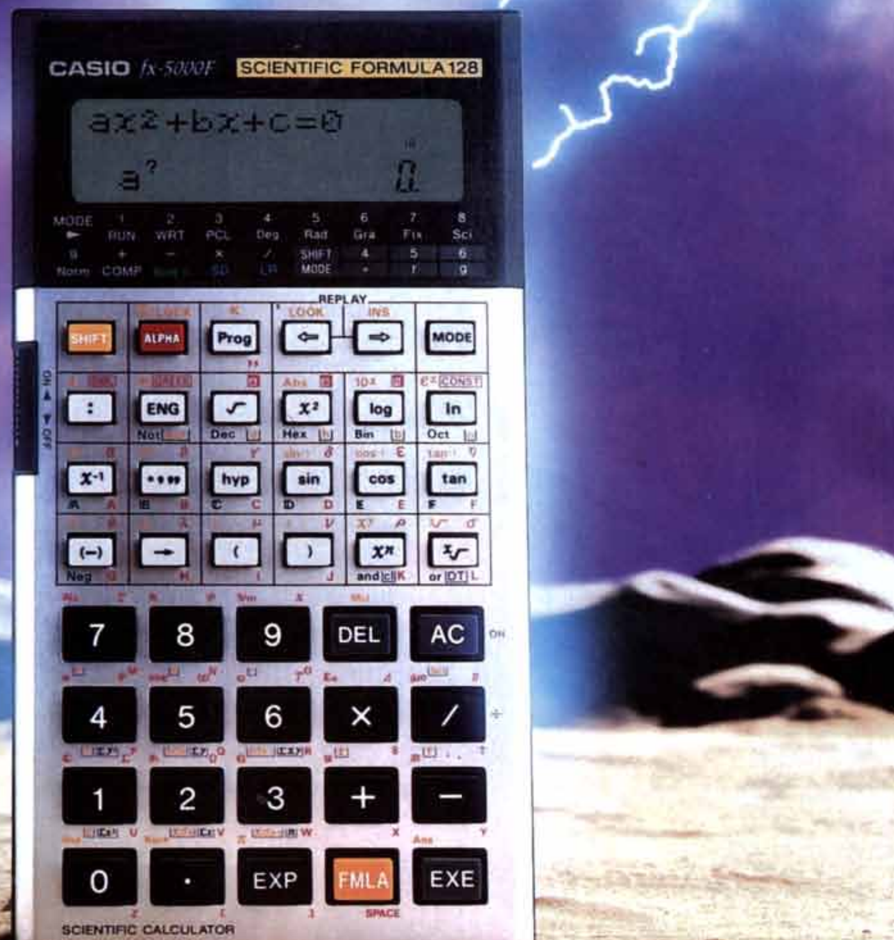
Page	Source
51	Dana L. Duke, courtesy of Federal Express Corp.
52–57	Andrew Christie
60–61	Christopher T. Russell, University of California at Los Angeles, and Laurence A. Soderblom, U.S. Geological Survey
62–68	George V. Kelvin
73	Peter A. Lawrence, Medical Research Council's Laboratory of Molecular Biology
74	Mark S. Bretscher
75	Tom Prentiss
76–89	Nichol Thomson and Mark S. Bretscher
90	Tom Prentiss
95	Quesada/Burke
96	Kitt Peak National Observatory
97	François R. Bouchet, University of California at Berkeley, and David P. Bennett, Fermi National Accelerator Laboratory and University of Chicago
100–102	George Retseck
105–112	Gabor Kiss
113	Melgar Photographers, Inc.
117	P. J. DeVries
118	Enid Kotschnig
119–121	Andrew Christie
123	Jon Brenneis
124–128	Ian Worpole
131	Courtesy of *The News*, Portsmouth, England
132–135	Hank Iken
136	Courtesy of the Newcastle Public Library, Newcastle, England
139	E. E. Donaldson and J. T. Dickinson
140	Michael Goodman
143	David Wiseman, University of Western Ontario
145	Edward Bell

SCIENTIFIC AMERICAN

Established 1845

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.


CASIO fx-5000F
SCIENTIFIC FORMULA 128
ax²+bx+c=0
SHIFT
ALPHA
Prog
MODE
ENG
log
ln
hyp
sin
cos
tan
7
8
9
DEL
AC
4
5
6
×
/
1
2
3
+
−
0
EXP
FMLA
EXE
SCIENTIFIC CALCULATOR


For gift delivery anywhere call 1-800-CHEER-UP (except where prohibited by law).


...and to all a Grand night.
Grand Marnier
Liqueur
PARIS FRANCE
Product of France. Made with fine cognac brandy 80 proof. ©1987 Carillon Importers, Ltd., Teaneck, N.J.

LETTERS

To the Editors:

As physicists working on gravitational wave detection we eagerly await the completion and successful operation of the LIGO system that was described in "Gravitational Wave Observatories," by Andrew D. Jeffries, Peter R. Saulson, Robert E. Spero and Michael E. Zucker [SCIENTIFIC AMERICAN, June]. Although such laser interferometric detectors hold great promise for the future of gravitational wave astronomy and may ultimately be the best means of reaching the sensitivity levels necessary to observe extragalactic sources, we should like to point out that cryogenic resonant-mass detectors, the most sensitive of existing detectors, will continue to play a significant and possibly decisive role in the effort to detect gravitational radiation.

The authors correctly point out that existing cryogenic bar detectors operate with strain sensitivities of 10^{-18}. This performance exceeds by about three orders of magnitude the energy sensitivity of earlier room-temperature bar detectors, an improvement chiefly due to the use of low temperatures to reduce the thermal noise of the instruments. The authors state that "it is possible to make more delicate measurements with the bars, but it will be difficult."

We believe this statement is misleading. After a decade of development and a total of nearly two years of operating bar detectors at strain sensitivities of 10^{-18}, we are confident that the next step in improving bar detectors is straightforward and will lead to a strain sensitivity approaching 10^{-20}, which represents an additional improvement in energy sensitivity of four orders of magnitude. The coming generation of bar detectors will include several innovations, the most important of them being the use of ultralow temperatures, which will reduce the thermal noise by as large a factor as the earlier step from room temperature to low temperature did. The new bar detectors will be able to sense the levels of radiation expected from gravitational-collapse events occurring anywhere in our galaxy.

We support a two-pronged attack on the problem of detecting gravitational radiation. In the next few years, at a cost of from $1 to $2 million per system, the cryogenic resonant-mass antennas should be improved to a strain sensitivity of about 10^{-20}. Meanwhile the more ambitious and more costly LIGO program should be pursued in the hope of eventually reaching a strain sensitivity exceeding 10^{-22}, the level required for detecting the more frequent gravitational signals from other galaxies.

The direct detection of gravitational radiation is one of the most challenging and important problems confronting 20th-century science. Ongoing research in this area has already greatly advanced the art of measuring tiny displacements and has produced antennas that are sensitive enough to detect the strongest predicted signals.

We are on the verge of opening a new astronomical window that will revolutionize our view of the cosmos. As the recent detection of a neutrino burst from supernova 1987A demonstrates, unambiguous detection of even a single event can be very important. We must be ready to seize such opportunities. In an age in which the cost of new instruments for investigating the frontiers of knowledge is often measured in billions of dollars, there should be no doubt that investing millions of dollars in what is perhaps Albert Einstein's most important legacy will pay generous dividends.

WILLIAM M. FAIRBANK

PETER F. MICHELSON

JOHN C. PRICE

ROBERT C. TABER

Stanford University
Stanford, Calif.

To the Editors:

In "Gravitational Wave Observatories" the authors propose that a laser interferometer be used as a gravitational wave detector. In the absence of gravitational waves the interferometer would be immersed in the "normal" gravitational field existing at the installation site: the arms would have some "normal" length. A gravitational wave would cause the arm lengths to deviate from "normal." I think this transient change in the gravitational field will change the speed of light so that the effects of different path lengths along the arms might be canceled.

Could this be another impasse of the kind encountered in the Michelson-Morley experiment?

ANTHONY V. RAINIS

Sudbury, Mass.

To the Editors:

We strongly support the efforts of our colleagues at Stanford and at other research centers to seek gravitational waves with resonant-mass detectors and look forward to the possibility that bar antennas will make the first confirmed discovery of gravitational waves. The discovery would have a great impact on astrophysics and would give a tremendous boost to the worldwide endeavor of experimental gravitational wave astronomy.

In reply to Mr. Rainis and other readers who raised the same point, we note that the general theory of relativity makes a distinction between physical effects, potentially observable by experiment, and descriptions of these effects that depend on the choice of reference frame. Specifically, the interaction of a gravitational wave with a detector of the type we have described can be viewed from two vantages: (1) the wave moves the test masses but does not affect the laser beams or (2) the wave changes the path followed by the laser beam through space-time without disturbing the masses. Either description, or even an intermediate one in which both the masses and the laser beams are affected, can be applied in predicting how a detector would respond to a gravitational wave. The predictions are the same: light waves launched in phase down orthogonal arms of the interferometer return slightly out of phase.

ANDREW D. JEFFRIES

PETER R. SAULSON

Massachusetts Institute of Technology
Cambridge, Mass.

ROBERT E. SPERO

MICHAEL E. ZUCKER

California Institute of Technology
Pasadena, Calif.

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

DECEMBER, 1937: "The threads of which all life fabric, every animate shape, is woven are the nitrogen-containing proteins. The center of life-control, the nucleus within every living cell, is rich in protein—therefore in nitrogen. Now heavy atoms of nitrogen can be used as tags or labels by which to distinguish certain molecules from others of the same kind found in the body. By such tags, molecules can be traced even into the most obscure of life's laboratories within the tissues."

"An automatic landing system for airplanes, wherein the plane, without human contact with its instruments, literally finds its own way to the landing field and makes a safe landing, has been designed and reduced to practice by a group of engineers cooperating with the Army Air Corps. The system makes use of radio guiding stations that control the heading and altitude of the plane through a Gyro-pilot and the speed through a small throttle engine."

"Thanks to the common sense and co-operativeness of 11 governments representing both the economic interests of the whaling industry and the broad interest of science in the preservation of our remaining fauna, that valuable and ever-fascinating aquatic mammal the whale, by far the hugest bulk of animate flesh this old earth has known at any period, is now probably saved. Game laws are now to be applied to whales. These game laws are to go into effect in the Antarctic regions on December 8, and whaling will continue only until March 7."

"One motor car company reports that black is gradually losing its place as the favorite color for automobiles, the indications being that the American public is going in for greater diversification of color."

"It would be quite proper to call 1936 the squarest year because: 1936 is the square of 44, 1 the square of 1, 9 the square of 3, 36 the square of 6, 16 the square of 4, 196 the square of 14, 361 the square of 19, 169 the square of 13, 961 the square of 31."

SCIENTIFIC AMERICAN

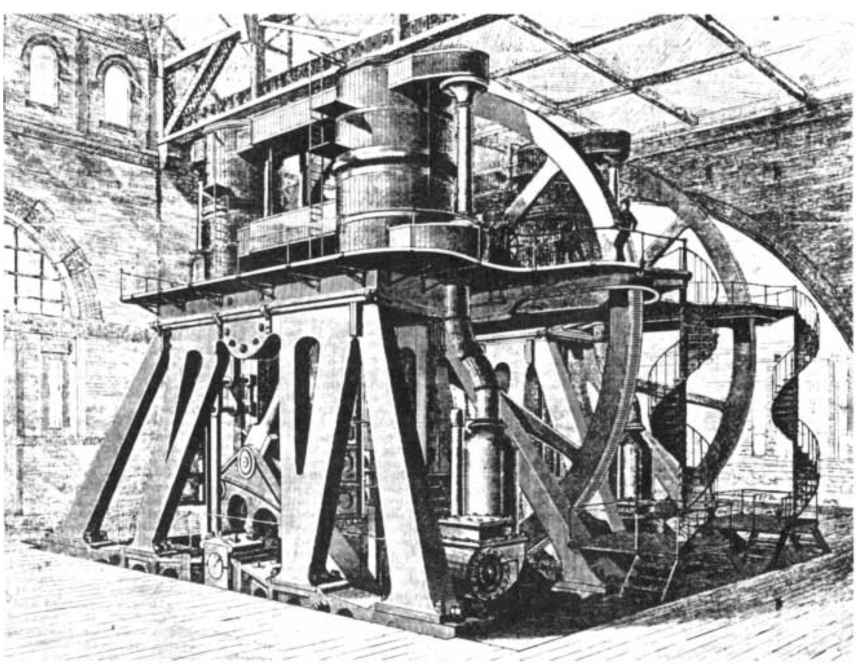
DECEMBER, 1887: "A very large area of the arable lands of Australia is now overrun by rabbits. So numerous and active are these animals that they destroy trees, grass, crops and everything that grows. They move in great armies and carry destruction in their path. The authorities now offer the handsome reward of $125,000 for a method of overcoming the evil."

"A giant raft of timber is now expected at this port. It left Nova Scotia on December 8, in tow of the steamship *Miranda.* In general shape it is a pointed cylindroid of elliptical section. The total length is 585 feet. The great structure is estimated to weigh 11,000 tons, or 2½ times as much as the *Great Eastern.*"

"A steam horse power is equal to three actual horses' power, and a living horse is equal to seven men. The steam engines of the world represent, therefore, approximately the work of 1,000,000,000 men, or more than double the working population of the earth, whose total population amounts to 1,455,923,000 inhabitants. Steam has accordingly trebled man's working power, enabling him to economize his physical strength while attending to his intellectual development."

"The new cellar-fire apparatus of the Paris fire department consists of a suit like that used by divers, which allows a fireman to enter a cellar in which the air has been rendered irrespirable by a conflagration. The fireproof suit consists of a leather blouse fastened at the waist and wrists with ligatures and provided with a hood and iron mask. The air necessary for respiration is introduced through an aperture in the back of the suit by means of a rubber tube of great length. The blouse is very roomy and allows of great liberty of motion."

"The city of Boston, Mass., has recently built and now has in full operation a system of sewage and drainage works that mark an important advance in sanitary engineering. Originally, the sewage was disposed of as in New York. It was allowed to run out into the water from numerous outlets. This was found objectionable. The water became contaminated, and the dock frontage was injured by deposits of sludge. It was determined to surround the city with an intercepting sewer, which receives the delivery from all the lines formerly discharging into the harbor and adjacent water. The sewage is taken to a distant point and, after proper clarifying, is discharged."



The main pumping engine of the new Boston sewage system

Nissan builds cars and trucks for the most important race of all.

ıan Race.

But above all, it should feel as if it were custom-built for you. At Nissan, we call this personal approach to automotive design Human Engineering.™

And it means a lot more than just radio knobs that are an arm's reach away.

It means everything about the vehicle works for you. From the way the mirrors adjust, to the way the transmission shifts, to

the way the jack works.

So no matter who you are, or how you use your Nissan, the end result is always the same. You feel good about your car or truck and what it does for you.

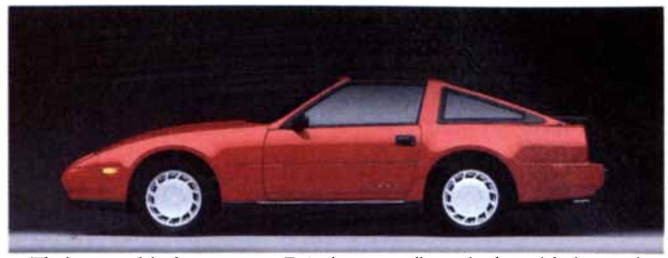
In striving to build these kinds of vehicles, Nissan faces some diverse and difficult engineering challenges.

In the case of the Newman-Sharp 300ZX Turbo, it's building the fastest car possible.

In the case of the Sentra,™ it's building the most economical car possible.

In the case of the modular Pulsar® NX, it's building a car that's actually more cars than you ever thought possible.

Yet, of all the things engineered into Nissan



The heritage of the famous Nissan Z-Car lives on in all its turbocharged, fuel-injected splendor with the new, sleeker-than-ever 300ZX.

Why pay a fortune for a performance sedan when the Nissan Maxima® lets you have your power. Your luxury. And your bank account.

Nissan's Pathfinder® blends the famous 4x4 Hardbody™ performance with sleek styling, exceptional comfort and uncommon practicality.

vehicles, we're most proud of the one thing that we haven't engineered out of them.

The pure pleasure of driving one.

Because at Nissan, we never forget why we build cars and trucks.

Or, who we build them for.

NISSAN®

Built for the Human Race.™

THE AUTHORS

JAMES BRIAN QUINN, JORDAN J. BARUCH and PENNY CUSHMAN PAQUETTE ("Technology in Services") share an interest in strategic planning and the management of technological change. Quinn is William and Josephine Buchanan Professor of Management at the Amos Tuck School of Business Administration at Dartmouth College. He earned a master's degree in business administration at Harvard University in 1951, after which he became a new-products analyst in industry and a marketing instructor. He moved to the Tuck School in 1957 and was awarded a doctorate in economics from Columbia University in 1958. Baruch is president of Jordan Baruch Associates, which advises industry and government on strategic planning and the integration of technology. He received a doctorate in electrical instrumentation from the Massachusetts Institute of Technology in 1950, remaining there to teach until 1971. He continued his academic career until 1977, initially at the Harvard University Graduate School of Business Administration and then at Dartmouth. He has also cofounded and directed two companies. From 1977 to 1981 Baruch was the Assistant Secretary of Commerce for Science and Technology. Paquette, who works as Quinn's research associate, holds a 1970 bachelor's degree from Smith College and a 1976 master's degree in business administration from the Tuck School. She has worked as a management consultant.

S. K. RUNCORN ("The Moon's Ancient Magnetism") is head of the school of physics at the University of Newcastle upon Tyne, where he has been a professor since 1956. He obtained a master's degree in mechanical sciences at Gonville and Caius College of the University of Cambridge in 1946 and a doctorate in physics from the University of Manchester in 1949. He was a fellow of Gonville and Caius College from 1948 to 1955 and assistant director of research in geophysics at Cambridge from 1950 to 1955.

MARK S. BRETSCHER ("How Animal Cells Move") is head of the division of cell biology at the Medical Research Council's Laboratory of Molecular Biology. He completed his Ph.D. at the University of Cambridge in 1965 and spent a year as a Fulbright scholar at Stanford University before returning to Cambridge to join the laboratory he now directs. He has been a visiting professor at Harvard (1974) and Stanford (1985) universities. This is Bretscher's second article for SCIENTIFIC AMERICAN; his first, "The Molecules of the Cell Membrane," appeared in October, 1985.

ALEXANDER VILENKIN ("Cosmic Strings"), who has been professor of physics at Tufts University since 1983, has devoted the past few years to the study of cosmic strings and the question of how the universe was created out of "nothing." A native of the Soviet Union, he received a bachelor's degree in physics from Kharkov State University in 1971. In the five years that followed he first served in the army and then worked at odd jobs while studying physics—and publishing papers—in his spare time. He came to the U.S. in 1976 and earned a Ph.D. the next year from the State University of New York at Buffalo. Vilenkin went to Tufts after postdoctoral studies at Case Western Reserve University.

DAVID W. TANK and JOHN J. HOPFIELD ("Collective Computation in Neuronlike Circuits") work together on the technical staff of the AT&T Bell Laboratories. Tank has a bachelor of science degree from Case Western Reserve University. He joined Bell Laboratories in 1983, the year he got his doctorate in physics from Cornell University. Hopfield has been at Bell Laboratories since 1973 and is also Roscoe Gilkey Dickinson professor of chemistry and biology at the California Institute of Technology. He got a Ph.D. in physics from Cornell University in 1958 and spent two years at Bell Laboratories before accepting a position as a research physicist at the École Normale Supérieure. He returned to the U.S. in 1961 to teach physics at the University of California at Berkeley, leaving there in 1964 for Princeton University. He took his post at Caltech in 1980.

DAVID CREWS ("Courtship in Unisexual Lizards: A Model for Brain Evolution") is professor of zoology and psychology at the University of Texas at Austin. He received a bachelor's degree in 1969 at the University of Maryland at College Park and a doctorate from Rutgers University in 1973. After two years as a research associate at the University of California at Berkeley he moved to Harvard University in 1975. Crews left there for Texas in 1982. He is currently studying the hormonal modulation of sexual differentiation and the factors that in garter snakes regulate adaptation to extreme environments.

TERRY A. MICHALSKE and BRUCE C. BUNKER ("The Fracturing of Glass") have collaborated in developing simple models to describe the molecular reactions that take place when glass cracks. Michalske has just assumed responsibility for managing the metals and ceramics science section at the Battelle—Pacific Northwest Laboratories. He got a doctorate in ceramic science in 1979 from Alfred University and then worked at the Center for Materials Science of the National Bureau of Standards until 1981, when he joined the technical staff of the Sandia National Laboratories. In six years with the company Michalske created a glass-ceramic material that has improved fracture resistance and developed a laser-based technique to repair cracks in glass. Bunker has been a member of the technical staff at Sandia since 1979, the year he completed his Ph.D. in inorganic chemistry at the University of Illinois at Urbana-Champaign. In addition to studying how glass fractures, he is developing chemical methods for the synthesis of ceramics. Bunker's research projects include the chemical synthesis of high-temperature semiconductors and oxide films for integrated-optics devices.

WALTER BROWNLEE ("H.M.S. *Warrior*") has been the full-time historian of the *Warrior* since 1983. He is also the author of a book about the ship, which he helped to restore. He entered the Merchant Navy of Great Britain in 1948 and served on cargo vessels and oil tankers throughout the world. After rising through the ranks from deck officer to chief officer, he earned a Master's Certificate of Competency in 1959. Two years later he changed careers, becoming a teacher specializing in history and geography. For many years Brownlee, who became a headmaster in 1968 and the warden of a teacher's center in 1973, has combined his teaching skill with a passion for maritime history to write books and articles that present history in a form palatable to nonhistorians.

We've just taken true supercomputing to an unheard-of level.

Departmental.

Thanks to ETA Systems, Control Data's supercomputer company, we've created the first true supercomputer a department can afford: The ETA10-P.

THE P STANDS FOR PERFORMANCE. ULTRA PERFORMANCE.

The air-cooled ETA10-P offers a processing speed of 375 megaflops.

That's as fast as the first generation Class VI machines. That's 30 times faster than a well-known, similarly configured mini-super. And over 100 times faster than today's most popular super-mini.

Yet an ETA10-P CPU is priced at an incredibly low $850,000. And a fully operational ETA10-P, including a CPU, disk, IOUs and software, is only $995,000.

You get a true supercomputer that's faster than a CRAY-2 CPU (based on Linpack 100x100, all FORTRAN) for under a million dollars.

INTRODUCING THE ETA10-P SUPER-COMPUTER.

AWESOME PERFORMANCE AT AN AFFORDABLE PRICE.

ETA SYSTEMS
A Control Data Company

AND THE ETA10-P IS JUST THE BEGINNING.

Because the ETA10-Q, an air-cooled supercomputer that's even faster than the P model, is also being announced.

Together with the present ETA10 machines—the liquid nitrogen cooled ETA10-E and the fastest supercomputer in the world, the ETA10-G—the new P and Q models help create the only complete line of supercomputers in the industry.

Each model is available with single or multiple processors. In fact, there are 44 different configurations.

The result is a continuous family of processors with an unprecedented 27 to 1 performance ratio—that's the largest operating range in the business.

And because they all share the same operating system, they provide you with what you need most:

A COMPLETELY COMPATIBLE GROWTH PATH TO TOMORROW.

Call your local Control Data sales office for complete information on the ETA10-P and the ETA10 family. Or call Hal Julsen, Vice President of Sales, at **800-553-2214**. Or write him at ETASystems, Incorporated, 1450 Energy Park Drive, St. Paul, MN 55108.

CONTROL DATA

SCIENCE AND THE CITIZEN

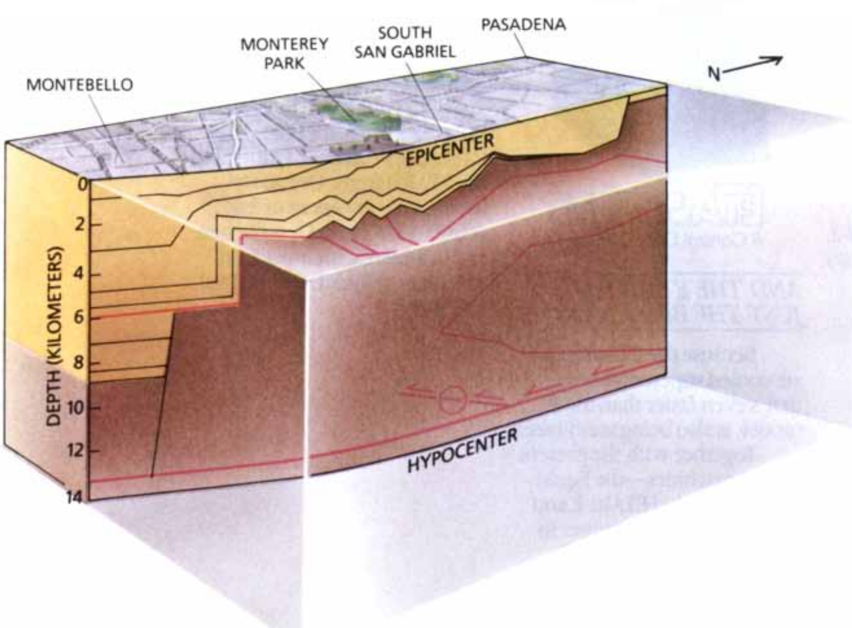
All Shook Up

Los Angeles earthquake reveals new signposts of seismic danger

Seismologists trying to predict where potentially destructive earthquakes will strike in California have usually relied on a somewhat comforting assumption: that the faults along which large quakes occur almost always extend all the way to the surface, where they can be detected and monitored. California is riddled with such highly visible faults. Most, including the notorious San Andreas, are strike-slip faults: vertical fissures between crustal plates moving in different directions.

Understandably, then, when a 6.1-magnitude earthquake rocked eastern Los Angeles on October 1—killing six people and causing more than $200 million in damage—most seismologists quickly blamed the nearest strike-slip fault marked on their maps: the Whittier fault, which runs southeast from the suburb of Whittier for about 30 miles. But in subsequent weeks analysis of seismic data pinpointed the epicenter of the quake several miles northwest of the end of the Whittier fault; moreover, the quake's motion seemed indicative not of strike-slip but of thrust faulting, which occurs when a highly compressed plate shears so that one section is lifted over another. Gradually investigators came to the disturbing conclusion that slippage along a thrust fault hidden deep below the surface—and probably unconnected to the Whittier fault or any other fault visible at the surface—had triggered the quake.

Folds—not just faults—near the earth's surface may mark regions threatened by earthquakes



THRUST FAULT in schist (brown) *some 12 kilometers below the earth's surface is thought to have caused a 6.1-magnitude earthquake that rocked eastern Los Angeles on October 1. Although South San Gabriel lies above the quake's hypocenter, towns to the south, the direction of slippage, sustained the heaviest damage. Previous faulting* (red) *in the schist, resulting from compression of the region, contributed to the folding of sedimentary rock* (yellow) *near the surface. The drawing is based on work done by Thomas L. Davis.*

Actually danger signs were there all the time in the form of folds near the surface of the earth, according to Thomas L. Davis, a geological consultant. For several years Davis has gathered core-sample data from oil wells dug throughout the Los Angeles basin. He found that long-term compression of the basin has caused sedimentary rock below its surface to buckle in numerous areas, including the site of the recent earthquake.

Correlating this information with seismic data and measurements of surface deformations, Davis concluded that the folding has come about not by slow "creep," as was commonly thought, but by the sudden rupture of thrust faults in layers of schist that underlie the sedimentary rock. The October quake, in other words, is simply the most recent of many quakes that have driven the folding over millions of years. Davis maintains that the earthquake threat in Los Angeles—and in other parts of the state—may be greater than seismologists have thought. "You see these young, active folds all over California," he says.

"I'm in complete agreement with Davis," Lucile M. Jones of the U.S. Geological Survey in Pasadena says. "His work, and this quake, show that we've got to take this threat more seriously." Jones thinks that thrust faults underlying Los Angeles probably cannot create quakes with a magnitude greater than 6.5 or 7. But she suggests that such earthquakes, originating directly under the city, could prove more devastating than an 8-magnitude quake along the San Andreas fault, which passes some 35 miles northwest of the city.

Ross S. Stein of the Geological Survey's Menlo Park office concurs that "our approach to earthquake-hazard reduction has been flawed. We thought that large quakes always cut the surface, and so the strategy we followed was to search for large faults that offset through recent deposits." Stein points out that an earth-

ible apply. Also, participating dealers back their customer-paid work with a free Lifetime Service Guarantee, good for as long as you own your vehicle. Ask to see these limited warranties when you visit your Ford Dealer.

Ford. Best-built American cars...seven years running.

Based on an average of owner-reported problems in a series of surveys of '81-'87 models designed and built in North America. At Ford, "Quality is Job 1."

Ford Escort GT.

quake that struck near the city of Coalinga, in south-central California, actually delivered the thrust-fold warning four years ago.

A deep anticline, or downward-bending fold, runs past Coalinga, but because the region lacked faults at the surface, workers had thought it was seismically inactive. "Geologists typically assumed that the earth was gradually progressing, not jumping, under these folds," Stein says. When the 6.7-magnitude quake deepened the bend in the anticline without rupturing the earth's surface, investigators realized that the anticline had been created by repeated slippages of a thrust fault deep underground.

Stein declares, "We can and must predict which folds may rupture." Once such folds have been identified, he says, workers must then try to forecast when the next quake might occur. The task is more straightforward if there has been a recent earthquake. Investigators can determine the increase in surface folding by comparing surveys of the earth's surface made before and after the quake. By comparing the recent increase with the total amount of folding that has occurred over geologic time, indicated by subsurface mapping, workers can estimate a "repeat time" for the quake. With this methodology, Stein says, he and another Geological Survey worker, Geoffrey C. P. King, derived a repeat time of roughly 1,000 years for the Coalinga quake. Stein is now trying to do the same thing for the October 1 quake. These efforts, he says, should help workers to take on a more difficult and important task: forecasting when a quake might strike sites where no quakes have been observed.

The key to forecasting, Davis says, is information about the subsurface structure of folds. Although oil companies gather such information, he notes, seismologists too often rely solely on seismic data and measurements of surface deformations. "That's just one aspect of what leads to a quake," he says, adding: "Geologists haven't been communicating that well." *—John Horgan*

Computing Cornucopia

Government is moving to exploit massively parallel machines

The White House's Office of Science and Technology Policy (OSTP) is putting the final touches on a plan for concerted Government action to ensure that scientists and engineers benefit from the potentially enormous gains in computing power afforded by the new generation of parallel machines. The National Research and Development Strategy for High-Performance Computing is largely a response to congressional requests; the plan is being drawn up for the OSTP by a committee of the Federal Coordinating Council for Science, Engineering and Technology and will go to Congress after a White House review.

In spite of the military and commercial promise of massively parallel computers—machines in which a host of processors attack a problem simultaneously—some officials are concerned that private industry will not act quickly enough to develop new architectures, software and networks to harness parallelism for key applications in science and engineering. Stephen L. Squires, computer-systems director at the Defense Advanced Research Projects Agency (DARPA), believes that without Federal efforts to catalyze innovation, market forces can "actually slow down" new technologies, because large corporations do best by exploiting their earlier investments.

Squires thinks that machines capable of 1,000 billion operations per second will be feasible by the early 1990's, and that the cost of computing will fall dramatically. To give investigators at universities and other scientific institutions routine access to such computing power, the plan would establish a national computer research network. Many Government agencies, such as the National Science Foundation, DARPA and the Department of Energy, already have their own networks, but the "bridges" between the different networks are seen as inadequate even for present needs. Officials think an integrated national network should be able to transmit up to 300 million bits of data per second—a capacity more than 1,000 times greater than that of existing networks.

Such speeds would make the network "a tool for major breakthroughs in science," according to Kenneth G. Wilson, director of Cornell University's Theory Center. It could readily cope with the huge amounts of data generated by high-speed scientific computation, making it possible, for example, to transmit moving color images representing scientific simulations in real time. Workers would also be able to access data generated at large experimental facilities and to collaborate on program development. For very large computing projects, the network might make it possible to pool the power of large machines at different sites. Paul G. Huray, who chairs the committee that produced the Federal plan, points out that many of the fiber-optic cables needed to build such a network are already in place, laid as redundant capacity by telecommunications companies.

Officials expect that, besides the research network, the plan will call for increased Federal investment in parallel computers themselves and in their software, perhaps through specialized computing-research centers similar to those now operated by the NSF. Development of software for parallel machines is widely seen as lagging seriously, particularly in scientific applications; Wilson says "the whole framework of scientific computation is hopelessly backward," primarily because "nobody has been able to get a committee of scientists together to address" how to translate abstract scientific ideas into software. The plan will also address concerns about a predicted shortage of trained manpower; one solution would be to increase research funds allocated to universities, thereby encouraging them to create new faculty positions.

Estimates of the cost of mounting a national computing initiative vary, but one official thinks that a good case could be made for spending $500 million per year. A study of the topic led by Harold J. Raveché of the Rensselaer Polytechnic Institute called earlier this year for an expenditure of $1.5 billion over the next five years. Specific funding proposals may be included in the Federal budget for fiscal 1989. *—Tim Beardsley*

Scientific Philanthropy

The Hughes legacy supports a booming research effort

After a decade of disputed claims, complete with forged Mormon wills and courtroom battles, much of the fortune of the late multimillionaire industrialist, aviator and recluse Howard Hughes is now supporting the largest scientific charity in the world. The Howard Hughes Medical Institute, which he established in 1953, has assets of approximately $5 billion, gained from the 1985 sale of

For Bill Demby, the difference means getting another shot.

When Bill Demby was in Vietnam, he used to dream of coming home and playing a little basketball with the guys.

A dream that all but died when he lost both his legs to a Viet Cong rocket.

But then, a group of researchers discovered that a remarkable Du Pont plastic could help make artificial limbs that were more resilient, more flexible, more like life itself.

Thanks to these efforts, Bill Demby is back. And some say, he hasn't lost a step.

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

Better things for better living. DU PONT

REG. U.S. PAT & TM OFF

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

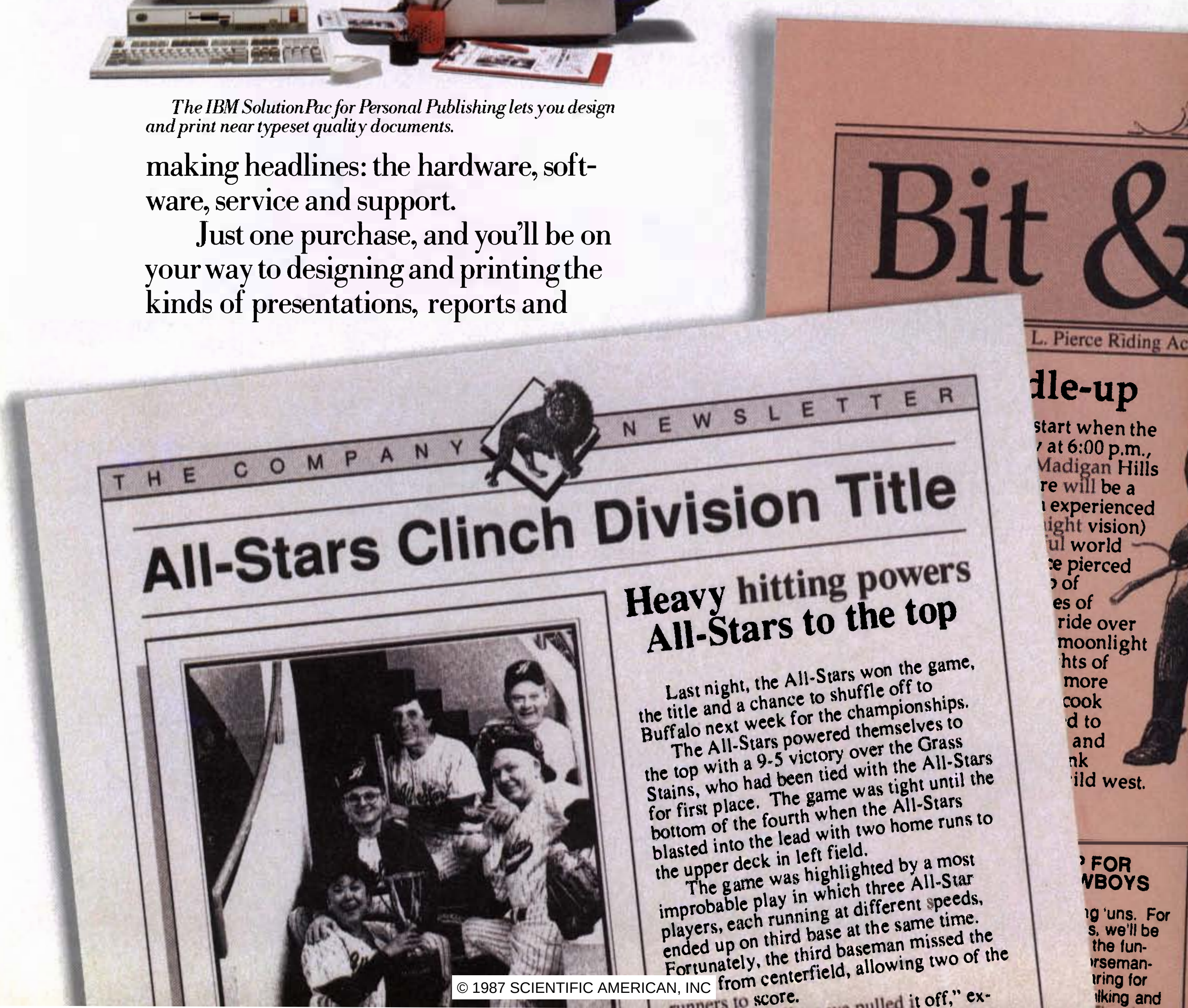
The IBM SolutionPac for Personal Publishing lets you design and print near typeset quality documents.

making headlines: the hardware, software, service and support.

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,™

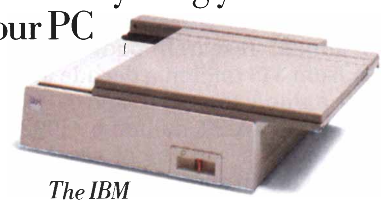


wanted to make headlines, publishing.

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.

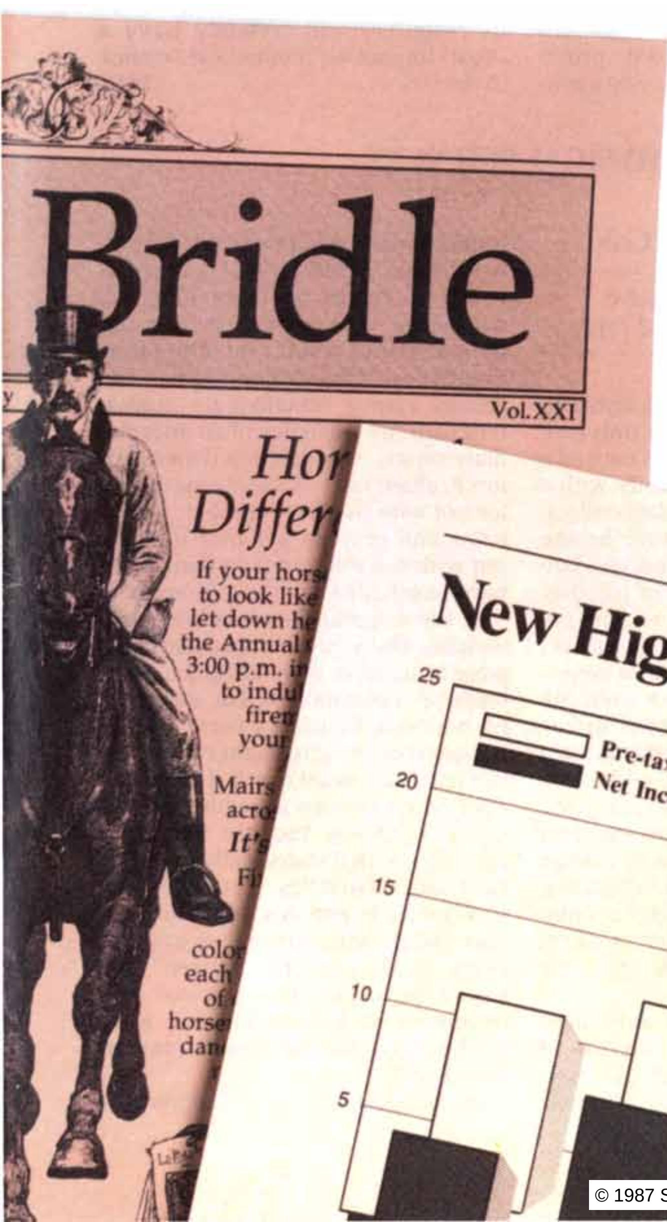
There's good news for IBM PC XT™/286 and Personal Computer AT® owners, too. The IBM Personal Publishing Option offers you everything you need to upgrade your PC into a cost-saving print shop.



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®



its major holding, the Hughes Aircraft Company, to the General Motors Corporation.

In its early years HHMI's operations were as obscure as those of its founder, and this year the institute was touched by scandal when its president, Donald S. Fredrickson, resigned after an investigation into his wife's use of HHMI funds. With a new president, Purnell W. Choppin, and an ambitious agenda of support for biomedical research and education, HHMI's prominence seems assured.

In order to preserve its tax-exempt status as a medical research organization HHMI has been obliged to spend at least 3.5 percent of the value of its endowment each year on medical research. As a result the institute's annual research budget has grown from $15 million a decade ago to $170 million this year and is expected to reach $235 million by 1992. What is more, a settlement reached early this year in a long-standing dispute with the Internal Revenue Service obliges HHMI to spend an additional $500 million for charitable purposes over the next 10 years. The institute has announced that the money will go to support biomedical education.

HHMI's research agenda focuses on five main areas: cell biology, genetics (including efforts to map the human genome), immunology, neurosciences and (most recently) the study of biological structures at the molecular level. Most Hughes research is fundamental, but much of it also has near-term clinical applications. Several Hughes investigators are studying the molecular basis of inherited diseases, for example, which is leading to better diagnoses and may suggest treatments; others are exploring immune-system activation, a process central to AIDS.

Traditionally HHMI's research has been done by its own salaried investigators at institutes set up within leading research universities and hospitals around the country. HHMI does not require long and time-consuming grant applications, a major bugbear for Federally supported researchers. Investigators can write their budget requests on a year-by-year basis, and they are entitled to seek research grants from other organizations as well.

The salaries HHMI funds are no higher than those funded by the National Institutes of Health, the major source of Government money for biomedical research, but Choppin thinks Hughes investigators have "a bit more flexibility and stability" than workers supported by Federal grants. Once chosen, an investigator can count on continuing support (providing his or her work is favorably reviewed) for a set period that can be as long as seven years. "We bet on individuals, not on projects," Choppin says, which enables HHMI to take "a more long-term approach" to problems.

There are now 170 Howard Hughes Investigators at 27 institutes nationwide; 13 more institutes are currently under construction, and HHMI expects to have a total of 270 investigators in five years' time. In a new emphasis, many of the investigators will be one-person "institutes." Choppin says that HHMI is considering "in depth" the possibility of supporting research abroad.

To its tradition of supporting its own investigators HHMI is now adding a program of educational funding. Grants that next year will total $30 million will go to colleges to strengthen education in biology and related disciplines; special provisions have been made to solicit applications from historically black institutions. In addition 60 doctoral fellowships—open to U.S. or foreign citizens—will be supported next year at U.S. universities, a number that will increase to 300 over five years. HHMI has underwritten a program to foster "greater public understanding of science and studies in health sciences policy" at the Institute of Medicine and a study of precollege biology education at the National Academy of Sciences. It has also awarded a total of $9 million to the Cold Spring Harbor Laboratory (for a neurosciences teaching facility) and the Jackson Laboratory at Bar Harbor (for genetics teaching).

Choppin modestly points out that HHMI's total assets, although "considerable," are less than the annual budget of the NIH and expresses the hope that HHMI's research will "complement, not compete with," NIH-funded work. Choppin's understatement notwithstanding, HHMI's continuing search for new ways to deploy all its resources will certainly have a broad impact on biomedical science in the U.S. *—T.M.B.*

PHYSICAL SCIENCES

Force of a Different Color

New results bolster the case for a "fifth" fundamental force

Nearly two years ago Ephraim Fischbach of Purdue University and his colleagues caused a stir in the physics community with a report that experimental data collected at the turn of the century by the Hungarian physicist Roland von Eötvös contained evidence of another fundamental force in nature. The putative weak, repulsive force was said to have an effective range of several hundred meters—longer than the range of two acknowledged forces (the strong and weak nuclear forces) but shorter than the range of the other two (the electromagnetic and gravitational forces). Its strength was held to depend not on the mass or charge of the interacting objects but on a feature that varies with chemical composition: the total number of baryons (protons and neutrons) per unit of mass.

Fischbach's hypothesis unleashed a flurry of experiments, certain of which seemed to support the existence of such a "fifth" force whereas others appeared to rule it out. Recently a group headed by Paul E. Boynton of the University of Washington has performed what is regarded as the most sensitive test for the fifth force. Coupled with the earlier findings, the results appear tantalizingly consistent with the existence of an intermediate-range, composition-dependent force, albeit one whose strength varies not with the total number of neutrons and protons per unit of mass but with the difference in their numbers—a quantity known as isospin.

In the original Eötvös experiment weights made of different materials were hung from the ends of a torsion balance (essentially a bar suspended horizontally from a wire). Eötvös hoped to detect any slight rotation of the bar that would result if the earth exerted an uneven attraction on the weights. Eötvös thought he found nothing beyond statistically insignificant inconsistencies. Yet according to Fischbach and his collaborators, these slight inconsistencies actually reveal a systematic pattern that would be expected if a force that depended on baryon number were working against the earth's gravitational pull.

In its original paper Fischbach's

Today it's almost impossible to communicate fully without text and graphics on the same page.

Lotus Manuscript™ makes it easy to put a column of text and a graphic element side by side.

You can import sophisticated graphics from Freelance® Plus to enhance the communications value of any written document.

Spreadsheets, graphics and charts from 1-2-3® and Symphony® can easily be mixed with text on the same page.

You can throw away your scissors and glue, cut and paste are a thing of the past.

Since early cave drawings, people have found graphics quite effective in communications. Yet in our information-driven society, graphics have taken a back seat to the written word. From typewriters to word processing, graphical elements have been treated like afterthoughts, relegated to "exhibit on next page" or "cut and paste" status.

Lotus Manuscript is the first word processor that is truly a complete document creation system. It's ideal for the needs of technical writers and writers of long complex documents.

Manuscript allows you to easily mix text on the same page as graphics; elements from 1-2-3 and Symphony, graphics from Freelance Plus, or diagrams and scanned images from other sources.

With our Document Preview feature you can see graphics and text on the same page before it's printed, with a zoom capability that lets you take a closer look for proofing your layouts or equations.

Try it like this. *Or, try it like this.*

Our intelligent print formatter gives you great control and flexibility over size and positioning of graphics on the page.

Manuscript is designed to work on most IBM® PCs and compatibles.* Its familiar 1-2-3 interface makes it easy to use. And our Manuscript evaluation kit makes it easy to try. For $10.00, you'll get a presentation disk, working software, and a tutorial manual. To get your evaluation kit, call 1-800-345-1043, ask for lot #BC-1450. Or, for more information, see your authorized Lotus Dealer, or write Lotus Development Corp., 90 Annex, Atlanta, GA 30390-0370.

Lotus Manuscript™

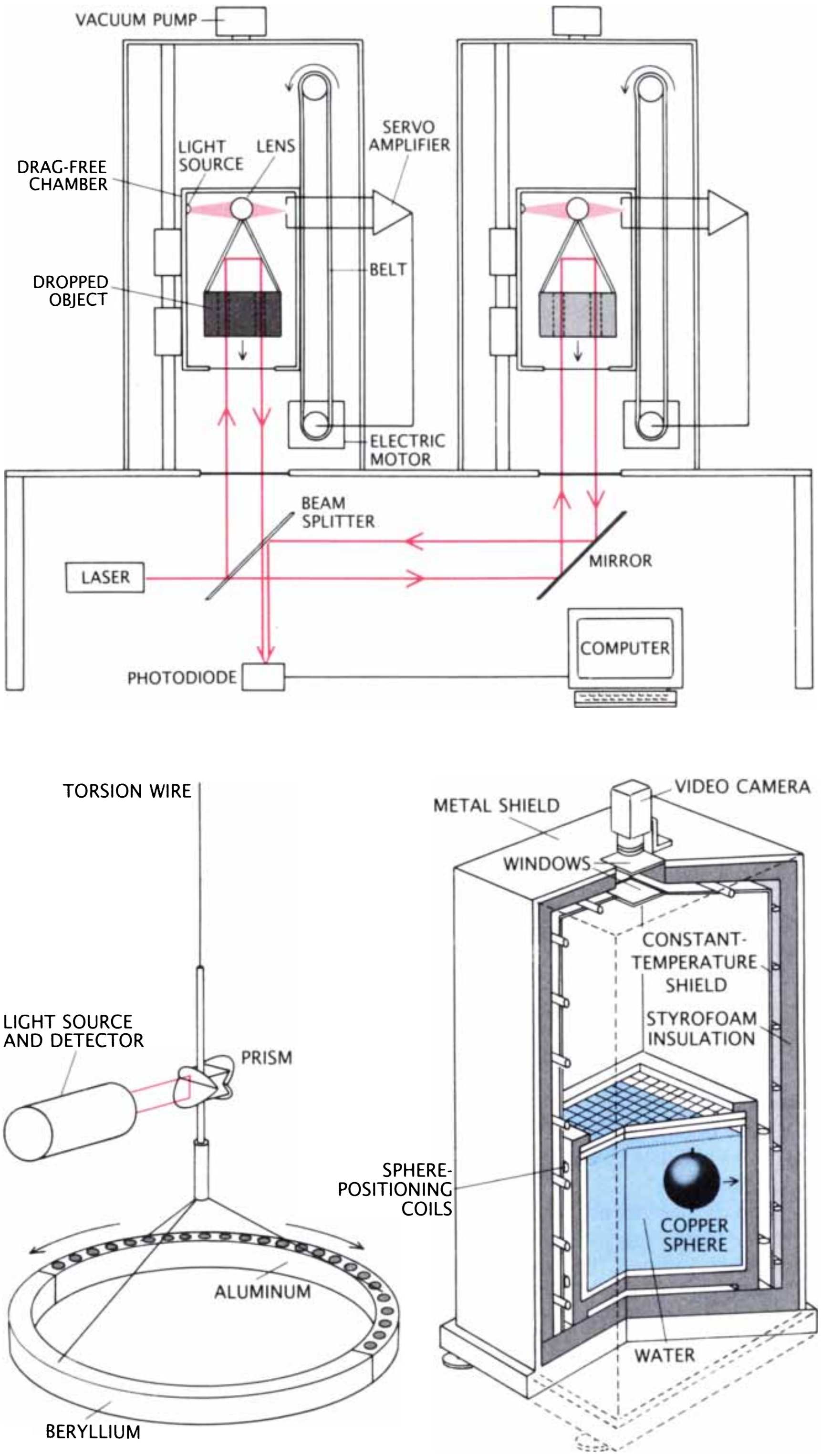
group called for more precise measurements. In due course a group at the University of Washington led by Christopher W. Stubbs and Eric G. Adelberger built a torsion balance and placed it on a hillside, hoping to detect any torque produced by the combination of the gravitational attraction of the hill and some additional effect on the test bodies. Within the sensitivity limits of their setup Stubbs and Adelberger found no evidence of the fifth force.

Timothy M. Niebauer, Martin P. McHugh and James E. Faller of the University of Colorado at Boulder then re-created the classic experiment Galileo is said to have performed from the tower of Pisa. They dropped two test masses of differing composition and measured the difference between their accelerations by means of a laser interferometer. The results of this test (which was less sensitive than the experiment of Stubbs and Adelberger to a short-range force) were negative: the bodies seemed to fall at the same rate.

Such results might well have discouraged other experimenters from pursuing the matter further if Peter Thieberger of the Brookhaven National Laboratory had not derived evidence for the fifth force from a third type of experiment. Since the number of baryons per unit of mass is different for copper and water, Thieberger reasoned that a buoyant copper sphere would move through water if the gravitational attraction of a nearby large mass (namely a cliff) were modified by the fifth force. The results of an experiment designed along these lines, Thieberger reported, were indeed "compatible with the existence of a medium-range, substance-dependent force which is more repulsive (or less attractive) for copper than for water."

Boynton's group has now entered the fray by publishing the results of its own experiment in *Physical Review Letters.* In place of the bar in the traditional torsion balance they employed a ring, half of which was made of beryllium and half of aluminum; instead of measuring small angular displacements they set the ring in motion, creating a torsional pendulum, and looked for changes in the period of the oscillations. When the experiment was done at the base of a granite cliff, the oscillations of the ring varied in a way that suggested "a static interaction of the cliff mass with some kind of asymmetry between the beryllium and aluminum halves of the pendulum ring."

A "fifth" force that counteracts gravity can be reconciled with seemingly contradictory findings

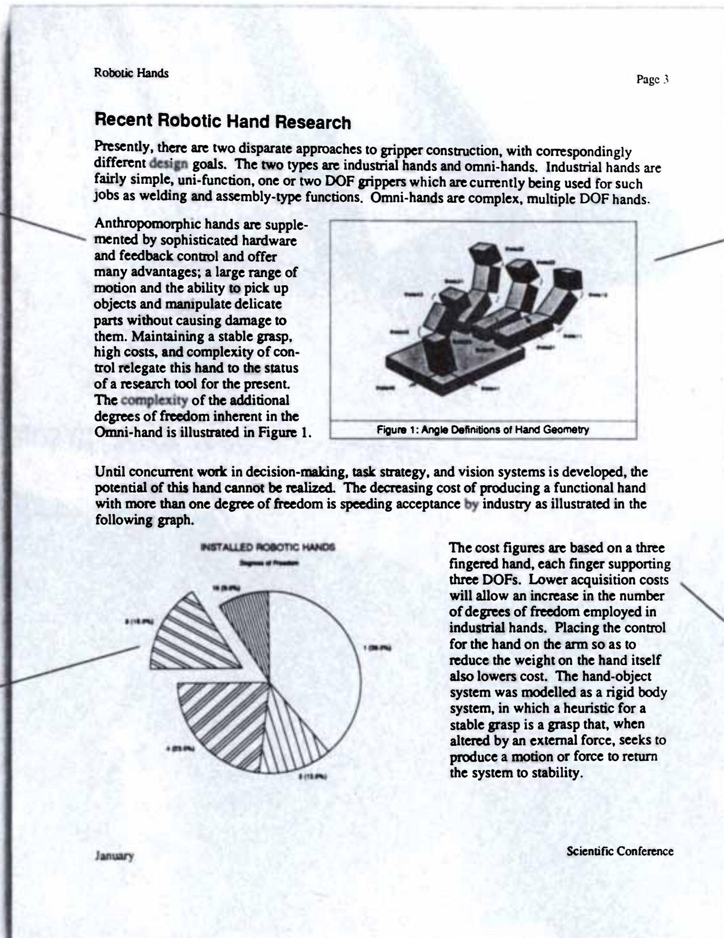


EVIDENCE IN FAVOR of a "fifth" force was inferred from the oscillations of a special torsion pendulum (top left) *and the movements of a buoyant copper sphere* (top right) *placed near a large mass: a cliff. A re-creation of Galileo's classic experiment* (bottom) *yielded negative results. As two masses of differing composition fell in evacuated tubes, a laser interferometer measured and compared their accelerations. To lessen drag, a servomechanism accelerated an inner chamber downward at the same rate as each falling mass.*

Today it's almost impossible to communicate fully without text and graphics on the same page.

Lotus Manuscript™ makes it easy to put a column of text and a graphic element side by side.

You can import sophisticated graphics from Freelance® Plus to enhance the communications value of any written document.



Robotic Hands — Page 3

Recent Robotic Hand Research

Presently, there are two disparate approaches to gripper construction, with correspondingly different design goals. The two types are industrial hands and omni-hands. Industrial hands are fairly simple, uni-function, one or two DOF grippers which are currently being used for such jobs as welding and assembly-type functions. Omni-hands are complex, multiple DOF hands.

Anthropomorphic hands are supplemented by sophisticated hardware and feedback control and offer many advantages; a large range of motion and the ability to pick up objects and manipulate delicate parts without causing damage to them. Maintaining a stable grasp, high costs, and complexity of control relegate this hand to the status of a research tool for the present. The complexity of the additional degrees of freedom inherent in the Omni-hand is illustrated in Figure 1.

Figure 1: Angle Definitions of Hand Geometry

Until concurrent work in decision-making, task strategy, and vision systems is developed, the potential of this hand cannot be realized. The decreasing cost of producing a functional hand with more than one degree of freedom is speeding acceptance by industry as illustrated in the following graph.

INSTALLED ROBOTIC HANDS

The cost figures are based on a three fingered hand, each finger supporting three DOFs. Lower acquisition costs will allow an increase in the number of degrees of freedom employed in industrial hands. Placing the control for the hand on the arm so as to reduce the weight on the hand itself also lowers cost. The hand-object system was modelled as a rigid body system, in which a heuristic for a stable grasp is a grasp that, when altered by an external force, seeks to produce a motion or force to return the system to stability.

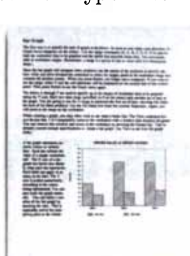
January — Scientific Conference

Spreadsheets, graphics and charts from 1-2-3® and Symphony® can easily be mixed with text on the same page.

You can throw away your scissors and glue, cut and paste are a thing of the past.

Since early cave drawings, people have found graphics quite effective in communications. Yet in our information-driven society, graphics have taken a back seat to the written word. From typewriters to word processing, graphical elements have been treated like afterthoughts, relegated to "exhibit on next page" or "cut and paste" status.

Try it like this.



Or, try it like this.

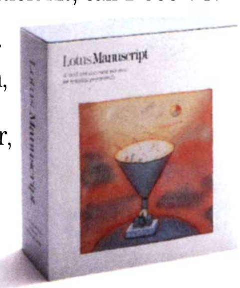
Our intelligent print formatter gives you great control and flexibility over size and positioning of graphics on the page.

Lotus Manuscript is the first word processor that is truly a complete document creation system. It's ideal for the needs of technical writers and writers of long complex documents.

Manuscript allows you to easily mix text on the same page as graphics; elements from 1-2-3 and Symphony, graphics from Freelance Plus, or diagrams and scanned images from other sources.

With our Document Preview feature you can see graphics and text on the same page before it's printed, with a zoom capability that lets you take a closer look for proofing your layouts or equations.

Manuscript is designed to work on most IBM® PCs and compatibles.* Its familiar 1-2-3 interface makes it easy to use. And our Manuscript evaluation kit makes it easy to try. For $10.00, you'll get a presentation disk, working software, and a tutorial manual. To get your evaluation kit, call 1-800-345-1043, ask for lot #BC-1450. Or, for more information, see your authorized Lotus Dealer, or write Lotus Development Corp., 90 Annex, Atlanta, GA 30390-0370.



Lotus Manuscript™

Another radical idea from Volk

Jetta Some carmakers are just now bringing out four-wheel steering.

A pretty good idea. So good in fact, our German engineers looked into it very carefully way back in the Seventies.

But it posed some real problems.

You see, "active four-wheel steering," as good as it is, requires extra moving parts. Extra weight. And extra expense. To you.

That's just not our way of building cars.

So we came up with a unique alternative.

Our patented, track-correcting rear axle.

©1987 Volkswagen Seatbelts save lives.

vagen. Rear wheels that turn.

It turns the rear wheels' alignment into a corner. Providing a smoother track. And reducing "rear steer" in hard cornering.

Another sensible way of making Volkswagens even more exhilarating to drive.

In this complex world, such a simple idea may seem like a radical idea to some.

But to us, it's just German engineering.

The Volkswagen way.

German engineering. The Volkswagen way.

How can these most recent results be reconciled with the other findings, which, if they do not rule out the fifth force altogether, set tight constraints on its strength and range? Boynton and his colleagues write that their results "may be made consistent with recent experimental constraints...if one assumes that the candidate interaction couples almost exclusively to [isospin]."

Robert H. Dicke of Princeton University acknowledges that Boynton's experiment is "an impressive piece of work" but believes such experiments "are notoriously difficult to carry out" without systematic errors. Sheldon Lee Glashow of Harvard University is more skeptical: "I have yet to see decisive and convincing evidence. Until then, I'm not interested."

One issue on which all observers seem to agree, however, is the need for more experimentation. Accordingly Boynton's group is replacing the ring of aluminum and beryllium with one of copper and polyethylene, materials with a greater difference in isospin. The group has also invited Stubbs and Adelberger to set up their experiment in the same area, in order to subject their apparatus to the pull (and push?) of the same cliff face. *—Gregory Greenwell*

Vapor Lock

Bubbles in Antarctic ice record carbon dioxide and climate

The concentration of carbon dioxide in the earth's atmosphere is increasing at a rate of .4 percent each year, primarily because of the combustion of fossil fuels. Carbon dioxide is a "greenhouse gas": it traps heat from the earth's surface that would otherwise be radiated into space. Its continuing increase is widely expected to cause a global warming, and so understanding its role in climate has assumed some importance. Hard evidence about the relation between atmospheric carbon dioxide and climate, however, must come from the past.

Fortunately nature has provided an extensive record of carbon dioxide and climatic change. The record consists of tiny bubbles of air trapped in the polar ice sheets—samples of prehistoric atmospheres preserved for millenniums—together with characteristics of the ice itself.

By recovering ice cores from the unprecedented depth of 2,083 meters, members of the Soviet Antarctic Expeditions at Vostok in East Antarctica have greatly extended the period for which that record is accessible. In collaboration with French investigators from the Laboratory of Isotopic Geochemistry and the Laboratory of Glaciology and Geophysics of the Environment, Soviet workers from the Arctic and Antarctic Research Institute in Leningrad and the Institute of Geography in Moscow derived from the cores a history of both carbon dioxide levels and global temperature over the past 160,000 years—a period beginning in the next-to-last ice age. They report the results in *Nature.*

The technique for estimating temperature relies on the relative abundances of hydrogen and its isotope deuterium in the water molecules locked in the ice. Ice deposited at lower temperatures contains relatively less deuterium, the heavier isotope; at lower temperatures water containing deuterium is more likely to condense and precipitate before it can reach polar regions. Analyses of isotopic ratios, elaborately cross-checked, enabled the investigators to conclude that average temperatures during the two most recent ice ages were nine degrees Celsius lower than today's temperatures; temperatures during the interglacial period about 130,000 years ago were two degrees higher.

When the workers analyzed the air bubbles, they found that during the postglacial warmings, which took only about 2,000 years, atmospheric carbon dioxide increased by about 40 percent. Indeed, carbon dioxide levels closely paralleled temperature over the entire 160,000 years, and so it now seems clear that the gas does play an important role in climate.

In the past, carbon dioxide probably only amplified the more fundamental effects that initiated and ended the ice ages. These are Milankovitch cycles: very slow oscillations in the shape of the earth's orbit and the inclination and direction of its spin axis, which cause slight variations in the amount of heat the planet receives from the sun. Subtle periodicities in the new temperature and carbon dioxide histories reflect such astronomical influences: one cycle has a period of 100,000 years, which matches the timing of recent ice ages, and another has a period of 21,000 years. A 40,000-year cycle is also apparent, but only in the temperature record.

Wallace S. Broecker of the Lamont-Doherty Geological Observatory comments that although the new measurements are "extremely important," their interpretation will be debated. Warming seems to coincide with the carbon dioxide increases, but cooling precedes the decreases in carbon dioxide. Broecker believes the changes in atmospheric carbon dioxide abundance could be explained by supposing that the astronomical changes somehow cause a "major reorganization" in ocean circulation patterns. Such a shift would affect both heat flow around the globe and the amount of carbon dioxide stored in the oceans, thereby amplifying the climatic effects of the orbital changes.

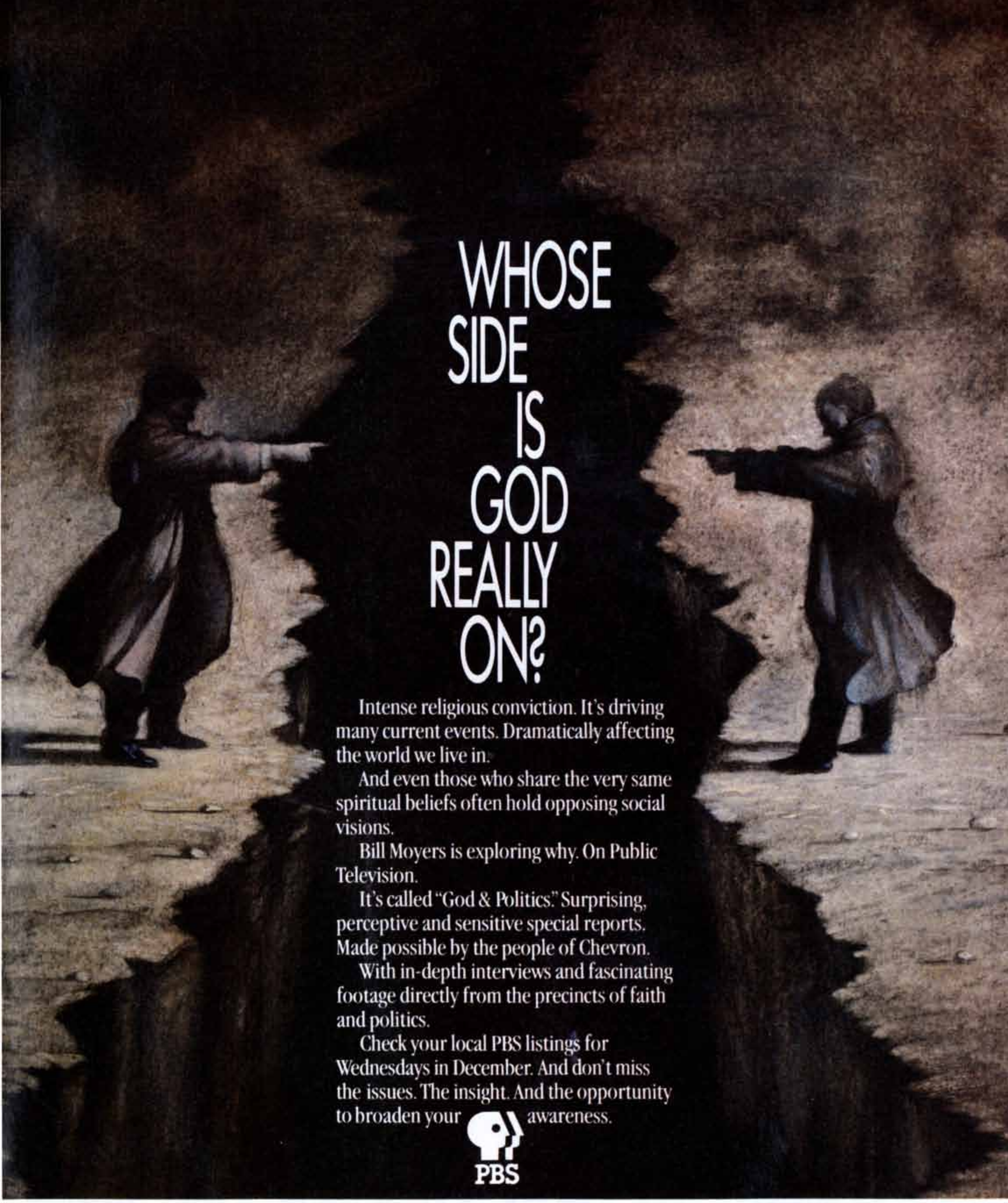
In their account the French and Soviet investigators sound a similar note. They believe the rapid warmings and the concomitant surges in carbon dioxide support the idea that "an abrupt modification of the deep oceanic circulation" can result from a slow external influence, such as the Milankovitch cycles. Unfortunately such ideas cannot yet predict the effects of rising carbon dioxide caused by human beings. *—T.M.B.*

TECHNOLOGY

Electronic Taskmasters

Does monitoring degrade the quality of working life?

Is Big Brother alive and well in the workplace? As more and more clerical tasks are performed at video-display terminals and over telephone lines, private employers are taking advantage of the equipment to monitor employee performance. Computers can be programmed to count keystrokes per minute or sales per hour; telephone conversations can be monitored routinely and silently to ensure that salespeople and operators remain courteous and follow their script. A recent study by Congress's Office of Technology Assessment (OTA), titled "The Electronic Supervisor," estimates that six million workers in industries such as telecommunications, financial services and airlines are routinely evaluated on the basis of monitoring; many more are monitored but do not


WHOSE SIDE IS GOD REALLY ON?
Intense religious conviction. It's driving many current events. Dramatically affecting the world we live in.
And even those who share the very same spiritual beliefs often hold opposing social visions.
Bill Moyers is exploring why. On Public Television.
It's called "God & Politics." Surprising, perceptive and sensitive special reports. Made possible by the people of Chevron.
With in-depth interviews and fascinating footage directly from the precincts of faith and politics.
Check your local PBS listings for Wednesdays in December. And don't miss the issues. The insight. And the opportunity to broaden your awareness.
PBS

RTED TASTE OF BOMBAY GIN.

IS ROOT) FROM ITALY   ALMONDS FROM INDOCHINA   LEMON PEEL FROM SPAIN   LICORICE FROM INDOCHINA get evaluations based on the results.

To the companies that practice it, among them the American Express Co. and AT&T, employee monitoring is sound management. They say that it helps to increase competitiveness and efficiency, since it enables them to provide feedback on employee performance. After all, they say, employees have been assessed for as long as there have been jobs. Unions disagree: the Communications Workers of America (CWA) believes new technology has brought changes that invade workers' privacy and intimidate them. The pressure that monitoring creates to complete transactions quickly can lead to inferior service, the CWA contends, and increases stress-related illnesses such as anxiety, high blood pressure and ulcers. The union has begun a campaign to limit the practice.

The OTA study found evidence that electronic monitoring exacerbates stress-related illnesses to be suggestive, not conclusive. It says the practice raises legitimate concerns, however, about fairness and quality of life, and it acknowledges "strong arguments" that monitoring could be abused. The OTA also believes eavesdropping on telephone calls raises concerns about the privacy of the customer.

The statutes that protect people against wiretapping in their private lives do not extend to employees (and customers) who are subject to such "service observation." The OTA notes, moreover, that "there is no legal right to be treated with dignity or as an autonomous person." But it points out that although the law has traditionally upheld the notion of "employment at will" (which, for the employee, might be roughly translated as "If you don't like it, lump it"), that view has been gradually eroded with the passage of antidiscrimination and safety legislation. Legislation to protect privacy in the workplace would therefore not be totally without precedent. Bills have been introduced in Congress that would prohibit service monitoring of telephone conversations without a beep tone to alert the participants.

The OTA report points out that similar threats to employee privacy are posed by other technologies. Polygraph screening and testing for drug residues are becoming widespread, and emerging biomedical technologies may soon enable organizations to screen employees for susceptibility to a large number of diseases. A system that is said to be capable of detecting drug abuse by analyzing brain electrical activity is already being marketed, and research in progress suggests that analysis of brain waves can uncover specific "guilty knowledge"—for example knowledge relating to a crime. —*T.M.B.*

A Mild Alternative?

Government takes a new look at a coal-gasification process

Oil has become the lifeblood of industrial society because its products, such as gasoline and fuel oil, are ready sources of energy; it is also a feedstock for the chemical industry. Coal can serve many of the same functions, but it is much harder to transport and process, largely because it is a solid. It has the great virtue, however, of being domestically abundant.

The search for better ways to exploit coal and its products has led the Department of Energy to reexamine the process it calls mild gasification, which was first discovered in the early 1800's. When coal is heated to between 1,100 and 1,500 degrees Fahrenheit in the absence of oxygen, it breaks down, producing a fuel gas, liquid hydrocarbons and a solid residue called char. (This process is distinct from the coal-gasification process employed in existing large-scale plants, in which coal is heated to much higher temperatures in an oxygen-rich atmosphere.) The Energy Department has recently awarded four contracts worth $13.5 million over the next three years for studies of mild gasification and its economic prospects. If the indications are good, the department will commission a demonstrator plant capable of processing 1,000 tons of coal a day.

Workers agree that mild gasification will be economical only if high-value applications are found for the products and the process is then adjusted to maximize returns. A mild-gasification plant built in New Jersey some years ago was a commercial flop because power companies were reluctant to buy the char for fuel: its high ash content makes it harder to burn than coal. The Electric Power Research Institute in Palo Alto studied the process more recently and also could not make it pay.

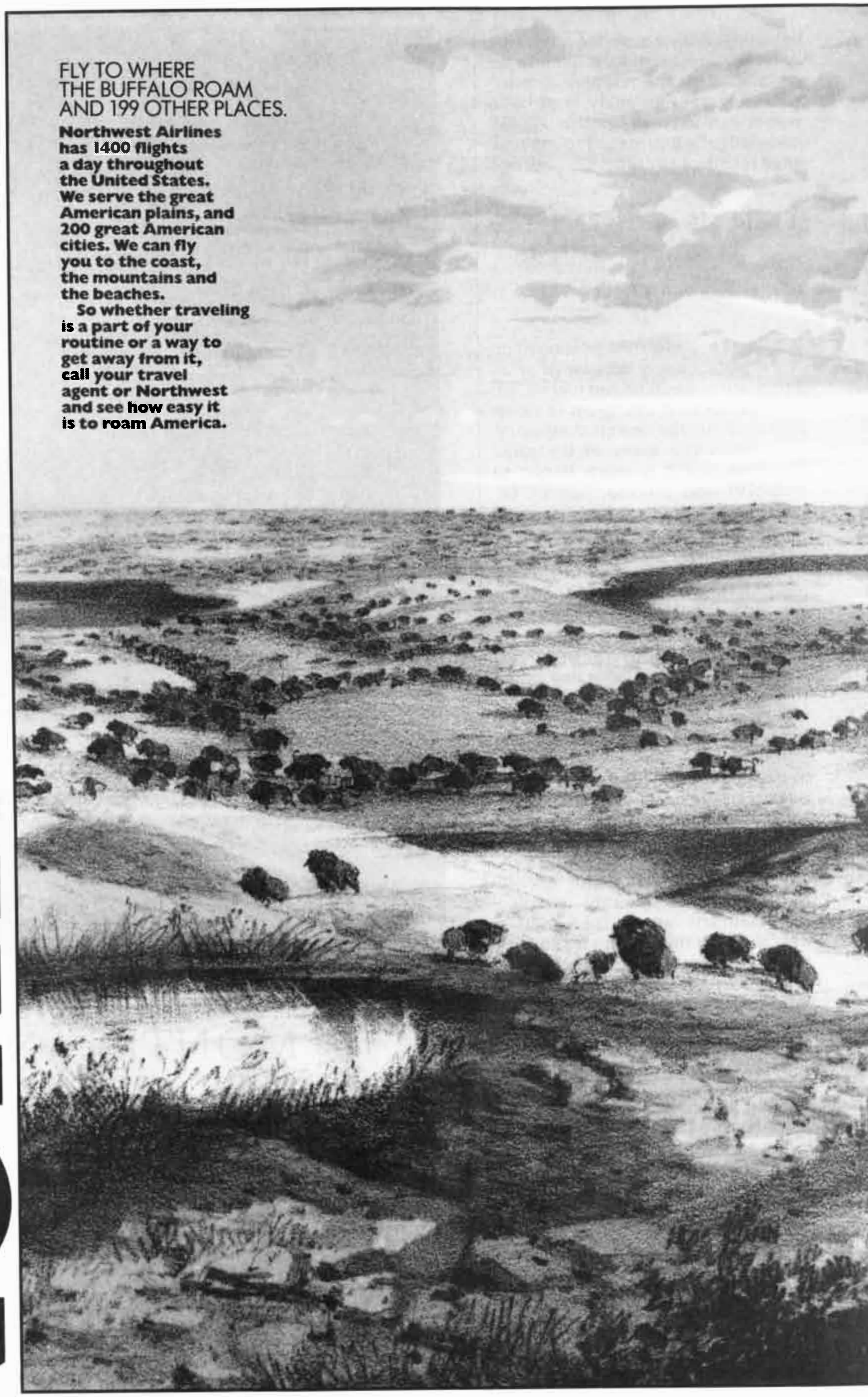
The gas product provides an industrial fuel. The liquid hydrocarbons could be valuable feedstocks for specialty chemical production, accord-

NORTHWEST

FLY TO WHERE THE BUFFALO ROAM AND 199 OTHER PLACES.

Northwest Airlines has 1400 flights a day throughout the United States. We serve the great American plains, and 200 great American cities. We can fly you to the coast, the mountains and the beaches.

So whether traveling is a part of your routine or a way to get away from it, call your travel agent or Northwest and see how easy it is to roam America.



LOOK TO US **NORTHWEST AIRLINES**

U.S. RESERVATIONS **1-800-225-2525**, INTERNATIONAL RESERVATIONS **1-800-447-4747**

THE BETTMANN ARCHIVE ©1987 Northwest Airlines, Inc.

ing to Robert O. Ness, Jr., of the Energy Research Center at the University of North Dakota. They might also serve as fuels; the Energy Department has shown that they can be converted into a high-energy jet fuel for the military and is investigating the feasibility of burning them in locomotives instead of diesel fuel. Mixed with char, the liquid products might also make a high-yield boiler fuel for industrial applications.

Refinements of the mild-gasification process could make it possible to convert as much as 40 percent by weight of the coal into hydrocarbon liquids, according to Madhav R. Ghate of the DOE's Energy Technology Center at Morgantown, W.Va.; past yields reached only 20 percent. Ghate has found that yields can be increased by forcing inert gases through the vessel to sweep out condensable products before they are broken down. Pretreating the coal with steam also improves yields by aiding chemical decomposition.

The char that is left behind is also finding new applications. John M. Lehto of Northern States Power, Inc., has demonstrated a process that employs it for smelting taconite, a low-grade iron ore; in addition the process yields heat, which could make it possible to produce iron and power at the same site. Ness believes Lehto's scheme could consume enough char to make a mild-gasification plant economical. Because char is rich in elemental carbon, it can also serve in the production of activated carbon, carbon fibers, graphite electrodes and hydrogen (made by reacting char with steam).

If the Energy Department shows mild gasification to be economically viable, plants might eventually be built at coal mines to minimize transportation costs, according to Ghate. The process does raise environmental concerns. Coal products contain high levels of sulfur and nitrogen, which are released in harmful emissions when the products are burned, but Ghate believes control technologies are feasible. —*T.M.B.*

BIOLOGICAL SCIENCES

Flocking Together

It may help ospreys to fish more efficiently

During bluefish season on Cape Cod, Mass., one often sees a few fishermen casting fruitlessly into the surf while many others look on. As soon as someone hooks a fish, the others—knowing that bluefish travel in schools—jump from their beach chairs and jeeps to try their own luck.

Ospreys, large fish-eating hawks found in North America, take advantage of one another's efforts in a similar way, according to Erick Greene of Princeton University. Greene spent five months watching a colony of ospreys nesting on a small, spruce-covered island in Cow Bay estuary, Nova Scotia; the colony, he reports in *Nature,* served as an "information centre" on good fishing spots.

ADULT OSPREY AND CHICK, about five weeks old, share a nest built high in a spruce tree on an island in Cow Bay estuary, Nova Scotia. The photograph was made by Erick Greene.

The 11 nesting couples built their nests high in the tallest spruces on the island, from which the birds had clear views of their fishing grounds. They fed largely on four fish species: winter flounder, pollock, alewife and smelt. Winter flounder live scattered on the bottom of shore waters and estuaries; the other three species commonly travel in large schools.

Male ospreys do most of the fishing during the breeding season, bringing their catches back to the nest and sharing them with their chicks and mates. Other ospreys can easily distinguish what kind of fish a returning bird clutches in its talons (the eyesight of ospreys is about 100 times keener than that of human beings). Greene observed that hawks on the island often flew toward a bird returning with a schooling fish and retraced its flight. If the bird carried a winter flounder, the other hawks usually stayed put.

Occasionally a successful forager seemed to communicate actively to others in the colony by persistently uttering shrill cries and rapidly swooping up and down. Greene noted that an osprey performed this display only if it had caught a schooling fish; moreover, such displays usually followed long periods when none of the ospreys had caught any fish.

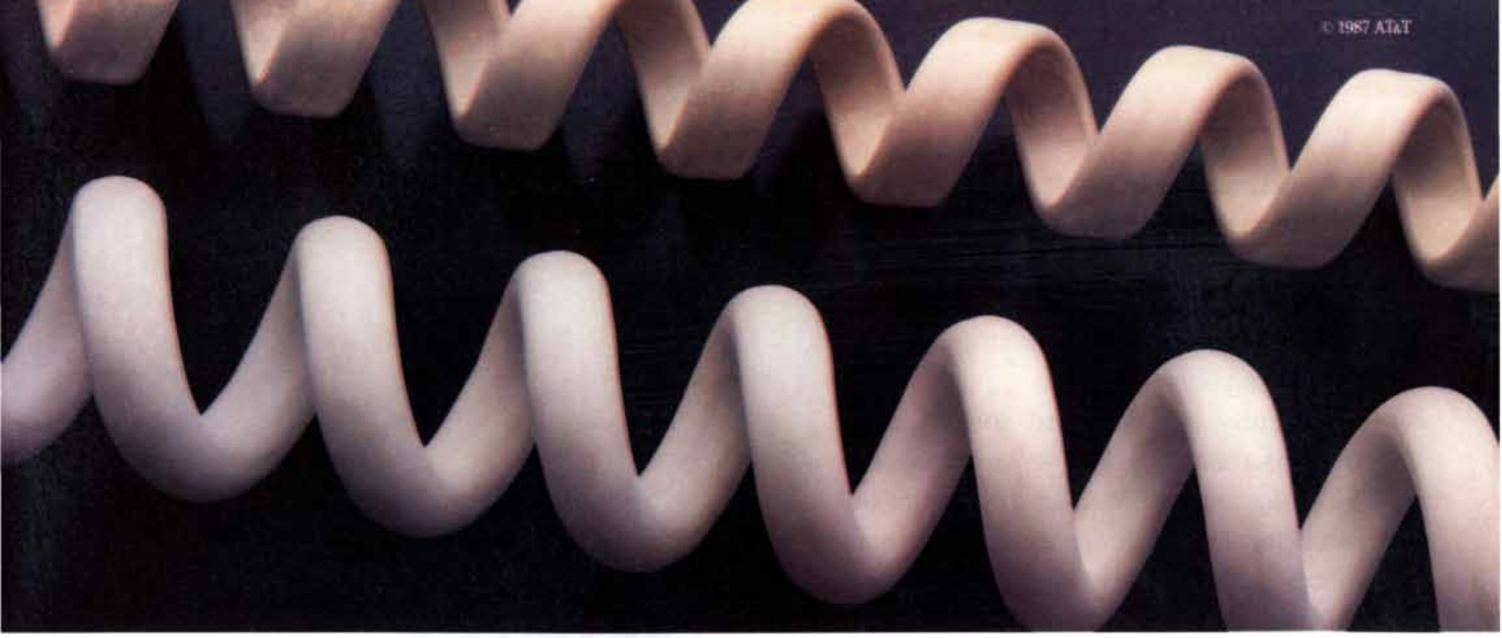
"Although these displays were uncommon," Greene recalls, "the response of other ospreys to this display was dramatic: all of the non-fishing birds at the colony immediately flew towards the displaying bird and began hunting near where the fish had been caught."

Greene speculates that kin selection may explain this actively altruistic behavior. He points out that whereas female ospreys range quite far in search of mates, male ospreys usually settle near their birthplace. In alerting males nesting nearby to the presence of a school of fish, an osprey might help its brothers or cousins to feed their offspring and so perpetuate the family genes. —*J.H.*

A Nervous Disposition

Some cells in an embryo's skin are "predisposed" to be nerves

A few hours after fertilization the embryo of a vertebrate is a smooth sphere with an outer layer of apparently identical cells called the



QUICK! WHICH IS THE PHONE, AND WHICH IS THE COMPUTER?

Once, a phone was a phone and a computer was a computer.

And anybody could tell the difference.

Today, however, telephones routinely boast computer memories, computer intelligence, even computer screens.

And computers are discovering the power of networking, as telephones already have.

The name of the game is getting the right information to the right people at the right time.

Because being able to do that easily will make our work more productive and our lives less hassled.

And that's what AT&T is all about today.

Our goal is to do for information what we've already done for conversation.

To accomplish that, the people at AT&T Bell Laboratories are working to combine everything you like about telephones with everything you expect from computers.

So computers will finally deliver what they have been promising for decades.

Consider some examples: Some day soon, instead of being limited to the new cars available in a dealer's inventory, you'll be able to sit down at a computer terminal in the dealer's showroom and enter the model you want, the engine, the options, the color, etc.

A data network will automatically translate your order into instructions to dozens of suppliers and plants in the production process. The result is a custom-made car delivered faster than you ever thought possible.

In banking, powerful and versatile data networks could mean loan approval in minutes instead of weeks. And account balances that are always current.

In retailing, a chain whose stores employ a data network could always have hot-selling items in the stores where they're selling hottest. Pleasing both the customer and the company controller.

The idea is networks which not only move information instantly, but which also interpret it, rearrange it and apply it in the most useful way. All automatically.

Just as the AT&T long distance network handles a telephone call, instantly, intelligently, automatically.

Which is to say, we're very close to the day when you won't be able to tell a phone from a computer, and won't even care.

But until then, the phone cord is the one on the top.

AT&T
The right choice.

Sometimes you're confronted with a situation that can seem overwhelming at first. How it turns out is solely determined by how you respond. And how you respond is a reflection of how strongly you believe in yourself.

Recently, we've faced some tough times as a company. But we've responded by being tougher. We had to in order to build the leaner, more efficient, better disciplined company we have become. We had to ask ourselves what kind of company did we want to be?

We decided we want to be an integrated domestic petroleum company with an emphasis on foreign exploration. For the time being, most of our capital will go to developing proven reserves. We are suspending high-risk, high-reward exploration and sticking to the more mature oil basins. And while we are spending less, we are spending smarter.

We will continue to support our main business through strong research and development as well as stronger reliance on high technology. It was this unshakeable commitment that resulted in our holding more active U.S. patents than any other oil company last year.

It's true we came face to face with adversity, but we'll come roaring back. Because we're ready to take on the future.

For more information please write to Patricia Marshall, Phillips Petroleum, 16D-4 Phillips Building, Bartlesville, Oklahoma 74004.

Phillips Petroleum Company
Performing to meet
the challenge of change.

PHILLIPS 66

TRUE CHARACTER IS REVEALED WHEN

YOU COME FACE TO FACE WITH ADVERSITY.

ONE-DAY-OLD EMBRYO of an African clawed frog already displays a dark neural plate in a micrograph by John B. Gurdon.

ectoderm. Grown alone in culture, ectoderm cells become skin cells, and in fact most of the ectoderm is destined to become skin. But during a process known as gastrulation a layer of cells within the embryo called the mesoderm pushes up against a strip of the ectoderm; this strip differentiates into a new layer of cells known as the neural plate. Visible initially only as a dark ridge along the dorsal, or top, side of the still spherical embryo, the neural plate is the precursor of the brain and spinal cord. What triggers its appearance?

Biologists have long thought that the external influence of the mesoderm spurs certain ectoderm cells to differentiate into neural tissue. This process of "induction" requires direct contact between the ectoderm and the mesoderm. Recent studies of the development of *Xenopus*, the African clawed frog, suggest that an internal program may also play a role. Colin R. Sharpe and John B. Gurdon of the University of Cambridge and Andreas Fritz and Eduardo M. De Robertis of the University of California at Los Angeles School of Medicine report in *Cell* that some ectoderm cells, even before they come in contact with the mesoderm, have a special "predisposition" toward becoming neural tissue.

A key step in the group's research was its isolation of a DNA fragment that is expressed in cells of *Xenopus'* developing neural plate. De Robertis notes that the fragment "is probably not only a marker of neural differentiation but also the causative agent." To stress this possibility the investigators call the marker a homeobox, a term first applied to a similar DNA sequence found to control the early development of the fruit fly *Drosophila*.

The workers took ectoderm cells from various regions of the undifferentiated embryo and combined them in vitro with mesoderm cells. The "surprising result," the team observes, was that ectoderm from the dorsal region of the embryo expressed the homeobox marker with greater frequency—that is, showed a greater sensitivity to induction—than ectoderm from the bottom, or ventral, region. The implication, according to Gurdon, is that even before gastrulation takes place changes have been wrought in the dorsal cells, preparing them for their transformation into neural tissue. Some earlier form of induction may be responsible, he says.

De Robertis hypothesizes that the ectoderm of mammals—including human beings—also differentiates into neural tissue in response to both induction and internal cues. The human genome, he notes, contains genetic sequences that resemble the homeoboxes of *Xenopus* and *Drosophila*. Confirming this hypothesis, however, will be difficult: in the first stages of differentiation mammalian embryos are thousands of times smaller than the embryos of amphibians like *Xenopus*. —*J.H.*

Smoking Gun?

Activation of an oncogene may be a cause of lung cancer

A decade of research points to a cohort of normal genes that, when altered, have some role in the development of cancer. These proto-oncogenes, as they are called, code for proteins that are vital to normal cellular processes. Structural abnormalities in or around the genes can activate them and convert them into oncogenes, which contribute to the malignant transformation of normal cells. The nature of that contribution is unclear, however, and clinical evidence about the role of oncogenes in human cancer is still scarce. Dutch investigators led by Sjoerd Rodenhuis of the Netherlands Cancer Institute have helped to fill that gap. Their data, published in the *New England Journal of Medicine*, implicate the K-*ras* oncogene in the development of five out of 10 studied cases of adenocarcinoma of the lung.

An important mechanism in the activation of proto-oncogenes is point mutation: a one-base substitution in the DNA that results in a corresponding one-amino-acid difference in the protein product of the gene. Three members of the *ras* gene family, H-*ras*, N-*ras* and K-*ras*, acquire carcinogenic potential by point mutations at positions 12, 13 or 61.

Rodenhuis' team examined 39 tumor specimens from patients with various types of untreated lung cancer. A highly sensitive hybridization assay revealed the significant point mutations in the three *ras* genes. Point mutations were observed only in K-*ras*, at position 12, and only in adenocarcinomas (which account for some 30 percent of lung cancers). The data support a previously observed connection between the K-*ras* gene and adenocarcinomas.

The *ras* genes code for proteins, termed p21, that are strikingly similar to G proteins, which help to pass growth signals from the cell surface to the nucleus. The *ras* proteins, like the G proteins, are at the inner side of the cell membrane, and they too seem to be part of a signal system. They are thought to have both "on" and "off" configurations. If a point mutation occurs at position 12, 13 or 61, Rodenhuis explains, the protein "sticks at 'on' and transduces a signal that in fact is not there." Such a mutation, Dennis J. Slamon of the University of California at Los Angeles School of Medicine writes in an editorial accompanying the Rodenhuis paper, is "clearly...a possible pathogenic event in lung adenocarcinomas."

Pinning down that possibility is difficult. A number of successive changes, perhaps including the activation of one or more oncogenes, seem to figure in the multistage initiation and promotion of cancer. To test where *ras* activation falls in the development of adenocarcinomas, Rodenhuis' team examined tumors in as early a stage as possible. The workers concluded that the activation of *ras* is a relatively early change in lung adenocarcinomas: the point mutations occurred in tumors less than two centimeters in size.

Many more patients will have to be studied, the workers write, in order to explore differences in the course of adenocarcinomas with an activated *ras* gene and those without one. Because all five patients with activated K-*ras* were heavy smokers, however, "it is tempting to speculate that activation of K-*ras* by a carcinogen in tobacco smoke may provide an 'alternative pathway' for the development of an adenocarcinoma, bypassing the need for one or more unidentified carcinogenic events unrelated to smoking." —*Elizabeth Collins*

Physiology or Medicine

Biologist is honored for cracking the mystery of antibody diversity

Good health is the result of an unremitting battle fought in the bloodstream. There the immune system wards off the continual threats posed by pathogens and toxins. For each invading agent white blood cells called *B* lymphocytes secrete specific protein tags, or antibodies, that bind to the invader and mark it for destruction. The *B* cells make a staggering variety of antibodies; they can respond to a virtually limitless array of invaders.

How the immune system generates its impressive repertoire of antibodies was a mystery just 12 years ago. Today the mechanism of antibody diversity is well understood. Much of the credit for the explanation goes to Susumu Tonegawa, a biologist at the Massachusetts Institute of Technology who recently won the Nobel prize in physiology or medicine.

The question Tonegawa faced in the early 1970's was one that had stumped immunologists for years: How can a finite set of genes dictate a seemingly infinite variety of antibodies? If each antibody were coded for by a separate gene, a *B* cell's genome would have to carry an improbably large number of antibody genes. Alternatively, the genes for the various components of antibody molecules might join together in different combinations in different white blood cells. At the time, however, no one knew a mechanism by which this rearrangement could occur.

In 1976, when Tonegawa was at the Basel Institute for Immunology, he and his colleague Nobumichi Hozumi obtained the first evidence that antibody genes are in fact shuffled during cell maturation [see "The Molecules of the Immune System," by Susumu Tonegawa; SCIENTIFIC AMERICAN, October, 1985]. The workers used techniques from the budding field of genetic engineering to reveal that the DNA sequences encoding different parts of a single antibody molecule are far apart in the genomes of embryonic mouse cells, whereas in mature antibody-secreting cells they are closer together.

Over the next few years Tonegawa and other investigators put together a dynamic profile of antibody genes, showing how as few as 1,000 coding regions can give rise to millions of specific antibodies. They described the way in which cellular enzymes "cut and paste" strands of DNA to vary the components of the resulting antibody molecule. When mutations were also demonstrated to be a source of antibody diversity, it became clear how a single immune system could produce more than a billion distinct antibodies.

Tonegawa's research has already furthered efforts to understand vaccine efficacy, autoimmune disorders and the rejection of transplanted organs. The broader significance of his discovery a decade ago lies in its rebuttal of the notion of an individual's genome as a static, immutable entity. The idea of a plastic genome—one that can change during an organism's lifetime—has ramifications far beyond immunology. *—Karen Wright*

Physics

IBM duo wins for seminal work in superconducting ceramics

Barely two years ago K. Alex Müller and J. Georg Bednorz of the IBM Zurich Research Laboratory were trying in vain to find materials that could conduct electricity without resistance at temperatures higher than 23.3 degrees Kelvin (degrees Celsius above absolute zero). Their subsequent success with a ceramic metal oxide transformed superconducting from an extremely costly, esoteric technology into one that promises to become integral to a host of industries. For their achievement the Swiss workers have received the Nobel prize for physics.

Before the breakthrough, research into superconducting materials had been stagnant. The previous holder of the record temperature for superconducting—an alloy composed of the metals niobium and germanium—had been discovered more than a decade earlier. Cooling niobium-germanium to below its transition temperature of 23.3 degrees K. requires bathing it either in liquid helium (boiling point 4.2 degrees K.), which is expensive and difficult to keep from evaporating, or in liquid hydrogen (boiling point 20.3 degrees), which is combustible.

Investigators hoped someday to find a material that could superconduct while being cooled by liquid nitrogen (boiling point 77 degrees K.), which is cheap, safe and easy to insulate. Many workers, frustrated by prolonged failure, decided they were confronting a fundamental limitation of nature and dropped out of the field. Most of those who remained explored two-metal alloys similar to niobium-germanium.

In a speech this past spring Müller recalled why several years ago he and Bednorz decided to examine a class of materials that other investigators considered unpromising: ceramic metal oxides. The laboratory had "a tradition of two decades of research" into oxide conductors, Müller said, and had found that many of these materials show strong "electron-phonon interactions." Phonons, which are vibrations in the crystalline lattice of a conductor, help to induce superconductivity.

Early in 1986 Müller and Bednorz found a mixture of lanthanum, barium and copper oxide that showed an abrupt drop in resistance to electricity at 30 degrees K. They published a somewhat cautious report in October, 1986, in the German physics journal *Zeitschrift für Physik.*

Müller noted that the results "at first were met with skepticism in the U.S. and Europe." When a group at the University of Tokyo reported at a meeting in November that it had confirmed the results, investigators worldwide began to experiment with other ceramic metal oxides. Within months researchers at the University of Houston led by Paul C. W. Chu developed a compound of yttrium, barium and copper oxide that superconducts at more than 90 degrees K. The race continues: recent, unconfirmed reports describe materials that superconduct at room temperature and above. *—J.H.*

Chemistry

Workers are lauded for making enzymelike "host" molecules

One of the most active fields in modern chemistry relies not just on apparatus that is visible on laboratory benches but also on a set of molecular tools: molecules that play "host" to other, "guest" molecules, selectively recognizing and coupling to them in much the same way as enzymes interact with specific substances in an organism. By means of host molecules other chemicals can be measured and manipulated; host-guest chemistry underlies precise assays, processes for capturing and eliminating toxic substances, and techniques for assembling new materials from elusive and unstable components. For their pioneering and independent achievements in this large and rapidly expanding field, Charles J. Pedersen, formerly of E. I. du Pont de Nemours & Company, Jean-Marie Lehn of the Louis Pasteur University in Strasbourg and the Collège de France in Paris and Donald J. Cram of the University of California at Los Angeles have won the Nobel prize in chemistry.

Pedersen is generally recognized as the founder of host-guest chemistry. In 1967, just two years before he retired from du Pont at the age of 65, he developed a class of molecules known as crown ethers. These two-dimensional loops of carbon, hydrogen and oxygen can couple to ions—charged particles that are normally highly reactive—to form a new, relatively nonreactive molecule. Lehn improved on crown ethers with a three-dimensional molecule, known as the cryptand, that completely "cages" the guest; cryptands are more selective than crown ethers and form stabler molecules. Cram developed yet another type of host, called a spherand, that is still stabler and more selective.

In recent years, according to James L. Dye of Michigan State University, workers around the world have synthesized thousands of new host molecules. "Natural enzymes," Dye notes, "are the most specific catalysts that exist, but they tend to be unstable and degrade. The goal is to make artificial enzymes that are robust for chemical processing." Workers can design a host molecule, for example, to capture a small, difficult-to-detect guest molecule in a solution. The product of the coupling is easier to detect and measure than the guest alone and is easier to separate from the solution. Hosts have been designed to bind to toxic or radioactive chemicals and either render them harmless or make them easier to eliminate from the environment. Workers have also developed hosts for research in biology. Lehn, for example, has synthesized a molecule that couples with the neurotransmitter acetylcholine. Other investigators have created new materials based on the host-guest interaction [see "Electrides," by James L. Dye; SCIENTIFIC AMERICAN, September, 1987]. *—J.H.*

The Clinical Spectrum

The Clinical Spectrum

Medical diagnoses sometimes depend on the ability to trace or detect minute amounts of biological species. Now researchers at the General Motors Research Laboratories have developed a method of spectrometry using a tunable diode laser that could lead to simpler, less costly, non-invasive diagnostic techniques.

Enhanced Resolution

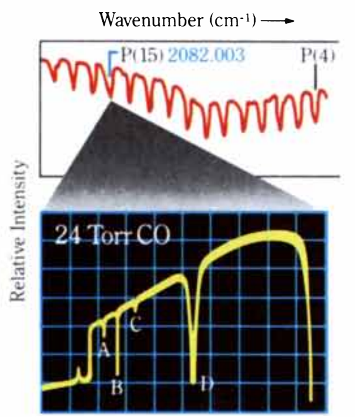


Figure 1: (Top) The absorption spectrum of CO obtained with a conventional spectrometer showing the P series rotation-vibration transitions separated by about 4 cm-1. (Bottom) The diode laser spectrum centered at $^{12}C^{16}O$ P(15) region showing the complete resolution of $^{12}C^{17}O$ P(9), $^{12}C^{18}O$ P(3) and $^{13}C^{16}O$ P(4) transitions.

Figure 2: The second harmonic detection of the $^{13}C^{16}O$ and $^{12}C^{16}O$ as naturally present in exhaled human breath.

In the process of living and growing, the body routinely takes in chemicals in the air we breathe and the food we eat, uses them, and converts them into other chemicals. These chemical activities, therefore, are often very good indicators of the health of the body or of its individual systems. The detection and measurement of particular chemical species is also of value in environmental, scientific and engineering studies.

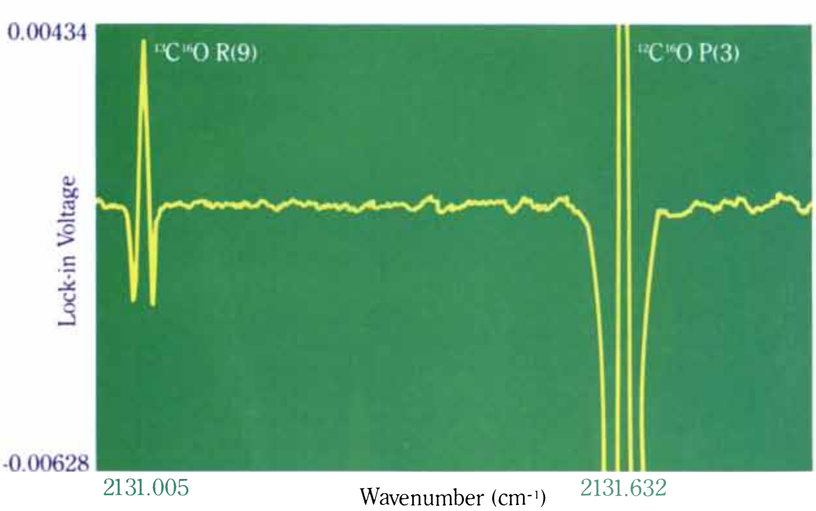
Radioactive isotopes of elements in these chemicals have been extensively used as tracers. Many investigations, however, preclude their use either because no suitable radioisotope is available, or because radiation exposure raises health or environmental concerns.

The use of stable, non-radioactive isotopes for detection and tracing predates that of radioisotopes. But routine application of stable isotopes has been hindered by the lack of a detection method as versatile and simple as the scintillation counting used for radioisotopes. Mass spectrometry is the traditional method of detection of stable isotopes, but it requires extensive sample preparation, expensive equipment, and a highly trained operator to distinguish and measure chemically different molecules of the same nominal mass—nitrogen gas $^{14}N^{14}N$ and carbon monoxide $^{12}C^{16}O$, for example.

It was this need for high resolution and greater versatility that prompted Dr. Peter S. Lee and Richard F. Majkowski to develop a system for stable isotopic tracer analysis based on the molecular absorption of infrared light. A tunable, single-mode diode laser, developed originally by the Physics Department of the General Motors Research Laboratories to measure automobile exhaust gases, was used as the IR emitting source in what has proved to be a remarkably sensitive spectrometer.

The infrared absorption spectrum of molecules normally consists of transitions between series of vibration-rotation energy levels. When an atom in a molecule is replaced by an isotope of the same element, there is a shift in the energy levels due to a change in mass. The resulting frequency shift in the transitions forms the basis of the laser spectroscopic analysis system.

In the case of carbon monoxide, for example, there are six possible forms of the molecule involving stable isotopes: $^{12}C^{16}O$, $^{12}C^{17}O$, $^{12}C^{18}O$, $^{13}C^{16}O$, $^{13}C^{17}O$, and $^{13}C^{18}O$. Consequently, there would be six sets of overlapping spectral lines. Within a region of 1 cm-1, there can be lines from several isotopic molecules, with as little as 0.1 cm-1 or



Carbon monoxide in exhaled human breath.
The scale has been expanded to show the excellent signal-to-noise ratio for $^{13}C^{16}O$.
Other than removal of water vapor, no specific sample preparation or separation was needed.

less between adjacent lines.

This adjacency presents no problem for a diode laser system. The spectral resolution (the laser linewidth) is typically better than 10^{-4} cm^{-1}, which is orders of magnitude less than the isotopic line spacings. Since the diode laser is tunable, it can be centered in a region where the absorption lines of several isotopic molecules can be scanned within a single longitudinal laser mode (Figure 1).

In the initial experimental system, the source of the monochromatic IR radiation was a diode laser, made out of a single crystal containing layers of doped lead telluride and a lead-europium-selenium-telluride alloy. The IR light was collimated through a cell containing the sample to be studied and then focused onto an IR detector.

The cell was designed to have two optical path lengths that can be varied so that isotopic molecules with vastly different abundances can be determined from the measurement of the incident and transmitted laser intensities. U.S. Patent 4,684,805 covers this spectroscopic detection system.

The laser system can be made extremely sensitive using wavelength modulation and harmonic detection. Figure 2 shows the detection of $^{13}C^{16}O$ in exhaled human breath, where $^{13}C^{16}O$ is naturally present at a typical level of 1 to 10 parts per 100 million.

The present system can be used to measure stable oxygen isotopes in biological and organic samples that can be converted into CO. However, the method is applicable to any sample that can be converted into a gas with a suitable IR absorption spectrum.

"The use of radioisotopes as tracers is already well established," says Dr. Lee. "The potential is just as great for stable isotopes if more versatile analytical methods are made readily accessible.

"Packaged as a simpler, relatively inexpensive instrument, a tunable laser IR system could be adapted to many clinical tests—for fat malabsorption, ileal dysfunction, small-intestine bacterial overgrowth, alcoholic cirrhosis and liver function, lung function, nutritional assessment, and diabetes, to name a few.

"Diabetes could be diagnosed from the lung exhalate of a subject who had been fed a stable isotopically tagged sugar sample. No taking blood, no long waits, no radiation health and safety concerns.

"Simpler isotopic tracer measurements could broaden the scope of tracer methodologies, could supplement some of the radioisotope studies now common, and could have significant economic implications."

General Motors

THE MEN BEHIND THE WORK

Dr. Peter S. Lee (right) is a Senior Staff Research Scientist in the Biomedical Science Department at the General Motors Research Laboratories. He received his undergraduate degree in Chemistry from the National Taiwan University. Dr. Lee also holds a Ph.D. in Physical Chemistry from the University of Illinois at Urbana-Champaign. His current research interests at GMRL include the study of biosensors and laser spectroscopy along with his work in stable isotopes. Dr. Lee came to GM in 1977 from the University of Illinois Medical Center in Chicago.

Richard F. Majkowski was, at the time of the work described here, a Staff Research Scientist in the GMRL Physics Department. Both his B.S. and M.S. degrees are from the University of Detroit in Physics and Mathematics. His research interests have included emission spectroscopy, coherent optics, holography and laser spectroscopy. Dick joined General Motors Research Laboratories in 1955 and retired in September, 1987, to become a Professor of Physics at Lawrence Institute of Technology.

SCIENTIFIC AMERICAN December 1987 Volume 257 Number 6

Technology in Services

Service industries can be very technology-intensive. They can stabilize U.S. employment, make U.S. manufacturing industries more competitive and support an ever higher standard of living

by James Brian Quinn, Jordan J. Baruch and Penny Cushman Paquette

In the decades since World War II the provision of services has displaced manufacturing as the largest element in the economy of virtually all advanced nations. In the U.S. the service industries, broadly defined, now account for 71 percent (or $2,996 billion) of the gross national product and 75 percent (or 81.4 million) of all jobs. The services sector continues to grow, although there are signs that its rate of growth may be slowing. In contrast, total employment in manufacturing has declined slightly over the past 15 years; although manufacturing output has continued to grow in real terms, some traditional and highly visible industries, such as basic steel and automaking, have experienced pronounced declines in both output and employment.

For the past two centuries it has been new technologies in agriculture and manufacturing that have driven economic growth and created steady increases in U.S. standards of living. What will be the long-term effect of technologies in the service industries? Can they have similar favorable consequences for the economy and for human welfare? To examine these questions we undertook a three-year study under the auspices of the National Academy of Engineering, with the support of the Bell & Howell Company, the Banker's Trust Company, the Royal Bank of Canada, Braxton Associates and the Bell Atlanticom Corporation.

This article is an early report on our findings. What may come as a surprise is the finding that many service industries are as large-scale, as capital-intensive and as thoroughly grounded in technology as manufacturing is. Our statistical data bases and case studies also demonstrate that new technologies can affect entire service industries intensely, and indeed restructure them, with consequences that radiate throughout the economy.

The technologies that are now transforming the service industries have profound implications for U.S. manufacturing, economic stability and growth, for national and regional job markets and for the position of the U.S. in world politics and international competition. Perhaps more important for the future is the fact that technology, properly applied, can enhance productivity, quality and economic output in the services sector just as it has in manufacturing. This means that a U.S. economy dominated by services can continue to support real increases in income and wealth for a very prolonged period.

What is meant by "services"? A clarification may help to dispel a number of misconceptions. Most authorities consider the services sector to include all economic activities whose output is not a physical product or construction, is generally consumed at the time it is produced and provides added value in forms (such as convenience, amusement, timeliness, comfort or health) that are essentially intangible concerns of its first purchaser. A raw material or manufactured product, in contrast, may retain its value when it is transported, stored or resold. The *Economist* has more simply defined services as "anything sold in trade that could not be dropped on your foot."

People alarmed at the growth of the services sector often caricature it as merely "making hamburgers" or "taking in laundry." Such services do fall within the broad definition, but so do activities that fulfill many much more basic needs: communications, transportation, finance, health care and education, to name only a few. Efficient and high-quality services are crucial not only to consumers but also to product manufacturers. Clearly an automobile manufacturer requires financial services to provide capital for its production facilities, communication and transportation services to coordinate and move its parts and finished products, and distribution and retail networks to present the products to consumers and to service them.

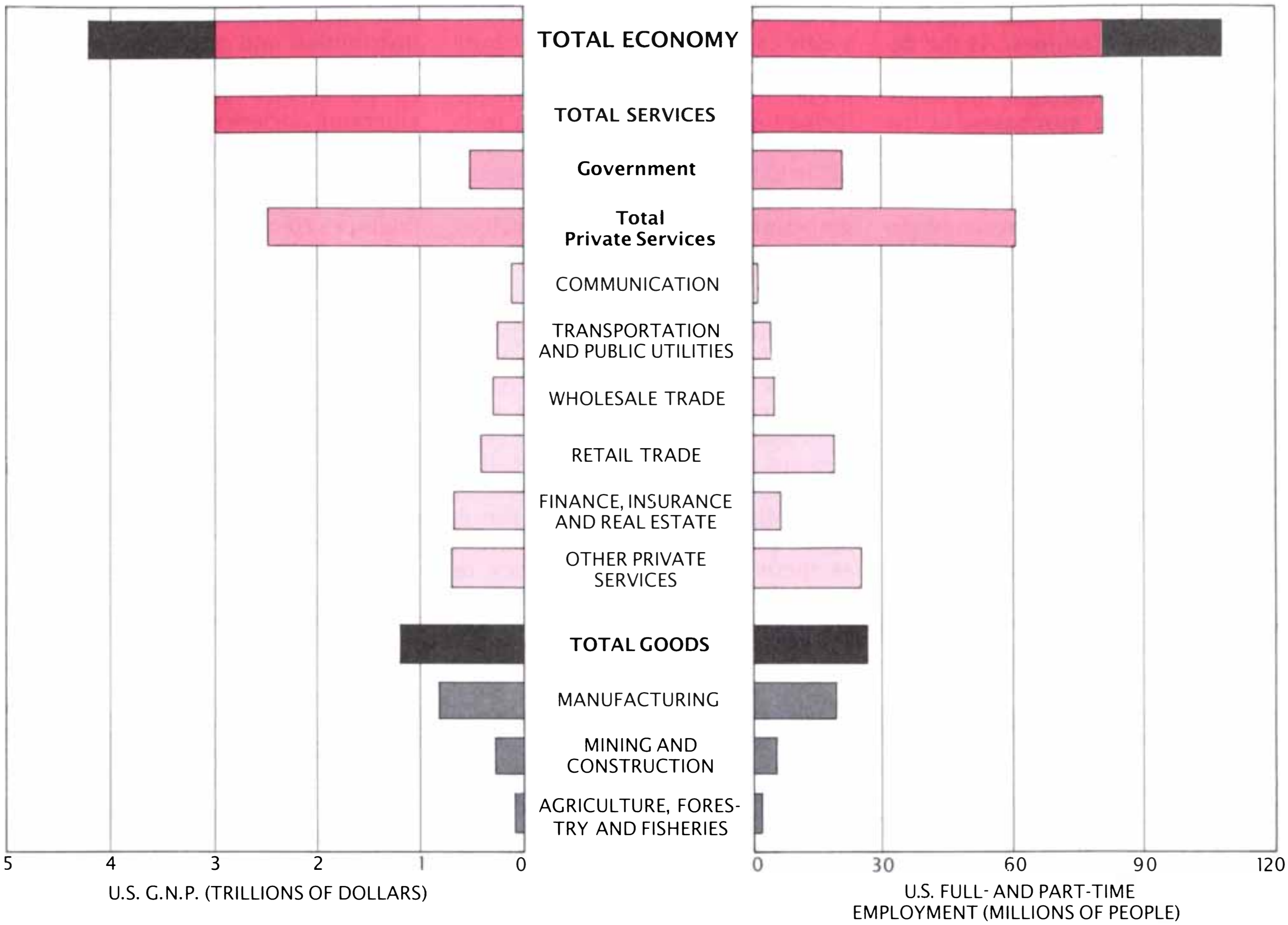
Why have services gained such importance in recent years? Steady productivity increases in manufacturing and agriculture—brought about largely by technology—have meant that it takes ever fewer hours of work to produce or buy a pound of food, an automobile, a piece of furniture or a home appliance. For example, in 1956 it took about 125 hours of work to buy a kitchen stove; in 1986 it took 41 hours. At the same time the demand for goods was capped somewhat: the average person can eat only so many pounds of food in a day and utilize only so many cars, sofas or washing machines. As the demand for individual products reached such constraints, the utility of other possible purchases, in the form of services, grew apace.

Growth in services, then, is a natural effect of increasing productivity in manufacturing, but constantly improving service technologies have also helped to shift economic activity toward services. They have expanded the versatility of existing services and made totally new types of services practical. Jet aircraft have radically improved the efficiency and convenience of long-distance travel and freight movements, opening entirely new markets and providing new options for siting production facilities. Revolutionary technologies for diagnosis and treatment have enlarged the dimensions of human health care. New methods for handling fragile, perishable or volatile goods have vastly extended the geographic horizons of international trade and the scope of its products. Electronic information and communication technologies have stimulated innovation in virtually all service areas—notably in retailing and wholesale trade, engineering design, financial services, communications and entertainment.

The burgeoning of the services sector disturbs many who hold a point of view—possibly first articulated by Adam Smith—that services are somehow less important than products on a scale of human needs. Perhaps in elemental societies it is true that production of food, shelter and clothing must at first take precedence over other demands. As soon as there is local self-sufficiency or a small surplus in a single product, however, any extra production has little value in the absence of storage, distribution and transportation systems for reaching new marketplaces—all service activities. In most emerging societies services such as health care, religion, banking, entertainment, trading, law and the arts quickly become more highly valued (high-priced or capable of generating great wealth) than basic production. And in affluent societies it is clear that a nation's wealth is measured as often by the level of its arts, education, public health and social services as by its sheer abundance of physical goods.

Even products themselves generally owe their value to the services they can deliver. An automobile delivers transportation and convenience: services. A television set has value because it can present broadcast entertainment: another service. A computer diskette's primary value

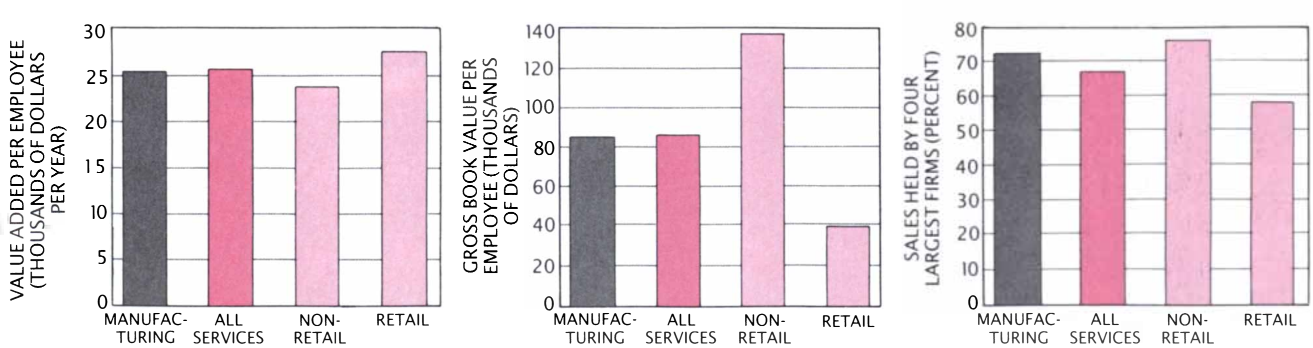
SORTING SYSTEM at the Memphis "superhub" of the Federal Express Corporation epitomizes the transformation of the services sector by technology. More than 650,000 packages and documents a night arrive from airports all over the U.S. at about midnight, are sorted and are dispatched to their destinations by about 4:00 A.M. Computer-controlled sorting equipment, now being adopted by other companies, has made overnight-delivery systems possible. The technology has thereby helped to create a new niche in the service economy and has lowered costs for many firms, including manufacturers shipping urgent packages.



SCALE OF SERVICES SECTOR is suggested by a comparison of the contribution to the gross national product (*left*) and employment (*right*) of service (*color*) and goods-producing (*gray*) industries in 1986, according to national-income accounts. Services have long surpassed manufacturing; now individual services-sector categories rival manufacturing as a whole in size.

lies not in its plastic base and magnetic coating but in the software programs or data sets it stores: more services. Indeed, in many cases products and services are interchangeable. A home washing machine (a product) can substitute for the use of a laundromat (a service), a frozen dinner for a meal in a restaurant or CAD/CAM software for additional production equipment.

At its core the professed alarm about the growing dominance of the services sector may reflect a genuine doubt that services can sustain the increases in real income and personal wealth that have been hallmarks of the industrial era. There is serious concern that the kinds of productivity increases automation has wrought in manufacturing cannot be replicat-



STEREOTYPED VIEW of services as being labor-intensive and small in scale is challenged by the evidence. The level of value added per employee (*left*) in large service companies is comparable to that of similar manufacturing enterprises, suggesting that services need be no more labor-intensive than manufacturing. In this survey services equaled or exceeded manufacturing in capital intensity (*middle*) and in concentration (*right*), a measure of the dominance of companies large enough to invest in new technology. All three indicators suggest that service companies utilize technology in a major way. Data from the PROFIT IMPACT OF MANAGEMENT STRATEGY (PIMS) data base of the Strategic Planning Institute were analyzed by Christopher E. Gagnon.

ed in services. Continuous increases in productivity will be required if there is to be noninflationary growth in incomes; otherwise higher wages will only increase the relative amount of money competing for available goods or services.

Can technology boost productivity in services as it has in manufacturing? To evaluate the potential one must understand the structure of service industries today. For new technology to have high and favorable impact the industries receiving it generally must have the scale, capital intensity and technical sophistication to be able to apply technology effectively. Services, in the popular stereotype, are small-scale, labor-intensive, relatively unsophisticated undertakings that make few large equipment investments.

This stereotype may be valid for some household-service and retail activities, but it clearly does not hold for the communications, transportation, pipeline, health-care and electric-utility systems that make up much of the U.S. services sector. Increasingly some other major service industries, including banking, entertainment, mass retailing, financial services, car rental and package or message delivery, are also investing heavily in technology.

Stephen Roach of Morgan Stanley and Company has found that total capital investment, in particular high-technology investment, per "information worker" has been rising rapidly since the mid-1960's and now exceeds that for workers in basic industrial activities. Similarly, in a sampling of 145 industries that were analyzed by Ronald E. Kutscher and Jerome A. Mark of the Bureau of Labor Statistics, nearly half of the 30 most capital-intensive industries were services. Our own analyses of the Strategic Planning Institute's PROFIT IMPACT OF MANAGEMENT STRATEGY (PIMS) data base confirm these findings.

PIMS and Fortune 500 data also suggest that major service enterprises are comparable in relative size and profitability to major manufacturers. The large service companies clearly have the financial power to buy technology as it becomes available and needed. In fact, many of these companies now play a crucial role in the creation and diffusion of new technologies for both products and services. Citicorp, for example, helped to develop and introduce the first automated teller machines; the Federal Express Corporation made major innovations in package-sorting, -handling and -tracking equipment that other express companies and materials-handling concerns have now begun to duplicate.

Can technology adequately improve productivity in the services sector, however? Productivity in services is notoriously difficult to measure. For most services it is harder to identify a unit of output than it is in manufacturing, not only because there are no physical goods to count or weigh but also because output must be defined with reference to quality too, and that is even more ephemeral in services than in manufacturing. How, for example, does one evaluate the productivity of a medical procedure that may consume fewer resources or take less time than its predecessor but may subject the patient to greater pain or risk? Is the number of letters delivered per postal worker an effective measure of productivity if such "productivity" means that more letters are lost or delayed? One should be extremely cautious in interpreting aggregate productivity data about services.

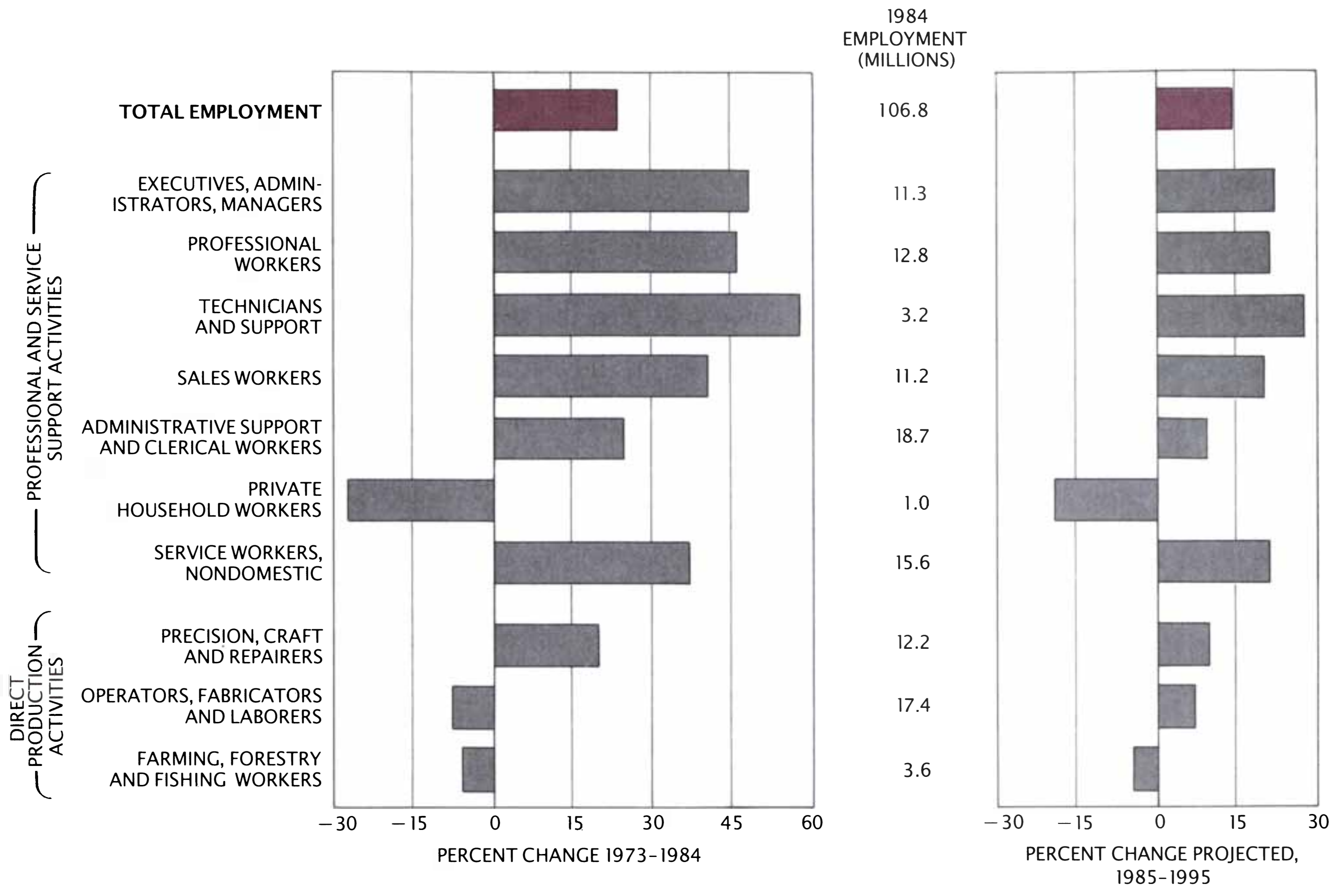
Nevertheless, if a customer has a choice between a given service and a competing service or between the service and a product that performs the same function, the revenue the service brings to its producers (its "sales value") offers a much better indication of both the quantity and the quality of output. Comparing this sales value with inputs of labor or capital can often provide a reasonable measure of an individual service industry's productivity. Bureau of Labor Statistics data for the past 35 years suggest that some individual service industries (such as communications, utilities or wholesale trade) can sustain rates of productivity growth as high as those in manufacturing for extended periods. Although data on U.S. services productivity growth from 1980 through 1985 are less encouraging in the aggregate, rates of productivity growth vary widely among individual service industries, just as they do in manufacturing. Some service industries have outperformed some manufacturing industries, and vice versa.

In a recent study of the British economy Richard Barras of the Technical Change Centre in London has shown that services-sector productivity (measured as the real value of output per employee) grew at 2.9 percent per year between 1960 and 1981; manufacturing productivity, according to the same measure, grew by less than 1 percent per year. The data suggest there is no inherent reason individual service industries cannot keep up with—or outperform—individual product industries in productivity increases.

Investment in technology has also soared among service producers since 1975. Companies in financial services and wholesale trade, for example, have made large purchases of computers and communications equipment; providers of health care have invested heavily in new medical technologies and information systems. Such technologies have precipitated major structural changes and employment adjustments that, after a period of adaptation, should substantially boost both the quality of output and productivity.

Other structural changes brought about by service technologies are already apparent. In industry after industry we examined there were similar patterns of change. Typically, major new service technologies at first created new economies of scale, generally favoring an industry's larger enterprises. Middle-size service enterprises, unable to afford such technologies themselves, were often forced to merge into larger companies, identify a protected niche in the marketplace or go out of business. Increases in the scale and relative power of larger companies in transportation, financial services, health care and banking readily attest to such trends. At the same time, however, smaller companies often flourish because of the countervailing effects of new data-management and communications systems. They make it possible for small companies to serve remote areas or specialized niches on their own, or to link up in coalitions or networks sharing data and resources with larger enterprises.

Once they have been properly installed, the same technologies that created the new economies of scale often generate a secondary effect that might be called "economies of scope." Enterprises exploiting the new technologies find they can handle a much wider array of data, variety of services and range of customers than before without significant cost increases. Banks (Citicorp), airlines (American Airlines), retailers (Sears) and bank-travel services (American Express) have used newly installed information and distribu-



JOB GROWTH in various employment categories is shown for the decade ending in 1984 and projected for the decade ending in 1995. Service-related jobs have grown more rapidly than blue-collar production jobs; the more highly paid professional categories have accounted for much of that growth. Many service-related jobs are actually in manufacturing and have been responsible for much of the increase in output value in the goods-producing sector. The trends shown here are expected to continue. The data are from a study done by George T. Silvestri and John M. Lukasiewicz of the U.S. Bureau of Labor Statistics.

tion facilities to extend their presence into a broad range of new activities—in the case of Sears, for example, into insurance, real estate, brokerage activities, financial services and credit cards.

The recent history of the securities industry exemplifies both major effects of new service technologies. In the early 1970's the volume of shares being traded (then between 10 and 12 million shares daily) began to overwhelm the securities trading houses, which were still handling and delivering each certificate physically. The major Wall Street firms formed what became the Depository Trust Company to collect and keep virtually all traded certificates under one roof and transfer their ownership electronically. The resulting technologies created new economies of scale for their sponsors but also forced many small and middle-size brokerage houses, which could not afford automation on their own, to merge or to affiliate with larger houses.

The surviving houses found that the in-house electronic systems they had developed for automated trading enabled them not only to cope with larger numbers of transactions (now more than 200 million per day) but also to provide novel specialized "products" and services (such as "cash management" accounts) for their customers. Since their installed electronics systems allowed these companies to present this greater variety of products at low additional cost (the essence of economies of scope), many of them started to decentralize again: to develop networks of affiliated local offices around the country to make the most of their new opportunities.

In the course of such restructurings, traditional industry boundaries have tended to blur or disappear. For example, as their new electronic technologies enabled the various financial institutions (banks, brokerage houses and insurance companies) to carry out easily functions traditionally reserved through regulation to their competitors, they began to seek new opportunities through deregulation. Soon consumers began to use financial-services houses quite interchangeably.

Similarly, airlines, hotels and tour operators have all begun to collaborate in selling vacation "packages" and replacing traditional middlemen by means of electronic technologies that make direct reservations and ticketing possible. Even the boundaries between public and private services are eroding, as town and local governments look to private companies for maintenance, health care, transportation, accounting and other functions they had long provided themselves.

What does growth and structural change in the services sector mean for employment and international trade? Perhaps the most pressing concern is jobs. Can the service industries provide both the quantity and the quality of jobs the U.S. needs for continuous growth in its economy and in its base of human skills? In recent years services have provided the engine for employment expansion throughout the country: between 1976 and 1986 they accounted for 85 percent of the new jobs in the private sector. No state registered a decline in service jobs, whereas manufacturing jobs shifted markedly among regions, states and industries. Today every state has more service jobs than it has manufacturing, mining, construction and agricultural jobs combined.

But what kinds of jobs are being created? On the average, hourly wages are somewhat lower in services than in manufacturing. Some observers, including Barry Bluestone and Bennett Harrison in a recent report to Congress, suggest that most new jobs in the U.S. economy have been in lower wage categories and that the number of workers in the middle-income brackets is declining. On the other hand, Neal Rosenthal of the Bureau of Labor Statistics has reported that during the decade from 1973 through 1982 (a period when service jobs represented a growing percentage of total employment) the number of middle-income jobs declined only slightly, whereas the number of lower-income jobs fell by a much larger amount.

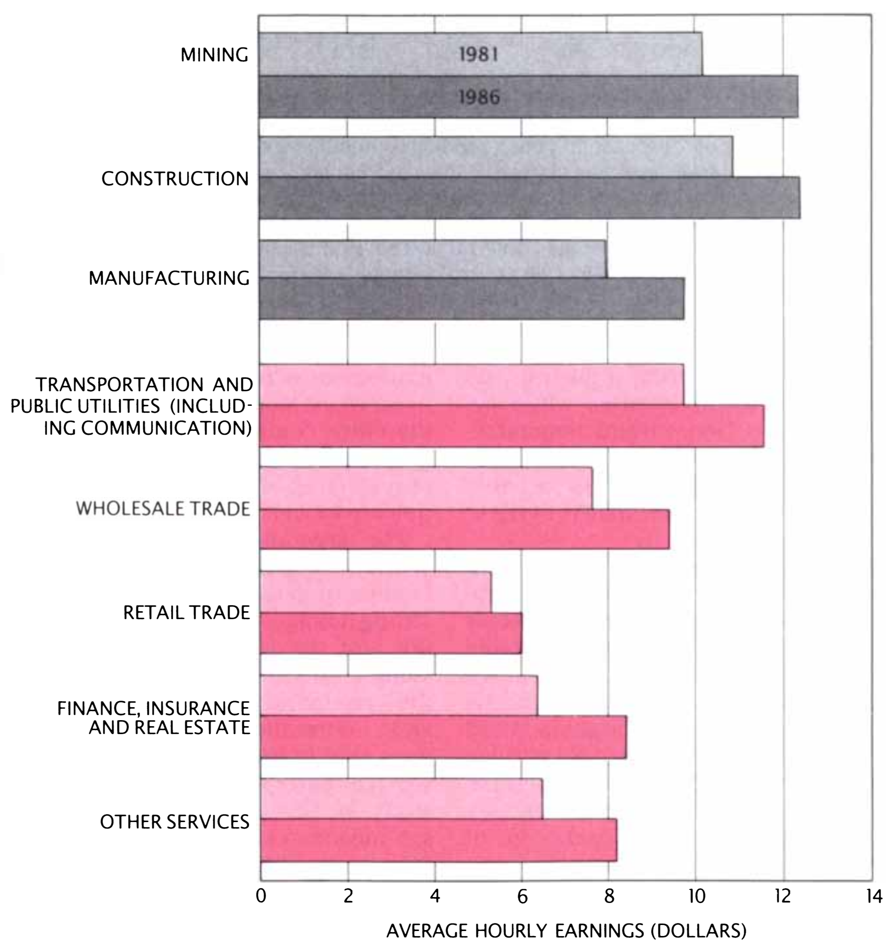
The evidence on recent wage trends, then, is somewhat ambiguous. More important is the future. The relation of pay scales in services to those in manufacturing will depend in large part on how rapidly technology can improve productivity in each sector. The more productivity grows in either sector, the more latitude that sector will have for higher wages. Rapid productivity increases and demand for higher skills will probably continue to keep wages in individual service industries (such as transportation or public utilities, including communications) higher than in some manufacturing fields. Moreover, the wages and work conditions many future service jobs offer should easily surpass the conditions employees in distressed manufacturing industries would have to accept to keep their companies from relocating production overseas. In either event, developing the services sector could be a better way to sustain a high U.S. standard of living than subsidizing certain outmoded manufacturing industries.

Services have a demonstrated advantage as a source of employment: they are far more resistant to recessions than goods-producing industries. In business cycles from 1948 through 1980, U.S. services employment actually advanced at an average annual rate of 2.1 percent during economic contractions. In these same recessions employment in the goods-producing sector declined by an average of 8.3 percent. In Canada the picture has been much the same: services employment declined only during the deepest recessions.

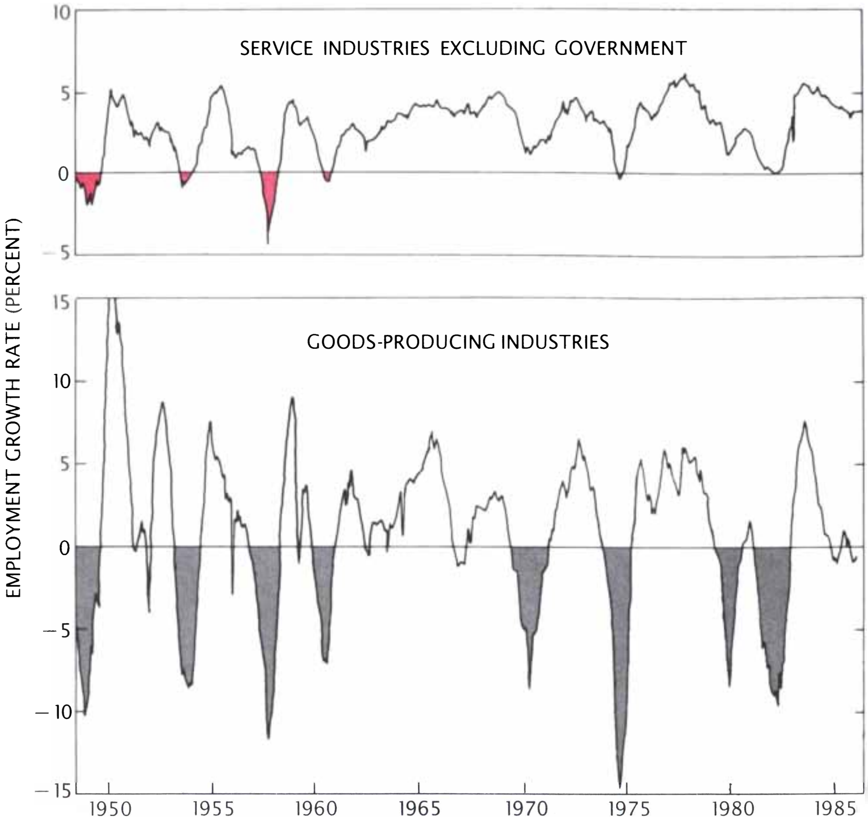
If services really are in some sense less vital than manufactures, one would expect people to give up their services first during a recession, thereby eliminating jobs in the sector. Actually much consumer spending for services seems to be less discretionary than spending for products. Although people may go to movies, restaurants or hairdressers less often during recessions, they seem reluctant to give up their telephone, health-care, education, banking or utility services. Instead they tend to postpone expenditures for durable goods.

The demand for certain services may even grow during recessions; people may forsake driving for public transportation, for example. And continuing population growth during these periods tends to keep increasing the demand for such services as education, health care and fire and police protection.

Even when demand for services does decline temporarily, employment can remain relatively stable. Because a substantial fraction of the income of many service workers comes from part-time work, tips, commissions or profit sharing, income levels can contract more easily during recessions without comparable contraction of the work force. In addition many service enterprises,



AVERAGE HOURLY EARNINGS in 1981 (*light*) and 1986 (*dark*) are compared for goods-producing industries (*gray*) and services (*color*). Average wages are higher in mining, construction and manufacturing than in services, but some service-industries wages equal or exceed those in manufacturing. The service industries paying the most tend to be ones in which the intensive application of technology has increased output value.



SERVICE-INDUSTRY JOBS are stabler than jobs in goods-producing industries. During four recessions between 1948 and 1961, employment in private services fell (*shaded areas*) an average of 1 percent; in goods-producing industries the average drop was 7.2 percent. During four recessions between 1969 and 1982, services employment actually rose an average of 1 percent; in goods-producing industries it dropped an average of 7.9 percent. Growth in services largely offset manufacturing declines, so that total nonfarm employment losses averaged only 1.6 percent in these last four recessions. Geoffrey H. Moore of the Columbia University Graduate School of Business did the study.

unlike manufacturing industries, do not build up inventories when demand slows. Hence when demand reasserts itself, there is no pause in service rehiring, whereas product sales merely deplete old inventories.

In addition to reducing the depth of recessions and their impact on employment levels, the services sector helps to make them shorter. Stable employment in the service industries provides wages that preserve a large, continuing market for manufactured consumer goods. Moreover, service industries provide major markets for complex technological products. According to Roach, the useful life of capital equipment in service businesses tends to be shorter than it is in manufacturing. One reason is that services rely heavily on communications and computer technologies, in which rapid technological advances quickly lead to obsolescence. Consequently the replacement of capital equipment is usually more constant in services. Many service enterprises therefore continue to invest during recessions, enabling manufacturers of capital goods to recover more quickly as well.

The large and stable markets that service companies create for manufactured goods is only one of several strong linkages between manufacturing and the services sector. Some studies suggest that 80 percent of the communications technologies and information-management systems sold in the U.S. in 1982 went to service enterprises. A comparable study in the U.K. showed that service businesses were the buyers of 70 percent of the computer systems sold there in 1984.

Similarly, manufacturing concerns are also major consumers of virtually all modern services. Manufacturing activities clearly depend heavily on communications, financing, transportation and waste-handling services, to name only a few essentials. Manufacturers also pay for their employees' health care, retirement and life insurance—either directly through insurance policies or indirectly through wages. When new technologies help to provide such services more efficiently, they obviously lower manufacturers' costs for directly purchased services; to the extent that efficiencies in services lower employees' living costs, they can help to reduce wage demands.

Yet the interdependence of manufacturing and services is even more intricate. Services are not simply external adjuncts to the manufacturing process, bought to keep a factory going or to get goods to consumers; they are intimately embodied in the manufactured object itself. For the manufacturing sector as a whole (according to studies done by the Office of the U.S. Trade Representative and the National Association of Accountants) service activities within the producing company itself create between 75 and 85 percent of all the value the average manufacturer adds in making its product.

This means that the price a product can command (whether it is a car, a processed food or a home appliance) reflects the product's content of raw materials and direct labor less than it does the characteristics, quality and availability of the product, which are created by research, product design, quality control and marketing. In industries such as automobiles or pharmaceuticals the cost of such services in a product can be from three to 10 times its direct labor costs and can provide virtually all the perceived distinction between it and competitive products. Clearly, therefore, an important way for U.S. manufacturers to improve their competitive position is to increase the effectiveness of their own internal service activities.

Computers, communications technologies and rapid-response inventory-control and distribution systems can also help manufacturers to exploit the expanding demand for differentiated and customized products. Affluent consumers increasingly want products that reflect their particular tastes, and they want immediate delivery rather than a wait of weeks or even months. Not surprisingly, both industrial and commercial customers also want a larger variety of products to be delivered more quickly and on a more precise schedule—a "just in time" approach, which can substantially lower inventory costs.

Ever more frequently, success in manufacturing requires not only flexible facilities for turning out a wide variety of products at low cost but also the capacity to acquire and respond to rapid feedback from the marketplace. The service technologies that make possible improved collection, analysis and transmittal of data about customer preferences have become crucial strategic weapons. For example, at the end of each day a national women's specialty retailer, Limited Stores, electronically gathers and aggregates that day's sales details (items, sizes, colors, styles and so on) from its entire network of stores. The information is converted into orders that are instantly transmitted to its manufacturer-suppliers all over the world. The orders specify precise deliveries, many of which depend on transpacific jumbo-jet flights—to defined distribution sites within a few days.

Today many of the factories supplying such large retail chains are in Asia or developing countries. Yet rising demand for faster response and more customized products offers potential advantages for U.S. manufacturers. A domestic manufacturer with sensitive links between its own flexible production system and its distribution and customer networks should be able both to respond more quickly to the market and to enjoy lower transportation costs than foreign competitors.

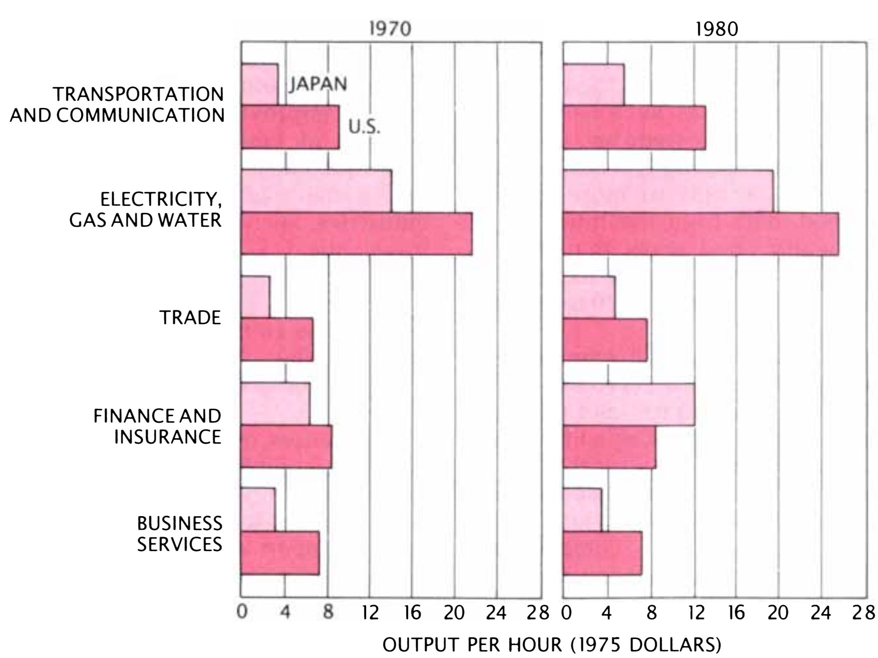
In the case of heavy, bulky or complex products such as automobiles, it could even become impossible for overseas producers of the more standardized lines to compete here against responsive, well-integrated U.S. manufacturing and retailing systems. As a case in point, the Honda Motor Company's decision to produce increasing numbers of cars in the U.S. coincided with its introduction of a much wider variety of styles and options to the U.S. market. Its U.S. manufacturing base was essential for the other key element in this strategy: fast delivery of its full product range, which would have been impossible if Honda had continued to make cars only in Japan.

Many other foreign investors are also seeing the potential of more direct connections with U.S. services. Technologies have made it possible for U.S. service companies to grow to a scale where they are attractive acquisitions for foreign competitors. The same technologies enable the acquirers to manage worldwide service networks from their home bases. Total foreign investment in the U.S. has ballooned 46.7 percent since 1982, much of it in services, notably in the areas of communications, finance and distribution.

Service technologies are also causing another profound change in the worldwide competitive situation. Computer and communications technologies are rapidly integrating the world's financial centers into a single financial marketplace. The flow of money from country to country has become largely independent of the movement of goods or trade. Whereas world trade in goods and services amounts to only between \$3 and \$4 trillion per year, the financial transactions handled by just one intermediary, the Clearing House for International Payments, totaled \$105 trillion in 1986, and transactions in early 1987 were proceeding at a rate of \$200 trillion per year. Instead of following goods, these vast sums now flow toward the highest available real interest rates or toward safer, more stable economic and political conditions.

Hence the value of a nation's currency increasingly depends less on its position in international trade than on the fiscal or monetary policies of many governments and on banking decisions and other world events that affect foreign investment. In recent years exchange rates for the dollar against even stable currencies such as the yen, the mark or the Swiss franc have fluctuated by close to 50 percent, not so much because of trade choices as because of fiscal and monetary decisions. The effect of such fluctuations on the relative cost of a product being sold in international trade have become so great that they can dwarf the impact of a producer's internal productivity improvements or other managerial decisions. Even well-run Japanese companies find it difficult to compete when the price of the yen—and hence their production costs in Japan—can jump 43 percent in 18 months in relation to their competitors' costs or their customers' buying power.

On the other hand, most governments, banks and large companies now have much freer access to all the major money markets. As a result it is becoming increasingly difficult for a single nation to maintain exceptionally low domestic interest rates in order to stimulate economic development—a policy followed by the U.S. in earlier years and by Japan



PRODUCTIVITY of U.S service industries, measured as the market value of output per hour of labor, has exceeded that of Japanese services except in finance and insurance. In addition to making service industries more competitive, service technologies reduce costs and improve quality in production, distribution and sale of manufactures.

more recently. Today an artificially low interest rate simply drives capital out of the country in search of higher interest rates—and at the same time it attracts foreign borrowers, whose demand tends to drive the manipulated interest rate up again. As variations in the cost of capital among nations are decreased by these forces, countries burdened by higher labor and materials costs, such as the U.S. and Japan, will feel even greater pressure from foreign manufacturers.

In this environment innovation in services offers one of the few strategies for keeping the domestic economy competitive. First, higher productivity in services will lower the cost or increase the value of the large part of the G.N.P. generated by the services sector. Second, more efficiently produced services will hold down the large component of each manufactured product's cost that reflects its services content. It is heartening to note that in most service industries U.S. productivity is higher than Japan's, but innovative and aggressive efforts will be necessary in order to stay ahead.

Service technologies, then, may help to support U.S. manufacturing against foreign incursions. Can they also enable services themselves to make a major contribution to the improvement of the U.S. trade balance? It is worth noting that unlike merchandise, in which there has been a trade deficit in half of the past 26 years (it reached $146 billion in 1986), services have consistently shown a positive trade balance. Since 1981, however, the U.S. advantage in services has become more tenuous. Indeed, data from the International Monetary Fund show that the U.S. share of world service exports fell from 23.8 percent in 1970 to 19.2 percent in 1985.

What does it mean to export a service? Relatively few services are produced in one country and then sold (as a software program is) for its full value abroad. More typical is the example of a U.S. bank's Middle Eastern branch participating in a large transaction there, or a U.S. company providing overnight package deliveries overseas. Much of the income generated by such service exports goes into salaries and overhead abroad and remains in the foreign country. In the case of an automobile export, trade data reflect the car's wholesale sales value; export data on services often recognize only the fees or profits that can be repatriated to the home office.

Hence the volume of U.S. international service transactions would have to expand enormously to eliminate the nation's huge merchandise trade deficit. A further complication is the fact that U.S. Government trade data recognize only about 40 categories of services (in contrast to some 10,000 merchandise items), so that the data may not capture many important service exports. Many observers now think the total volume of U.S. services trade, both export and import, has been seriously understated. Attempting to improve on the official figures, the Office of Technology Assessment estimated that in 1984 the U.S. exported between $69 and $91 billion in services and imported between $57 and $74 billion; the official figures were $43.8 and $41.5 billion respectively. Even interpreted optimistically, these figures still do not come close to covering current merchandise trade deficits.

In individual categories the U.S. services-trade situation varies greatly. In some areas, such as public accounting, law, communications and international finance, U.S. companies enjoy a strong position in the global market. Some individual U.S. service companies, such as Citicorp, AT&T and Federal Express, enjoy economies of scale and scope that few international competitors can equal. Their aggressive application of proprietary technologies has forced domestic competitors to follow suit, thereby improving the comparative position of both their own industries and those they support in international trade. In other service industries, such as international air travel, the U.S. has been faltering. The once dominant U.S. carriers, Pan American and TWA, have done badly, whereas competitors such as Japan Air Lines, Swissair and Singapore Airlines have thrived.

The causes of such shifts are always complicated, but much of the credit for the foreign airlines' gains must go to their governments' strong support and their own long-term investment in equipment and exceptional attention to the quality of customer care on flights. Much thoughtful customer-oriented management and serious long-term investment in technology will be essential if U.S. service enterprises are to forestall and defeat further competitive incursions into both their domestic and their international markets. With more than 70 percent of its total economy at stake, this is a battle the U.S. cannot afford to lose.

Enlightened Government policies can also help. U.S. deregulation of services has already stimulated substantial restructuring and application of technology in U.S. service industries, although this has sometimes been at the cost of extensive customer inconvenience. Perhaps in time the more open competition will prompt U.S. service companies to avoid the complacency, concern with short-term gains, inattention to quality and emphasis on economies of scale (rather than on customers' concerns) that earlier undercut U.S. competitiveness in manufacturing.

The fact is that deregulation itself has created a need for further social innovations. Already new requirements for product labeling, financial disclosure and self-reporting have been found to be necessary to replace, respectively, the direct regulation of some over-the-counter drugs, SEC registrations and detailed FAA approvals of airline routes and schedules. Future dynamics will undoubtedly call for even more imaginative solutions.

The overwhelming evidence is that the U.S. can bolster its standard of living, the stability of its job markets and some aspects of its international competitive position by developing its services sector effectively. That conclusion still leaves some nagging concerns, notably the sense that an economy dominated by services could weaken the U.S. in world affairs. There is no doubt that a drop in manufacturing and in raw-materials production beyond some point could decrease U.S. flexibility and capabilities in defense and impair the nation's bargaining power in world affairs.

Nonetheless, over the next two decades most of the nation's growth and many of its greatest opportunities for entrepreneurship and application of new technologies will arise in the services sector, as they have in the past 20 years. We should not fear an economy dominated by services, or deride it. Rather we should fear the possibility that we may misunderstand the services sector, underdevelop it or mismanage it, at the same time attempting to shore up certain troubled manufacturing areas at great corporate and national cost.

TOYOTA CAMRY

POWERFULLY SATISFYING.
Pour from a satisfying reserve of power: Camry's twin cam, 16-valve EFI engine shrinks distance with 115 horsepower.

PERFECTLY SATISFYING.

The 1988 Toyota Camry is an inspired blend of refined power, graceful style, and plush comfort. In 1987, Camry was ranked #1 in its segment in customer satisfaction.* The heritage continues. Enhanced in 1988, Camry LE's sophisticated, fuel-efficient** multi-valve engine delivers tireless cruising capabilities, quick acceleration, and plenty of passing power. Camry's ride is blissfully smooth and quiet, and front-wheel drive is an ally in any weather. Its flush, aerodynamic styling is a perfect finishing touch. And the new Camry Wagon is an extended version of the same satisfying story. 1988 Camry. Satisfaction, perfected.

Get More From Life . . . Buckle Up!

GRATIFYING CONTROL.
The quality feel of control, handling, and performance is a positive pleasure for anyone who enjoys driving a fine sedan.

*J.D. Power & Associates 1987 Customer Satisfaction with Product Quality and Dealer Service (for 1986 vehicles).
**EPA estimated 25 city/32 highway MPG for Camry LE with ECT 4-speed automatic overdrive transmission.
© 1987 Toyota Motor Sales, U.S.A., Inc.

TOYOTA QUALITY
WHO COULD ASK FOR ANYTHING MORE!

The Moon's Ancient Magnetism

The moon is now a dead body, but it seems once to have generated its own magnetic field. Since then it has been shifted with respect to its spin axis—perhaps by collisions with moon-orbiting satellites

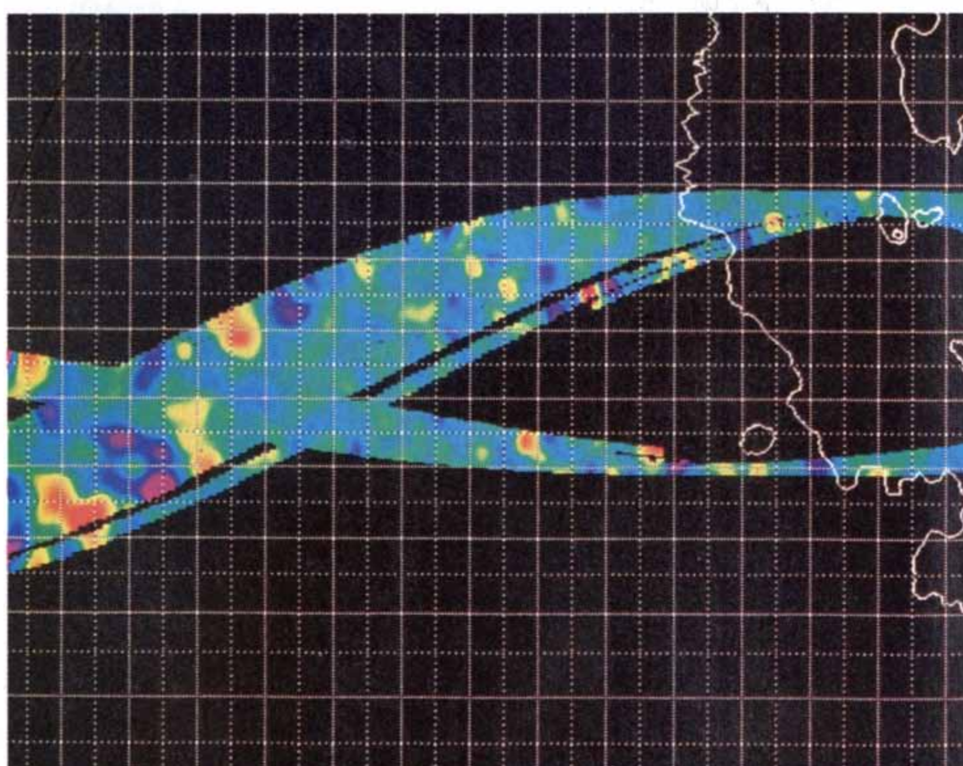
by S. K. Runcorn

At the time of the *Apollo 11* landing in July, 1969, it was generally agreed that the moon was a dead body. Its surface features had been shown to be due largely to external processes and not, as is the case on the earth, to internal processes. For example, it had been shown that the vast majority of lunar craters are due not to volcanoes—as Galileo and most geologists since his time had supposed—but to bombardment by meteorites over billions of years. No phenomenon involving large crustal displacements (such as the continental drift that takes place on the earth) had happened on the moon. The moon was thought not to possess an iron core, because the mean lunar density is very close to that of the ferromagnesium silicates that constitute the upper mantle of the earth.

Certainly an inherently dead moon seemed to be what one would expect from theoretical considerations. Such a small body as the moon would have lost its internal heat more rapidly than the earth, because its ratio of surface area to volume is much greater than the earth's. It was therefore not surprising that the moon displayed neither a magnetic field nor plate tectonics, even though in the earth there is still plentiful heat energy to drive the motions of the molten core that produce the magnetic field and to drive the much slower flow of the viscous mantle (the layer between the core and the lithosphere, or rigid surface layer), which moves the tectonic plates. The model of the moon as an undifferentiated, internally inactive body seemed to have been confirmed in September, 1959, when the Soviet *Luna 2* probe showed that the moon had no detectable magnetic field.

Yet close examination of the lunar samples returned to the earth by the Apollo missions began to reveal a different story. Some of the samples were magnetized, implying that they had been exposed to magnetic fields when they were first forming. In addition, lunar probes began to detect magnetic anomalies—regions of magnetic field arising from rock formations—on the moon. Analysis of these discoveries by my colleagues and me and by other groups suggested a surprising hypothesis: that the moon once had its own magnetic field, and a remarkably strong one. At one time the moon's field may have been nearly twice as strong as the present-day magnetic field of the earth.

It seems that at different times in the moon's history the magnetic field has pointed in different directions with respect to the crust. My co-workers and I have calculated the strengths and directions of the



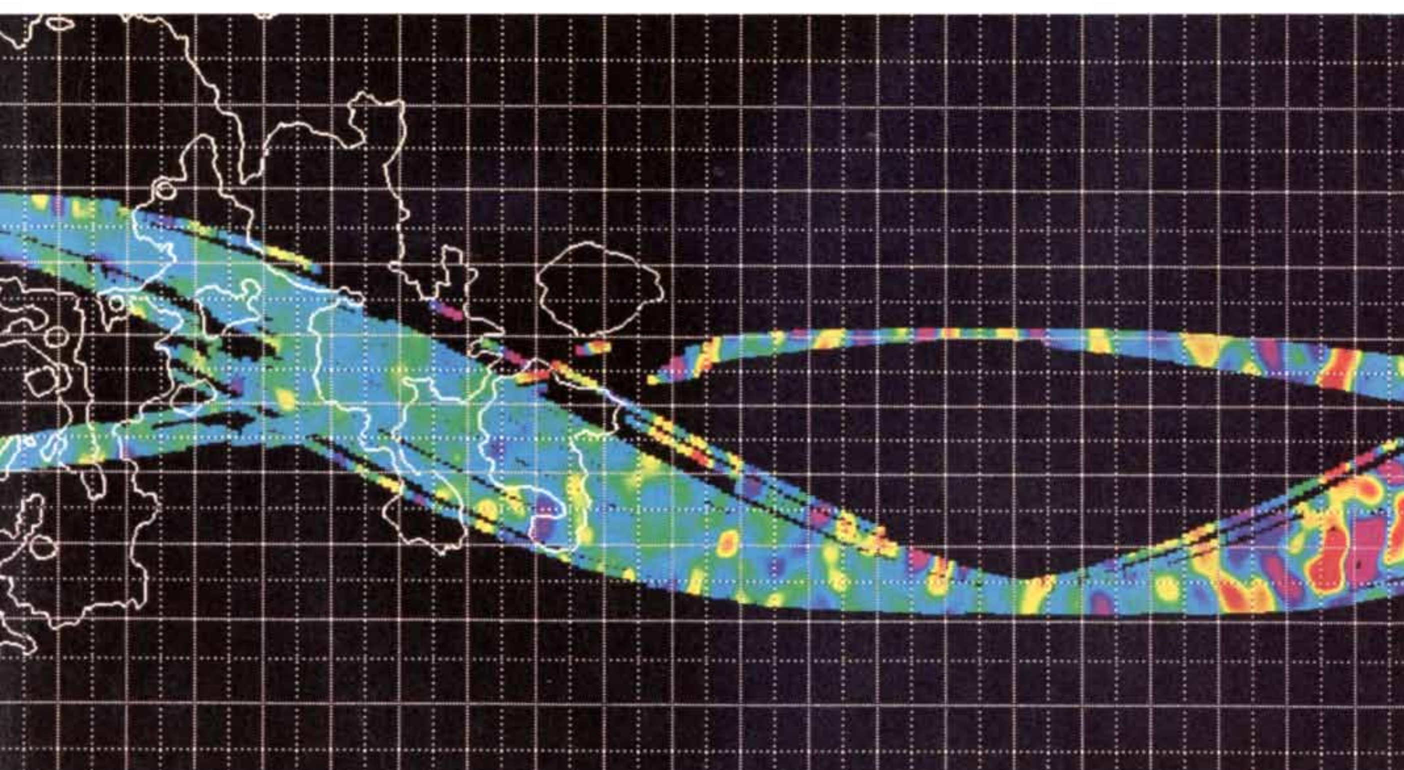
MAGNETIC MAP OF THE MOON, based on data from *Apollo 15* and *Apollo 16* orbiters, depicts fields that arise from magnetized strata of crust. According to the author's hypothesis, the strata were magnetized between 3.6 and 4.2 billion years ago by a field generated within the moon. This computer image shows the radial (vertical) component of the magnetization. Green, yellow and red contours indicate regions where the field points outward from the surface; red indicates the strongest field (about .5 nano-

moon's ancient magnetism, and the results have striking implications concerning the nature of the moon itself and the history of the earth-moon system. For example, it seems that the body of the moon as a whole has shifted several times in relation to its own axis of spin. The spin axis has preserved its orientation and position in space, but the moon itself has rotated in such a way that regions that were once at the poles (where the spin axis intersects the surface) are now closer to the equator. Such a process is called polar wandering, because from the point of view of the surface it is the positions of the poles that seem to have changed.

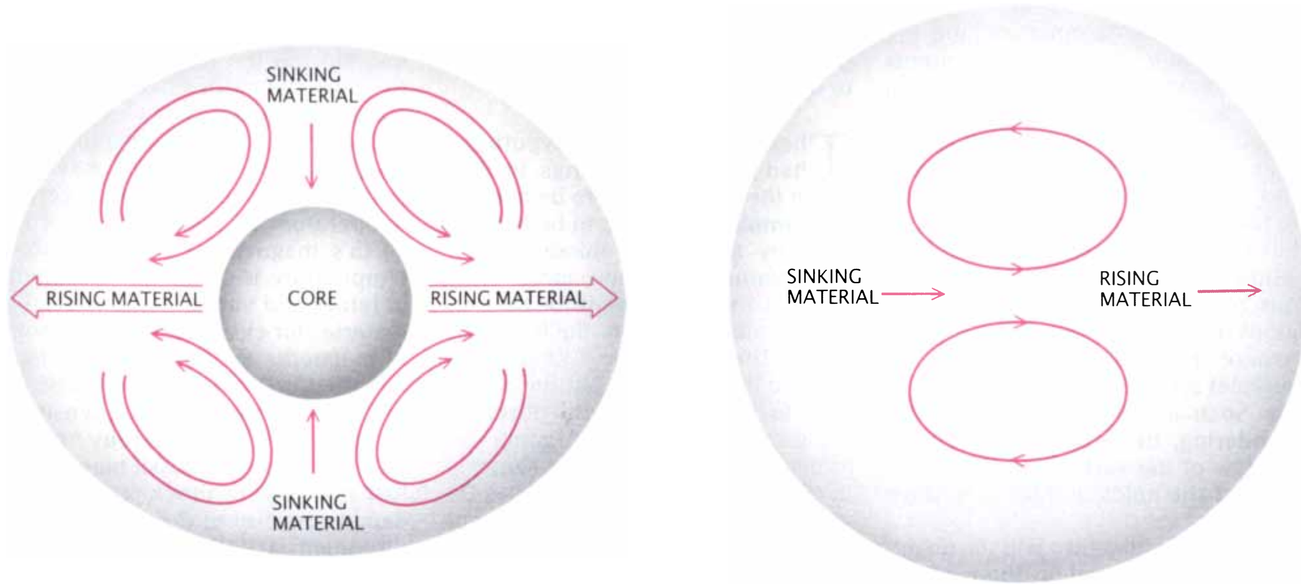
This result, together with an analysis of the shapes and positions of certain large craters, has led to what may be the most startling suggestion of all: that the earth-moon system once contained several other large bodies, all of them in orbit around the moon. As their orbits decayed, the objects broke up, each of them eventually leaving a trail of impacts near what was then the moon's equator. Indeed, it was probably these impacts that caused the moon's poles to wander.

These revolutionary hypotheses had humble beginnings. In 1965, when the experiments to be done on the samples that were to be brought back by the Apollo missions were first planned, no one gave much thought to testing them for their inherent magnetization. Such properties as their magnetic susceptibility were to be measured, but the results were to be used, along with other petrological data, mainly to gauge the mineral composition of the samples. In 1969, however, my colleagues David W. Collinson, Alan Stephenson and I at the University of Newcastle upon Tyne did test the magnetization of a few samples, as we had so often done for terrestrial rocks. We found that the lavas, and those breccia (fragments of rock welded together during impacts) that had been heated to very high temperatures during their formation, were indeed magnetized. In many cases the magnetization was of a kind that could not have been acquired after their removal from the moon, by exposure to either the earth's field or a strong laboratory field.

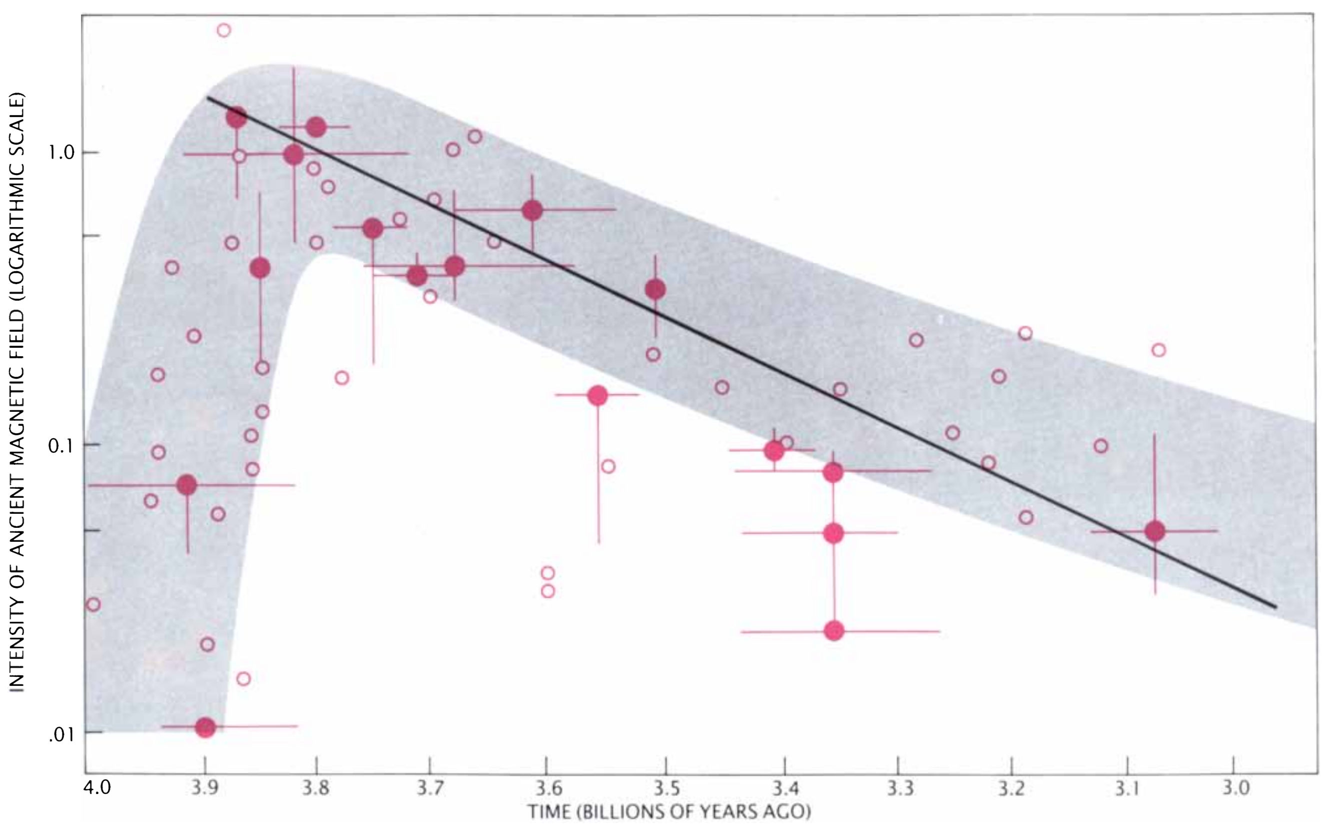
Lavas on the earth acquire their magnetization after they have solidified, by cooling below a certain temperature in the presence of the earth's magnetic field. (The crucial temperature is called the Curie temperature and varies from material to material; for example, the Curie point of magnetite is 580 degrees Celsius and the Curie point of iron is 780 degrees.) The direction and intensity of the earth's field are thereby "frozen into" the rock as a fossil magnetism. Just as a fossil in a rock provides evidence concerning the biological environment at the time of the rock's formation, so fossil magnetism provides evidence concerning the ambient magnetic field at the time the rock was formed. We at Newcastle and investigators in other laboratories (particularly David W. Strangway of the University of Toronto and Michael D. Fuller, now at the University of California at Santa Barbara) soon concluded that the moon rocks



tesla, or 1/100,000 of the field at the earth's surface). Violet and magenta indicate regions where the field points toward the moon's center; magenta indicates the strongest field (again, about .5 nanotesla). The geologic features indicated (*white lines*) are lunar maria. Mare Imbrium is at the top center, to the left and sloping downward is Oceanus Procellarum and at the bottom center are Mare Humorum (*left*) and Mare Nubium (*center*). Farther to the right, and forming a diagonal line, are Mare Serenitatis (*top*), Mare Tranquillitatis (*middle*) and Mare Fecunditatis (*bottom*). At the far right is Mare Crisium. The image was made by Christopher T. Russell of the University of California at Los Angeles and Laurence A. Soderblom of the U.S. Geological Survey.



CONVECTION PATTERNS illustrate an early argument supporting the suggestion that the moon has an iron core. The moon is known to have a bulge many times larger than the equatorial bulge that would be caused by the centrifugal forces due to its spin. The author suggested in 1962 that the moon's bulge might be caused by a process known as two-cell convection (*left*), a circulation pattern in which hot material rises in two columns toward the lunar surface, where it cools and sinks in two columns toward the moon's center. (The moving material is actually the moon's solid mantle—the region between the lithosphere, or rigid outer layer, and the purported core—and it flows by a slow process known as solid-state creep.) Upward pressure from the rising hot columns creates the bulge. If the moon did not have a dense core (*right*), the two-cell pattern of convection would not be possible. One-cell convection—one rising column and one sinking column—could occur, but it would not create a bulge.



INTENSITY of the hypothesized lunar magnetic field seems to have declined with time, as it would have if the field had been generated by motions in a molten iron core that cooled slowly. Data points are based on the magnetization found in samples of lunar rock whose ages were determined by radioactive dating. If one assumes the samples were magnetized by cooling in the presence of the moon's magnetic field at the time of their formation, then each sample's degree of magnetization should indicate the intensity of the field that magnetized it and hence the intensity of the lunar field at the time. Circles indicate data based on relatively imprecise methods of determining ancient intensities, whereas dots indicate data based on more precise methods. The sharp rise in intensity before 3.95 billion years ago, which is based on relatively uncertain data, is unexplained.

had probably acquired their magnetization in a similar way: by cooling below their Curie temperature in the presence of a magnetic field. But what was the magnetic field? This question soon became one of the most intriguing and hotly debated questions in lunar science.

At Newcastle we were disposed to answer the question by saying that the moon does in fact have a small iron core, in which a magnetic field could have been generated during the period when the lavas were being extruded (between 3.2 and 3.6 billion years ago, according to radioactive dating of the samples). To generate a field the core would have had to be molten at that time. The molten core would have undergone convection: a circulation pattern in which columns of hot material rise to the surface of the core, where they cool and then return downward in a colder, denser column. The circulating molten iron would have acted as a dynamo, generating a dipole field much like the field of the present-day earth. Other planets (Mercury, Jupiter, Saturn and Uranus) are now known to have fields that are probably generated in fluid, electrically conducting cores.

Actually the suggestion that the moon may have an iron core dates back to 1962, when I was trying to determine what modifications had to be made to the generally accepted models of the earth's interior in order to take continental drift into account. The earth was known to have a core 3,500 kilometers in radius, which is surrounded by a 2,900-kilometer mantle made of iron magnesium silicates and a lithosphere, also made of silicates, from about 50 to 100 kilometers thick. Seismologists had shown, by analyzing the earth's vibrations during and after earthquakes, that both the mantle and the lithosphere were solid, although the core was fluid. Yet in order to explain how the independent sections of lithosphere making up continental plates could move, it had to be postulated that the mantle underlying the plates could flow, carrying the plates with it.

Fortunately the field of solid-state physics, which was developing rapidly, provided a mechanism whereby flow could be possible in the mantle. The mechanism, known as solid-state creep, involves the slow deformation of a solid object as the result of small stresses that act over long periods. One example of solid-state creep can be seen in airplane engines, in which metal fan blades, under the influence of centrifugal force, can slowly stretch and deform—in effect flowing, like a highly viscous and sticky molasses—as they rotate. Solid-state creep becomes important above a certain critical temperature, which in the earth is reached at the boundary that divides the warm mantle from the cold lithosphere.

I suggested that the force driving the solid-state creep of the earth's mantle is convection. In other words, I suggested that in transporting heat from the surface of the core to the lithosphere, the "solid" mantle itself flows in an extremely slow-moving circulation pattern. The slow circulation of the mantle in turn drives the horizontal plate movements that were first recognized as continental drift by Alfred Wegener 75 years ago.

I then suggested that if the moon had an iron core, a similar kind of solid-state convection could resolve a 300-year-old mystery concerning the forces underlying the geometry of the moon's orbit. In 1693 Jean Dominique Cassini, the first director of the Paris observatory, discovered a set of three laws that describe the moon's rotation. According to one of these laws, the moon behaves in some ways like a slightly tilted gyroscope: its spin axis precesses (revolves in space) at a regular rate. A century later Pierre-Simon de Laplace showed that this law could hold only if the moon has a large equatorial bulge on which the earth pulls, providing the necessary gyroscopic torque. Laplace showed that the bulge would have to be about 17 times the size of the bulge that would be present as a result of the centrifugal force due to the moon's spin. Over the past century the presence of the equatorial bulge has been confirmed and its size has been measured by earth-based telescopes that operate according to the principle of the stereoscope.

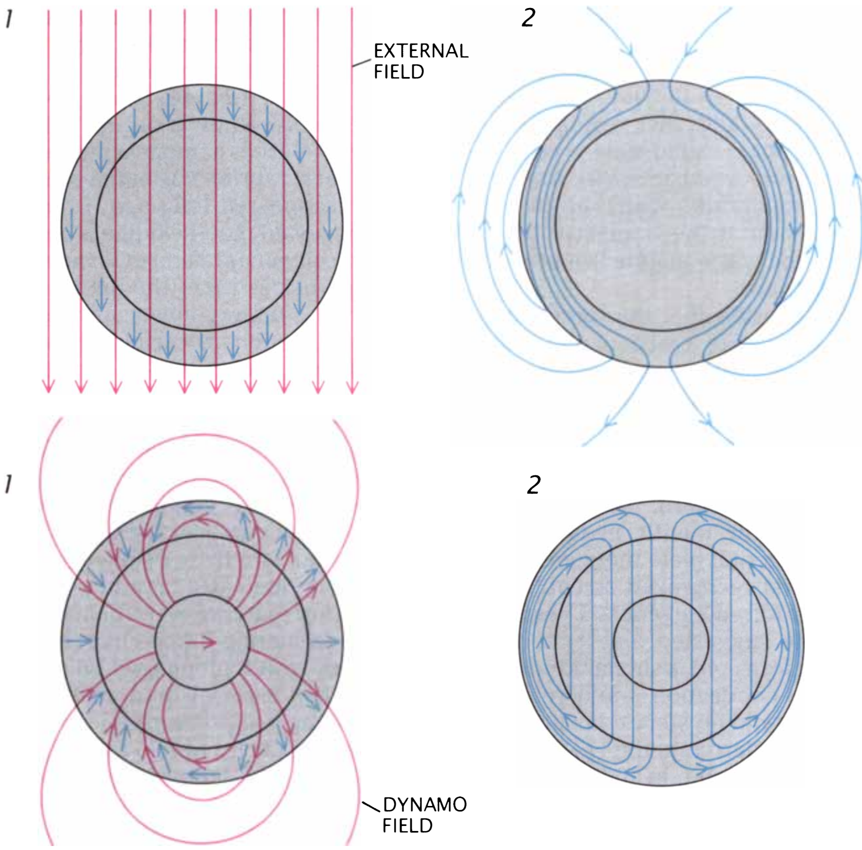
What caused the bulge? Laplace supposed that various strains might have developed when the moon initially cooled from a molten state, and that the shape resulting from those strains had been retained after solidification. Later Sir Harold Jeffreys suggested that the strain could have been caused by the tidal action of the earth. By this explanation the bulge would have had to be frozen in when the moon was about 40 percent of its current distance from the earth.

Both Laplace and Jeffreys assumed the rigidity of the moon would have guaranteed that an early distortion would remain to the present day. Yet if solid-state creep occurred in the moon at one-trillionth the rate at which we now know it occurs in laboratory materials at modest temperatures, such a primeval bulge would have disappeared long ago. I therefore suggested that the bulge is being dynamically maintained even now.

I hypothesized that the interior of the moon is undergoing a slow solid-state convection. A certain convection pattern, called two-cell convection, could indeed produce the right kind of bulge. In areas where material is rising the moon's shape would be stretched outward, and in areas where material is sinking the shape would pull inward [*see top illustration on opposite page*]. If the rising material were along the earth-moon axis and the sinking material were below the limb, the resulting shape—like that of a football with its long axis pointing toward the earth—would resemble the existing bulge. (Borehole measurements made by Apollo astronauts showed that the temperature within the moon increases with increasing depth, probably because of the decay of radioactive elements, and so the moon has sufficient internal heat to drive convection.)

In order for such two-cell convection to occur in a body, the body must have a dense core that does not take part in the convection. In the case of the moon I calculated that the core would have to be between about 300 and 500 kilometers in radius. The core would consist of dense material, and the commonest dense element is iron. On that and other geophysical grounds I therefore assumed that the moon's core, like the core of the earth, is made of iron. Hence when magnetized rocks were returned by the Apollo missions, it then seemed reasonable to assume that such an iron core, when it was molten, could have been the source of the early lunar magnetic field that magnetized the rocks.

Unfortunately the rocks returned were pieces of lava that had been broken from the bedrock by impacts and were lying loosely on the surface; their original orientations were not known, and so it was impossible to determine the directions in which they had been lying when they were magnetized or the direction of the field that had magnetized them. Instead work was begun on the more difficult task of determining the strength of the field that had been



SIMPLE ARGUMENT addresses the observation that the total field due to magnetized areas of lunar crust does not resemble a dipole field (a field like that of a bar magnet). If the moon's crust had been magnetized by an external field, such as the earth's (*top, 1*), the crust would have been magnetized uniformly. After the external field vanished (*top, 2*) the crust would create a small dipole field of its own. If, on the other hand, the crust had been magnetized by an internal dynamo in the core (*bottom, 1*), the direction and strength of the crust's magnetization would vary from place to place. When the dynamo action ceased (*bottom, 2*), the crust's magnetization would lead to no overall external dipole field. Thus the fact that the field due to magnetized crust is not a dipole field supports the author's hypothesis that the crust was magnetized by an internal dynamo.

responsible for the magnetization.

The method for determining the original field strength is in principle very simple: magnetize a sample in a laboratory field, and assume that the ratio of the applied laboratory field to the resulting magnetization is roughly the same as the ratio of the original lunar field to the original magnetization of the rock when it was found on the lunar surface.

In practice the work involves great difficulties. There are three general techniques by which it can be accomplished. The simplest technique is to heat a sample above its Curie temperature and then cool the sample in the presence of a magnetic field. This method gives a reasonable simulation of the conditions in which the original magnetization occurred, but it is invasive and may change the chemistry of the rock.

In a second technique (developed largely in our laboratory by Stephenson), which is somewhat less invasive, the sample is exposed to an intense alternating magnetic field and a weak constant magnetic field. The intense alternating field agitates the rock's magnetic structure in much the same way as heating the rock would, and it makes the rock susceptible to being magnetized by the weak constant field. The resulting magnetization can then be compared with the original magnetization of the sample to determine the ratio of the weak laboratory field to the original lunar field.

In the third technique (developed by Stan Cisowski and Fuller at Santa Barbara), which is still less invasive, one applies an extremely strong constant laboratory field without heating the rock or applying an alternating field. The aim is to determine how much magnetism could possibly be induced in the sample under controlled conditions. In a sense the experimenter is calibrating the rock to determine its natural capacity for being magnetized. It is then possible to estimate how close the sample's original magnetization was to its maximum possible magnetization, and thereby to estimate the strength of the original magnetizing field. This technique makes it possible to analyze a great many rock samples, but it is not as exact as the other two methods.

The results of these investigations were startling. Studies of rocks 3.9 billion years old indicated that the moon's magnetic field at the time the rocks were formed was about one gauss, almost twice the polar field of the earth today. In addition the moon's field seems to have decreased exponentially between 3.9 and 3.2 billion years ago, as one might expect if one assumes that the heat sources available to drive dynamo action in the core gradually declined over that period. The field's subsequent disappearance is also to be expected, because the heat sources would eventually have fallen below the critical value necessary to sustain dynamo action, or even to maintain the core's temperature above the melting point of iron.

Our results were received with some skepticism, in particular the suggestion that the moon, whose core would have to be much smaller than the earth's, could have produced such an intense field. Nevertheless, the fact that results from all three of our independent methods were in good agreement makes the results themselves as secure as they can be in a historical science: one cannot now reproduce exactly the conditions under which the lunar rocks originally formed and were magnetized, and so one must proceed by testing whether different lines of evidence all point to the same conclusion.

Laboratory examination of returned samples is not the only way to find clues to the nature of the moon's ancient magnetism. Many manifestations of the present magnetization of crustal rocks were observed by the various Apollo and Luna experiments. The experiments involved magnetometers placed on the surface by Apollo astronauts or carried on the Soviet wheeled vehicle *Lunakhod,* as well as magnetic surveys carried out by subsatellites launched into low-altitude lunar orbits (100 kilometers above the surface) by the *Apollo 15* and *Apollo 16* spacecraft.

One observation was that charged particles in the solar wind were occasionally deflected from areas of the moon's surface. Early observations of this effect were made by the *Explorer* lunar satellite, which from time to time noted that magnetic disturbances in the solar wind were deflected into the shadow cast by the moon. Even more interesting were the later observations by the Apollo spacecraft that low-energy electrons in the solar wind were sometimes reflected from the moon's surface. Both effects are due to regions of strong magnetic field just above the moon's surface caused by magnetic anomalies, whose presence indicates the existence of strongly magnetized strata of crustal rock. Such observations enabled Robert P. Lin and Kinsey A. Anderson of the University of California at Berkeley to map the strengths (but not the directions) of the surface magnetic field due to regions of magnetized crust.

For certain strips of the lunar surface, maps can be made of the direction as well as the strength of the surface magnetic field. Working with data from the *Apollo 15* and *Apollo 16* magnetometers, Paul J. Coleman, Jr., Christopher T. Russell and Lonnie L. Hood of the University of California at Los Angeles have drawn contour maps of the vertical component of the field at the crust as well as maps of the northward and eastward horizontal components of the field. The maps show that the moon does not at present have a significant dipole field (a field resembling that of a bar magnet). There is no dynamo within the moon, and the total magnetic field due to individual magnetized regions of the crust does not lead to even the smallest dipole field above the surface of the moon.

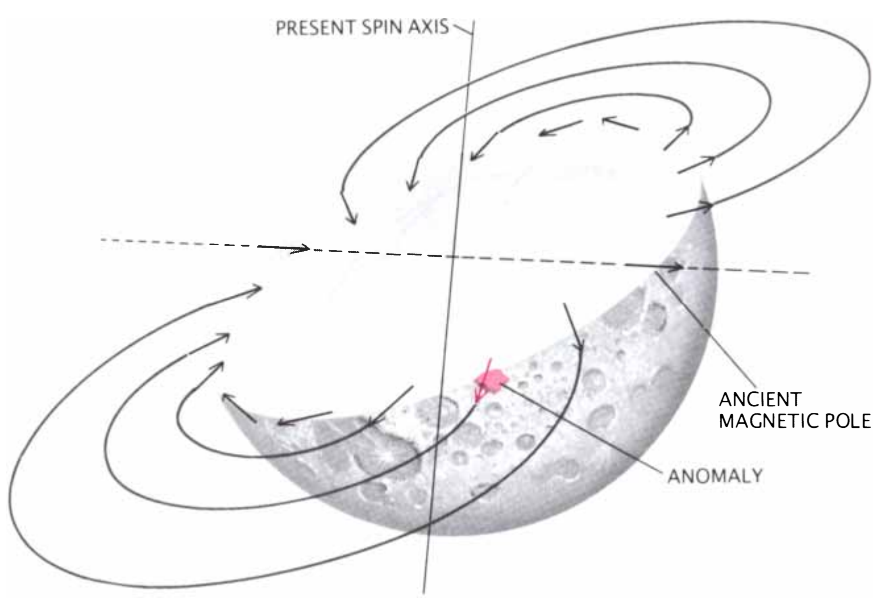
This result is actually a good argument in favor of the theory that the moon once had an internal dynamo. Many investigators who are uncomfortable with the idea that the moon had its own magnetic field have suggested that the lunar crust was magnetized by exposure to some strong external field (perhaps the earth's field when the earth and moon were closer). I have shown, however, by a relatively simple argument that such a uniformly magnetized outer shell would have an overall field resembling that of a dipole [*see illustration on opposite page*]. On the other hand, I showed that a crust that had been magnetized by an internal dynamo field that had since vanished would not have any dipole field above its surface. The moon fits this second model.

Coleman, Russell and Hood tried to determine the shapes and locations of the magnetized material responsible for each anomaly. They first assumed that the source of each anomaly was a simple dipole, but they found that in order to produce anomalies whose characteristics (such as the strength and direction of the field at a number of points in space) match those of the observed anomalies, such dipoles would have to be some 50 to 60 kilometers below the lunar surface. At that depth the ambient temperature is greater than the Curie temperature of crust minerals, and so any dipoles there would not remain magnetized.

On the other hand, it was easily shown that the anomalies could also be simulated as uniformly magnetized disks of crust material on or near the lunar surface. By adjusting the strengths and magnetization directions of hypothetical disks, and by moving them about on a computer-modeled lunar surface, Coleman, Russell and Hood found a good fit with the observed anomalies. In order for the anomalies to produce the observed strong fields at heights of from 20 to 100 kilometers above the lunar surface, the disks of magnetized material would have to be quite large—as much as 100 kilometers in diameter—and the magnetization of each would have to be uniform in direction over a considerable area. There is no problem with such a model if the magnetization was originally due to a lunar dipole. On the other hand, such wide areas of uniform magnetization direction would eliminate one contrary hypothesis: that local releases of energy, such as meteorite impacts, could have generated transient fields, magnetizing small regions of the crust (as a lightning bolt can magnetize a rock outcrop) and creating the anomalies. Such small local fields cannot match the data from the magnetic surveys.

To interpret these results, I calculated how the lunar field would have had to be aligned when the anomalies were created. In other words, I tried to find the direction of the magnetic poles of the field that had magnetized each anomaly.

I found that the anomalies fell into three groups. Within each group the projected north magnetic poles fall into two clusters that are antipodal to each other (that is, they are on opposite sides of the moon). Each pair of clusters thus defines an axis of the moon. The axis defined by each group points in a very different direction from the axes defined by the other two groups and from the moon's present axis of rotation: the moon seems to have had three separate sets of magnetic poles. The reason ancient north poles are found at both ends of each axis is probably that the lunar field switched its polarity without changing the orientation of its



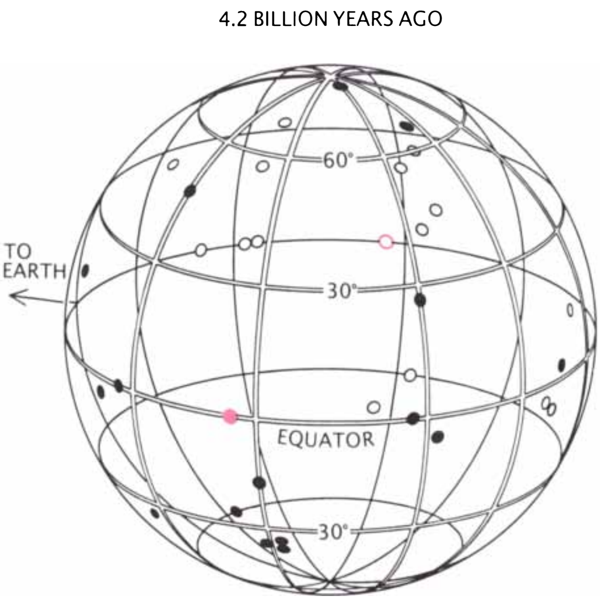
ORIENTATION of the moon's ancient magnetic field can be deduced from the magnetization of a magnetic anomaly (a magnetized region of the crust). If one knows the direction of the anomaly's magnetization (*color*) and assumes that the anomaly was originally magnetized by a dipole field, one can reconstruct the field that was present when the anomaly was magnetized (*black lines*) and find where the magnetic poles were then.

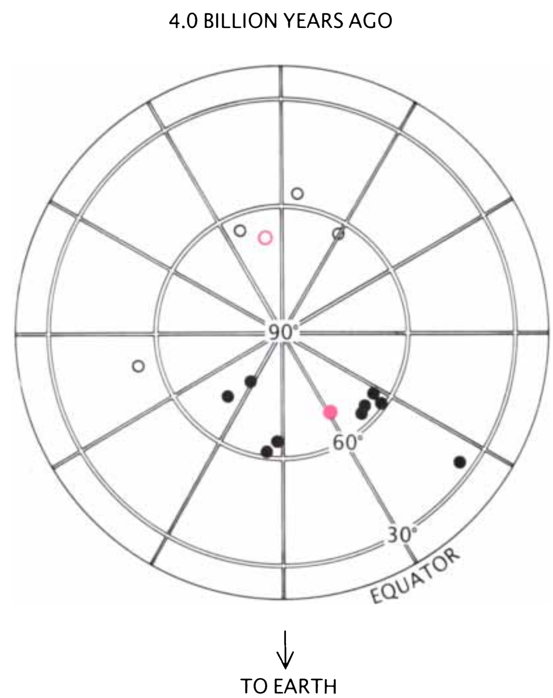
axis, a phenomenon that has taken place hundreds of times on the earth. But how is one to explain the three independent axes?

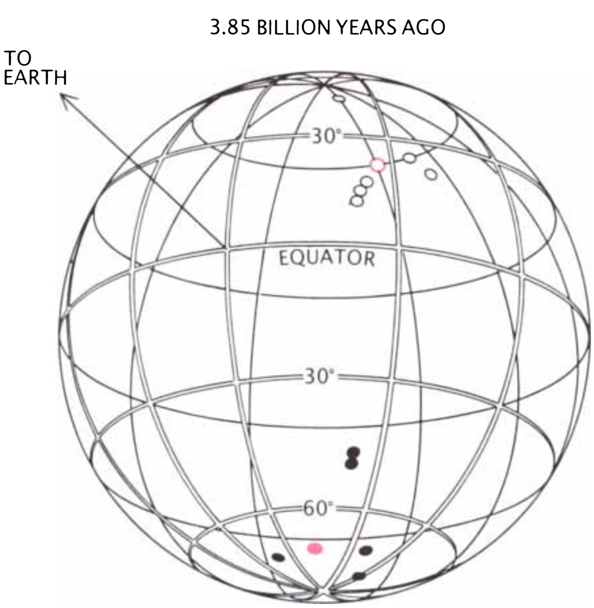
A clue comes from the ages of the anomalies. It is difficult to establish the time at which each anomaly was magnetized, because the anomalies are generally in regions from which samples have not been returned, but one can estimate their approximate dates by geological mapping of the moon (work undertaken largely by Don E. Wilhelms of the U.S. Geological Survey). Such analysis indicates that the anomalies within each group were formed at roughly the same time as one another but that the three separate groups of magnetic anomalies were formed in different periods of the moon's history. The earliest period was about 4.2 billion years ago, the intermediate period was about 4.0 billion years ago and the most recent period was about 3.85 billion years ago.

Thus the axis of the moon's magnetic field pointed in different directions at different times. If one assumes that the axis of the moon's magnetic field was always roughly aligned with the moon's spin axis (the assumption is justified by studies showing the overwhelming role of the Coriolis force—a complex force due to the spin of the earth—in the magnetohydrodynamics of the earth's core), then the implication is that at different times the moon's spin axis must have pointed in different directions with respect to the surface. How is this possible?

The answer, polar wandering, is actually an old idea that was first proposed late in the 19th century to explain how parts of India, Australia and Africa, which are now near the earth's equator, could have experi-

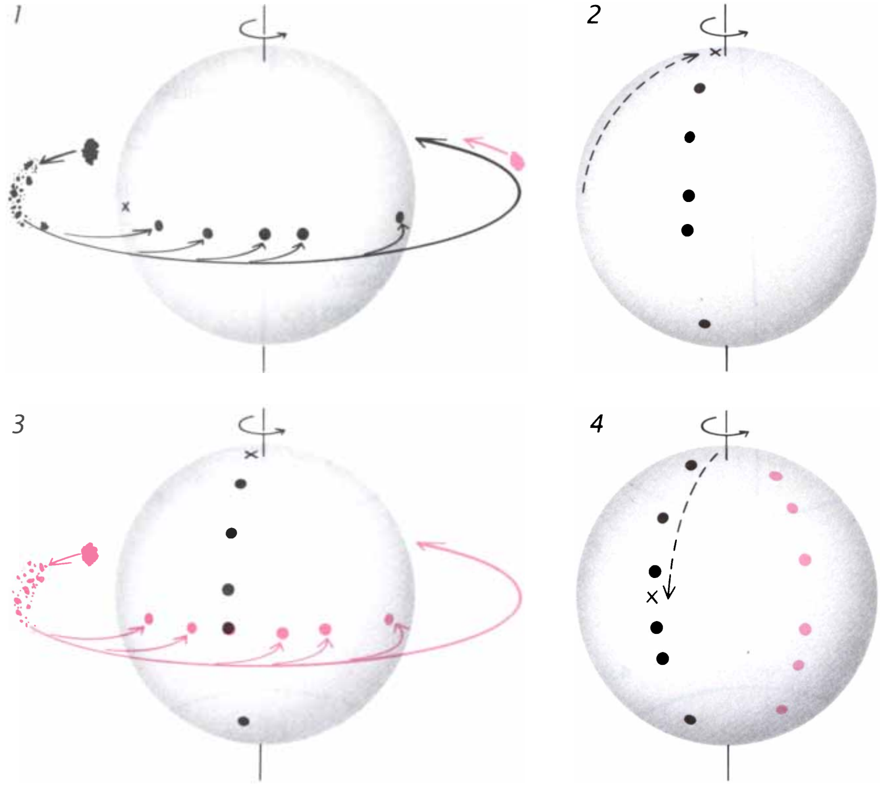






ANCIENT NORTH POLES of the moon's magnetic field, as deduced from magnetic anomalies, fall into three groups. Each pole position was deduced from a single anomaly. Within each group the anomalies from which the pole positions were deduced are about the same age; the oldest group is roughly 4.2 billion years old, the intermediate group about 4.0 billion years old and the youngest group about 3.85 billion years old. In each group (particularly the youngest two) the projected magnetic axis is about the same for every anomaly, but in some cases the north pole is at one end of the axis and in others it is at the other end. (Here open circles represent points on the far side of the moon as it appears on the page and solid dots represent points on the near side; colored circles represent the average position of the projected poles on the far side and colored dots the average position of projected poles on the near side.) The moon's magnetic field seems to have reversed its polarity several times.

enced glaciations 200 million years ago. In the analysis of the earth's history polar wandering has largely been replaced by the idea of continental drift.



HYPOTHETICAL HISTORY OF THE MOON was deduced from the positions of the moon's ancient magnetic poles and from the positions, shapes and ages of certain large impact craters. In this model the moon was originally orbited by several satellites. One of the satellites, on being drawn inward by tidal forces, broke into fragments that all struck the lunar surface within a comparatively short time (*1*), leaving a trail of impact basins near the equator. The consequent redistribution of mass on the moon's surface caused the moon to shift its orientation with respect to its own spin axis, so that many of the basins were far from the equator (*2*). Then another lunar satellite was drawn in and fragmented (*3*), leaving a trail of basins near the new equator, and the moon reoriented again (*4*). A similar sequence of events probably happened a third time also.

Polar wandering is based on fundamental principles in the physics of spinning bodies [see "Polar Wandering on Mars," by Peter H. Schultz; SCIENTIFIC AMERICAN, December, 1985]. As Leonhard Euler showed, bodies spin most stably when the maximum possible amount of mass is far from the axis of spin. (The idea makes intuitive sense: when a body is spinning, centrifugal force will tend to pull the massive parts of the body outward, away from the spin axis.) If a body is spinning stably and its distribution of mass is suddenly altered, the body will tend to wobble and realign itself until it is once again spinning with the maximum possible amount of mass far from the axis of rotation.

According to the law of conservation of angular momentum, the position and orientation in space of the spin axis itself cannot change. Rather, the body reorients itself so that the spin axis goes through a different part of the body. Hence one could say that it is really the surface that wanders in relation to the pole (the point in space where the spin axis intersects the surface). Regions that had been at the poles move closer to the equator, and regions at the equator move closer to the poles. If polar wandering had occurred on the moon, then at different times the magnetic poles would indeed have been at different locations on the surface and the magnetic axis would have had different orientations in relation to the moon as a whole.

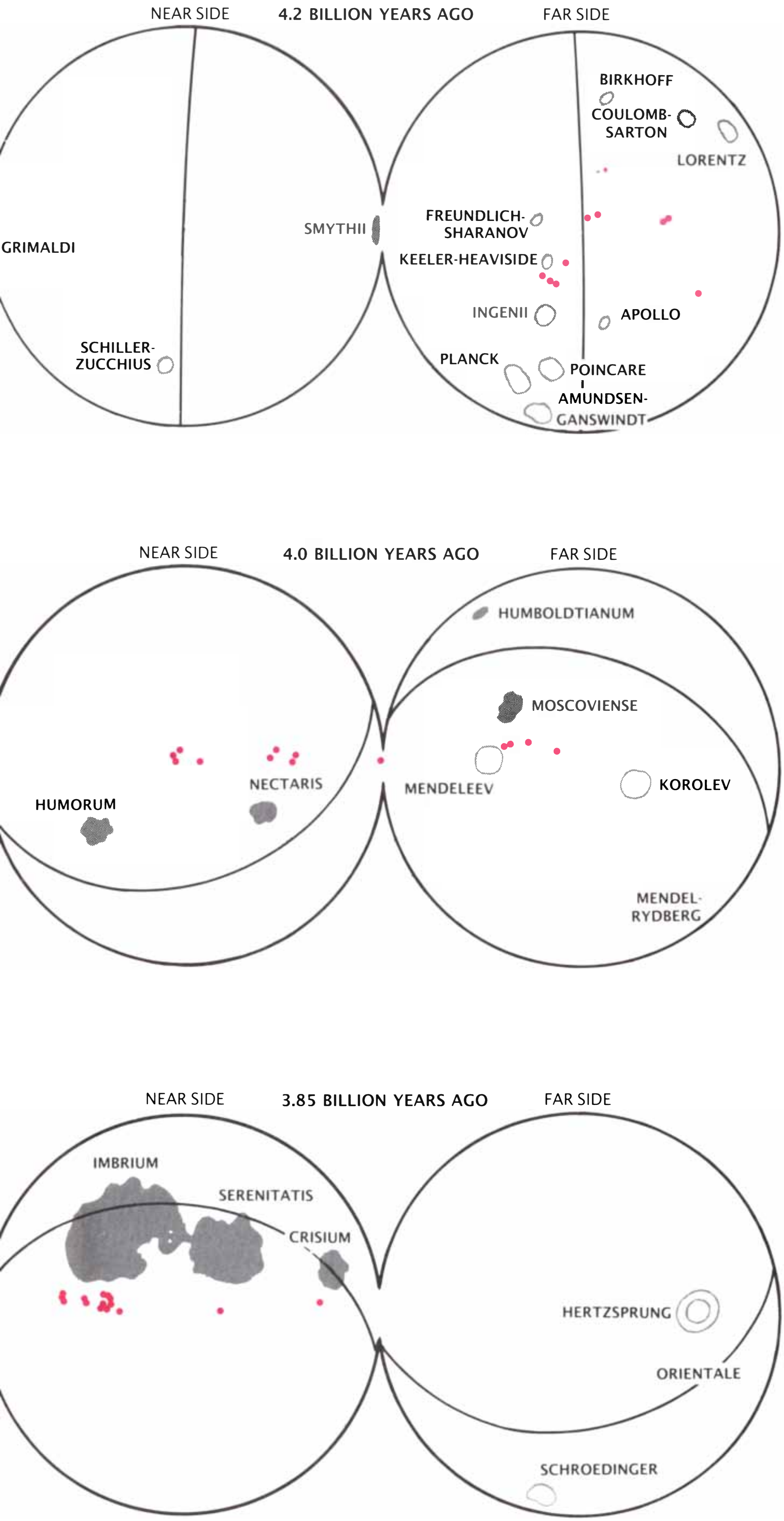
What could have caused polar wandering on the moon? The answer to this question leads to some of the most exciting possibilities raised in my investigation. When I plotted where the moon's equator would have been in each of the three periods during which the magnetic anomalies were formed, I noticed that many of the large impact basins formed in each period were near what was then the moon's equator. This turned out to be the clue to two important discoveries.

First of all, it suggests that the moon was reoriented at least three times after planetesimals (small objects that were the original bodies in the solar system) had collided with it near its equator. Every impact would have created a basin—a region of low mass—near the equator, far from the axis of spin. That is not a stable configuration. The moon would therefore have reoriented itself so that more massive regions were near the equator and the basins lay close to the pole. Actually this seems to be what happened each time [*see illustration on next page*]. The picture is a little oversimplified; there are a number of basins near each paleoequator, and so during each period there must have been a series of impacts separated from one another by relatively short intervals.

Secondly, the fact that the planetesimals collided with the moon near its equator suggests they were not asteroids or comets orbiting the sun, which would have struck the moon anywhere on its surface. Instead they were probably satellites of the moon itself, orbiting the moon in its equatorial plane (much as the satellites of Mars, Jupiter, Saturn and Uranus orbit those planets nearly in their equatorial planes).

These satellites would gradually have been drawn in toward the moon by tidal action (as Phobos is now being drawn toward Mars). I suggest that there were at least three satellites. As the first satellite was drawn toward the moon it would have been broken apart by tidal forces. Its fragments would have struck the moon within a short time of one another, leaving a trail of impact basins and causing the moon to reorient itself. Some 200 million years later the next satellite would have been broken apart by tidal forces, and so on.

Analysis of the shapes of the impact basins reinforces this hypothesis. The fragmented satellites would have approached the lunar surface at a very small angle as their orbits slowly decayed. Experiments on hypervelocity impacts done by Donald E. Gault of the National Aeronautics and Space Administration's Ames Re-



search Center have shown that at angles of approach smaller than about five degrees the debris from an impact tends to be shot outward in a butterfly-shaped pattern and the impact basin tends to be elliptical. Such patterns are obvious in the photographs of several of the great basins on the moon. In addition, the long axes of these elliptical basins are roughly parallel to the line followed by the equator during the period when the basins were made, indicating that the impacting objects were traveling in or near the plane of the lunar equator at the time of impact.

I thus conclude that the moon had a system of satellites early in its history. This first evidence for the existence of other bodies in the earth-moon system should carry valuable information concerning the origin of the earth-moon system and of the planets generally.

Investigators have a reasonably clear picture of the early evolution of the solar system: a cloud of interstellar gas and dust collapsed from a diameter of about 100,000 astronomical units (an astronomical unit is the mean distance from the earth to the sun, about 93 million miles) to a diameter of about 50 A.U.'s, flattening into a disk as it collapsed. Our picture of the later stages is quite sketchy, however. Small grains of silicates and iron must have accreted somehow to form millions of small bodies, which must themselves have accreted to form larger bodies. The final stages of this accretion are visible in the many impact craters with which such surfaces as those of the moon and Mercury are peppered, but there is almost no evidence concerning earlier stages of the accretion process. The suggestion that there were rather large satellites in the early earth-moon system may well make it possible to reconstruct some of this intermediate phase in the formation of the solar system.

ANCIENT EQUATORS of the moon are marked on images of the moon as it is aligned today. Solid black lines represent the equator in each of the three periods depicted. Dots show the positions of the magnetic anomalies whose directions of magnetization led to knowledge of the ancient pole positions from which these ancient equators were derived. The major basins whose names are given, which all lie near ancient lunar equators, may have been caused by the impact of fragments of satellites that once orbited the moon.

INTRODUCING
DODGE ARIES AMERICA

$6995.* IF YOU'RE LOOKING FOR THE BEST VALUE IN ITS CLASS,** A CAR THAT GIVES YOU 2.2 LITERS OF FUEL-INJECTED POWER, ROAD-GRIPPING FRONT-WHEEL DRIVE, OVER 30 OTHER STANDARD FEATURES, PLUS OUR UNBEATABLE 7/70 PROTECTION PLAN† ALL FOR A NEW LOW PRICE, THEN LOOK NO FURTHER...IT'S GOTTA BE THE 1988 DODGE ARIES AMERICA.

7/70

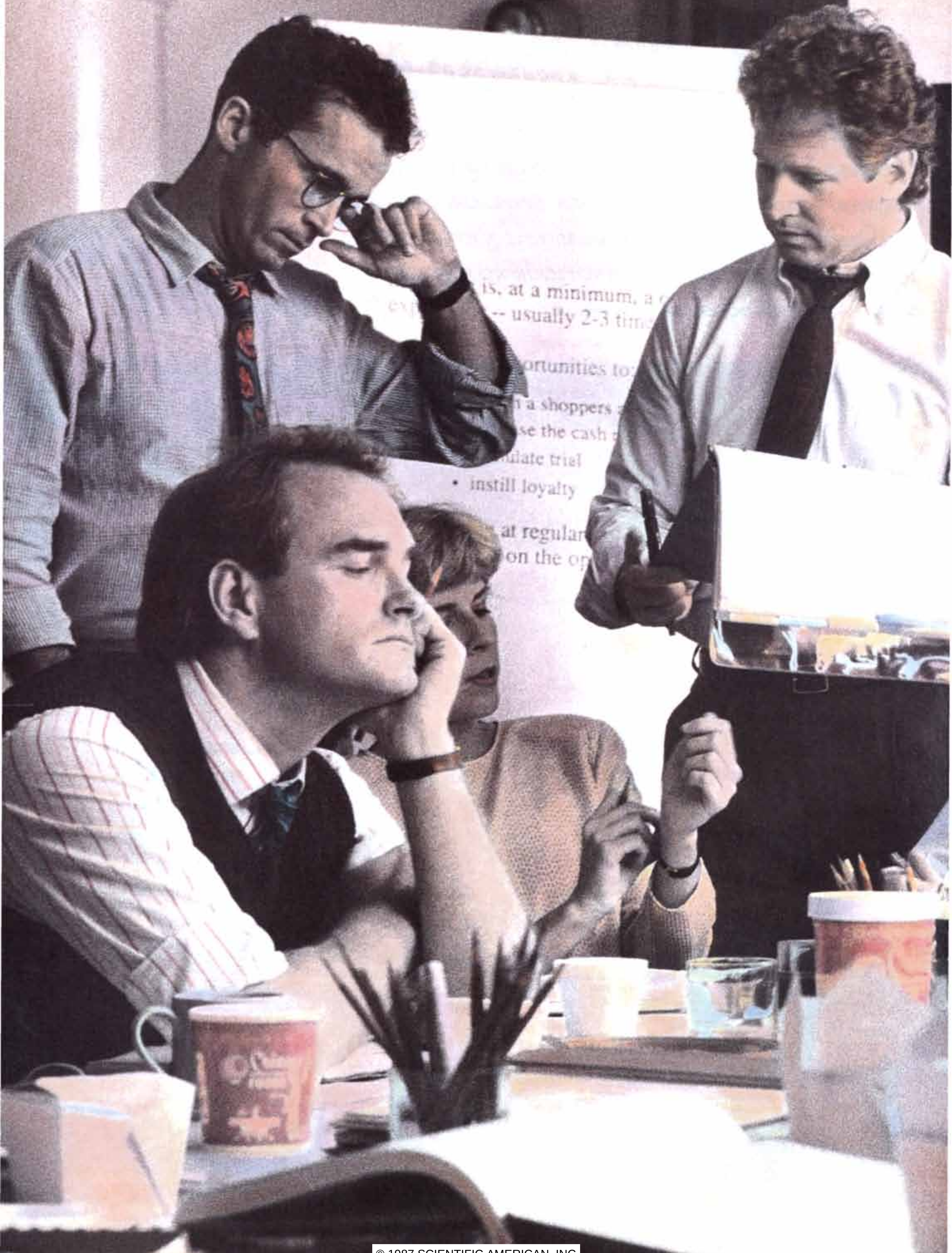
Dodge

IT'S GOTTA BE A DODGE.

DIVISION OF CHRYSLER MOTORS

*Sticker price. Tax & destination charge extra. **Based on comparably equipped sticker price comparisons to 1987 competitive mid-size models (1988 competitive data incomplete at time of printing).
†Based on warranty comparison with competitively priced models at time of printing. See 7/70 powertrain & 7/100 outer body rust-through limited warranty at dealer. Restrictions apply.

BUCKLE UP FOR SAFETY.


• instill loyalty

In the good old days, there were no free lunches. But a lot of free dinners.

We practically lived in that conference room, and did everything face to face. 'Working small' we called it. Until we started getting bigger and nearly grew ourselves into a company of strangers. Well, we got quickly re-acquainted, after we talked to AT&T. They gave us an easy way to plug in new people and their PCs so we could function as one. As if we were all still sitting around one big table. Now we're 'working small' again. But on a bigger scale. Which just goes to show. Success breeds success, and

AT&T comes through.

The new AT&T System 75 XE

Flexible in both size and function, this expandable voice and data switch is designed to grow as your business grows. Serving 40 users today, up to 600 tomorrow.

AT&T's modular architecture and once-and-for-all wiring scheme make it easy to add voice or data stations the same way you add staff: as needed.

Software upgrade packages make it just as easy to add features – AT&T's AUDIX message-recording system, for example – or to use System 75 as the hub of a company-wide network.

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.

How Animal Cells Move

They do so by bringing pieces of the outer membrane into the cytoplasm and then recycling them to the surface in a directed way. Nutrients are brought into the cell by the same process

by Mark S. Bretscher

Movement is a fundamental property of living things. Indeed, along with such activities as respiration and reproduction, locomotion is one of the qualities that distinguish organisms from nonliving things. For an animal the capacity of its cells to move may be crucial in both health and disease. White blood cells must be able to move through the body to fight pathogens; in response to injury, fibroblasts migrate into a wound to take part in the healing process. Cell movement can also have deadly results. Individual cancer cells sometimes break loose from a tumor and travel to other parts of the body to establish new tumors in the process called metastasis. Cell locomotion, then, has a fundamental role in life, and sometimes in death. Yet until recently little was known about the details of how an animal cell actually moves.

Over the past decade, however, investigators have learned a great deal about the processes underlying locomotion. Here I shall describe how a moving animal cell brings pieces of its outer membrane into the cell and later returns them to the surface in the process called the endocytic cycle. The cycle has a dual function: it brings nutrient molecules into the cell and it is essential for cell locomotion. There are two aspects to its role in locomotion. The cycle provides the means whereby a cell extends itself forward; it may also generate the motive force that propels the cell. Whether it does in fact provide the motive force is still unproved, but I shall show how these two aspects of locomotion are inseparable.

Except for sperm, all animal cells are believed to move by the same mechanism. Sperm cells have an active flagellum that operates much like the tail of an eel, enabling the sperm to "swim" forward. All other animal cells "crawl" forward over a substratum rather than swimming through a liquid medium, and it is the great majority of crawling cells that form the subject of this article. In particular I shall concentrate on the fibroblastic cells that migrate into a wound and later help to pull the torn edges together. Of all the mammalian cells that move, fibroblasts have been studied the most extensively, largely because they are easy to grow in laboratory culture.

Although fibroblasts are often stationary in culture, they may begin to move when they become crowded together. When a fibroblast is in motion, it is often roughly triangular in top view: one side of the triangle forms the front, or leading edge, and the cell tapers off behind to a narrow trailing edge. Fibroblasts move slowly by macroscopic standards—a centimeter a day—and their motion is best observed in time-lapse photography. Seen in that way the leading edge is much more active than the rest of the cell. Its activity does not follow a predictable pattern, however. One region may briefly extend forward over the substratum; then, when the activity subsides, another region along the leading edge may spread out. The sequence is repeated until the cell has advanced along its entire front.

Sometimes the irregular forward spreading of the cell goes astray, because the new extension of the leading edge has not attached itself firmly enough to the substratum. When that happens, the detached segment rises up and curls back over the cell; eventually the extension is resorbed into the cell body and disappears, and a new one takes its place. In view of all this disorderly lifting and fluttering, the leading edge of a fibroblast is often called the ruffling edge. The ruffling is seen only on certain substrates. On other substances the leading edge attaches itself with such efficiency that there is no obvious ruffling, and then the cell oozes smoothly forward.

Many years of observation of fibroblasts led biologists to conclude that the ruffling edge is where the action is in cell movement. A series of experiments beginning in 1970 have confirmed that the leading edge is indeed the site of a crucial piece of the action. Those experiments have also helped to show how the events at the leading edge are linked to another process that takes place throughout the cell.

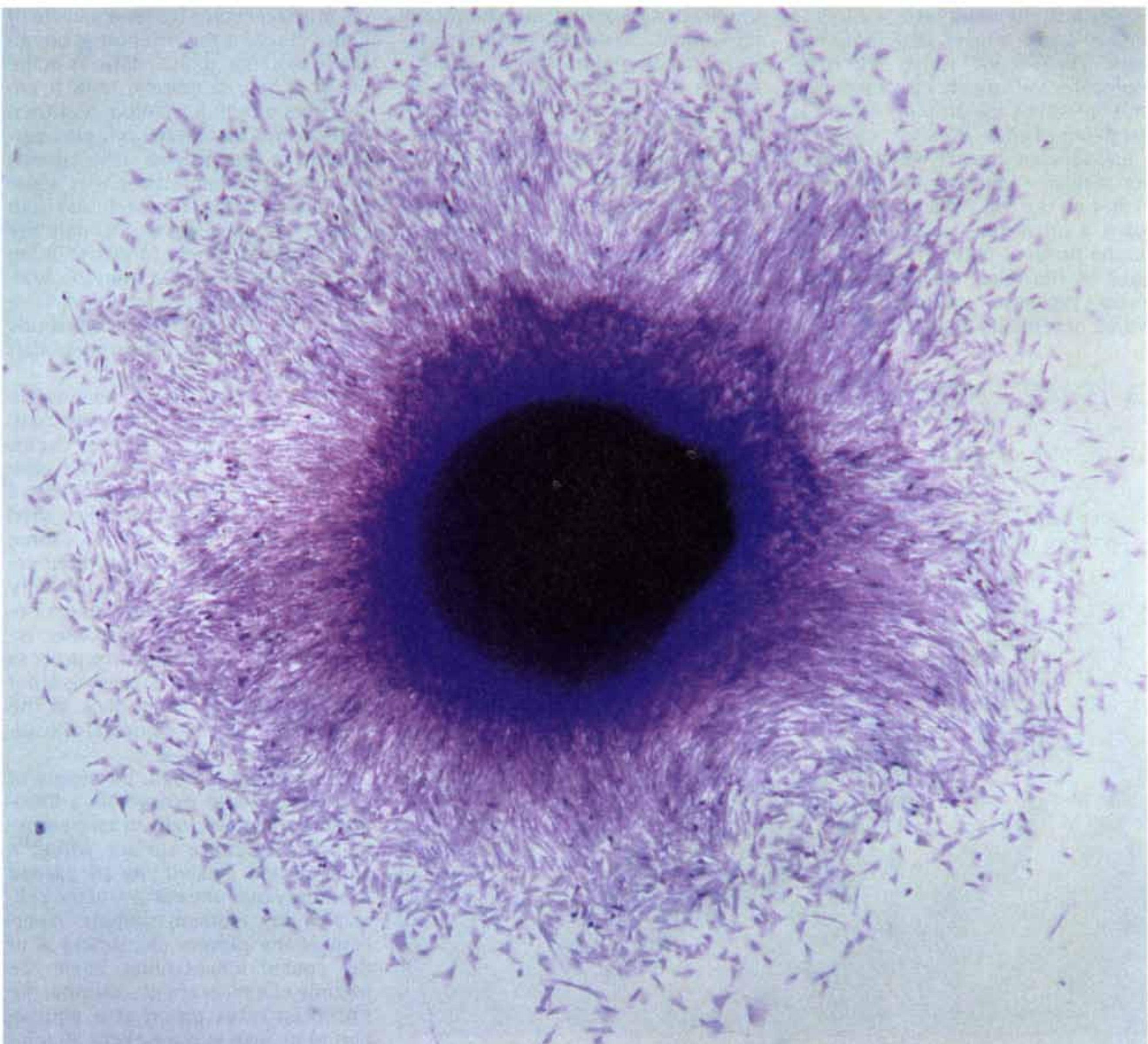
In 1970 a key experiment was carried out by Michael Abercrombie, Joan E. M. Heaysman and Susan M. Pegrum, then at University College London, who observed what happened when moving fibroblasts encountered microscopic pieces of charcoal scattered on the substratum. Sometimes the cell simply crawled over the charcoal. At other times, however, the particle became attached to the upper surface of the leading edge. When that happened, the particle was observed to move backward on the surface of the cell as the cell made its way forward. The motion of the particle followed a more or less straight line, and eventually the little black dot came to rest near the trailing edge.

Abercrombie and his co-workers proposed an illuminating explanation for these striking results. They suggested that since the particle is moving backward and is attached to the plasma (outer) membrane of the cell, the plasma membrane itself must also be moving backward: there must be a continuous flow of membrane from the leading edge of the cell to its trailing edge. To account for the flow, Abercrombie's group proposed that membrane material is somehow resorbed at the back of the cell and then transferred through the cell to the front, where it is reinserted in the plasma membrane. They saw that such a membrane-transfer cycle would enable the cell to extend itself forward over the substratum.

The cycle proposed by Abercrombie resembles the propulsive mechanism of a tank. In his model the plasma membrane was seen as a continuous flexible sheet analogous to the tread of the tank. The membrane covering the cell's lower surface, adjacent to the substratum, corresponds to the lower tread of the tank, which is in contact with the ground and moves backward with respect to the tank. The intracellular transport of membrane from the rear of the cell to the front corresponds to the tank's upper tread.

Experiments in many laboratories have shown that the Abercrombie model is correct in outline, but since its formulation major advances have been made in understanding the structure of cell membranes and the process by which membrane is recycled in the cell [see "The Molecules of the Cell Membrane," by Mark S. Bretscher; SCIENTIFIC AMERICAN, October, 1985]. Such advances have at the same time confirmed Abercrombie's overall model and shown where it was lacking.

The basis of all cellular membranes is a double layer of lipid (fatty) molecules. Into this lipid-bilayer matrix are inserted a host of proteins with different functions. A key property of the bilayer is that although it is a stable planar structure, within the plane it is a liquid: lipids and proteins are free to diffuse sideways in any direction. The distance traveled by a molecule depends on



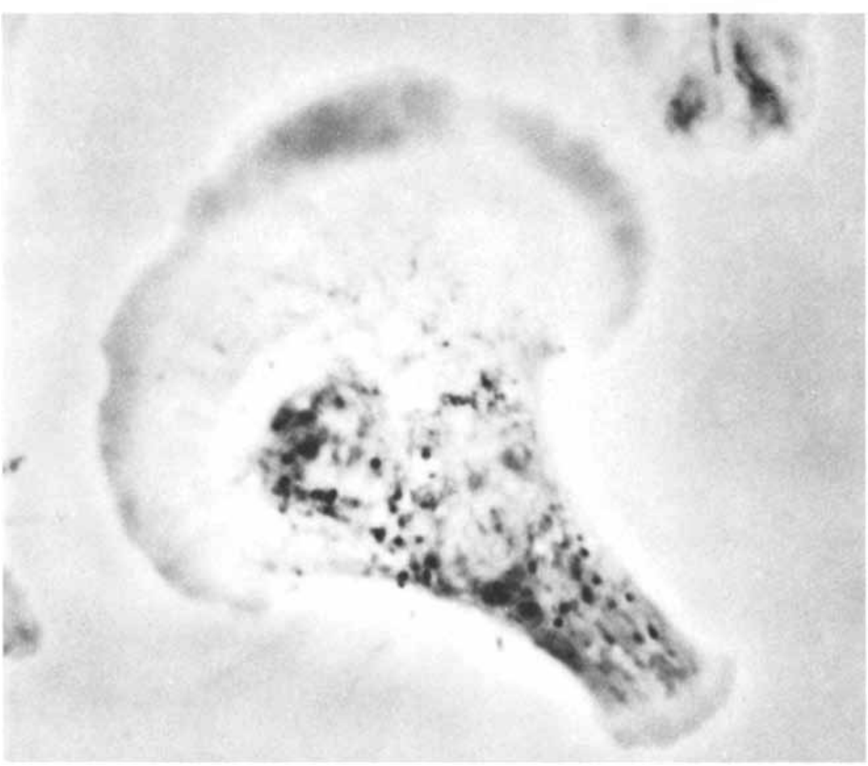
FIBROBLASTS MIGRATE out and away from a piece of embryonic chicken-heart tissue placed on a glass slide. Most animal cells are thought to have the capacity for movement, but not all of them put it to use; fibroblasts, which participate in the healing of wounds, are among the small class of cells that need to move if they are to carry out their normal functions. The photomicrograph was made by Peter A. Lawrence of the Medical Research Council's Laboratory of Molecular Biology in Cambridge.

the square root of the length of time it moves and on its size and shape. Because lipids are smaller, they move more quickly than proteins. A typical protein might move 10 micrometers (millionths of a meter) in one minute. Since a fibroblast is, say, 50 micrometers long, an average protein could diffuse from one end to the other in half an hour.

The membrane and its associated proteins are heavily implicated in the two routes by which nutrients can enter the cell. The first route (which is not directly relevant here) is by means of a variety of protein pumps and channels that span the plasma membrane. Such pumps and channels transport a variety of small molecules into the cytoplasm, including amino acids, sugars and inorganic ions. The second route, taken by molecules too large to be transported into the cell by the protein carriers, is by means of endocytosis.

Endocytosis begins when a nutrient molecule binds to its specific receptor on the cell surface. The receptor is a membrane protein, and the bound nutrient is referred to as a ligand. A fibroblast has some 50 different types of receptor, each of which accepts its ligand and no other. After the ligands bind to their receptors they are rounded up—about 1,000 at a time—at a spot where the plasma membrane has begun to dip inward. The indentation is called a coated pit, because of the characteristic protein layer found on its cytoplasmic side. The coated pit buds inward and pinches off in the cytoplasm as a little balloon of membrane called a coated vesicle. It is the coated vesicle that brings the ligands into the cell.

Much has been learned in recent years about coated pits, coated vesicles and the receptors they transport, and the new knowledge, from many groups of investigators, has implications reaching far beyond cell movement. Coated pits and vesicles were discovered in 1964 by Thomas F. Roth and Keith R. Porter of Harvard University. In 1975 Barbara M. F. Pearse of the Medical Research Council's Laboratory of Molecular Biology in Cambridge showed that the furry coat of the pits and vesicles is a basketlike network whose raw material is a protein to which Pearse gave the name clathrin. Within the basket of clathrin there is a sphere of membrane; in the membrane there are a variety of proteins, among them the joined receptors and ligands.



FIBROBLAST IN MOTION displays the vaguely triangular shape characteristic of these cells when they move. The convex gray border is the cell's large leading edge, or front. From the front the cell tapers to a narrow rear, or trailing edge. The fibroblast, enlarged here some 2,000 diameters, "crawls" forward by extending first one part and then another of its leading edge. The pace is slow: the cell covers only a centimeter a day.

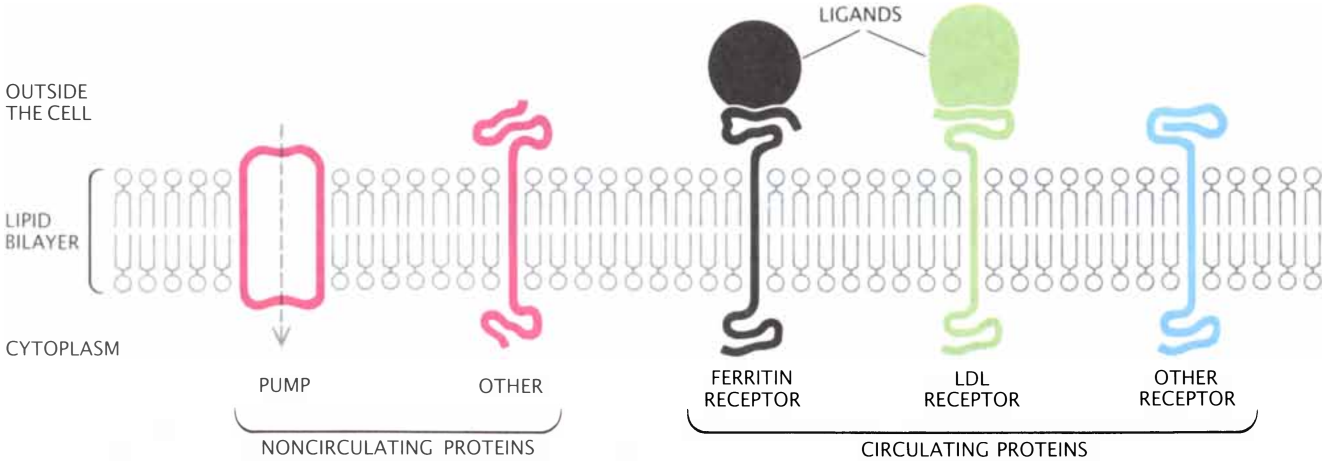
The receptor that has been best described is the one for low-density lipoprotein (LDL), which has been studied in exquisite detail by Michael S. Brown and Joseph L. Goldstein of the University of Texas Health Science Center at Dallas. The function of the LDL receptor is to bring cholesterol-laden LDL from the blood plasma into the cell, where cholesterol is utilized for making new membranes. Other well-known receptors include those for iron-bearing proteins called transferrin and ferritin. The transferrin and ferritin receptors were crucial for my work on cell locomotion.

The binding of the ligand to an unoccupied receptor initiates a cycle of endocytosis. If the receptor is not already in a coated pit, it diffuses in the plane of the membrane until it encounters one in formation, and then joins it. Not all surface proteins participate in this process. One current mystery of endocytosis is why some proteins (circulating proteins) join coated pits and cycle through the cell, whereas others (noncirculating proteins) are excluded from the coated pits. When enough receptors have joined the coated pit, the pit buds into the cytoplasm in a process that takes about a minute.

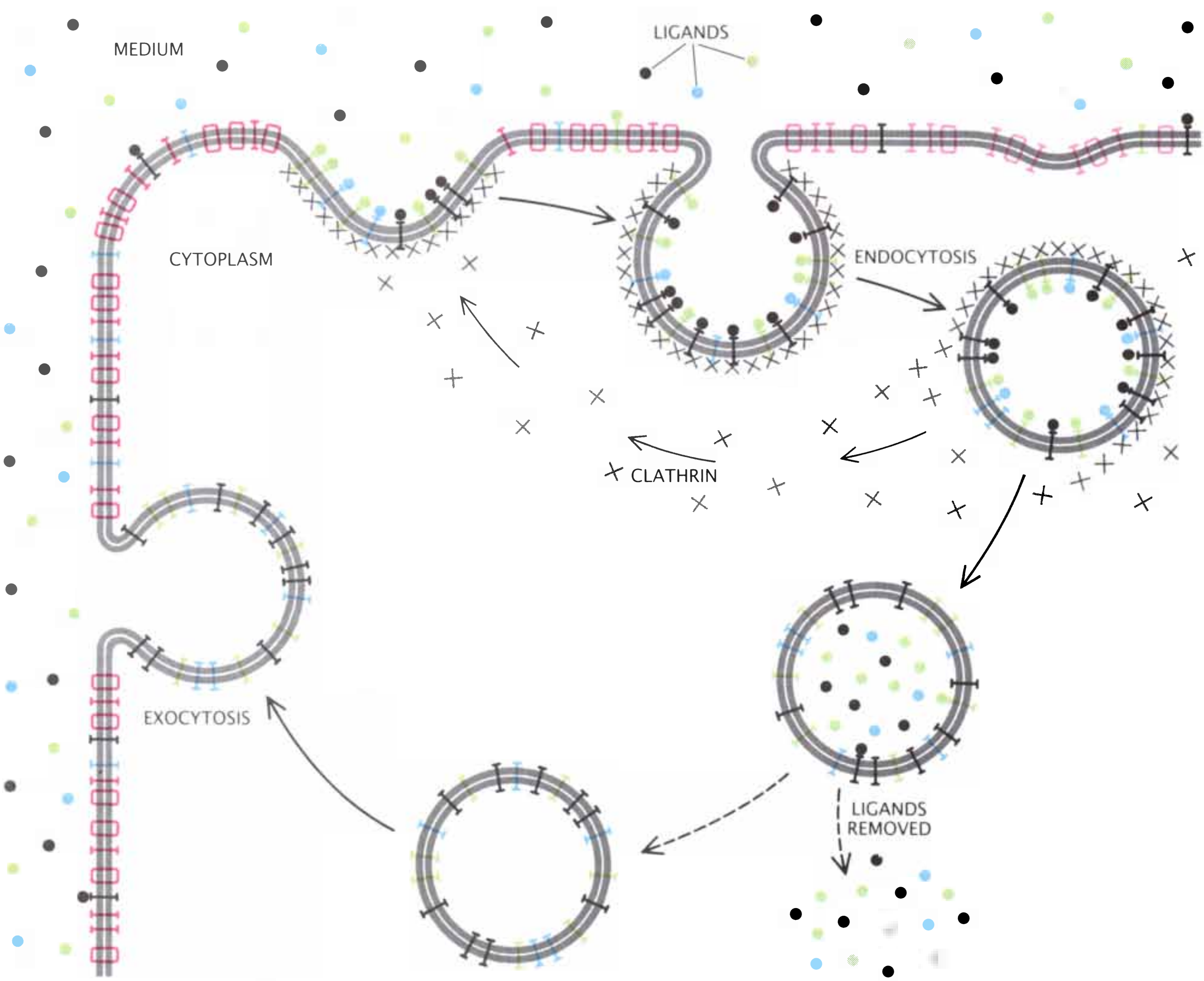
Once it is inside the cell, the vesicle almost immediately sheds its coat, releasing clathrin for further cycles of endocytosis. The uncloaked vesicle—including ligands, receptors and lipids—fuses with an organelle called an endosome. Inside the endosome ligands are separated from their receptors in a process that is not fully understood. The receptors and associated lipid molecules are later returned to the membrane in a process called exocytosis, which is also not fully understood. Once back in the membrane, the receptors are available for the next cycle.

Remarkably enough, by means of endocytosis and exocytosis a fibroblast is able to recycle an area equivalent to its entire surface within a short period. Coated pits are spread randomly over the surface of the cell, and at any moment roughly 2 percent of the plasma membrane is in the coated indentations. Since the lifetime of a pit is about a minute, the fibroblast takes up an area equivalent to its own surface every 50 minutes. It is this dramatic recirculation of the cellular envelope that drives cell movement.

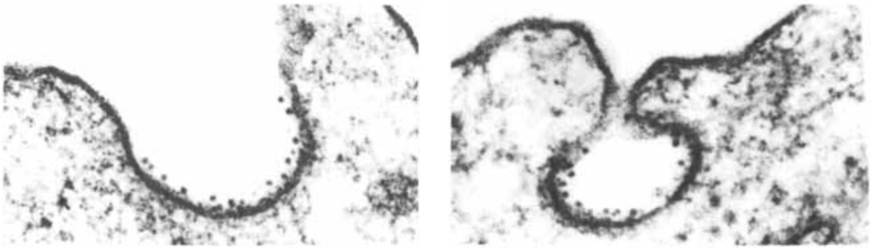
Now, in Abercrombie's model of that movement the membrane car-



PROTEIN MOLECULES are inserted into the plasma (outer) membrane of the animal cell. The matrix of the membrane is a bilayer of lipid (fatty) molecules. The inserted proteins have a variety of functions. Some of them are pumps for small molecules. Others are receptors that bind nutrients (often called ligands), which are then brought into the cell through the process of endocytosis. Some surface proteins, including receptors, circulate through the cell in the endocytic cycle and others do not.



ENDOCYTIC CYCLE draws membrane into the cell and later returns it to the surface. The receptors, usually carrying a bound ligand, join a "coated pit," an indentation on the cell surface that has a layer of a protein called clathrin on its cytoplasmic (inner) side. The pit then balloons into the cytoplasm and pinches off to form a coated vesicle, which soon sheds its clathrin coat. Next the ligands are separated from their receptors and removed by complex processes that are not well understood. Eventually the membrane lipids and receptors rejoin the cell membrane in a process known as exocytosis, which also is little understood.



COATED PITS form and bud into the cytoplasm in electron micrographs made by Nichol Thomson and the author. The pits' dense border is the clathrin coat. The black dots in the pits are molecules of electron-dense ferritin that are bound to ferritin receptors.

ried through the cycle of endocytosis is returned to the cellular surface at the leading edge. But is that actually the case? Is the recycled membrane reinjected at the front of the cell? The question is a key one but is difficult to test directly. All is not lost, however. Knowing that the circulating membrane is limited to a handful of receptors and lipids gives rise to a prediction that is subject to test. The prediction concerns the distribution of receptors on the surface of the cell, and in 1983, at the Laboratory of Molecular Biology, Nichol Thomson and I decided to put it to the test.

If membrane is exocytosed only at the front of a large cell, it follows that the receptors should be concentrated there. The reasoning is as follows. Let us say that between rounds of endocytosis a receptor spends four minutes on the cell surface before it is captured by a forming coated pit. Like other proteins, receptors are free to diffuse in the plane of the membrane, and in its four minutes on the surface a typical receptor might diffuse, say, 20 micrometers. On a smallish cell—10 micrometers long—the receptors would diffuse all the way from front to back in less than four minutes, and so at any instant the cell would have an even surface distribution of these receptors. On a 100-micrometer cell, however, a typical receptor would diffuse only a fifth of the way back before being recaptured and returned to the front. As a result most receptors would be bunched toward the leading edge.

Such calculations led Thomson and me to expect that the receptors would be concentrated at the leading edge, an expectation that is readily tested. We decided to examine the distribution of ferritin receptors on a large cell with a ruffling edge. Ferritin receptors were chosen for two reasons: they are circulating proteins and they are easily detected in the electron microscope. For convenience we worked with giant HeLa cells, which can be induced to grow from the HeLa cell line originally derived from a human tumor. Giant HeLa cells resemble fried eggs. They are round and flat, with a nucleus in the middle where the yolk would be and with a ruffling edge all the way around. Because the ruffling edge fringes the entire cell, the cell tries to move in all directions at once, with the result that it does not translocate. It seems likely, however, that the mechanism of formation of the ruffling edge is identical with the mechanism in cells that do translocate.

Thomson and I briefly exposed the giant HeLa cells to ferritin at zero degrees Celsius, a temperature low enough to slow diffusion and prevent endocytosis. Under those conditions ferritin receptors on the surface of the cell would bind ferritin molecules, but the distribution of ferritin receptors would not be changed by diffusion or endocytosis. The result would therefore be a "snapshot" of the distribution of receptors at one moment. After the cells had been labeled with ferritin they were fixed, sliced into thin sections and examined in the electron microscope for ferritin molecules bound to their receptors. As predicted, the ferritin receptors were indeed highly concentrated at the cell's leading edge. A similar distribution was found for LDL receptors and transferrin receptors on giant HeLa cells. Noncirculating proteins, on the other hand, were found to be distributed more or less at random on the surface.

Similar observations were made at about the same time on moving fibroblasts by Peter Ekblom, Irma Thesleff, Veli-Pekka Lehto and Ismo Virtanen of the University of Helsinki. Taken together, these experiments clearly imply that exocytosis occurs at the front edge of a motile cell. Moreover, straightforward calculations show that the process of exocytosis directed toward the leading edge is sufficient to account for the actual forward movement seen in fibroblasts.

As I mentioned above, endocytosis leads to the resorption of about 2 percent of the fibroblast's surface in a minute. Exocytosis of all that membrane at the front of the cell would yield a forward extension of about 1 percent of the cell's total length (since half of the exocytosed membrane goes to the cell's upper surface and half to the lower surface). If the fibroblast were 100 micrometers long, its leading edge would extend forward at a rate of about one micrometer per minute, and an object attached to the membrane (such as a charcoal particle) would move backward at the same pace—which agrees nicely with the rate observed in Abercrombie's work with charcoal particles.

So far the "tank" analogy has provided a fairly apt description of membrane flow. Yet if it is pushed further, the tank metaphor breaks down. One reason is that the entire tank tread is carried along in its cycle, whereas only some membrane components (the lipids and the circulating proteins) participate in endocytosis. Another significant difference is that although the tank tread moves at an even pace throughout its cycle, the membrane flow decelerates toward the rear of the cell. The reason for the deceleration becomes clear if one considers a carbon particle attached to the surface of the cell. The particle moves backward only when a piece of membrane that is behind it (toward the trailing edge) is endocytosed and carried to the front of the cell. The part of the endocytosis that takes place in front of the particle results in no rearward displacement [*see illustration on page 90*].

Now, as the particle moves backward on the cell, the surface area between it and the trailing edge decreases. Therefore, since the coated pits are distributed randomly on the surface, less endocytosis can take place behind the particle and correspondingly more endocytosis takes place in front of it. As a result its rate of backward movement falls. If the rearward velocity is one micrometer per minute at the front of the cell, it will be half a micrometer in the middle. By the time the particle has reached the trailing edge no endocytosis is taking place behind it, and its motion stops.

Most of my discussion so far has touched on circulating proteins, but


3.5

THE BMW 735i. NEVER HAS SO MUCH ENGINEERING INTELLIGENCE BEEN LAVISHED ON ONE AUTOMOBILE.

All automobile engineers dream of "a clean sheet of paper." A chance to create a car that embodies their best thinking, regardless of the demands of mass production.

For the engineers of BMW, the 735i is that dream come true.

Over the course of 7 years, they devoted 2 million hours and $1 billion in development funds to literally rethinking every aspect of the automobile.

As a result, the 735i "represents the best sedan in the world" (Automobile Magazine) and provides "our clearest glimpse at next-century car design" (Automotive Industries) wherever you focus attention.

Observe the sleek, muscular body. While its lines produce a stable 0.32 drag coefficient and a subdued airflow, its 20 percent fewer parts make it immensely strong and virtually rattle-free.

Close the hand-fitted doors and flush windows. Their unique hermetic seals combine with the car's aerodynamics, rigid construction, and doubly-insulated sus-

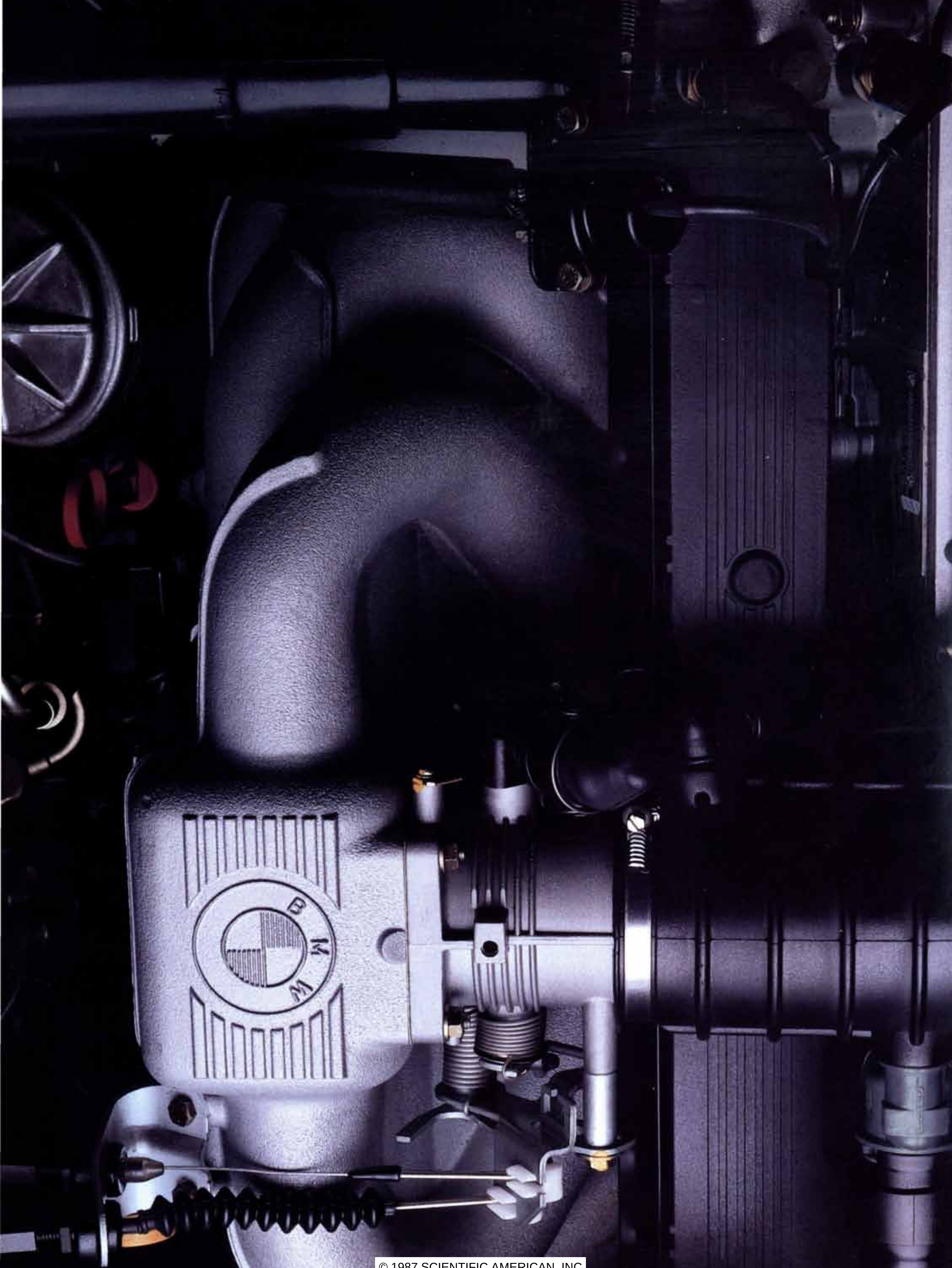
A THOUGHTFULLY PLANNED, METICULOUSLY CRAFTED RESPITE FROM THE OUTSIDE WORLD.

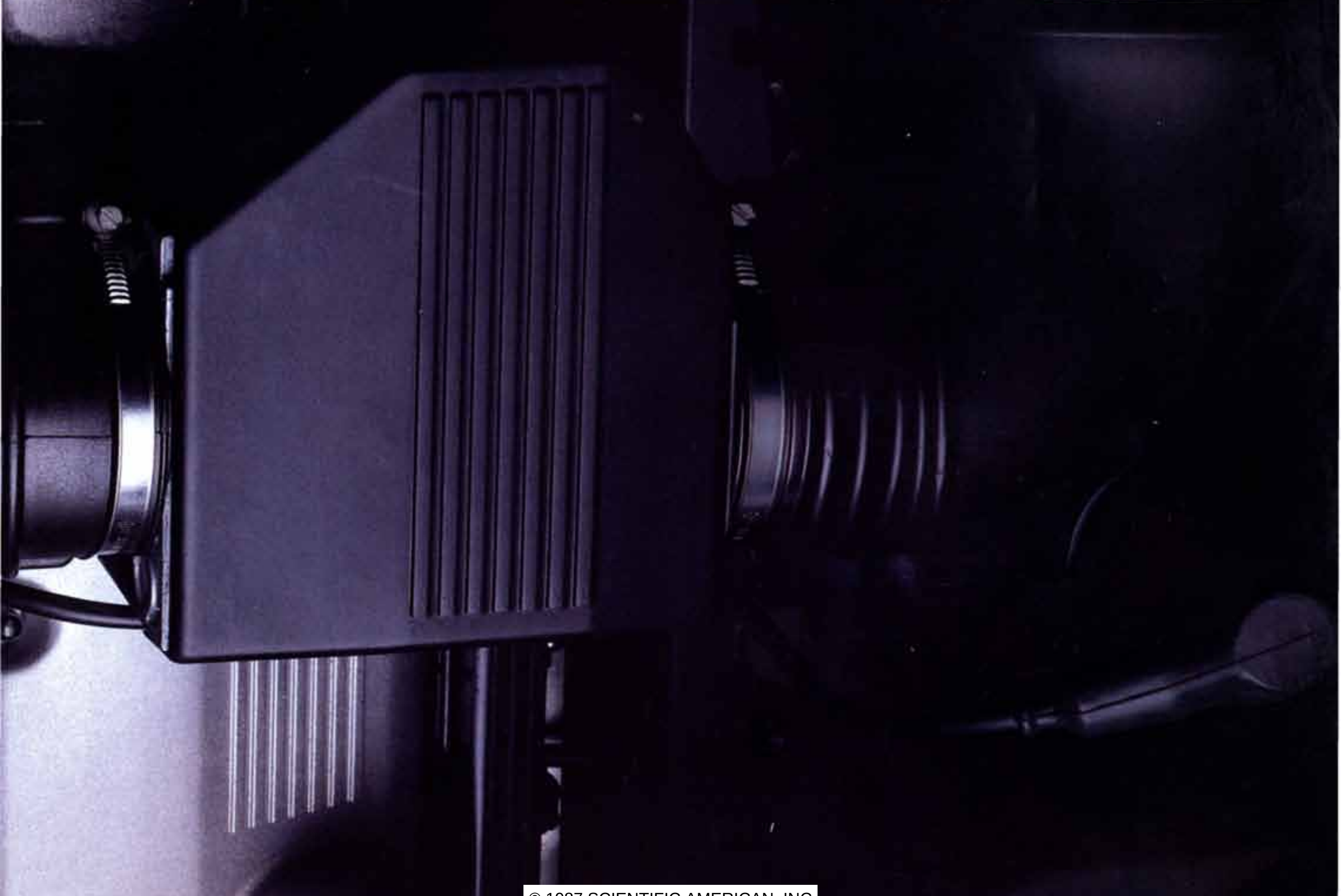
The quiet cockpit of the 735i is, as one demanding British automotive journalist has put it, "soothing, comfortable, and quite elegant." A sanctuary in which supple Nappa leather and rich wood trim provide aesthetic satisfaction commensurate with superb performance.

It is also, in the words of Automobile Magazine, an "ergonomically perfect" environment that can make a good driver masterful.

Intelligibly marked controls are within ready reach. Analog instruments are conveniently clustered in the curving console. The BMW Active Check Control advises you on the status of some 26 vital functions, and the supportive seating and faultless driving position help ensure freedom from fatigue over long distances.

Distances which may be further beguiled by the electronic cruise control and the bass-rich, 10-speaker anti-theft stereo system.


BMW




DN
RW
71
67
63
AUTO

A SYSTEM THAT ALLOWS THE HOT-BLOODED AND THE COLD-BLOODED TO PEACEFULLY COEXIST.

Among the technological amenities that make the 735i "radically different and innovative" (Automobile Magazine) is a most accommodating climate-control system.

Proven at the Arctic Circle and in California's Death Valley, this system includes air conditioning that is the world's "strongest and most automatic" (Auto Motor und Sport), plus individual thermostat controls for both the driver and front passenger.

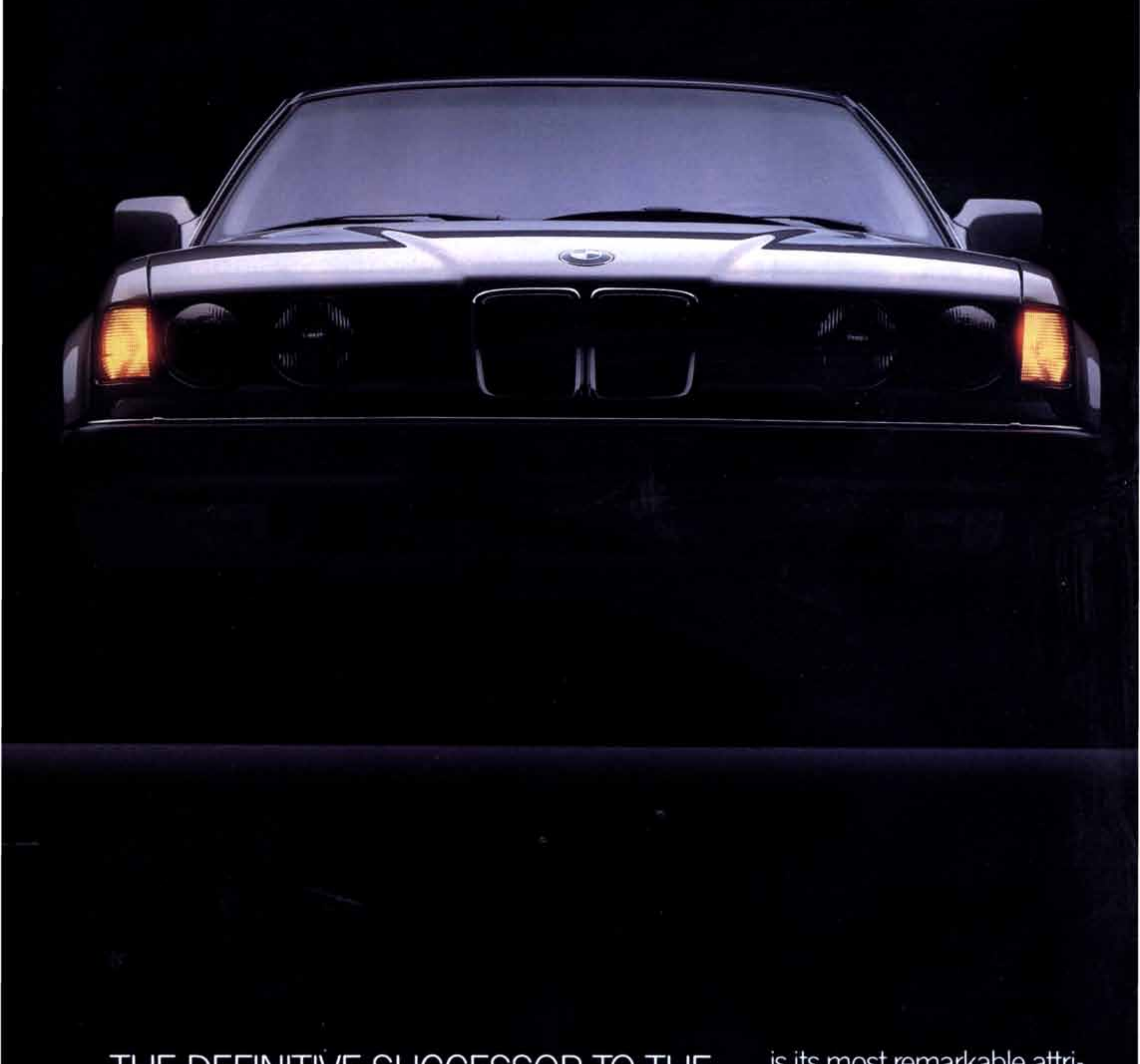
With equal thoughtfulness, the 735i provides an easy-to-program Onboard Computer that calculates distance on remaining fuel and warns you when road temperatures approach freezing.

There's also a special keylock function that lets you close the windows and two-way power sunroof from outside the car. And the right outside mirror automatically tilts down when you shift into reverse, so you can watch the curb as you park.

These and other features must acquit themselves in a series of final inspections that few car makers would have the time, or perhaps the courage, to duplicate.

Because each 735i is painstakingly scrutinized against a 4,800-item checklist. And each is tested at up to 100 mph and then given a final tuning.


BRUTE STRENGTH AND AGILITY COMBINE TO BRING NEW MEANING TO THE WORD "SECURE."
It has always been BMW's contention that true automotive safety is the ability to endure mishaps combined with performance features that help you to avoid them in the first place.
No BMW advances this definition more inventively than the 735i.
Its body shell and integrated bumpers are designed not merely to meet the U.S. crashworthiness standards, but to surpass them. Its twist-proof seatbelts tighten on impact, and its steering-wheel airbag provides full-blown protection in 67 milliseconds.
On the other hand, the 735i's responsiveness and handling are abetted by a fast, sure antilock braking system. It helps to prevent the wheels from seizing when you apply the big, thick, four-wheel disc brakes, so that you can steer the car around potential accidents.
And along with vision-enhancing ellipsoid lights and vigilant Active Check Control, there is a foresightful Service Interval Indicator. It calculates the effects of actual driving on 17 systems, to avoid unwise delays in routine maintenance.



membrane flow also has some interesting implications for noncirculating surface proteins. Proteins that do not circulate are continuously swept backward by the flow, and because they do not enter coated pits, they are not carried forward by endocytosis. Yet the noncirculating proteins are also subject to diffusion, which tends to randomize their distribution and so works against the unidirectional backward flow. Does flow or diffusion win out? The answer is that it depends on the size and shape of the protein, which affect the rate at which it diffuses. A protein that diffuses rapidly should barely be swept backward by the flow, a prediction that has been verified by various observations.

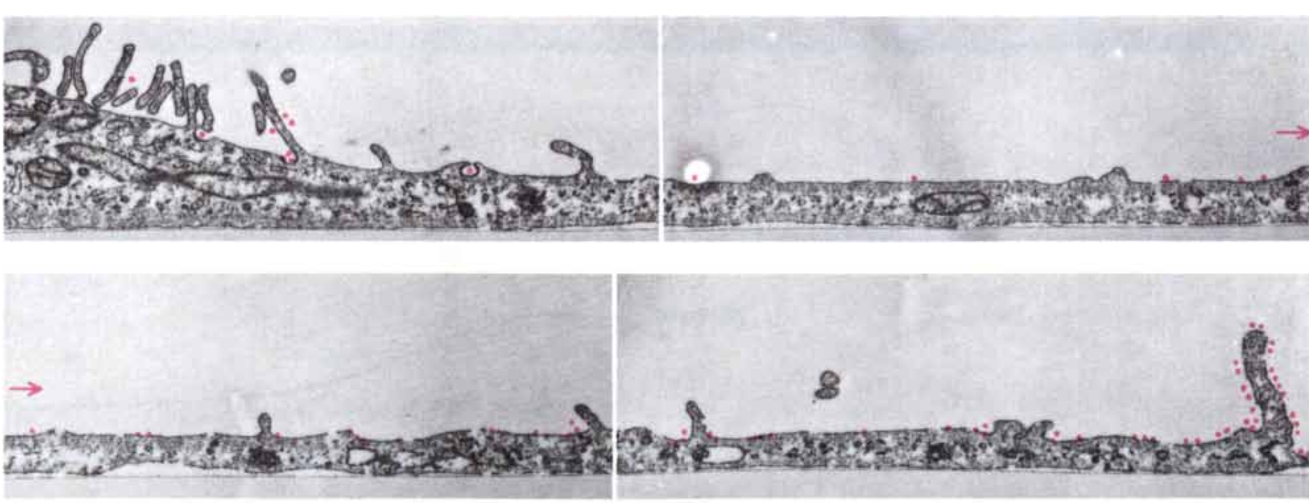
A protein that diffuses slowly, however, might be swept toward the back of the cell and accumulate there. Such proteins were unknown until recently, when Akira Ishihara, Bruce F. Holifield and Kenneth A. Jacobson of the University of North Carolina School of Medicine examined the distribution of a protein they call GP80 on the surface of moving fibroblasts. Several years earlier Pearse, Thomson and I had shown that GP80 does not enter coated pits and therefore is noncirculating. Ishihara and his co-workers went on to show that the protein diffuses quite slowly. On observing its distribution on moving fibroblasts, they found 20 times as much GP80 at the rear of the cell as at the front. It was clear that slowly diffusing proteins are swept relentlessly to the back of a motile cell by the flow of membrane.

Actually the same result had been observed much earlier in a different context. In 1971 Roger B. Taylor and W. Philip H. Duffus of Bristol University, along with Martin C. Raff and Stefanello de Petris of the Medical Research Council's National Institute for Medical Research, had made a remarkable discovery. They were studying the behavior of antibodies on the surface of mouse *B* lymphocytes (a type of white blood cell). The mouse cells were exposed to rabbit proteins that have the effect of cross-linking the antibodies into large clusters on the lymphocyte's surface. When the cells were examined by a fluorescent technique, there were bright patches: islands of these cross-linked antibodies. What is remarkable is that when the cells were warmed up from the low temperature at which the rabbit proteins were added, the bright patches migrated to the rear of the cell, where they form a "cap."

In 1971 little was understood about the endocytic cycle, and so the phenomenon of cap formation could not be linked to that process. Now it is known that capping results from the same forces that cause GP80 to accumulate at the rear of the moving cell. Whereas individual, unlinked antibody molecules diffuse fast enough to prevent their being swept to the rear of the cell, the cross-linked mass diffuses much more slowly, and it is therefore swept to the back of the cell, just as GP80 is. The similarity of the two processes is underlined by the finding that capping takes place on all motile cells—but not on nonmotile ones.

Although much progress has been made in understanding capping, the associated membrane flow and the general process of cell locomotion, there remains the question of how the overall process of locomotion is switched on and off. It is generally thought that all animal cells have an intrinsic capacity for movement, but the capacity is not exploited at all times. Many cells, including fibroblasts, are generally stationary in laboratory culture. When they are put in culture, the cells spread out and may remain motionless. They display no ruffling edges; cross-linked proteins do not cap on their surface, indicating there is no membrane flow. Nevertheless, it can be shown that endocytosis is just as active in nonmoving cells as in moving ones. One must conclude that in nonmotile cells the membrane drawn into the coated pits is not returned to the surface at a single location—otherwise the cell would move forward from that point, and there would be a rearward membrane flow.

Indeed, that difference constitutes what is perhaps the most important single message about how cells move. In motile cells exocytosis is a directed process: membrane internalized by endocytosis is exocytosed only at the leading edge of the cell. In stationary cells, on the other hand, exocytosis is not directed: en-



GIANT HELA CELL, part of which is seen here in two cross sections, provided information about the distribution of ferritin receptors and other circulating proteins on the surface of motile cells. Only the giant cell's leading lamella, about a fifth of the length of the cell, is shown; the cell's leading edge, often called the ruffling edge because of its appearance, is at the bottom right. The ferritin receptors are labeled by bound ferritin molecules. The receptors (each indicated here by a colored dot) are seen to be concentrated at the ruffling edge. The experiment helped to confirm that in motile cells the recycling of membrane is a directional process: although coated pits form at random on the cell surface, exocytosis takes place only at the leading edge.

docytosed membrane is reinserted randomly over the surface of the cell.

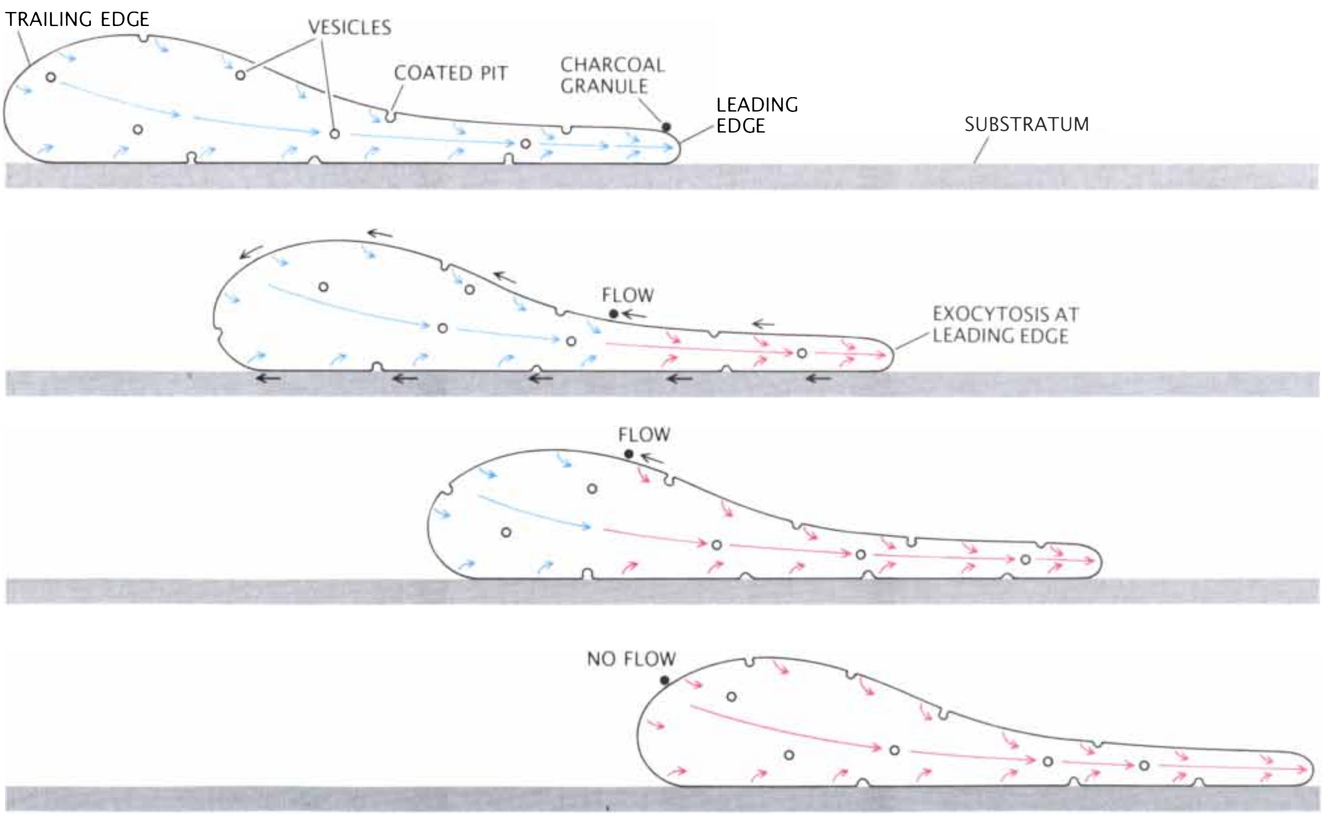
Surprisingly, this crucial difference is readily reversible. Under the right conditions of cell culture it takes only a few minutes to make a stationary fibroblast move. Somehow the random exocytosis present in the nonmotile cell has been directed toward the leading edge, as is required for movement. Synthesis of new proteins is not necessary for the change to occur: drugs that block protein synthesis do not interfere with it. It appears that a stationary fibroblast possesses all the biological machinery needed for locomotion but that the machinery is not in the right configuration. What the correct configuration is and how it can be induced are questions for the future.

On a level more amenable to experiment is a problem with which I am currently much preoccupied: how the moving cell interacts with its substratum. It is known that in their natural environment cells crawl over a complex network of molecules known as the extracellular matrix. The matrix consists largely of the protein collagen and various long-chain polysaccharides. Included in the network are some specialized proteins—generally very large—that the cell recognizes as it moves along. Among them is one called fibronectin, and it is to fibronectin that the moving fibroblast attaches [see "Fibronectins," by Richard O. Hynes; SCIENTIFIC AMERICAN, June, 1986].

The cell's attachment is mediated by a receptor for fibronectin in the plasma membrane. The receptor recognizes fibronectin and binds to it, forming a "foot" that provides the thrust for forward motion. Not much is known about how the receptor-foot works, but I have developed a working hypothesis based on the assumption that the receptors are circulating proteins.

As the fibroblast moves along, the fibronectin receptors are continually binding and releasing fibronectin. At the front of the cell the receptors bind fibronectin. While they are so bound the receptors cannot diffuse in the membrane, because they are indirectly attached to the rigid extracellular matrix. Hence the membrane flowing backward pushes against them, thereby generating the locomotor thrust. As membrane is added at the front of the cell, the feet are moved back. Whenever a receptor is released from fibronectin, it can diffuse into a coated pit, whence it is carried to the leading edge of the cell to bind fibronectin again.

For the moment this is merely a hypothesis, and how the feet work is one unanswered question among many that might still be asked about the movement of animal cells. What is the source of energy for the cell's movement? What distinguishes circulating proteins from noncirculating ones? The most challenging questions have to do with the control of exocytosis. What distinguishes the leading edge of a motile cell so that exocytosis takes place there and not elsewhere? Even more important for understanding metastasis and for the treatment of cancer, what triggers a stationary cell to become motile?



MEMBRANE-FLOW MODEL for animal-cell locomotion is based on directional endocytosis. If a moving cell encounters a charcoal particle, the particle may stick to the cell's leading edge (*1*). Exocytosis of membrane at the leading edge displaces membrane backward, yielding a flow that carries the charcoal particle along with it (*2*); the particle thus serves as a marker of membrane flow. By the time the particle is near the middle of the cell, its motion has slowed considerably (*3*). The reason is that only coated pits originating behind the particle can displace it backward on reaching the leading edge, and as the particle moves farther back, progressively fewer pits originate behind it. By the time it is at the trailing edge the particle is no longer moving (*4*).


BUSINESS GRAPHICS
SALES

Cosmic Strings

Why are stars and galaxies clumped rather than spread out evenly in space? What drew them together? Thin strings of energy created during the birth of the universe may have provided the attraction

by Alexander Vilenkin

The universe is rather lumpy: stars gather into galaxies, and galaxies in turn form clusters. With time the universe gets lumpier as the gravitational pull of galaxy clusters attracts other galaxies from neighboring regions. Modern theories of galaxy formation assume that in the past the universe was much smoother than it is today and that all galaxies and clusters have grown out of small incongruities in an otherwise nearly uniform distribution of matter. The implications of these theories have been studied in great detail, but one basic question still looms above the others: What are the initial incongruities, and where did they come from?

Enter cosmic strings, exotic, invisible entities spun by theories of particle physics. Strings are threads remaining from the fabric of the newborn universe. They are incredibly dense and incredibly energetic; they travel at almost the speed of light and bend the space around them. Created in the second following the big bang, strings in tangled, infinite stretches give rise to loops that wiggle violently and gradually exhaust their energy through their own gyrations.

No one knows for certain whether cosmic strings exist. If they do, many physicists believe they might account for the lumpy distribution of matter in the universe. Particularly massive loops of string could provide the gravitational attraction necessary to nucleate galaxies and clusters of galaxies. But such loops also have short lifetimes; even if they once pervaded the universe, most of them will have vanished by now.

Less massive strings could still exist, but they are hard to detect. With diligence and ever more sensitive equipment astronomers might nonetheless be able to reject or confirm the existence of cosmic strings within a few years. Their search is full of suspense because the discovery of a string would open windows on the elementary nature of matter as well as the birth of the universe. In order to understand how, the notion of cosmic strings must be considered in the context of both particle physics and cosmology.

Broken Symmetry

It is now fairly well established that the universe began some 15 billion years ago in a grand explosion commonly known as the big bang. The universe is still expanding from the force of that explosion; distant galaxies are moving away from the earth at very high speeds. By combining astronomical observations with the verified laws of particle physics, physicists can trace the history of the universe back to the fraction of a second immediately following the big bang. In those moments there were no stars or galaxies or atoms; the universe was merely a hot, dense fireball of elementary particles such as electrons and photons.

Underlying the particles and determining their interactions is the vacuum. Far from connoting "nothing," a vacuum to the physicist is a state of minimum energy obtained in the absence of all particles. The relation between elementary particles and the vacuum is similar to the relation between sound waves and the material in which they propagate: the types of waves and the speed of propagation are different in different materials. Because the properties of the vacuum have not always been the same, the properties and interactions of elementary particles have also changed.

The early vacuum had an enormously high energy as well as a high degree of symmetry; that is to say, there were no distinctions between the interactions of fundamental particles. The electromagnetic, weak and strong nuclear forces were manifested as parts of a single, unified force. Today the vacuum energy is zero and the elementary forces are distinct both in strength and in character; there is little left of the original unity. How was the early symmetry broken?

While the universe expanded and cooled after the big bang the vacuum went through a rapid succession of changes known as phase transitions. The most familiar phase transitions are those water undergoes when it is cooled from steam to liquid and finally to ice. Phase transitions can also be described in terms of symmetry breaking: they often reduce symmetrical states to asymmetrical states. A crystal, for example, is less symmetrical than a liquid: a liquid "looks the same" in all directions, whereas different directions in a crystalline lattice are not all equivalent.

No one knows exactly how many phase transitions occurred in the young vacuum, but all of them probably took place within the first second after the big bang. And like phase transitions in more familiar materials, cosmological phase transitions can give rise to defects. Inside the defects symmetry is not broken and the earlier, more symmetrical vacuum is trapped. Different types of defects are predicted by different particle theories. Some theories predict that the defects will assume the form of surfaces; others predict lines or points. These defect types are called domain walls, strings and monopoles respectively.

Lines and Loops

Cosmic strings are therefore only one of three possible types of flaws in the continuity of the vacuum. Why have they been singled out in theories of galaxy formation? Ironically, one reason is that they are the least conspicuous of the defects. The high-energy vacuum trapped inside defects is, by the Einstein mass-energy relation, extremely massive. Consequently defects can have a profound influence on the evolution of the universe. A single domain wall stretching across the present-day universe would have far more mass than all the matter in the universe combined; it would induce galaxies to cluster much more than they actually do. On the other hand, although a single monopole might escape detection, the theories that predict monopoles predict them in very large numbers. If monopoles existed, the universe would be swarming with them; they would be difficult to ignore. Yet neither domain walls nor monopoles have been observed.

No one has ever seen a cosmic string either, but physicists do not expect strings to be as obvious. Work on cosmic strings was pioneered 11 years ago by T. W. B. Kibble of the Imperial College of Science and Technology in London. Kibble studied how strings might be formed in the early universe and in a 1976 paper discussed some aspects of their evolution. The notion of strings, however, did not attract much attention

DEFECTS IN ICE arise during freezing, when different regions of water crystallize with different orientations. Some theories of particle physics suggest that an analogous process may have generated linear defects in the vacuum of space as the universe cooled after the big bang. So far there is no empirical evidence that the defects, which are called cosmic strings, actually exist.

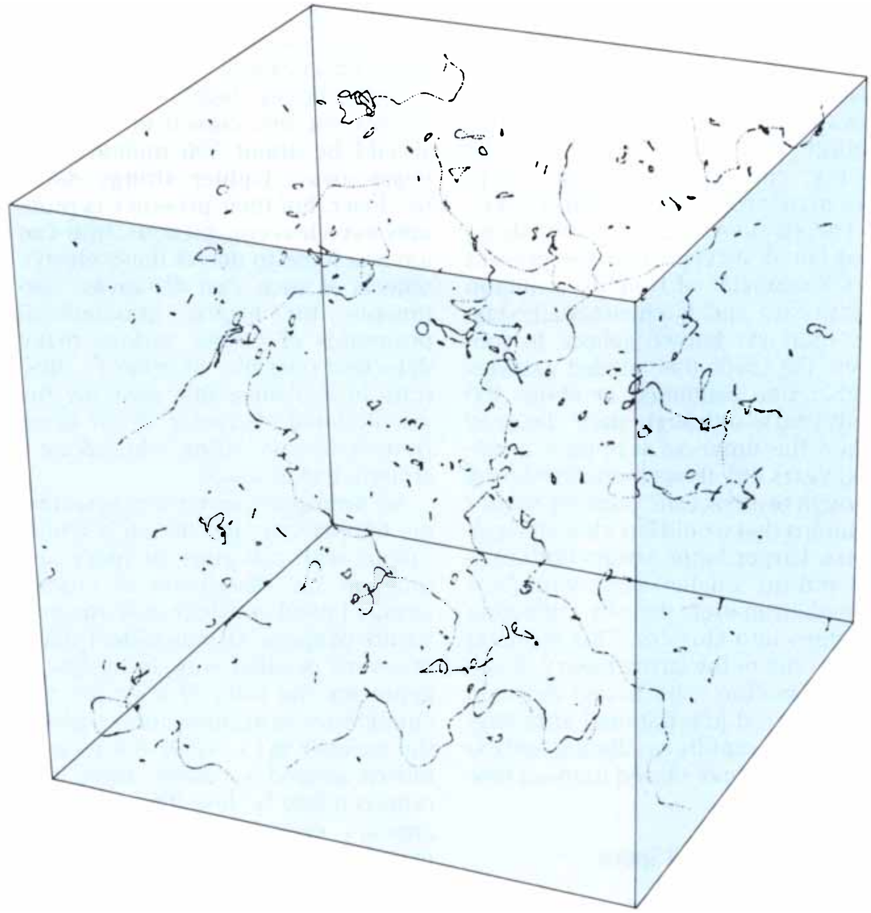
VIRGO CLUSTER consists of thousands of galaxies, some of which are pictured here (*white blurs*). The author proposes that clusters and galaxies alike were seeded by massive loops of cosmic string that swept matter from space into localized areas.

until four or five years later, when Yakov B. Zel'dovich of the Institute of Physical Problems in Moscow and I independently realized that strings might be able to explain the "clumping" of matter in the universe. Our ideas inspired a small group of investigators to explore string theories in greater detail.

The physical properties of strings turned out to be fascinating and unique. Cosmic string theory quickly developed an attraction for physicists akin to the attraction the strings themselves are said to exert on stars and galaxies. Last year an avalanche of papers on cosmic strings descended on the technical literature, and yet no direct experimental evidence for their existence has been found. Even in the absence of empirical data, physicists have managed to put together an impressive profile of the strings' properties. Some properties depend on the particle theory used to derive them, whereas others are common to all theories.

Cosmic strings are thin tubes of symmetrical, high-energy vacuum. They do not have ends; they either form closed loops or extend to infinity. A string's physical character is determined by the energy of the vacuum trapped inside it. The strings with the most symmetrical vacuum, in which strong, weak and electromagnetic forces are united, are the thinnest and most massive. These are also the most conspicuous of strings, and they are of the greatest cosmological interest because they could be responsible for the formation of galaxies.

Such strings have a thickness on the order of 10^{-30} centimeters and an astonishingly large mass: one inch of string would weigh 10 million billion tons! The tension in these strings is a match for their mass. The tension causes closed loops of string to oscillate fiercely at speeds approaching the speed of light. For example, a loop one light-year long would complete an oscillation in a little more than a year. (A light-year, a unit of length, is the distance light travels in one year: about five trillion miles.)



STRING NETWORK has been modeled on a computer so that the formation and evolution of strings can be simulated. The network includes both closed loops and long pieces of convoluted string that stretch across the universe and give rise to more loops.

Making Waves

The strings produced during a phase transition weave a tangled network that pervades the entire universe. The evolution of this cosmic web is rich in physical processes. Although on the average the network is uniform, its individual strings are quite irregular and convoluted. Wiggling violently from tension, curved strings often cross themselves and one another; they break at the point of intersection and join again in different configurations. A closed loop, for example, splits in two when it twists on itself. Long, coiled strings cross themselves many times over and closed loops get lopped off at the intersections.

It takes longer to make large loops than it does to make small ones because a string must fold back on itself to shed a loop, and it must fold back farther for a large loop. The size of the loop that can be liberated at any moment is limited by the amount of time that has elapsed since the big bang. In particular, given that the strings move at about the speed of light, the loop can be no larger than the distance light has traveled since the birth of the universe, which is known as the horizon length. Hence smaller loops of cosmic string are characteristic of a younger universe, whereas the loops created today are much bigger.

This does not mean that the current network of cosmic strings looks very different from the one that was initially established. Indeed, the evolution of strings includes an interesting feature called self-similarity that preserves the statistical constancy of the network over time. If the string network were to be photographed at two different times, the main difference between the two shots would be in overall scale, which is set by the horizon length. Magnifying the first photograph by the ratio of the two horizon lengths would yield a picture quite similar to the second.

The network is continuously producing copious amounts of closed loops; if they were allowed to accumulate, the universe would be swimming with them. What, then, happens to the loops? Theoretical analysis shows that, as they oscillate, loops generate rhythmic pulses of gravitational energy that propagate at the speed of light. These pulses are known as gravitational waves, and they sap the energy of a loop until it shrinks and eventually disappears. The lifetime of a typical string, regardless of its size, is about 10,000 oscillations. Because the period of a single oscillation is greater for larger loops, they live longer than smaller ones. Likewise loops of lighter, low-energy string last longer than the heavy, energetic loops that swept raw matter into galaxies and clusters.

The smallest loop of heavy string that could survive until the present has a diameter of roughly a million light-years and is substantially bigger than any known galaxy. In contrast, the loops that seeded galaxies had a size estimated at about 100 light-years; although they decayed when the universe was only a million years old, those loops lived long enough to attract surrounding matter in lumps that would later become galaxies. Larger loops attract both matter and the smaller loops with their attendant matter, thereby collecting galaxies into clusters. This scenario is the crux of the string theory of galaxy formation. The model rests on the localized gravitational attraction exerted by rapidly oscillating and extremely massive closed loops of cosmic string.

Conical Space

Development of the cosmic string scenario began in earnest only recently. Early in 1985 Andreas Albrecht and Neil Turok of the Fermi National Accelerator Laboratory devised a computer program to simulate the evolution of cosmic strings. Later that year Turok used the simulation to calculate the number and distribution of galaxy clusters predicted by the string theory. His results agree with what is actually observed: clusters are not located randomly on the sky but instead tend to come in clusters themselves. This clumping is described mathematically by a correlation function. There is a striking similarity between the correlation functions derived from astronomical observations and those derived from Turok's computer simulations.

In spite of its initial success the string theory is a long way from being a completely satisfactory explanation of galaxy formation. Recent observations of the large-scale distribution of galaxies revealed filamentary and sheetlike patterns as well as enormous voids having almost no galaxies at all. Cosmologists are now trying to find out whether or not the string model can account for these features. Even if it can, physicists will never *really* believe in cosmic strings until their existence is confirmed by direct observation.

Of the heavy, highly symmetrical strings, the one closest to the earth should be about 300 million light-years away. Lighter strings might be closer, but their presence is probably even less conspicuous. How can anyone hope to detect these elusive objects at such vast distances? Fortunately the bizarre gravitational properties of cosmic strings make detection possible. In order to illustrate how, I must first describe the gravitational character of an idealized, stationary string lying along a straight line in space.

According to Einstein's general theory of relativity, gravitation is synonymous with curvature of space and time; in my discussion of cosmic strings I need consider only the curvature of space. Strings distort space in a very peculiar way. In Euclidean geometry the ratio of a circle's circumference to its diameter is equal to the number pi (3.14159). For a circle drawn around a cosmic string this ratio is a tiny bit less (the difference appears only in the fourth decimal place). The space around a string has a conical nature. To visualize it, imagine cutting a small angular wedge from the Euclidean space with its tip at the string; then glue the exposed surfaces together, not by stretching but by bending the space around them. The result is that all planes perpendicular to the string become cones.

The angle of the wedge thus removed is called the deficit angle and in strings it corresponds to a few seconds of arc. All objects passing by the string—photons, atoms, stars—will be deflected from their original direction of motion by an angle comparable to the deficit angle. Two objects moving along symmetrical, parallel paths on opposite sides of the string will collide after they pass the string. To a person sitting on one of these objects the other object would initially appear to be at rest; when the string passes in front of it, it would suddenly start moving toward the hapless observer at a speed equal to .00002 times the velocity of the string. Since strings move at nearly the speed of light, they induce velocities of about four miles per second.

What would happen if a string passed through a person? The effect is not difficult to picture. As the string cuts across the individual's waist, his head and feet start moving toward each other at a speed of four miles per second. This experience would, of course, be unhealthy, but there is no reason to panic: the probability of a string's traveling through the solar system is very small indeed.

The conical-space rendering of cosmic strings' gravitational properties applies only to straight strings. The gravitational effects of curved strings and closed loops are much more complicated. Small segments of such strings can be thought of as being approximately straight, however, and by combining the analyses of many small segments it can be shown that at great distances from an oscillating closed loop the average effect of all segments is an ordinary gravitational attraction like that associated with the earth or the sun.

A Light Divided

The distortions that cosmic strings induce in space might betray their presence. For example, because they bend space, strings can act as gravitational lenses, meaning they will deflect the light from a distant galaxy so that it reaches the earth by two different paths. As a result observers on the earth see two images of the same galaxy, separated by an angle comparable to the deficit angle of the string. Astronomers have in fact found several pairs of galaxies and of extremely bright, distant objects called quasars in which the members show a compelling resemblance to one other and which are therefore considered to be double images of the same object.

Ordinary galaxies or galaxy clusters can also act as gravitational lenses and so an additional test must be applied to multiple images to ascertain their cause. Nick Kaiser of the University of Cambridge and Albert Stebbins of Fermilab have pointed out that cosmic strings should have a rather unusual effect on the cosmic microwave radiation. This radiation is a kind of afterglow from the big bang; it fills the entire universe and comes from all directions with equal intensity. In the wake of a string, however, some radiation would gain extra momentum in the direction of the earth and would therefore approach the planet with a greater intensity. And whereas other gravitational entities can cause smooth changes in microwave intensity, the change wrought by a cosmic string would be quite abrupt. This sudden change of intensity should occur


A BLACK HOLE THAT HELPS MILITARY HELICOPTERS ESCAPE HOT PURSUIT.
Because of hot engine exhaust, military helicopters can be easy prey for heat-seeking missiles. The best way to escape this threat and still accomplish the mission: Cool the exhaust. To do this, our engineers developed the Black Hole Infrared Suppression System, currently in use on the U.S. Army Apache.
Engine exhaust is mixed with outside air, routed through a series of pipes and expelled through finned nozzles. The exhaust is so cool it is ignored by heat-seeking missiles.
Because it stays cool, the Black Hole exhaust system can be made of inexpensive metals. And because the system has no moving parts, it requires little maintenance.
We're giving America its money's worth with breakthroughs in aviation, space and information systems.
We're McDonnell Douglas.
For more information, write: Apache, McDonnell Douglas, Box 14526, St. Louis, MO 63178.
MCDONNELL DOUGLAS
HELICOPTERS SPACECRAFT & MISSILES MILITARY & COMMERCIAL AIRCRAFT INFORMATION SYSTEMS TRAINING SYSTEMS HEALTH CARE FINANCING

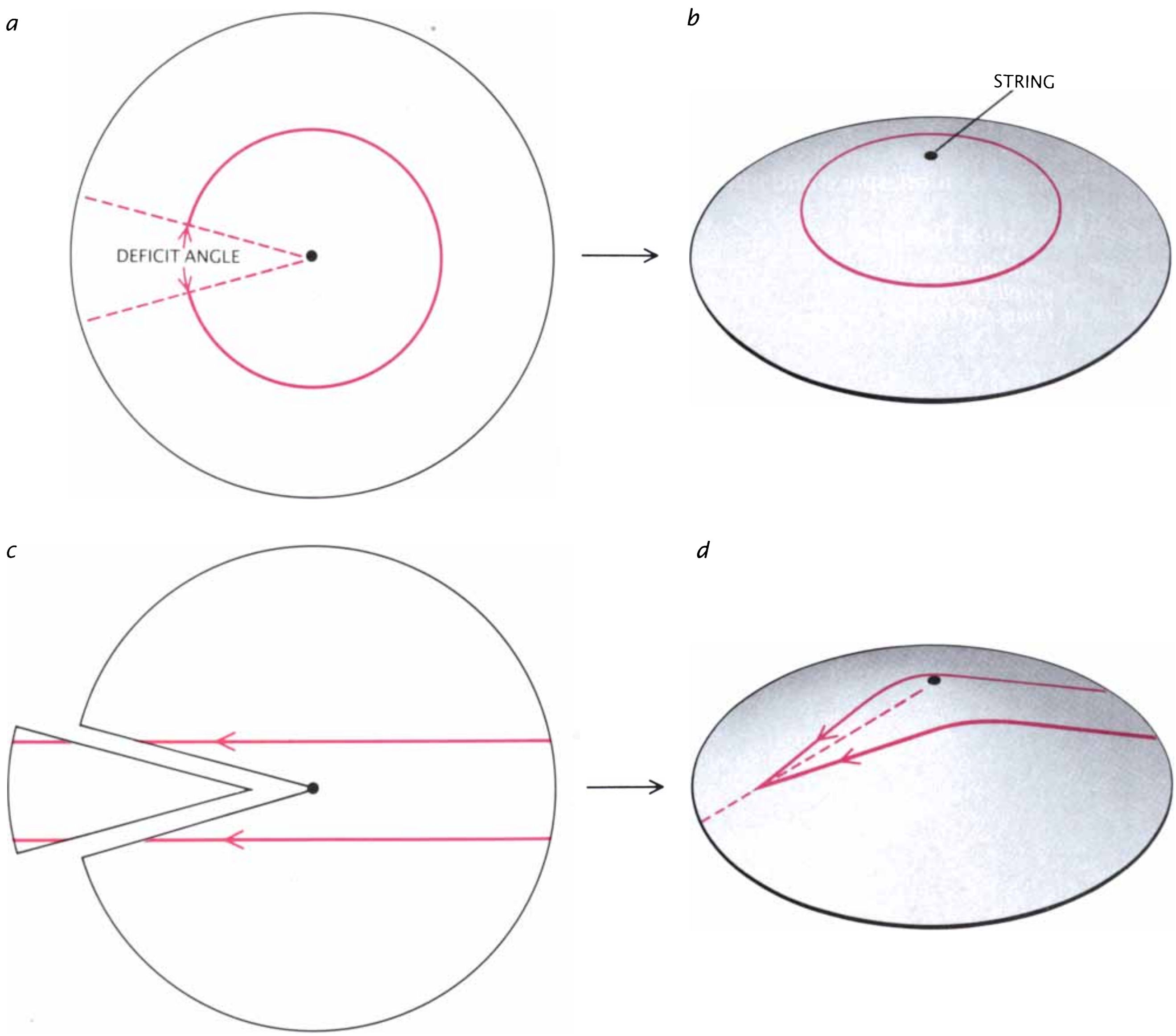
along a line drawn between the two images representing a single galaxy. The magnitude of the change may be only one part in 100,000, but detection, although difficult, would not be impossible.

In addition, it might someday be possible to detect cosmic strings by looking for evidence of their gravitational waves. Waves from loops of all shapes and sizes add up to a background of gravitational noise, the collective swan song of dead and dying strings. The intensity of this noise is high compared with that of the gravitational waves emanating from other sources. Gravity, however, happens to be the weakest of all forces in nature, and the predicted level of noise from the strings is still slightly below current observational limits.

Cosmic Superconductors

Until recently physicists assumed that cosmic strings could interact with matter only through gravitational forces. Two years ago, however, Edward Witten of Princeton University proposed that cosmic strings might be endowed with a property highly regarded here on earth: superconductivity. Witten showed that some theories of elementary particles suggest strings have unusual electromagnetic qualities of the kind that would make them behave like superconducting wires. Later it was shown that this property could produce dramatic cosmological effects.

Strings could be superconducting because the symmetrical vacuum trapped inside them changes the behavior of particles. In particular, some charged particles, like electrons, might have no mass inside a cosmic string. Therefore it is possible with very little energy to create particle-antiparticle pairs in which the members have opposite charge and travel in opposite directions. The total charge and total momentum of a pair are equal to zero; the only energy input required is that necessary to impart motion. Because the particles are massless, they move at the speed of light and cannot venture outside



CONES OF SPACE around heavy cosmic strings illustrate their peculiar gravitational effects. This drawing shows the distortion of space caused by an idealized straight string. The distortion can be represented by cutting a wedge out of a plane of space perpendicular to the string (*a*) and pulling the two edges of the plane together to form a cone (*b*). Hence if two objects traveling along parallel paths pass a string on opposite sides (*c*), they will be deflected and will collide on the far side of the string (*d*).

the string, where their mass would be greater than zero. Hence particles zoom through the trapped vacuum carrying an electric current and encountering no resistance: the hallmarks of superconductivity.

Jeremiah P. Ostriker of Princeton and his student Christopher Thompson collaborated with Witten to put forward an alternative string model of galaxy formation. The current in a superconducting string produces electric and magnetic fields that in empty space would propagate away from the string as electromagnetic waves. But the interstellar and intergalactic space is not exactly empty: it is filled with a dilute gas of electrons and charged atoms that prevent waves from leaving the vicinity of the string. As the energy of the radiation accumulates, it develops tremendous pressure and starts blowing a bubble, sweeping surrounding matter into a hot, expanding shell of gas. The expanding bubble is not much different from a huge explosion. In this scenario, galaxies are formed where bubbles collide.

In a sense the explosive theory is antithetical to the gravitational model of clustering: matter is blown away from the string instead of being attracted to it. The theory also gives a more natural explanation for the gaping voids in galaxy distribution and predicts that galaxies should occur in thin, sheetlike arrangements, a picture not unlike the one astronomers observe. What other empirical trials could be used to test the explosive scenario?

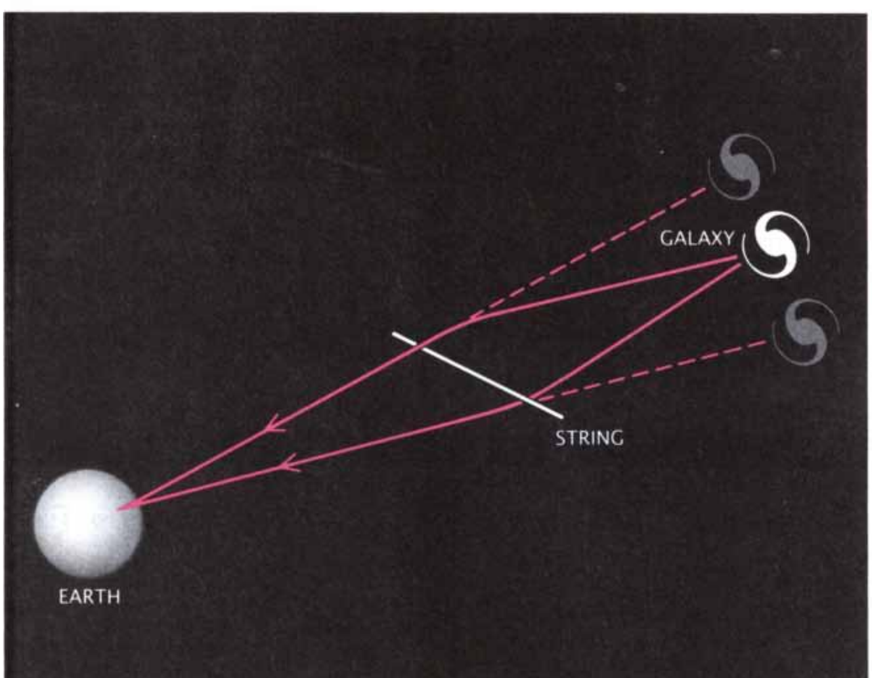
Ostriker and his colleagues have to posit that the universe was magnetized soon after the big bang, because a magnetic field had to be available to launch the particle-antiparticle pairs. No one knows for certain how this magnetization might originate, but if it was present, residual magnetization should persist in the intergalactic space today. The theory can therefore be tested by looking for evidence of weak, delocalized magnetization in the universe. The hot gas on the boundaries of the bubbles might also emit characteristic radio waves that could be detected here on the earth with radio telescopes.

Threads of Evidence?

Aside from its role in theories of galaxy formation, the concept of superconducting cosmic strings suggests another way the strings might be found. Even if magnetic fields were absent from the early universe, they are eventually generated by the rotational energy of the galaxies through a phenomenon known as the galactic dynamo effect. Today the strength of the magnetic field in a typical galaxy is just one-millionth of the earth's magnetic field. The current such a field would induce in a superconducting loop of string is too weak to prompt an exploding bubble of radiation; however, calculations I did with Eugene M. Chudnovsky of Tufts University, George B. Field of the Harvard-Smithsonian Center for Astrophysics and David N. Spergel of the Institute for Advanced Study demonstrate that the interaction of this current with the charged particles in interstellar space can generate radio waves.

In December, 1985, Mark Morris of the University of California at Los Angeles and Farhad Yusef-Zadeh of the National Aeronautics and Space Administration's Goddard Space Flight Center discovered at the center of the Milky Way several stringlike radio sources they call threads. These threads could be light, low-energy cosmic strings, and if they are, it should be possible to see them move. On the sky the speed with which such strings move would translate into a few seconds of arc per year. Preliminary measurements have already established an upper limit on the motion of 1.5 arc seconds per year. Although this figure is less than physicists expect, it does not rigorously eliminate galactic threads from consideration because much of their motion could be in the direction of the line of sight rather than perpendicular to it. Motion along the line of sight could not be detected.

Even as empirical trials of cosmic string theory begin, physicists are tempted to use the rich and unusual properties of the hypothetical strings to account for all kinds of mysterious phenomena. Strings have already been suggested as possible sources of cosmic rays, ubiquitous but unexplained streams of energetic particles in space. They might also be the origin of powerful bursts of gamma rays that are regularly observed but poorly understood. Strings are even suspected of being the power engines behind quasars. The rationales given for these attributions are not particularly compelling; most will probably turn out to be wrong. Nevertheless, theorists are having great fun exploring the potential of cosmic strings—and nature, after all, will give the final verdict on their work.



DOUBLE IMAGES of a single galaxy result from the gravitational distortion of an intervening string. The string bends light from a single source to give the impression of two sources to an observer on the earth. If two galaxies look unusually similar, physicists can check for evidence of a string or of another gravitational lens between the pair. Strings can be distinguished from other such lenses because they cause an abrupt change in background microwave radiation; so far, however, none have been found.

Collective Computation in Neuronlike Circuits

Electronic circuits based on neurobiological models are able to solve complex problems rapidly. Their computational properties emerge from the collective interaction of many parts linked together in a network

by David W. Tank and John J. Hopfield

Modern digital computers are latecomers to the world of computation. Biological computers—the brain and nervous system of animals and human beings—have existed for millions of years, and they are marvelously effective in processing sensory information and controlling the interactions of animals with their environment. Tasks such as reaching for a sandwich, recognizing a face or remembering things associated with the taste of madeleines are computations just as much as multiplication and running video games are.

The fact that biological computation is so effective suggests that it may be possible to attain similar capabilities in artificial devices based on the design principles of neural systems. We have studied a number of "neural network" electronic circuits that can carry out significant computations. Such simple models have only a metaphorical resemblance to nature's computers, but they offer an elegant, different way of thinking about machine computation, which is inspiring new microelectronic chip and computer designs. They may also provide fresh insights into the biological systems.

Current research on this subject builds on a long history of efforts to capture the principles of biological computation in mathematical models. The effort began with the pioneering investigations of neurons as logical devices by Warren S. McCulloch and Walter H. Pitts in 1943. In the 1960's Frank Rosenblatt of Cornell University and Bernard Widrow, who is now at Stanford University, created "adaptive neurons" and simple networks that learn. Widrow's Adaline (short for adaptive linear element) is a single-neuron system that can learn to recognize a pattern such as a letter regardless of its orientation or size. Through the 1960's and 1970's a small number of investigators such as Shunichi Amari, Leon N. Cooper, Kunihiko Fukushima and Stephen Grossberg attempted to model the behavior of real neurons in computational networks more closely and to develop mathematics and architectures for extracting features from patterns, for classifying patterns and for "associative memory," in which pieces of the stored information itself serve to retrieve an entire memory.

The 1980's have seen an extraordinary growth of interest in neural models and their computational properties. Many factors converged to bring this about: neurobiologists were gaining more understanding of how information is processed in nature, cheap computer power made it possible to analyze the models in detail and there was growing interest in parallel computation and analog VLSI (very-large-scale integration), which lend themselves to implementations of neuronlike circuits. New concepts in the mathematics of neural models accompanied these developments. Our work has focused on the principles that give rise to computational behavior in a particular type of neuronlike circuit.

Neurons, or nerve cells, are complex, but even a highly simplified model of a neuron, when it is connected with others in an appropriate network, can do significant computations. A biological neuron receives information from other neurons through synaptic connections and passes on signals to as many as a thousand other neurons. The synapse, or connection between neurons, mediates the "strength" with which a signal crosses from one neuron to another. One can readily build artificial "neural" circuits from simple electronic components: operational amplifiers replace the neurons, and wires, resistors and capacitors replace the synaptic connections. The output voltage of the amplifier represents the activity of the model neuron, and currents through the wires and resistors represent the flow of information in the network.

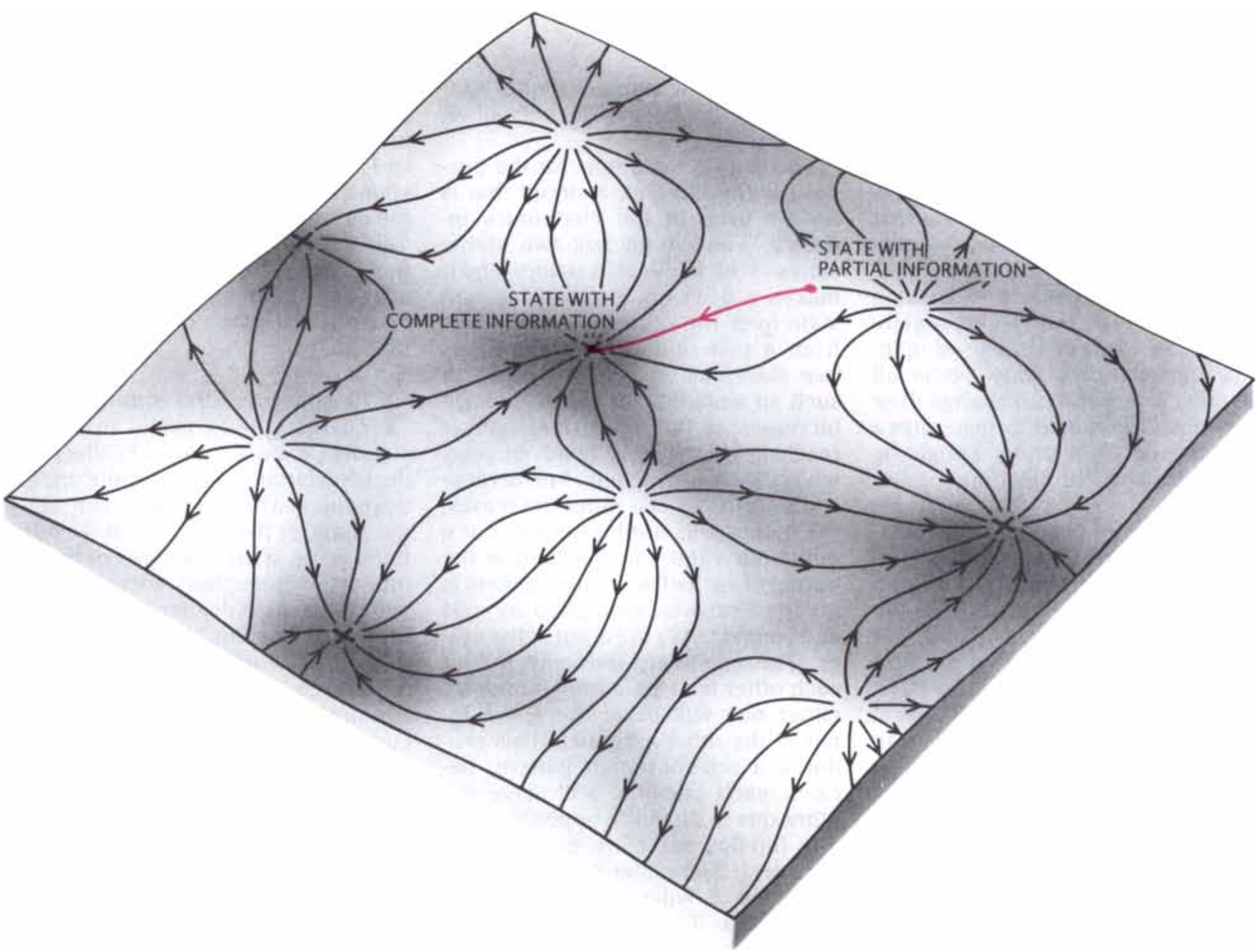
Strikingly, both the simplified biological model and the artificial network share a common mathematical formulation as a dynamical system—a system of several interacting parts whose state evolves continuously with time. The manner in which a dynamical system evolves depends on the form of the interactions. In any neural network the interactions result from the effects one "neuron" has on another by virtue of the connection between them. Thus it is not surprising that the behavior of the neural circuits depends critically on the details of the connections. The particular circuits we have studied have connection patterns appropriate for computing solutions to optimization problems, a class of mathematical problems that involve finding a "best solution" from among a very large number of choices.

The computational behavior exhibited by such circuits is a collective property that results from having many computing elements act on one another in a richly interconnected system. The collective properties can be studied using simplified model neurons, in much the same way as it is possible to understand other large physical systems by greatly reducing the details of their basic components. For example, to study the origin of collective laws of fluid motion, one can simplify the description of complex molecular collisions and produce a tractable model that captures collective features such as temperature and viscosity. Similarly, in seeking to develop a tractable model of the computations carried out by a large number of model neurons, we de-emphasized the details of the processing that goes on at the level of the individual cells and synapses. By simplifying in this way, we were able to discover the general principles by which one can understand collective computation in these circuits.

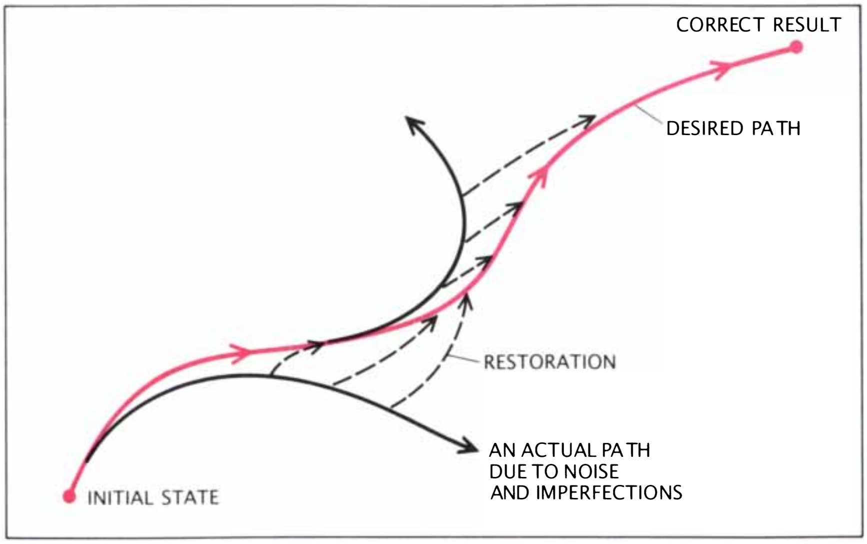
To comprehend how collective circuits work, it helps to take a very broad view of the essence of computation. Any computing entity, whether it is a digital or analog device or a collection of nerve cells, begins with an initial state and moves through a series of changes to arrive at a state that corresponds to an "answer." The process can be pictured as a path, from initial state to answer, through the physical "configuration space" of the computer as it evolves with time. In a digital computer, for example, the configuration space is defined by the set of voltages for its devices. The input data and program provide initial values for these voltage settings, which change as the computation proceeds and eventually reach a final configuration, which is reported to an output device, such as a screen or a printer.

For any computer there are two critical questions: How does it determine the overall path? And how does it restore itself to that path when physical fluctuations and "noise" cause the computation to drift hopelessly off course? In a digital computer the path is broken down into logical steps that are embodied in the computer's program. In addition, each computing unit protects against voltage fluctuations by treating a range of voltages, rather than just the exact voltage, as being equal to a nominal value; for example, signals between .8 volt and 1.2 volts can all be restored to 1.0 volt after each logical step in the computation.

In collective-decision circuits the process of computation is significantly different. The overall progress of the computation is determined not by step-by-step instructions but by the rich structure of connections among computing devices. Instead of advancing and then restoring the computational path at discrete intervals, the circuit channels or focuses it in one continuous process. These



COMPUTATIONAL ENERGY of a collective-decision circuit can be pictured as a landscape of hills and valleys. The connection pattern and other physical characteristics of the circuit determine its contours. The circuit computes by following a path that decreases the computational energy until the path reaches the bottom of a valley, just as a raindrop moves downhill to minimize its gravitational potential energy. The surface shown here could represent an associative memory, in which the valleys correspond to memories that are stored as associated sets of information (*x's*). If the circuit is started out with approximate or incomplete information, it follows a path downhill (*colored arrow*) to the nearest valley, which contains the complete information.



COMPUTATIONAL PATH describes how the physical state of a computer changes as it computes. In a digital computer the path is a sequence of discrete steps controlled by the lines of code in a computer program. After each step the computation is restored to the desired path (*red line*). In collective-decision circuits the computational path is continuously focused in a way determined by the pattern of connections in the circuit.

two styles of computation are rather like two different approaches by which a committee makes decisions. In a digital-computer-style committee the members vote yes or no in sequence; each member knows about only a few preceding votes and cannot change a vote once it is cast. In contrast, in a collective-decision committee the members vote together and can express a range of opinions; the members know about all the other votes and can change their opinions. The committee generates a collective decision, or what might be called a sense of the meeting.

The nature of collective computation suggests that it might be particularly effective for problems that involve global interaction between different parts of the problem. We have designed circuits that perform this type of computation to solve certain optimization problems. A typical example is the task-assignment problem, which poses the question: If you have a certain number of assistants and a certain number of tasks, and each assistant does each task at a different rate, how should you assign the tasks so that the corresponding rates add up to the largest total rate? The neural-network circuit that can solve the problem has many interconnected amplifiers that process the data in parallel. It is able to follow the computational path to a solution rapidly. Because this is a rather complicated circuit, it is helpful to first examine some simple circuits that illuminate the basic principles of all such circuits.

The simplest example for the purpose is the flip-flop, a circuit that is widely used in the electronics industry. The circuit has two stable states—which give it its name—and it makes a decision by choosing one state over the other. It can be built from a pair of saturable amplifiers [*see illustration on opposite page*]. In such an amplifier the output voltage increases as the input rises until it reaches a saturation level, beyond which it will not change. The reverse is also true: as the input decreases, the output falls until it saturates at a minimum value. In the flip-flop the output of each of the two amplifiers is inverted (that is, multiplied by -1) and connected to the input of the other. The amplifiers mutually inhibit each other because a high output by either one will drive down the input of the other amplifier. This produces a self-consistent pattern, because each amplifier will drive the other one to be in the opposite state. The flip-flop therefore has two stable states: if amplifier A is putting out $+1$, then B will put out -1, and vice versa. The significant feature of this circuit is that the pattern of the connections is the key to its stability and determines the form of its stable states.

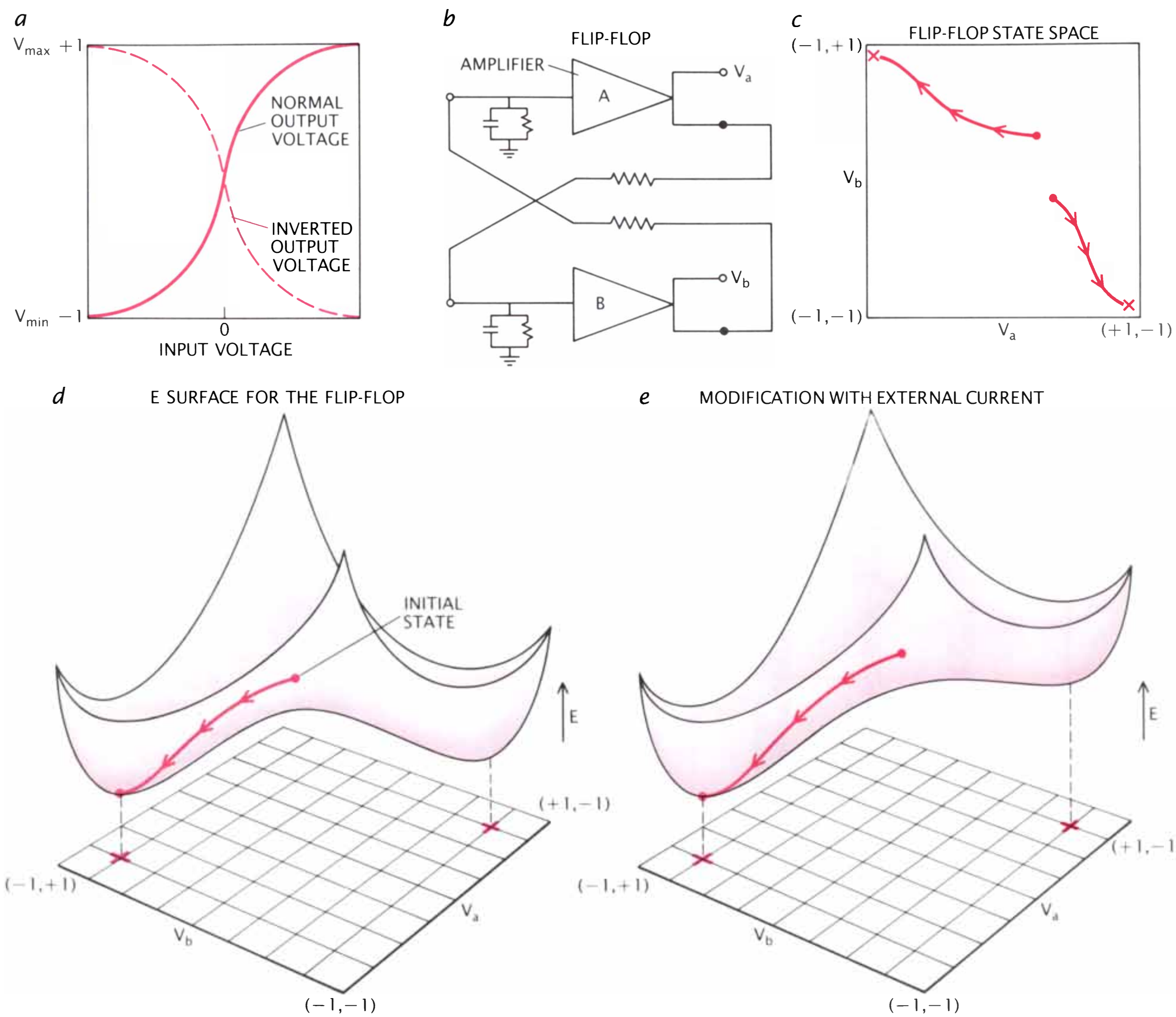
A seemingly remarkable feature of the flip-flop is that no matter what initial inputs are supplied to the circuit when it is turned on, it will make a rapid trajectory to one of the stable states. To understand the phenomenon, picture what happens when a raindrop lands on a terrain of hills and valleys. The drop moves generally downhill until it ends up at the bottom of a nearby valley. The path taken is one in which the gravitational potential energy of the raindrop is continuously decreasing. Similarly, the flip-flop's trajectory is associated with a mathematical quantity we call the computational energy E, which can be visualized as a terrain on which the flip-flop's voltage state moves continuously downhill.

E is defined by an explicit mathematical formula that depends on the characteristics of the amplifiers, the strength of the excitatory and inhibitory connections between them, and any external inputs. For fixed inputs in a particular circuit, if E is calculated for each possible configuration of amplifier voltages, it defines a continuous surface. For the flip-flop E can be plotted on a three-dimensional graph [*see illustration on opposite page*]. The surface contains two valleys near the voltage configurations $(+1,-1)$ and $(-1,+1)$, which correspond to the two stable states. When the circuit is operating, the changing voltages will describe a downhill motion along the E surface, and eventually the circuit's configuration will come to rest at the bottom of one of the valleys.

The concept of the computational energy proves useful in understanding many features of collective-decision circuits. For example, modifications of the flip-flop circuit alter the shape of the E surface in well-defined ways. If the strengths of the inhibitory connections increase, the valleys become deeper in relation to the "neutral point," or saddle point, in the middle of the E surface. External sources of current also alter the contours of the surface; if a positive current is supplied to the input of one of the amplifiers, it will tend to drive the amplifier to the $+1$ output state. The valley corresponding to this stable configuration will become deeper, and the change will be accompanied by an increase in the size of the "basin of attraction," the area within which any starting point will settle into the stable state at the bottom of the basin. If the external current is large enough, the basin of attraction will fill the entire flip-flop space, eliminating the valley corresponding to the other stable state and leaving only one stable state for the circuit.

The simple flip-flop circuit illustrates how the process of following the trajectory can be interpreted as a process of decision making. For example, the circuit can decide which of two numbers is larger if the amplifiers are given two external input currents that are proportional to the numbers. The amplifier with the larger input will then have a deeper valley at the stable state for which its output is +1, and its basin of attraction would expand to include the "neutral point." When the computation is begun by setting the voltages at this point, the circuit state would follow a downhill path to the deeper valley. When the circuit stabilizes, one can note which of the two amplifiers is in the +1 state and so determine which number is the larger. For any pair of numbers the corresponding input currents will cause the *E* surface to change in an appropriate way, thereby ensuring that the path will lead to the correct answer.

For more complicated collective-decision circuits the corresponding *E* surface acquires so many dimensions that it becomes impossible to draw. Nonetheless, one can understand the general features of the sur-



FLIP-FLOP CIRCUIT is built from two saturable amplifiers. In a saturable amplifier, as the input voltage increases or decreases, the output saturates at maximum and minimum voltages. The output can be normal or inverted (*a*). A resistor connects the output of one amplifier to the input of the other; its resistance determines the strength of the connection. The normal output terminal (*open circle*) can be used to make an excitatory connection. In the flip-flop the inverted output terminal (*filled circle*) is employed instead to make inhibitory connections. A capacitor and a resistor are connected in parallel at each input to store the charge flowing to the terminal and produce an input voltage and to allow a discharge current to flow (*b*). If the minimum and maximum outputs are +1 and −1 and amplifier *A* is saturated at +1, *B*'s input will be driven down and *B*'s output will saturate at −1. *B*'s output will in turn be inverted and drive up the input to *A*, thus keeping the output of *A* saturated at +1. The reverse situation, in which *A* is saturated at −1 and *B* at +1, is also stable. The configuration of the amplifier voltages is represented as a point on a two-dimensional plane (*c*). Each axis represents the output of one of the amplifiers, from −1 to +1. The circuit will always move to one of the two stable points near (+1,−1) and (−1,+1), no matter what the initial voltages were. A third axis represents the value of the computational energy *E* for each voltage configuration (*d*). The two stable points appear as valleys in the *E* surface. The edges of the surface rise steeply, because it is impossible to exceed the minimum and maximum outputs. If an external current is given to one of the amplifiers, this will deepen the valley that corresponds to that amplifier's be ng in the +1 state (*e*).

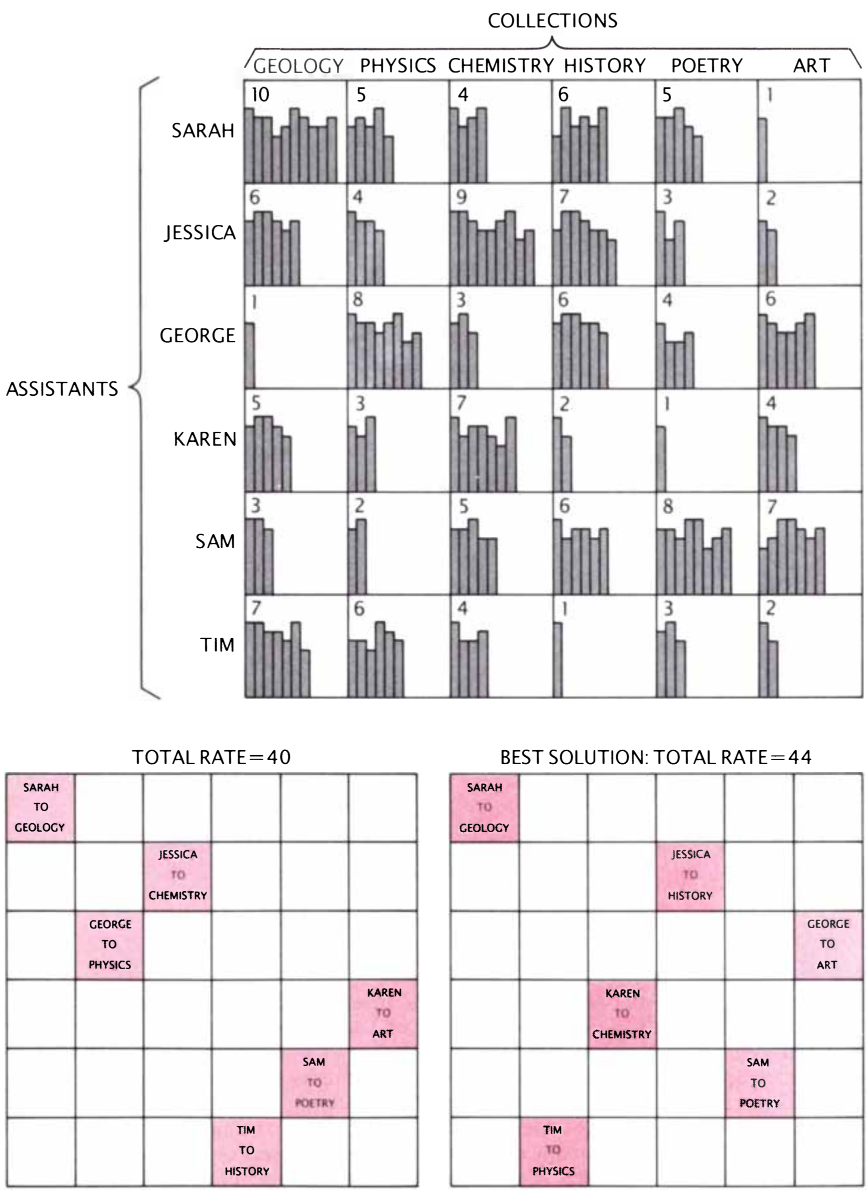
face and use these as a guide to designing and understanding the circuits. For example, we can generalize from the *E* surface of the flip-flop to devise a collective-decision circuit that can solve the slightly more difficult problem of determining the largest number in a set of *n* numbers. The circuit can be thought of as an *n*-flop, consisting of *n* amplifiers, each of which is connected to all others with inhibitory connections of equal value. It would have *n* stable states and its *E* surface would have *n* valleys. When a set of input currents is supplied to its amplifiers, the deepest valley would develop for the state that has a +1 output for the amplifier receiving the largest input.

In both the flip-flop and the *n*-flop there is a one-to-one relation between the number of amplifiers and the possible solutions, so that as the number of solutions gets larger, the size of the circuit does too. Is it possible to design a collective circuit that can represent a greater number of solutions than there are amplifiers? Such circuits do indeed exist. They have stable states that consist of *configurations* of amplifiers in the +1 state. This is a more economical use of amplifiers, just as the Roman alphabet is more economical than Chinese ideographs in its use of symbols to encode words.

In 1984 we discovered that networks of this type could rapidly compute good solutions to optimization problems such as the task-assignment problem mentioned above. As an example, imagine you are supervising the job of reshelving books for a large library. You have a number of assistants to do this for you. Each one is familiar with each collection—history, physics and so on—but to varying degrees. Thus Jessica can shelve six books per minute in geology, four per minute in physics and so on, whereas George can shelve one book per minute in geology, eight per minute in physics and so on. You must assign one category to each assistant. How should you assign the tasks so that the total rate of reshelving the books is as high as possible?

One could attack the problem by brute force, trying out every possible combination in sequence. But if there are many assistants and collections, one would soon be overwhelmed by the number of possibilities. If there are *n* categories of books and *n* assistants, the number of possible solutions would be the factorial of *n*, or $n(n-1)(n-2)\ldots1$. There are better iterative digital-computer algorithms that can arrive at an answer in a time period proportional to *n* cubed. The computation could be done even faster, however, if one could take full advantage of the problem's essence: the fact that the proper assignment of each worker depends on the capabilities of every other worker. Ideally the mutual dependencies should be considered simultaneously. It is precisely this kind of computation that can be done quickly and efficiently by a collective-decision circuit.

The data in the task-assignment problem consist of the set of shelving rates. These data can be arranged in a table, in which each row contains the rates for an individual assistant and each column represents a book category. An assignment of tasks can then be thought of as the choice of *n* elements in the table, with the constraint that there can be one and only one element chosen in each row and column, because only one assistant can be assigned to each category. The best solution has the highest sum of rates for the chosen assistants.



TASK-ASSIGNMENT PROBLEM requires assigning each assistant to one collection of books. The rates at which books are shelved per minute are represented in a table (*a*). For this six-by-six problem there are 720 ways to assign the tasks. The pink squares show two possibilities (*b*). The best assignment has the largest sum of shelving rates.

Tandy Computers: Because there is no better value.™

Desktop Publishing

Now you can create professional-looking documents with an affordable PC system.

Now you can design, lay out and produce top-quality printed materials right at your desk. PageMaker software turns the new Tandy 4000 computer into a sophisticated publishing system, so simple to learn and use that anyone can get terrific results.

With PageMaker desktop publishing software, you can produce almost any printed piece imaginable—from newsletters and brochures to catalogs and annual reports. The finished product is so professional most people will think you had it typeset.

A Tandy computer system using PageMaker software will save you time and money on virtually any printed work you produce by eliminating typesetting delays and charges.

Our finest system includes the 80386-based Tandy 4000 with a 20 MB hard disk, LP 1000 Laser Printer, EGM-1 Enhanced Graphics Monitor, plus PageMaker software, all for a surprisingly affordable price.

Step up to typeset-quality desktop publishing today. Visit your nearby Radio Shack Computer Center for a demonstration.

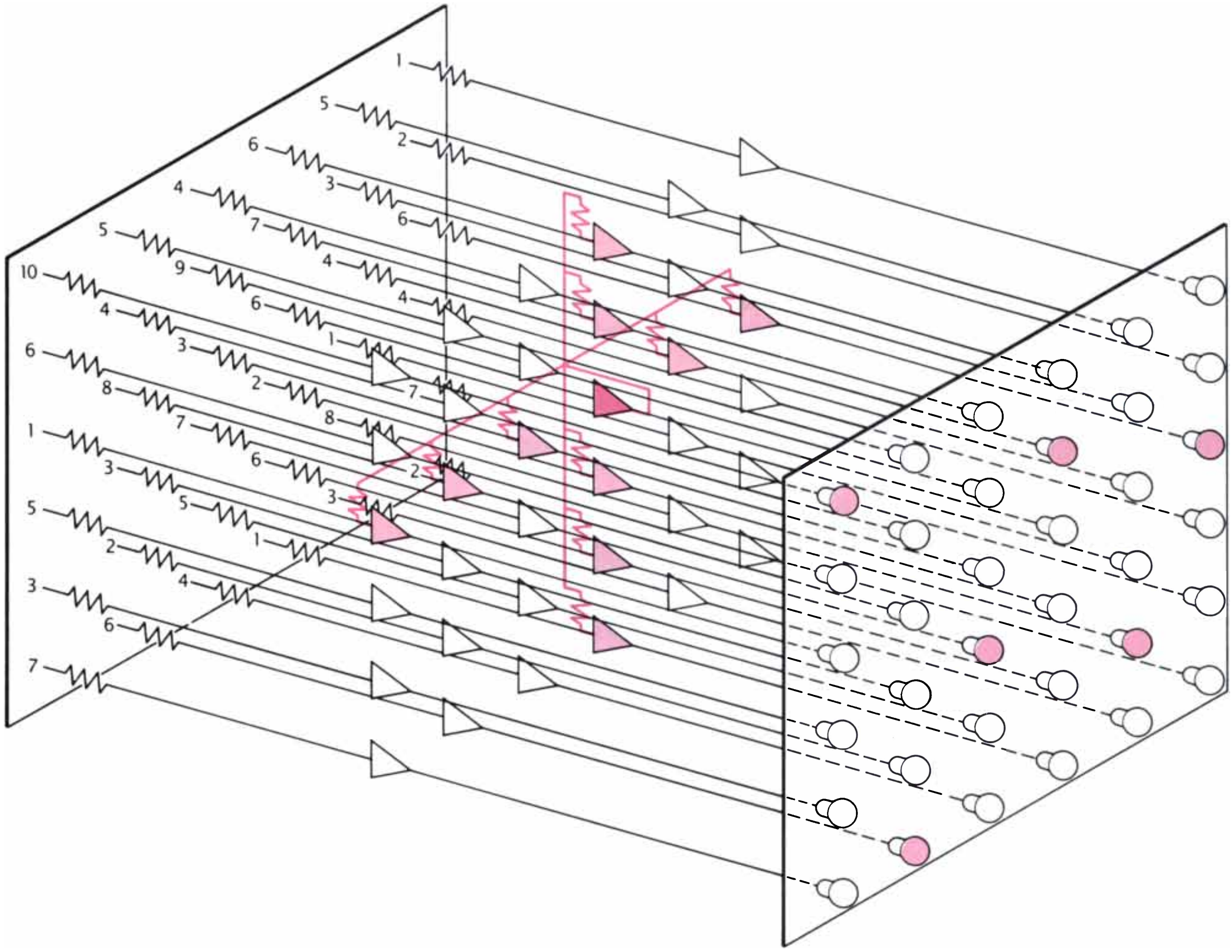
Send me a new 1988 computer catalog.

Mail To: Radio Shack
Dept. 88-A-910
300 One Tandy Center
Fort Worth, TX 76102

Name ____________________
Company ____________________
Address ____________________
City ____________ State ______
ZIP ____________________
Phone ____________________

Radio Shack®
COMPUTER CENTERS
A DIVISION OF TANDY CORPORATION

PageMaker/Reg. TM Aldus Corp. AT/Reg. TM IBM Corp.



OPTIMIZING CIRCUIT to solve the task-assignment problem consists of a network of interconnected *n*-flops. The amplifiers in each row and column are linked by inhibitory connections, which provide the constraint that only one amplifier in any given row and column can be in the +1 state. Because each of the 36 amplifiers in this network inhibits 10 other amplifiers, there are 360 connections altogether. The diagram depicts the connections for one of the amplifiers. The amplifiers receive input currents proportional to the shelving rates. The amplifiers that correspond to the best solution—the combination of inputs that add up to the largest sum—put out a +1 and the rest put out a 0. The +1 outputs can drive a display, such as a light-bulb array.

We solved the problem by building an *n*-by-*n* array of amplifiers in which each row corresponds to an assistant and each amplifier in the row corresponds to a different task. The amplifiers in each row and column are linked by mutually inhibitory connections; this provides the constraint that only one assistant can be assigned to each collection, because if one of the amplifiers has a +1 output, the other amplifiers are inhibited. Another way of looking at the circuit is that each row and column is an *n*-flop. These *n*-flops cannot function independently, however, because each amplifier belongs to two such *n*-flops. As you will see below, this pattern of connections is the key to the circuit: it ensures that the circuit will have self-consistent stable states that correspond to possible solutions to the problem.

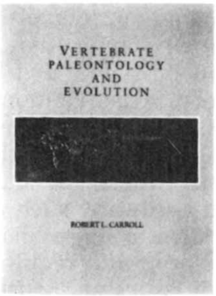
What are the stable states of this network and what does its *E* surface look like? The stable states consist of configurations of 36 amplifiers in which there are six amplifiers with +1 outputs, with one and only one such amplifier in any row or column. In a six-by-six array the number of these stable states is 720, or 6 factorial. The *E* surface for the circuit has valleys of equal depth for each of the 720 possibilities. An input current proportional to the shelving rate of each assistant for each collection is fed to the corresponding amplifier. The valley for each possible solution becomes deeper by an amount proportional to the sum of its corresponding shelving rates.

The network carries out the computation by following a trajectory down the *E* surface. In the final configuration the circuit usually settles into the deepest valley, which is the correct choice because it corresponds to the task assignment that has the highest total shelving rate. In simulation studies we have shown that this particular circuit will almost always find the best solution to the problem, and a slightly more complex circuit will always find the best solution.

One reason we are interested in studying this type of circuit is that perceptual problems can often be expressed as an optimization. Our senses gather a large set of information about the external world—information that is inevitably imprecise

Save 20% on these new titles from W. H. Freeman and Company

The book publishing arm of *Scientific American*



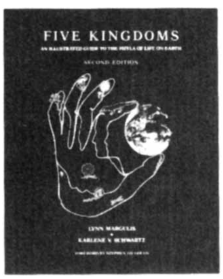
Vertebrate Paleontology and Evolution

Robert Carroll,
REDPATH MUSEUM, McGILL UNIVERSITY

"A real tour de force."
—Léo Laporte,
University of California, Santa Cruz

Vertebrate Paleontology and Evolution is the first detailed, authoritative review of the fossil record in twenty years. Incorporating the latest discoveries and newest ideas about evolutionary processes, it provides complete, current data on all fossil groups and examines each in terms of classification, skeletal and functional anatomy, way of life, distribution in time and space, and interrelationships with other groups. Featuring more than 1700 illustrations and a massive bibliography containing nearly 1500 citations, this is an indispensable reference to the fascinating history of fossil vertebrates through 600 million years of evolution.

698 pages, 1779 illustrations
Regular Price $52.95 **YOUR PRICE $42.59**



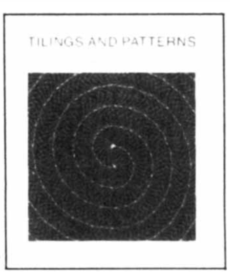
Five Kingdoms An Illustrated Guide to the Phyla of Life on Earth

Second Edition

Lynn Margulis, BOSTON UNIVERSITY
Karlene V. Schwartz,
UNIVERSITY OF MASSACHUSETTS, BOSTON
Foreword by Stephen Jay Gould

This profusely illustrated guide classifies all life on Earth into the 92 phyla that make up nature's five great kindgoms: Prokaryotae, Protoctista, Fungi, Animalia, and Plantae. The Second Edition is completely up-to-date and includes three newly-recognized phyla—including the first new animal phylum to be acknowledged in 30 years—and incorporates new ideas in both protoctistan and bacterial relationships. The Second Edition includes bibliographies, an appendix listing the phylum and common name of some 1000 genera, a glossary of technical terms, and an extensive index.

416 pages, 421 illustrations
Regular Price $35.95 **YOUR PRICE $28.79**



Tilings and Patterns

Branko Grünbaum,
UNIVERSITY OF WASHINGTON
Geoffrey C. Shephard,
UNIVERSITY OF EAST ANGLIA, ENGLAND

This is the first comprehensive and systematic approach to the classification and enumeration of tilings. "I am enormously impressed.... Will be the definitive reference on tiling theory for many decades. Not only does the book bring together older results that have not been brought together before, but it contains a wealth of new material.... I know of no comparable book." *—Martin Gardner*

700 pages, 1,406 illustrations
Regular Price $59.95 **YOUR PRICE $47.99**

Molecular Cell Biology

James Darnell, ROCKEFELLER UNIVERSITY
Harvey Lodish, INSTITUTE FOR BIOMEDICAL RESEARCH and M. I. T.
David Baltimore, DIRECTOR, WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH, and M. I. T.

Darnell-Lodish-Baltimore's *Molecular Cell Biology* is the new standard in the field, offering the most authoritative and up-to-date treatment of cell biology and molecular genetics available today. "A didactically effective and intellectually elegant work.... An exceptionally versatile book [that] will meet the needs of both students and more advanced investigators and teachers...." —Thoru Pederson, in *Cell* (12/5/86)

1,187 pages, 1,248 illustrations
Regular Price $42.95 **YOUR PRICE $34.59**

Please send me:

Quan	ISBN	Title	Special Low Price	Amount
	1822-7	Vertebrate Paleontology	$42.59	
	1885-5	Five Kingdoms	$28.79	
	1193-1	Tilings and Patterns	$47.99	
	1448-5	Molecular Cell Biology	$34.59	
			TOTAL	
			Add $1.50 for shipping and handling, plus $1.00 for each additional book	
			Sales tax (NY, NJ, CA, and UT)	
			FINAL TOTAL	

☐ I enclose a check made payable to W. H. Freeman and Company

☐ Charge my:

☐ MasterCard ☐ VISA Account #__________

Expiration date___/___ Signature__________
(Credit card orders must be signed)

Name__________

Address__________

City / State / Zip

Mail to: **W. H. Freeman and Company**
4419 West 1980 South, Salt Lake City, Utah 84104

and "noisy." The edge of an object might be hidden behind another object, for example. We know, however, that the edges of objects are continuous, and just because we cannot see an edge does not make us wonder whether the object has changed its shape. Our interpretation of the information is constrained by what we already know.

This knowledge can often be represented as a set of constraints, similar to those in the task-assignment problem, and express it in an E function. The perceptual problem then becomes equivalent to finding the deepest valley in the E surface. For example, Cristof Koch, Jose Marroquin and Alan Yuille, who were then at the Massachusetts Institute of Technology, showed how several important problems in computer vision could be cast as an optimization problem and solved by a collective-decision circuit in which knowledge of the real world had been imposed as a set of constraints. Their circuit was able to take incomplete depth information of a three-dimensional world and reconstruct missing information such as the locations of the edges of objects.

Another particularly interesting application for collective-decision circuits is associative memory, which is a form of optimization problem. An associative memory is different in principle from a digital-computer memory. A conventional computer stores information by assigning addresses, which identify the physical locations where the data will be stored in hardware, such as a sector or track on a floppy disk. When the central processor requires a piece of data, it issues an instruction to read the data at a particular address. The address itself contains no information about the nature of the data stored there.

Now reflect for a moment about your own memories. If you think of a particular friend, you will remember many facts—name, age, hair color, height, job, hobbies, schooling, family, house, shared experiences and so forth. These facts are somehow combined to form your memory of the individual. There is no notion of storage address in the way you retrieve such information from your memory. Instead pieces of the information itself are used in place of an address.

Associative memory is an idea that came from psychology, not electrical engineering. Fruit flies and garden slugs have associative memories. Indeed, the fact that such relatively simple nervous systems display the phenomenon suggests that it must be a natural—almost spontaneous—property of neuron ensembles. It seems reasonable to ask whether associative memory could also be achieved in networks of artificial neuronlike devices. In the 1970's a number of investigators, including James A. Anderson of Brown University and Teuvo Kohonen of the University of Helsinki, developed mathematical models of associative memory. The concept of the E surface provides a means to understand and study associative-memory circuits built of saturable amplifiers.

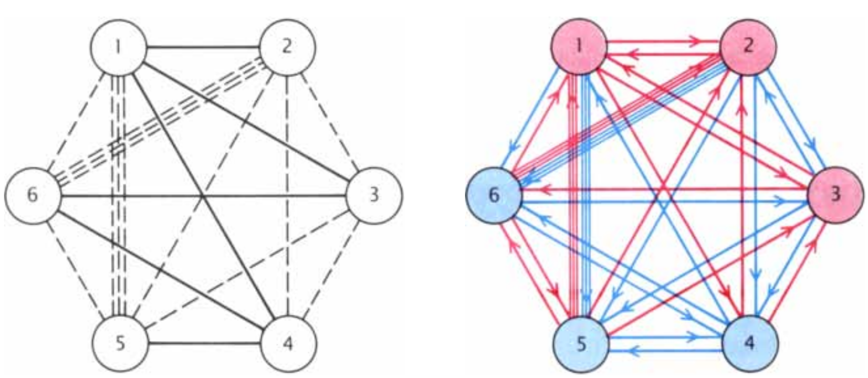
How would one make a collective-decision circuit behave like an associative memory? Consider a space of many Cartesian coordinates in which each axis is labeled with some attribute a person might have. One axis might refer to height, one to hair color, one to weight, one to sailing experience, one to the first name of the individual, one to city of residence and so on. Any point in the space de-

FEATURES ASSIGNED TO NODES

	1 NAME	2 HEIGHT	3 AGE	4 WEIGHT	5 HAIR	6 EYES
−1	SMITH	TALL	OLD	THIN	BROWN	BLUE
+1	JONES	SHORT	YOUNG	FAT	BLOND	BROWN

NODES

MEMORIES	1	2	3	4	5	6
A	+1	+1	+1	−1	−1	−1
B	+1	−1	+1	+1	−1	+1
C	+1	+1	−1	+1	−1	−1



ASSOCIATIVE MEMORY with six nodes, or "neurons," is linked by excitatory (*solid line*) and inhibitory (*broken line*) connections. The number of lines in each link represents the strength of the connection; each solid line represents a connection strength of +1 and each broken line represents a strength of −1. Each node might represent a characteristic of a person, as is shown in the table (*a*). Suppose one wants to store three memories, or sets of characteristics (*b*). The nodes that are supposed to be in the +1 state are given an excitatory link to the other +1 nodes and an inhibitory link to the −1 nodes. To store information about all three memories one simply adds up the connections (*c*). For example, the link between nodes 2 and 4 is (−1) + (−1) + (+1), or −1. When the circuit is turned on for memory *A*, the network produces (*d*) the correct pattern of nodes in the +1 state (*red*) and −1 state (*blue*). The pattern is self-consistent: at each node the positive (*red*) and negative (*blue*) incoming currents always add up to have the same sign as the node itself. If the network is given partial data—about a thin, short Jones, for example—it will go into a stable state from which one can retrieve the entire memory.

scribes the characteristics of a hypothetical possible individual. Each of your friends is represented by a particular point in the space. Because you have very few friends compared with the set of all possible individuals, if you put a mark at the position of each of the people you know, you will have marked a very few points in a large space. When someone gives you partial information about a person—for example color of hair and weight but not name—this describes an approximate location in the space of possible people. The idea of an associative memory is to find the friend who best matches the partial data.

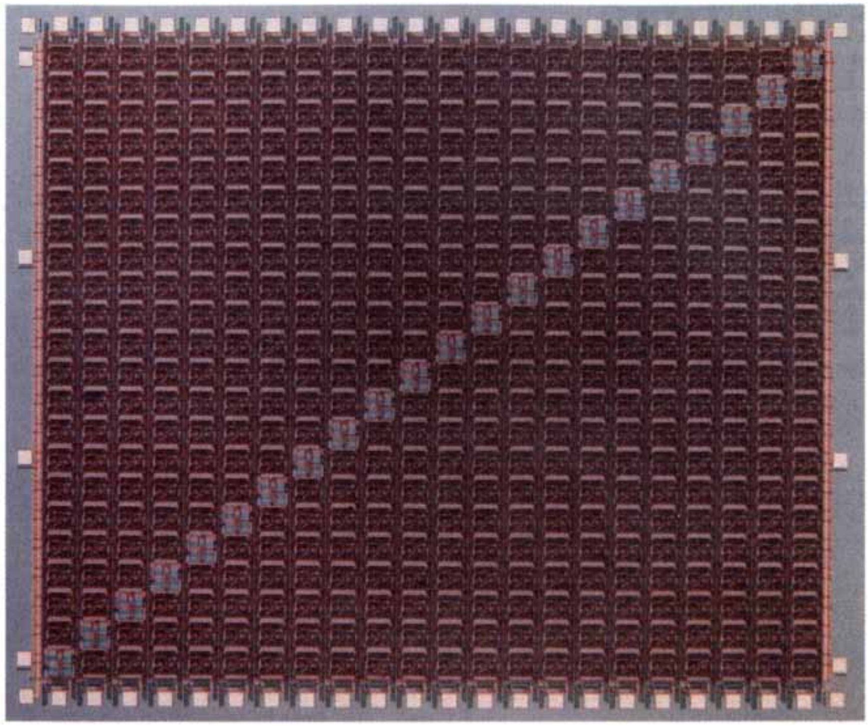
A collective-decision circuit such as the one described for the task-assignment problem could perform as an associative memory if the *E* surface can be shaped to have valleys, or stable points, at the places that correspond to particular memories. A pattern of input voltages corresponding to a partial memory would be supplied to the amplifiers and the circuit would then follow a trajectory to the bottom of a local valley in the *E* terrain and read out the output state of the amplifiers as the stored memory. Unlike the task-assignment circuit, in which the connections are highly regular because of the simple global rules that constrain the problem, in an associative memory the connections are irregular and the stable points are scattered somewhat at random because the memories need not have any particular relationship among themselves. To construct an associative memory, therefore, one must find connections between amplifiers such that the many desired memories are represented simultaneously by the circuit's stable states.

A simple associative memory of six interconnected amplifiers illustrates how information can be stored in such a network [*see illustration on opposite page*]. The memory states of the system could be described as six-bit binary words, in which each bit corresponds to one of the two possible saturated output states of an amplifier, $+1$ and -1. For example, memory *A* is $(+1,+1,+1,-1,-1,-1)$. As with the flip-flop circuit, a state can be stable only if it is self-consistent. This is accomplished by ensuring that each amplifier with a $+1$ output has an excitatory connection to the input of every other amplifier that has a $+1$ output and an inhibitory connection to the input of each amplifier that has a -1 output, and vice versa for amplifiers that have -1 outputs. All the inputs to an amplifier are added up to give a big signal with the correct sign. If one looks at the *E* surface for this associative memory, one will find that the connections have created a valley at the location of the memory.

Because the data are distributed in the pattern of the connections in the circuit, many other memories can be overlaid in the same circuit. It is merely necessary to calculate the connections separately for each memory and add them to the connections for the memories already in storage. This simple additive rule works quite effectively as long as not too many of the same connections are shared among many memories. Problems arise if memories are too similar or too numerous; the valleys on the *E* surface get too close and begin to interact. (The number of unrelated memories that can be stored effectively is about 15 percent of the number of "neurons" in the circuit.) There are cleverer schemes that can store a larger number of memories or memories that are more similar.

The associative memory described above requires only local information about two linked "neurons" in order to modify the strength of the existing connection between them. This is appealing because it offers a theory of associative memory that is consistent with a biological model proposed more than 30 years ago by Donald O. Hebb. Hebb postulated that biological associative memory must reside in the synaptic connections between nerve cells and that the process of learning and memory storage involves changes in the strength with which nerve signals are transmitted across individual synapses. According to his theory, synapses linking pairs of neurons that are simultaneously active become stronger, thereby reinforcing those pathways in the brain that are excited by specific experiences. As in our associative-memory model, this involves local instead of global changes in the connections. The Hebbian synapse had long eluded actual observation, but recently several investigators have reported evidence for such mechanisms in the brain.

Many laboratories are now exploring how to fabricate and use devices for collective computation. A variety of prototypes have already been built with microelectronic and



VLSI COLLECTIVE-DECISION CIRCUIT was designed in 1985 at the California Institute of Technology by Massimo Sivilotti, Michael R. Emerling and Carver A. Mead. It contains 22 amplifiers, which are the lighter-color components along the diagonal. The devices filling the rest of the square provide the connections, which can be programmed to make the chip an associative memory. The size of the chip is six by six millimeters.

SCIENTIFIC AMERICAN

In Other Languages

LE SCIENZE
L. 3,500/copy L. 35,000/year L. 45,000/(abroad)
Editorial, subscription correspondence:
Le Scienze S.p.A., Via G. De Alessandri, 11
20144 Milano, Italy
Advertising correspondence:
Publietas, S.p.A., Via Cino de Duca, 5,
20122 Milano, Italy

サイエンス
Y880/copy Y9600/year Y13,000/(abroad)
Editorial, subscription, advertising correspondence:
Nikkei Science, Inc.
No. 9-5, 1-Chome, Otemachi
Chiyoda-ku, Tokyo, Japan

INVESTIGACION Y CIENCIA
450 Ptas/copy 4950Ptas/year $35/(abroad)
Editorial, subscription, advertising correspondence:
Prensa Cientifica S.A.,
Calabria, 235-239
08029 Barcelona, Spain

POUR LA SCIENCE
27FF/copy 265FF/year 315FF/year (abroad)
Editorial, subscription, advertising correspondence:
Pour la Science S.A.R.L.,
8, rue Férou,
75006 Paris, France

Spektrum DER WISSENSCHAFT
9.80 DM/copy 99 DM/year 112.20 DM/(abroad)
Editorial, subscription correspondence:
Spektrum der Wissenschaft GmbH & Co.
Moenchhofstrasse, 15
D-6900 Heidelberg,
Federal Republic of Germany
Advertising correspondence:
Gesellschaft Für Wirtschaftspublizistik
Kasernenstrasse 67
D-4000 Duesseldorf,
Federal Republic of Germany

科学
1.40RMB/copy 16RMB/year $24/(abroad)
Editorial, subscription correspondence:
ISTIC-Chongqing Branch, P.O. Box 2104,
Chongqing, People's Republic of China

В МИРЕ НАУКИ
2R/copy 24R/year $70/(abroad)
Editorial correspondence:
MIR Publishers
2, Pervy Rizhsky Pereulok
129820 Moscow U.S.S.R.
Subscription correspondence:
Victor Kamkin, Inc.
12224 Parklawn Drive,
Rockville, MD 20852, USA

TUDOMÁNY
98Ft/copy 1,176Ft/year 2,100Ft/(abroad)
Editorial correspondence:
TUDOMÁNY
H-1536 Budapest, Pf 338
Hungary
Subscription correspondence:
"KULTURA"
H-3891 Budapest, Pf. 149
Hungary

العلوم
1KD/copy 10KD/year $40/(abroad)
Editorial, subscription, advertising correspondence:
MAJALLAT AL-OLOOM
P.O. BOX 20856 Safat,
13069 – Kuwait

Advertising correspondence all editions:
SCIENTIFIC AMERICAN, Inc.
415 Madison Avenue
New York, NY 10017
Telephone: (212) 754-0550 Telex: 236115

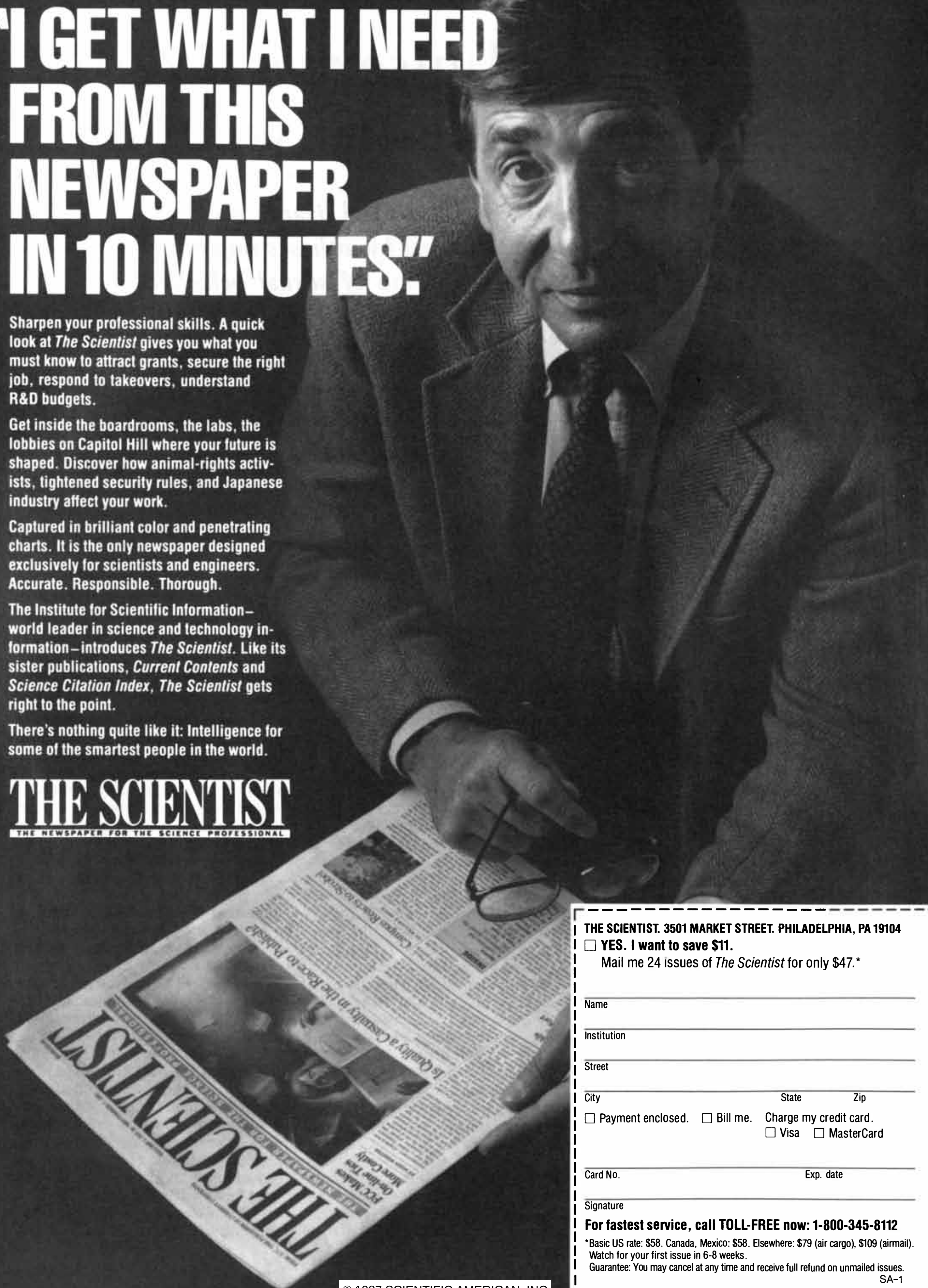
optical hardware. To be useful, circuits will have to be large, with hundreds or thousands of "neurons," and because these may be densely interconnected, the circuits may contain tens of thousands or even millions of connections. In addition, in order to build a general-purpose circuit such as an associative-memory chip one would need a simple method for modifying the connection strengths.

John J. Lambe and his collaborators at the Jet Propulsion Laboratory constructed from integrated-circuit amplifiers the first associative-memory network of the type we have described. The connections between each pair of amplifiers were chosen through a mechanical switch. The network was expanded to contain 32 amplifiers, with microcomputer-controlled transistor switches replacing the mechanical ones. These pioneering circuits work as predicted but are too cumbersome to be of practical use. The first VLSI version was fabricated by Massimo Sivilotti, Michael R. Emerling and Carver A. Mead of the California Institute of Technology. The circuit reduced a 22-amplifier network with 462 interconnections to an area smaller than a square centimeter. The chip functioned as an associative memory when the connection matrix was appropriately programmed. Similar VLSI circuits with 54 amplifiers have been built by Lawrence D. Jackel, Richard E. Howard and Hans Peter Graf of the AT&T Bell Laboratories. One of the attractive features of collective-decision circuits is that they converge on a good solution rapidly, typically in a few multiples of the characteristic response time of the computing devices. In several of the microelectronic implementations this convergence has occurred in less than one microsecond. Mead's group has recently built a VLSI "artificial retina" chip for image processing, using collective-computation principles in the design.

Advanced optics provides another promising medium for building collective-decision circuits. In that approach light beams would replace the wires. Because light beams can pass through one another without interaction, this raises the possibility of implementing complicated network topologies that might be difficult to achieve in VLSI. Demetri Psaltis of Caltech and Nabil Farhat of the University of Pennsylvania have built working prototypes of optical collective circuits [see "Optical Neural Computers," by Yaser S. Abu-Mostafa and Demetri Psaltis; SCIENTIFIC AMERICAN, March].

Many investigators are studying neuronlike circuits different from the ones we have described. A popular model is the feedforward Perceptron, which has been shown to be effective for a broad range of applications, such as pattern recognition. This model consists of simple processing units arranged in several layers. Information is passed into the network through an input layer, and the result of the network's computation is read out at the output layer. There are connections between the layers, and information flows forward only. Such feedforward networks have simplified dynamical behavior and reduced computational capability. On the other hand, many useful learning rules have been devised for such circuits that make it easy to find the appropriate connection pattern. One well-known example, called back-propagation, has been independently derived by David Parker, by David Rumelhart, Geoffrey Hinton and Ronald Williams, and by Paul J. Werbos. One goal of current research is to understand how similar learning algorithms might be applied to networks that have the richer dynamical behavior produced by the kind of feedback employed in the circuits we have discussed.

The study of collective computation in neuronlike circuits has shown that such networks can carry out computations that are not trivial. Computations that are more complicated may require having many simple decisions interact collectively to produce a complex decision. Another feature of many complex decisions is that they must combine information arriving over an extended period of time. Suppose, for example, one wants to identify someone from a distance by the way he walks. One must first make simple decisions about the positions of limbs, combine these over time to determine a sequence of movements and from these form a complex pattern that can be associated with a particular individual. The study of such hierarchical and time-varying collective-decision systems has just begun, but we believe that, as in the case of the circuit-design principles we have described, the research will be propelled by the architectures and design rules of nature's computers.



THE SCIENTIST
Is Quality a Casualty in the Race to Publish?

Courtship in Unisexual Lizards: A Model for Brain Evolution

An all-female species of whiptail lizards presents a unique opportunity to test hypotheses regarding the nature and the evolution of sexual behavior

by David Crews

Courtship of females by males is a necessary part of life for most animal species; if successful, it ends in copulation, thereby ensuring fertilization and the perpetuation of the species. In the majority of animals studied so far, however, courtship serves a function unrelated to fertilization: it stimulates ovarian development and reproductive activity on the part of the female and increases the overall efficiency of reproduction [see "The Reproductive Behavior of Ring Doves," by Daniel S. Lehrman; SCIENTIFIC AMERICAN, November, 1964]. Thus it may not seem odd that in species that lack males, male-like behavior continues to be exhibited.

This is true for an unusual species of whiptail lizard, *Cnemidophorus uniparens,* which inhabits the southwestern U.S. The species is derived from a "bisexual" ancestor (one that consists of males and females) but is itself "unisexual" (there are no males). The females reproduce by parthenogenesis, that is, reproduction without fertilization, and therefore copulation between its members is not directly related to the production of offspring. Nevertheless, the females in this species actively engage in lengthy courtship behaviors that are virtually identical with those observed between male and female whiptail lizards.

What mechanisms would allow such a behavior to persist in the absence of males? The brain, which controls mating behavior in males and females, not only has adapted to a new set of stimuli in this species but also has mediated a switch to females of behavioral patterns that are normally associated with males. This reinforces the observation that the brain is equipped with neural circuits for both male and female behavioral repertoires, regardless of biological sex. By investigating the manner in which that has come about, using unisexual lizards as my model, I have gained insight into the ability of the brain to adjust to changing conditions during the course of evolution.

I became interested in the subject almost 10 years ago while studying the influence of hormones on the courtship of anole lizards [see "The Hormonal Control of Behavior in a Lizard," by David Crews; SCIENTIFIC AMERICAN, August, 1979]. I had been sent six adults of *C. uniparens* and wanted to compare their hormone cycles with those of the male and female anoles I had been studying for several years.

While working in my laboratory I happened to glance at the cage containing the whiptail lizards and noticed two individuals engaged in courtship with each other. The observation, made purely by accident, stunned and intrigued me. Over several weeks the behavior was repeated many times, following the same pattern on each occasion. This clearly was not random interaction between individuals but a highly ritualized form of behavior. How might it be explained? It seemed unlikely that parthenogenetic females would expend so much time and energy on an activity that had no apparent purpose. Was the behavior a useless vestige of the species' sexual history or did it have some real biological significance?

In order to answer that question I decided to compare the pseudosexual behaviors of these unisexual lizards with the sexual behavior of their closest bisexual relatives. I needed to do this for two reasons: first, I wanted to determine what behavioral differences, if any, existed between these species and how they might have been altered in the switch from bisexuality to unisexuality; second, I wanted to ascertain whether or not courtship between male and female whiptails serves a function other than insemination. If it does, this would suggest that a similar function might be served by the courtship seen between parthenogenetic females.

One way to test this hypothesis would be to compare the number of offspring produced by females that engage in courtship behavior with the number produced by females that do not. If courtship among unisexual females could be shown to result in greater reproductive output, I would have reason to believe it is not merely a useless vestige of the past but a strategy critical to the continuing success of the species.

To understand the mechanisms by which pseudosexual activity among unisexual females evolved requires a look at the neuronal and hormonal factors that mediate it. It is well known that courtship is controlled by the central nervous system: sex hormones synthesized by males and females act directly on certain areas of the brain to trigger sex-typical behaviors. How is it that male-typical courtship behavior (which is mediated by male hormones) shifted from males to females during the course of evolution? Two possible explanations seem reasonable. Either the mechanisms that control male behavior in the ancestral sexual species were retained but modified in some way in the descendant species, or an entirely new and unknown mechanism had evolved to control male-like behavior in unisexual females.

The whiptail lizards (genus *Cnemidophorus*) have several unique characteristics that make them ideal subjects for a study of this kind. There are 45 species in the genus; of these 15 have been identified as all-female species derived from existing bisexual ones. The bisexual species consist of males and females in a normal one-to-one ratio and reproduce sexually (as the vast majority of vertebrates do). The unisexual species are formed from the hybridization of closely related species; they consist solely of females that reproduce parthenogenetically. The offspring in such species are clones: they are genetically identical with their mothers and with one another [see "Unisexual Lizards," by Charles J. Cole; SCIENTIFIC AMERICAN, January, 1984].

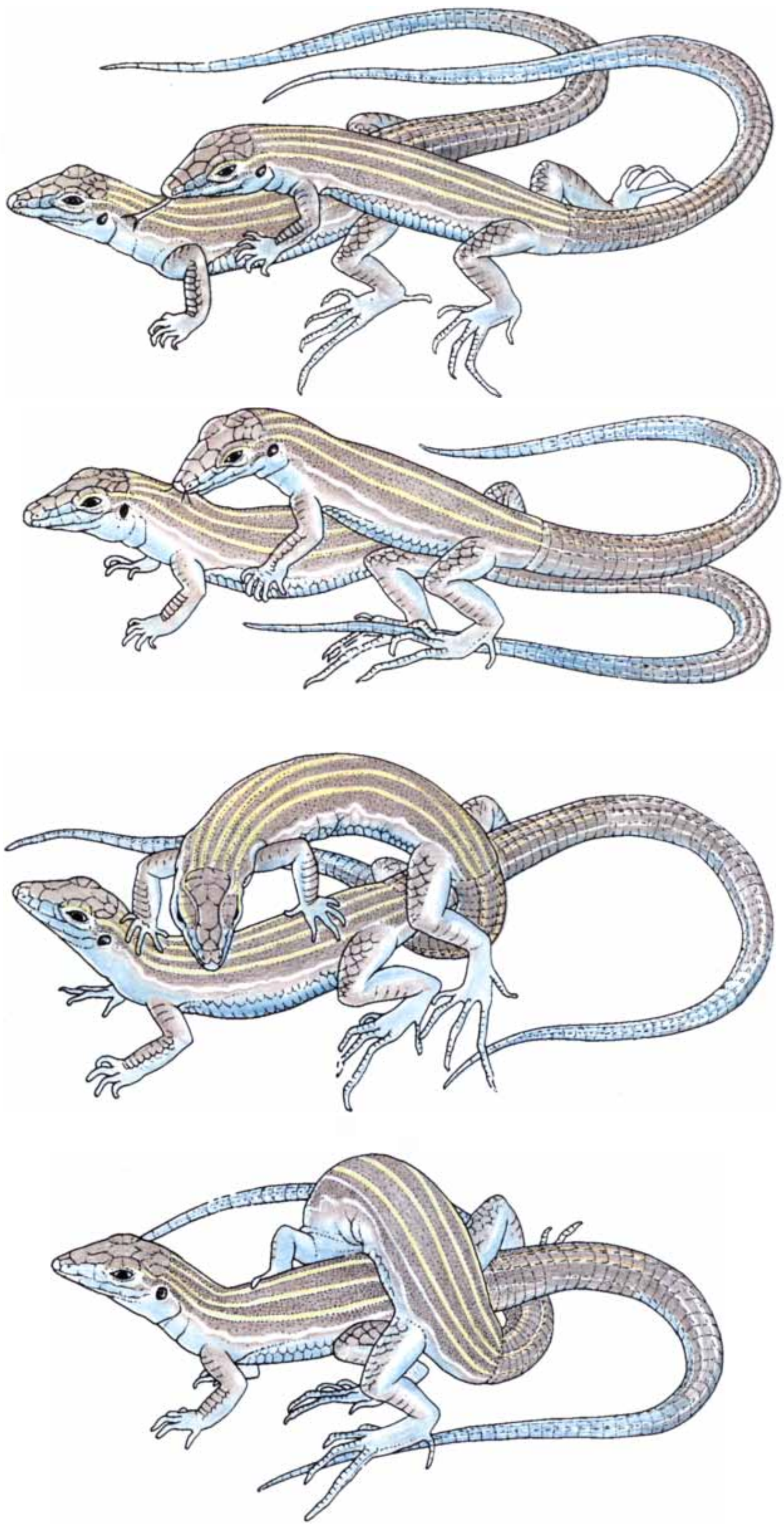
Because these unisexual lizards are the result of a hybrid union, formed by the fusion of gametes from two existing species, both the ancestral (that is, the parental) and the descendant (the unisexual) species can be identified using molecular techniques. This is a subtle but critical point. The existence of an ancestral species provides a control group against which the reproductive strategies of its descendant species can be compared. By comparing the DNA sequences of various whiptail lizards, Llewellyn D. Densmore III, Craig C. Moritz and Wesley M. Brown of the University of Michigan were able to determine that the maternal ancestor of *C. uniparens* is the bisexual species *C. inornatus*. Although they were unable to identify the paternal ancestor, the identification of one parental ancestor was sufficient for the purposes of this study.

I began my investigation by comparing mating in both species. In *C. inornatus*, as in *C. uniparens*, courtship follows a well-defined series of steps: the male typically approaches a female and probes her body with his tongue; if she is sexually receptive, he grips her with his jaws by grabbing the skin on either her neck or her foreleg. That act appears to pacify the female and provides the male with an opportunity to climb on her back, straddling her with his legs. Once mounted, the male scratches the sides of the female with his fore and hind legs, pressing her body into the substrate as he does so. After several minutes the male maneuvers his tail under that of his female partner, thereby bringing his cloaca into contact with hers.

During mating one of the lizard's two hemipenes (lizards, like snakes, have paired penises) is inserted into the female's genital opening. On intromission the male shifts his jaw grip from the female's neck to her pelvic region, assuming a contorted position I refer to as the doughnut posture. He maintains this position for between five and 10 minutes, during which time ejaculation occurs. The male then rapidly dismounts and leaves the female.

Members of the descendant unisexual species *C. uniparens* exhibit a courtship ritual almost identical with that of their bisexual ancestors, except that some females assume the role of the male (I call this behavior male-like) whereas others behave like typical females. Male-

WHIPTAIL LIZARD, *Cnemidophorus uniparens*, from the southwestern U.S. is one of 15 species in the genus that are known to consist only of females. In six of the species females have been observed participating in courtship behavior with one another.



COURTSHIP RITUAL between females in the unisexual species *C. uniparens* is virtually identical with that of its bisexual ancestor, *C. inornatus.* The "male" (the female that displays male-like behavior) will approach a female (one displaying typical female-like behavior) (*top*) and investigate her with its tongue. If the female is receptive, the "male" will mount her (*top center*), grabbing her by the skin of her neck as it does so. After a few minutes the "male" will swing its tail under that of the female (*bottom center*), a move that places their cloacal openings in apposition. The "male" then shifts its jaw grip to the pelvic region of the female (*bottom*), forming the characteristic doughnut posture. This step, also known as pseudocopulation, lasts between five and 10 minutes.

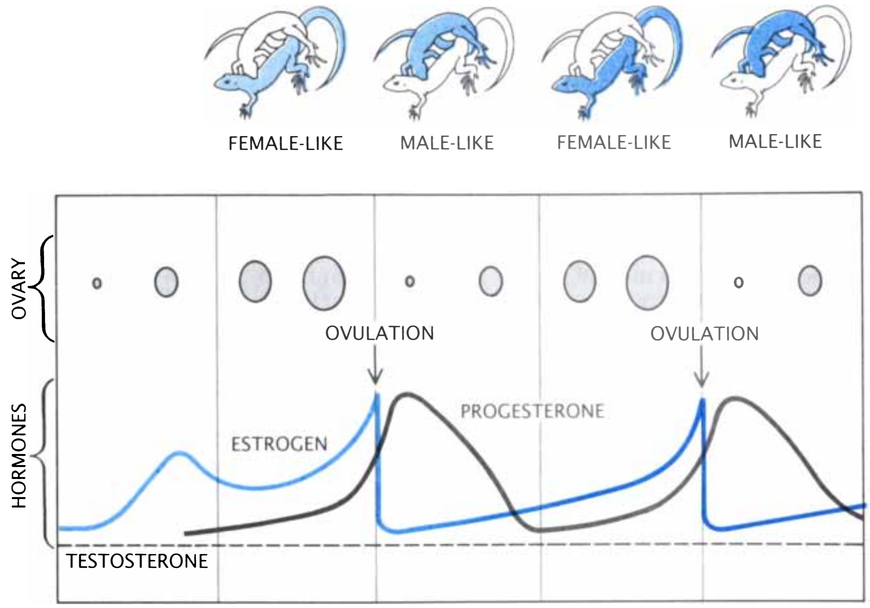
like females approach and mount sexually receptive females, gripping them with their jaws as they do so; after a few minutes they swing their tails under those of their partners and shift their jaw grips to form the characteristic doughnut posture (I call this behavior pseudocopulation). The only difference between pseudocopulation and true copulation is that the unisexual lizards are morphologically female (they lack hemipenes), and so intromission cannot occur between them.

Having determined that the courtship displays of these two species of whiptail lizards are virtually identical, I needed to determine what impact, if any, courtship (without insemination) has on reproduction. Among many species, including human beings, courtship has a primer effect: it serves to synchronize the reproductive physiologies of the male and female and to regulate normal ovarian development in females. If I could find evidence that courtship by males stimulates ovarian development in *C. inornatus,* I would have reason to suspect that a similar phenomenon might be taking place in *C. uniparens,* only with females substituting for males.

With Jonathan K. Lindzey, a graduate student in my laboratory, I began examining rates of ovulation in *C. inornatus* females reared under three different conditions: alone, with other females, or with males. The results were consistent with those found for other vertebrates: these females will not ovulate in isolation; if caged with other females, only a few will ovulate and they produce fewer eggs. In the presence of males, however, their ovaries grow much more rapidly, so that there is more frequent ovulation and greater total egg production. This represented a major finding: the results clearly indicate that courtship behavior in the ancestral species is a potent physiological force and is critical to the reproductive success of the species.

I wanted to see if a similar physiological effect might be operating in *C. uniparens.* In its natural habitat a female of this species will lay two or three egg clutches, each containing an average of two or three eggs, during the course of the breeding season (which lasts from spring until midsummer). I was able to demonstrate, with the help of Jill E. Gustafson, an undergraduate in my laboratory, that females reared in isolation laid an average of only .8 clutch over the course of the breeding season, whereas those reared in the presence of other sexually active females (with whom they engaged in pseudocopulation) produced more than three times that number, or an average of 2.5 clutches over the same period.

Thus we successfully proved that courtship behavior enhances reproduction in both bisexual and unisexual lizards. One major question remained unanswered, however: What underlying physiological mechanisms control sexual activity in these two species? If I could answer that question, I would know whether or not the same hormones responsible for male behavior in *C. inornatus* trigger male-like behavior in *C. uniparens*. This in turn would enable me to understand the role of the brain in sexual behavior and how it may have evolved in response to unisexuality.



SEXUAL ACTIVITY in *C. uniparens* fluctuates in accordance with the ovarian cycle. Females that are preovulatory display female-like courtship behavior; those that are postovulatory exhibit behavior that is male-like. The change from female to male sex roles is mediated by changes in the circulating levels of sex hormones: high levels of estrogen are associated with female-like behavior; a sharp drop in estrogen (following ovulation) and a rapid increase in progesterone are associated with the switch to male-like behavior. The cycle is repeated two or three times during the breeding season.

I sought first to test the response of *C. uniparens* females to elevated levels of the male sex hormones testosterone and dihydrotestosterone (collectively called androgens). I operated on these lizards to remove their ovaries (the primary source of sex hormones in females) and then administered androgens to them. By doing so I found I could consistently induce male-like behavior in ovariectomized females. Although this seemed unusual (androgens are normally synthesized by the testes), the implications were intriguing. If male hormones were in fact responsible for male-like behavior, then the hormone profiles of these females must have been radically altered over the course of evolution. Had the ovaries, which normally produce female sex hormones, assumed the role of the testes? Were they producing male hormones to compensate for the absence of males? To prove such a hypothesis would require a thorough investigation into the physiological basis for sexual activity in both the unisexual and the bisexual species.

I began that undertaking by mapping the ovarian cycle, the focal point of reproduction in females and the process by which eggs are produced. We had already established, on the basis of egg production, that *Cnemidophorus* females undergo two or three discrete ovarian cycles (each approximately from three to four weeks in duration) during the three or four months of the mating season. Each of these cycles in turn can be divided into two distinct phases: the follicular, or preovulatory, phase, and the luteal, or postovulatory, phase. During the follicular phase yolk critical to the development of the embryo is deposited in the follicles, which reach a maximum diameter of 10 millimeters in six to eight days. The luteal phase follows ovulation when corpora lutea (the remnants of the follicles) form in the ovary and the eggs pass into the oviducts. Once in the oviduct the eggs acquire a hard outer shell and are usually laid between seven and 14 days after ovulation.

Having mapped the ovarian cycle, we now needed to compare it with fluctuations in sexual activity. Lindzey found that females of the ancestral species, *C. inornatus*, are receptive to males only during the follicular phase, when they are preovulatory. During the luteal phase, when they are postovulatory, they will aggressively reject the advances of males. We found a similar pattern of activity in *C. uniparens*. Female-like behavior is only expressed by females in the preovulatory phase, whereas male-like behavior appears consistently in females that have recently ovulated and are in the postovulatory phase.

As with sexual activity, patterns of hormone secretion vary according to the ovarian cycle. The female hormones, estrogen and progesterone (which are produced by the follicles and the corpora lutea respectively), fluctuate in these lizards much as they do in most vertebrate females. Michael C. Moore, a postdoctoral fellow in my laboratory, found that in both *C. uniparens* and *C. inornatus* circulating concentrations of estrogen increase during the follicular phase, and reach peak levels at the time of ovulation. Progesterone levels, which are low during most of the follicular phase, are at their highest during the luteal phase and then decline rapidly when the corpora lutea are resorbed in from seven to 10 days. A new cycle begins in both species about a week later. The similarities between the estrogen and progesterone cycles in the two species indicate that their hormone profiles have remained stable over the course of evolution.

Contrary to what I had expected based on the results of my early experiment with ovariectomized females, we found that circulating levels of the male hormones testosterone and dihydrotestosterone in male-like females were uniformly low and in most instances could not be detected in our radioimmunoassay tests. There was no indication, even during or after pseudocopulation, of transient surges in their concentration. During the entire reproductive cycle of *C. uniparens,* androgen levels were at least 1,000 times lower than is characteristic of sexually active *C. inornatus* males. Although this suggested that the hormone profiles of this all-female lizard species had not been profoundly altered, it failed to explain why the administration of androgens had triggered male-like behavior in females whose ovaries had been removed.

Analysis of the males of *C. inornatus* reveals that testosterone and dihydrotestosterone are present in normal levels in these animals. When sexual activity begins during the spring and early summer months, both these hormones are elevated in the males' bloodstream. The levels increase significantly when the females emerge from hibernation about two weeks after the emergence of the males, a correlation suggesting that the presence of females increases hormonal levels in male whiptails. Similarly, concentrations of the female hormones estrogen and progesterone are low, a pattern that conforms to the one found in most males of bisexual species.

I still needed to prove that testicular hormones control courtship and copulation in the males of *C. inornatus.* To do this, Lindzey and I surgically removed the testes from a group of sexually active males and then placed the males in cages with sexually receptive females. The results were unambiguous: castrated males court females significantly less often than intact males do. Moreover, by administering sex hormones to the castrated males, we were able to obtain direct evidence that the decline in sexual activity was linked to the loss of hormones secreted by the testes. Those that received hormones renewed courtship activity in the presence of sexually active females.

It was at this point that we made an exciting discovery. Some of the castrated males treated with progesterone responded in precisely the same way as all the males did to androgens: both groups became sexually active and copulated with the same degree of frequency. We later discovered that the different hormones also trigger an identical response in males near the end of the breeding season. Normally courtship behavior declines markedly at this time, but males given either progesterone or an androgen continue to display high levels of sexual activity in the presence of females.

These findings are exciting for two reasons. First, they are the opposite of what has been found in mammals and birds (in those animals progesterone is a potent suppressor of sexual activity in males; here progesterone stimulates sexual activity), and second, they suggest that this sensitivity to the female hormone progesterone may have set the stage for the evolution of pseudocopulation in unisexual whiptails.

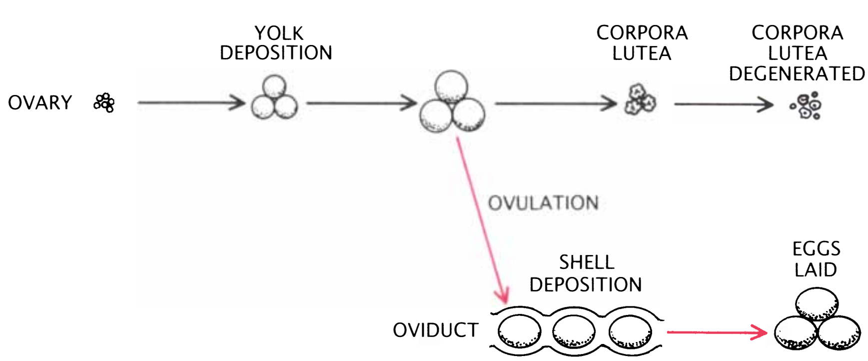
Notwithstanding, pseudocopulation remained a mystery. We knew that the loss of males in *C. uniparens* had been accompanied by a significant loss of testosterone and dihydrotestosterone, but we did not know what had taken their place as regulators of male-like behavior. The fact that male-like courtship was retained in this species indicated that the neural circuits underlying male behavior had been retained, but the hormones had changed.

We knew that androgens trigger courtship behavior in *C. inornatus* males, but we could find no evidence for either testosterone or dihydrotestosterone during any phase of the ovarian cycle in *C. uniparens.* Thus we suspected that male-like behavior in that species is triggered by another cue, most likely a hormone produced by the ovary. We had several reasons for believing this was the case. Pseudocopulation had never been observed in females that were either reproductively inactive or had had their ovaries removed. Moreover, female- and male-like behaviors are clearly defined by the ovarian cycle: female-like behavior is limited to the follicular phase, male-like behavior to the luteal phase.

In many animals, transitions in sexual behavior correspond to hormonal changes. In these lizards, during the transition from female- to male-like behavior, circulating levels of estrogen drop by a factor of three, whereas levels of progesterone increase by a factor of nine. Could this abrupt shift in hormone concentration play a crucial role in the expression of pseudosexual behavior? Mark Grassman, a postdoctoral fellow in my laboratory, and I hypothesized this was so. We guessed that progesterone might be the endogenous stimulus responsible for male-like behavior and estrogen the stimulus for female-like behavior.

Furthermore, we surmised that if pseudocopulation were to operate as an effective reproductive strategy, complementarity would have to exist, that is, at any one time during the breeding season some females would be in the follicular phase and others would be in the luteal phase. We could test this by removing the ovaries from a number of females and then implanting capsules in them that contained either estrogen, progesterone or nothing. By pairing them in different combinations and monitoring their ensuing sexual activity, we could correlate behavior with specific hormones.

The results were unequivocal. When females that received progesterone were paired with females that received estrogen, pseudocopulation ensued, with the former individuals consistently performing the role of males and the latter the role of females. Pairs in which one individual received an empty capsule and the other a capsule containing

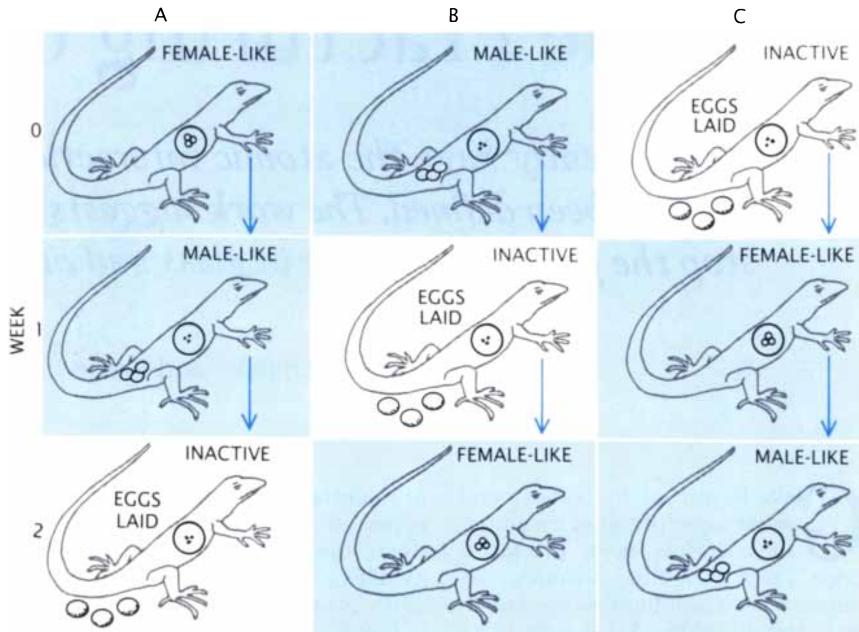


SUCCESSIVE CHANGES in the ovaries of whiptail lizards take place during the course of the ovarian cycle. At the beginning the ovaries are small, but they become progressively larger as the eggs inside them develop and acquire greater amounts of yolk. When the eggs are mature, they are released into the oviducts (this step is called ovulation), where they acquire hard shells, a process that lasts for from seven to 14 days. During this period the corpora lutea in the ovaries degenerate, a step that readies the ovaries for another round of egg production. The entire cycle, which lasts for from three to four weeks, is repeated several days after the first batch of eggs is laid.

either progesterone or estrogen did not engage in sexual activity. I had previously determined that individuals housed together quickly establish and maintain this type of complementarity in their reproductive states. By alternating sex roles they maximize fecundity and increase the efficiency of reproduction.

The data from these experiments fit nicely into existing notions about the ability of the brain to adjust to different hormonal influences. It has often been said that the brain is initially bisexual and it is only later in development that it becomes sexually dimorphic. We wanted to explore this notion further, believing that in doing so we might finally solve the mystery of unisexual courtship.

We knew from studies on bisexual vertebrates that two distinct neural circuits mediate sexual behavior between males and females: one includes the anterior hypothalamus–preoptic area (AH–POA) of the brain and controls mounting and copulation in males; the other includes the ventromedial hypothalamus (VMH) and controls sexual receptivity in females. Moreover, both these areas of the brain have receptors for the corresponding male or female sex hormones.



COMPLEMENTARITY in the ovarian cycles of *C. uniparens* females makes courtship in this species possible. During the breeding season females will synchronize their activities so that when some individuals are preovulatory, others are postovulatory. This is represented here by three females whose behavior was observed over the course of several weeks. As a female changes from a state of preovulation (indicated by enlarged ovaries) to one of postovulation (indicated by the presence of eggs in the oviduct) her behavior changes from female-like to male-like. Because these lizards are reproductively out of phase (when one is female-like, her partner is male-like), courtship activity is continuous during the mating period. Sexually inactive females (those that have just laid their eggs) resume female-like behavior at the start of their next ovarian cycle. Individuals engaged in courtship with one another during a given week are shown in color.

Would the same distinctions hold true for unisexual lizards? With help from Mark Mayo, an undergraduate, and Juli Wade, a graduate student in my laboratory, I sought to determine which areas of the brain control male- and female-like behavior in *C. uniparens*. We implanted minute quantities of sex hormones directly into the brain of individuals that were sexually inactive (we had surgically removed their ovaries). The results were clear-cut. Androgens and progesterone implanted into the AH–POA activated male-typical pseudosexual behavior (estrogen did not); conversely, estrogen implanted into the VMH stimulated female-typical behavior (androgen and progesterone did not).

Thus we knew the neuronal circuits that mediate sexual behavior in *C. uniparens* were the same ones that control sexual behavior in bisexual vertebrates. We hypothesized that progesterone was able to stimulate male behavior in the AH–POA by binding to androgen receptors present in the neural circuits. Because these animals lack androgen almost entirely, there would be little competition for the binding sites and progesterone would bind readily to them.

Experiments Lindzey and I have carried out with Kathleen Matt of Arizona State University on bisexual lizards suggest that is the case. By radioactively labeling androgens and progesterone we were able to determine that progesterone is effective in displacing androgen from hormone receptors in the AH–POA. These experiments are only suggestive, however, and do not conclusively prove that progesterone binds to androgen receptors in the brain of *C. uniparens* females.

Some other possible mechanisms by which male-like behavior may be expressed by females must be considered. One is that progesterone is converted in the brain into androgen. This is unlikely, however, because we have recently found that R5020, a synthetic progesterone analogue, which cannot be converted into androgen in the brain, will stimulate sexual behavior in castrated *C. inornatus* males. Another possibility is that progesterone binds to progesterone receptors that are functionally linked to the AH–POA. But this is also unlikely, because it suggests there is a fundamental difference between the progesterone receptor of whiptail lizards and all other vertebrates studied to date.

By dissecting the elements of pseudosexual behavior in a unisexual species of whiptail lizards we have shown how behavior critical to ovarian development and reproductive success can be retained when the conditions under which it originally evolved change. In the case of *C. uniparens* the loss of males meant the loss of male hormones that normally control male-typical mating behavior. Nevertheless, the persistence of male-like behavior in a unisexual lizard was possible because particular features of the brain of its ancestor, *C. inornatus* (namely the presence of dual neural circuits mediating male- and female-typical behaviors, and the sensitivity of its androgen receptors to progesterone), were coopted to serve new functions in the absence of males. In this way courtship behavior between females has provided new insights into behavioral evolution and the means by which the neuroendocrine mechanisms that control behavior adapt to changing conditions.

The Fracturing of Glass

Only recently have the atomic interactions underlying glass fracture been defined. The work suggests ways to slow or even stop the growth of cracks in glass and other brittle materials

by Terry A. Michalske and Bruce C. Bunker

Glass is one of the oldest and most widespread of commodities: it has been produced since 7000 B.C. and currently accounts for a $20-billion-a-year industry in the U.S. alone. Yet in spite of its impressive history of meeting a remarkable range of consumer needs, glass has always been limited in its applications by its tendency to fracture. This inherent weakness poses an increasing challenge to engineers. Young and promising technologies that require fiber-optic networks, ceramic bone replacements or novel optical and electronic components have placed a high premium on glasses and ceramics that will prove particularly resistant to cracking. A transoceanic fiber-optic cable, for example, must have a long service life to make laying the cable on the ocean floor a viable enterprise.

Until recently very little was understood about the mechanism by which glass cracks. In the mid-1960's, for instance, when investigators reported precise measurements showing that the stress needed to crack glass decreases with increasing exposure to water, they helped to explain why water aids glass cutters but did little to illuminate precisely how it does. An answer to the question of how glass cracks started to emerge in 1979. In that year the two of us, together with our colleagues at the Sandia National Laboratories and, independently, Stephen W. Freiman of the National Bureau of Standards, began to develop mathematical and chemical models to describe the fracture of glass at the atomic level. We are still developing and refining the models, but our efforts have already been rewarded. The physical and chemical interactions that control the rupture of interatomic bonds at the tip of a crack provide a fascinating link between the atomic structure of materials and the real-world concerns of product reliability. Moreover, the fundamental knowledge evolving from atomistic studies of crack propagation enhance the ability to utilize glass and other oxide materials in many of the demanding applications now being planned.

Most of us are introduced to the topic of glass fracture at an early age. The lesson may involve a shattered milk glass or the crash of a window struck by an errant baseball. When glass shatters, cracks seem to appear instantaneously. High-speed photography reveals that cracks can spread through glass at speeds of hundreds of meters per second, or roughly half the speed of sound in glass.

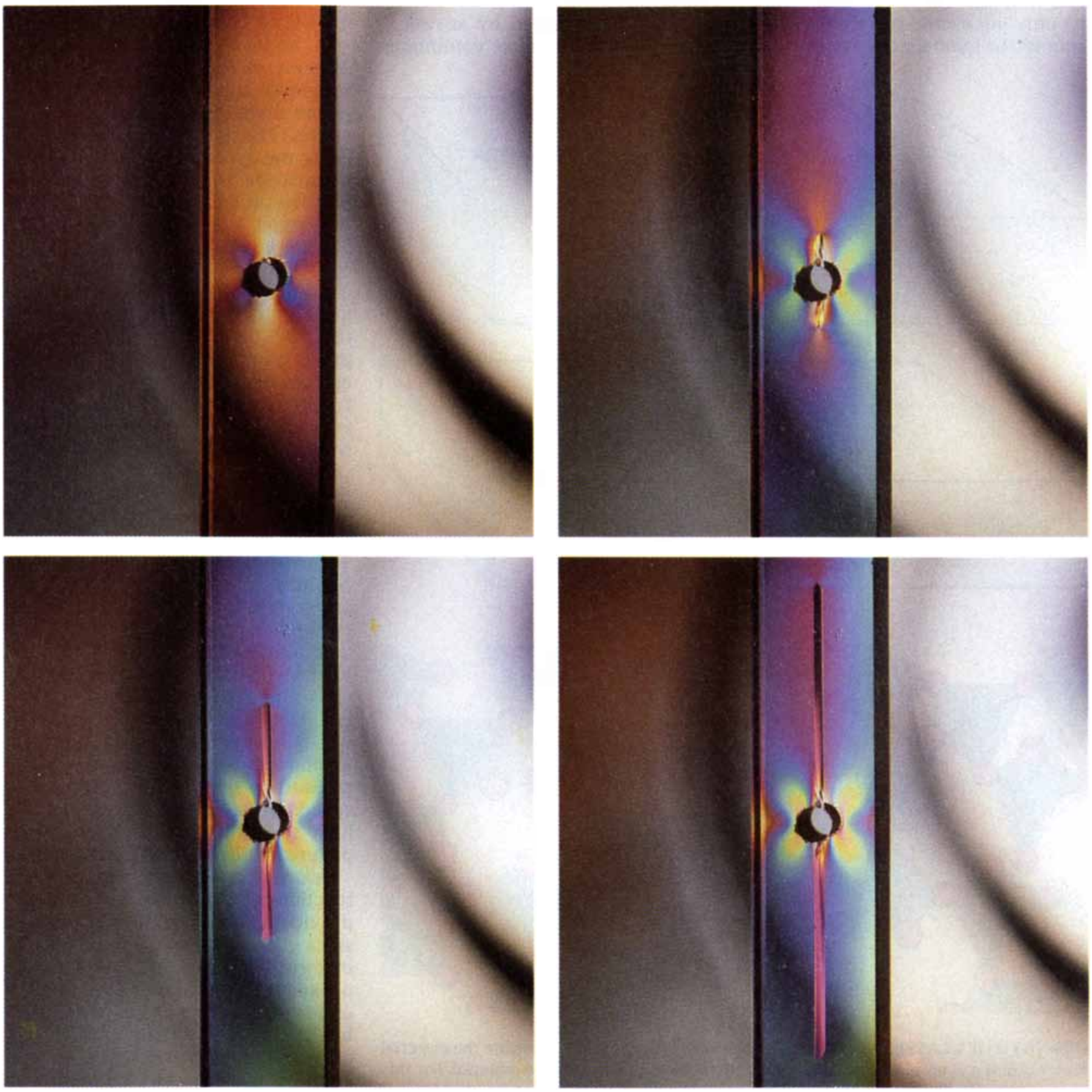
Even though the fracture of glass can be a dramatic event, many failures are preceded by the slow extension of preexisting cracks. A good example of a slowly spreading crack is often found in the windshield of an automobile. The extension of a small crack, which may have started from the impact of a stone, can be followed day by day as the crack gradually propagates across the entire windshield. In other cases small, unnoticed surface cracks can grow during an incubation period and cause a catastrophic failure when they reach a critical size. Cracks in glass can grow at speeds of less than one-trillionth of an inch per hour, and under these conditions the incubation period can span several years before the catastrophic failure is observed. On an atomic scale the slow growth of cracks corresponds to the sequential rupturing of interatomic bonds at rates as low as one bond rupture per hour. The wide range of rates over which glass can fracture—varying by 12 orders of magnitude (factors of 10) from the fastest shatter to the slowest creep—makes the investigation of crack growth a particularly engaging enterprise.

Surprisingly, glass is intrinsically one of the strongest materials known. Under high-vacuum conditions flaw-free glass can withstand tensile loads greater than two million pounds per square inch, which is 10 times the strength of most commercial metal alloys. In normal service, however, glass surfaces are exposed to abrasives and chemicals that create small surface cracks and promote their growth, ultimately reducing the strength of the glass. The effects of chemicals on the growth rate of cracks pose the greatest challenge to the design engineer, since these effects not only reduce the immediate strength of the glass but also can result in the failure of a structure that has supported a stress load for several years.

One of the most potent chemical agents is water, which presents a particularly severe threat because it is always in the atmosphere. Water can accelerate the rate of crack growth more than a million times by attacking the structure of the glass at the very root of the crack. Much of our work has centered on understanding from an atomistic point of view how water and other chemicals accelerate crack growth and how the acceleration can be slowed, stopped and even reversed.

Artisans, of course, have long recognized and utilized the ability of water to cause glass to crack more easily. Evidence suggests that early American Indians exploited the effect of water in making arrowheads from flint, a form of silica (silicon dioxide) related to glass, sand and quartz. Indians in the Catahoula Lake area of Louisiana, for instance, performed a ceremony in which they steamed flints before knapping (fracturing) them. Recent experiments on similar materials confirm that knapping is made easier by soaking flint in water. Today glaziers often apply water, usually in the form of saliva, to the shallow crack produced by their scribing tool. The water decreases the stress required to propagate the initial crack and causes the glass to break more smoothly.

A scientific basis for determining the conditions necessary for crack growth and fracture began to evolve 60 years ago with the pioneering work of A. A. Griffith of the Royal Aircraft Establishment. Griffith sought to calculate the minimum energy necessary to make a crack grow. He started with the well-known observation that atoms at the surface of a solid do not mesh with their neighboring atoms in the same way as atoms in the interior do: the atoms sit rather uncomfortably on the surface of the solid, and so their energy is higher than the energy of the interior atoms. As a consequence, anytime a new surface is created, energy must be supplied. Griffith reasoned that a crack can grow in glass only when the applied mechanical energy, or stress, is greater than the energy of the new surfaces created by the fracture. (Until the applied stress exceeds the minimum, the energy is stored in



GROWTH OF CRACKS in a glass bar with a hole drilled in its center has been monitored in a series of experiments done by the authors and their colleagues at the Sandia National Laboratories. Compressive load applied along the long direction of the bar creates tensile stress that propagates two cracks, one above the hole and one below it. The rate at which the cracks grow is measured with a microscope that sits on top of the apparatus. The entire assembly is enclosed in a high-vacuum chamber so that the chemical environment can be carefully controlled. Here a large amount of water vapor is present, which speeds the growth of the cracks. The photographs, made using polarized light, show that stresses are concentrated at the tip of each advancing crack.

the glass as it would be in a spring.) By applying his knowledge of the surface energy of glass and using existing calculations for the stresses around surface cracks, Griffith determined the breaking load for cracked plates. He triumphantly confirmed his prediction in actual experiments with glass tubes.

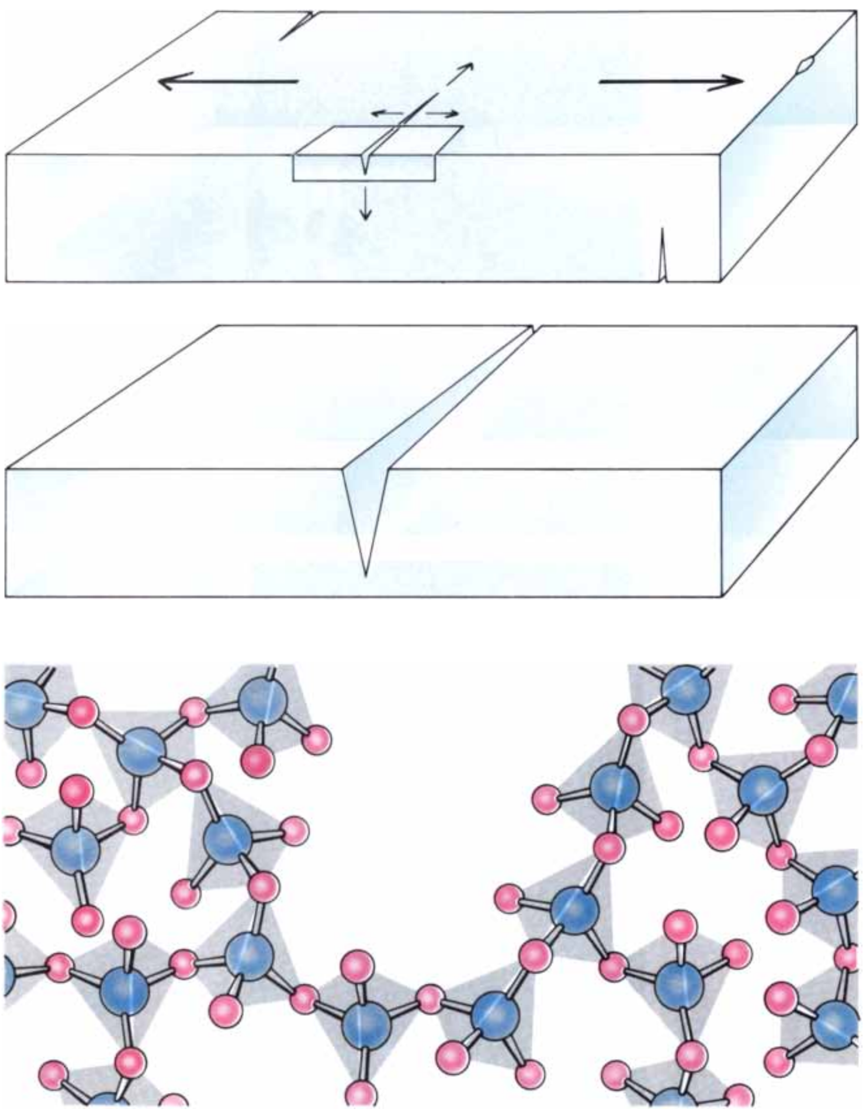
Griffith also determined that the smaller the initial crack in a piece of glass is, the greater the applied stress must be to extend it. This explains why pristine fibers of glass having only minuscule surface flaws are from 100 to 1,000 times stronger than ordinary window glass, which has usually acquired relatively large surface flaws in the course of handling. Griffith's energy-balance approach to strength and fracture also suggested the importance of surface chemistry in the mechanical behavior of brittle materials. Chemicals, such as water, that lower the surface energy of a solid ultimately will reduce the strength of the material.

The 1950's and 1960's saw a flurry of activity in the experimental study of glass strength. Much of the new enthusiasm was spurred by advances in the ability to produce continuous filaments of glass (fiber glass) that could serve to reinforce plastics. Various workers confirmed that water reduces the strength of glass and showed that the breaking load also depends on how long a stressed fiber is exposed to water. They found that load durations of only two weeks reduce the strength of glass threefold. Moreover, it appeared that given enough time under a load the strength of the glass would actually approach zero. Obviously these observations were most disturbing to the glass-manufacturing community.

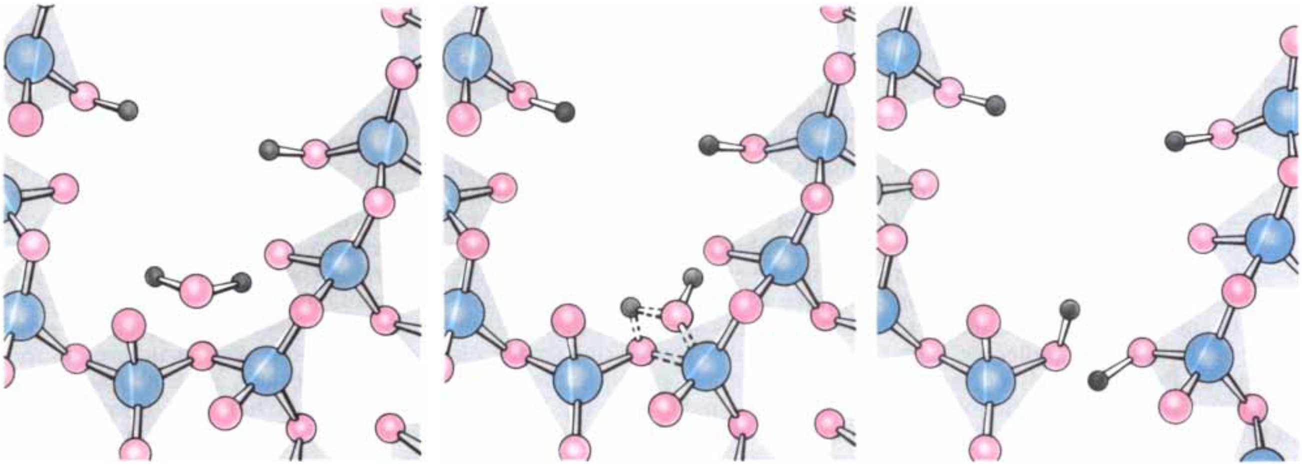
In the mid-1960's Sheldon M. Wiederhorn of the National Bureau of Standards and several other investigators studied just how the strength of glass changes with time. By measuring the growth of cracks under a microscope while carefully controlling the amount of stress and environmental conditions, they made several important discoveries. First, cracks in glass grow continuously at controlled rates; fracture is not simply an off-on phenomenon. Second, the rate of crack propagation depends on the applied stress and the amount of water in the environment. Finally, one can predict the time of failure by knowing how fast small surface flaws grow under stress; as small cracks in stressed glass slowly extend, more stress is concentrated, until rapid failure occurs.

Unfortunately many failures that may occur 10 years from now can result from small cracks that begin growing today at speeds of less than one-trillionth of an inch per hour—which, as we have mentioned, corresponds to the sequential rupturing of interatomic bonds at the rate of one bond per hour. Current experimental techniques are not capable of measuring these very slow crack speeds. It is for this reason that we have sought to develop an atomistic picture of the events at the crack tip.

In order to describe atomic-scale reactions at a crack tip we first must provide a picture of the structure of silica glass. The basic building block of most forms of silica is a close-packed tetrahedral unit that consists of a central silicon atom surrounded by four oxygen atoms. Each corner oxygen atom is shared by the silicon atoms of two adjacent tetrahedrons, so that each tetrahedron is connected to four neighbors to form a space-filling network. In crystalline forms of silica such as quartz the tetrahedrons occupy regular positions that are repeated throughout the crystal-



STRENGTH OF GLASS is controlled by the growth of cracks that penetrate the material (*top*). When the glass is stressed, the crack tip (*middle*) penetrates the material. For this reason the authors have sought to describe the growth of crack tips at the atomic level (*bottom*). The basic unit of silica glass is a close-packed tetrahedron consisting of a central silicon atom (*blue*) surrounded by four oxygen atoms (*red*). Each oxygen atom is shared by the silicon atoms of two adjacent tetrahedrons, so that each tetrahedron is connected to four neighbors. The tetrahedrons form a network of interconnected rings, each containing from five to seven tetrahedrons. For clarity the oxygen atoms are drawn smaller than the silicon atoms and not all bonds to tetrahedrons are shown.



WATER can react with glass and cause it to crack more easily. Here a water molecule enters a crack (*left*) and adsorbs onto the tip (*middle*). The molecule causes a concerted chemical reaction (*right*) in which a silicon-oxygen bond at the crack tip and an oxygen-hydrogen bond in the water molecule are both cleaved, producing two silanol groups (hydroxyl groups attached to silicon). In the process the length of the crack increases by one bond rupture. The reaction with water lowers the energy needed to break the silicon-oxygen bonds by a factor of 20, and so the bond-rupture reaction allows cracks to grow faster.

line structure. In silica glass the tetrahedrons form a random network of interconnected rings, each of which usually contains from five to seven tetrahedrons [*see bottom of illustration on opposite page*]. As a crack grows, tetrahedral units are torn apart by the rupture of individual silicon-oxygen bonds. The crack tip reflects the dimensions of a single ring structure that has been pried open on one side, exposing the next silicon-oxygen bond to be broken. The smallest incremental distance the crack can move is the diameter of the silicate ring, which is from .4 to .5 nanometer (billionth of a meter). The precise value depends on the number of tetrahedrons in the ring.

Our results show that the amount of energy required to rupture the silicon-oxygen bond between two silicate tetrahedrons drops steeply—by a factor of nearly 20—in the presence of water. Specifically, if a piece of silica glass is kept in a near-perfect vacuum, the bonds between the silicate tetrahedrons are highly stable. They are so stable, in fact, that 1,300 calories of energy would have to be expended to destroy the silicon-oxygen bonds in one gram of the silica. (In comparison, 75 calories of energy are necessary to heat one gram of water from room temperature to boiling.) When water is present, however, a chemical reaction can take place between a water molecule and a silicon-oxygen bond, making the tetrahedral units much easier to separate.

A detailed picture of the interaction can be divided into three steps. First, a water molecule makes its way along the opening of the crack and adsorbs to the crack tip. Spare electrons from the oxygen atom in the water molecule begin to form a bond with the unoccupied electron orbitals of a silicon atom. Meanwhile one of the hydrogen atoms in the water molecule is attracted to an oxygen atom in the silicon-oxygen chain. In the next step of the process the newly formed bonds strengthen while the original oxygen bond is weakened. Eventually the hydrogen from the water molecule is transferred to the oxygen in the chain and the silicon-oxygen chain ruptures. Finally, the water molecule and the original silicon-oxygen bond split apart and are replaced by two surface silanol groups (hydroxyl groups attached to silicon). The crack has advanced by one atomic step. The entire process is called dissociative chemisorption.

In short, the chemical reaction between the silica and the water reduces the amount of energy that must be supplied to make the crack extend. The highly stable silicon-oxygen bond has been replaced with two nearly equally stable reaction products, the surface silanol groups. Since the energy consumed by a chemical reaction is equal to the difference between the energies of the starting and ending complexes, it can be shown that the rupture of silicon-oxygen bonds by water requires only 78 calories per gram, as opposed to the 1,300 calories per gram required in a vacuum.

The model of dissociative chemisorption for crack-tip bond rupture has enabled us to predict which chemicals will make slow cracks grow in silica. Such chemicals must be able to donate electrons to the formation of a bond with the silicon atom at the crack tip and also to donate a positively charged hydrogen atom to bind with the oxygen atom that was once attached to the silicon atom. In addition a single molecule of the chemical must be small enough to fit into the crack tip in such a way that the breaking and forming of the bonds can take place simultaneously. Ammonia and methanol, for example, satisfy both requirements, and we have found that both chemicals speed up the rate at which cracks grow in silica. In fact, the effect of ammonia, whose molecules are roughly the same size as water molecules, is nearly identical with that of water.

The rate of crack growth depends not only on the chemical environment but also on the magnitude of the applied stress. The development of a complete model for the kinetics of fracture requires an understanding of how stress accelerates the bond-rupture reaction.

In the absence of stress, silica reacts very slowly with water. The dissociative reaction we have discussed causes the surface of silica glass to dissolve in water at a rate of 10^{-17} meter per second; stress-free silicon-oxygen surface bonds are so unreactive toward water that they will not even adsorb water vapor. Yet the application of stress can cause cracks to grow at speeds greater than one millimeter per second.

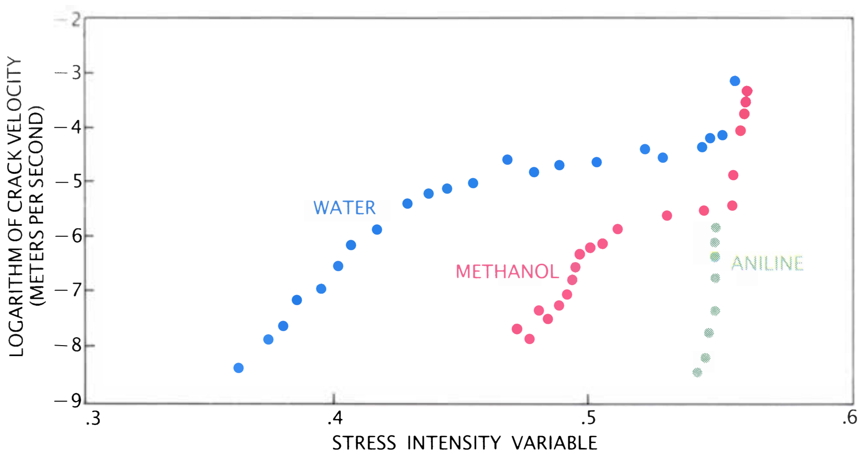
A crack tip focuses stress in much the same way as a needle point focuses an electric field. The closer one approaches the crack tip, the higher the stress becomes, until it reaches several million pounds per square inch within a few atomic dimensions of the tip. Exposed to these large stresses, the atomic structure of silica becomes distorted from its normal bonding configuration. Theoretical calculations show that if a silicate tetrahedron is distorted by pulling on the corner oxygen atoms, the silicon atom in the center is more likely to bond with a water molecule. Furthermore, the chemical interaction with water will decrease the force required to distort the silicon-oxygen bonds further.

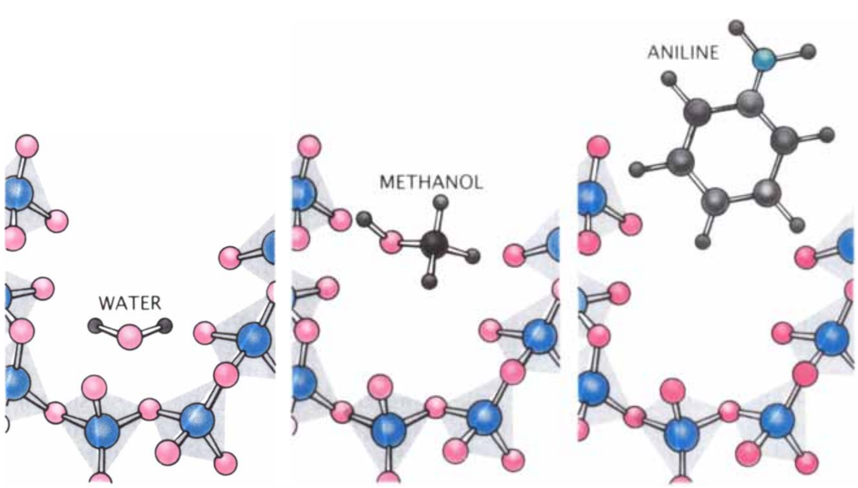
With the results of the theoretical calculations to guide our selection of a model chemical system, we have experimentally explored how stress enhances reactivity. We have had to work with a simple model system because the experimental techniques currently available do not enable us to observe directly chemical reactions along the line of atoms at the crack tip. Our model system consists of small rings of silicon and oxygen atoms (two silicon and two oxygen atoms per ring). Such so-called edge-sharing rings, which result when two tetrahedral units bond together along one edge, can be formed on the surface of silica powders heated to temperatures above 900 degrees Celsius. The edge-shared ring structures are practical for studying the effects of bond strain, because the bond angles and distances between atoms are greatly distorted from the distances and angles found in normal silica glass.

In order to examine the mechanism and kinetics of reactions between water (and other chemicals) and the edge-shared silicate rings, we have employed a technique called Fourier-transform infrared spectroscopy. In infrared spectroscopy a sample is bathed in infrared radiation. The radiation transmitted through the sample is separated into individual frequencies, and the amount of radiation transmitted at each frequency is measured. At certain frequencies there will be a big drop in the amount of radiation received by the detector. The effect arises because each distinct molecular structure in the sample vibrates at a characteristic frequency. When that frequency is matched by the frequency of the infrared radiation, there is a "resonance" effect and most of the radiation is absorbed by the sample, leaving little to reach the detector. By noting both the frequency and the relative amount of absorbed radiation, one can determine what kinds of molecular structures make up the sample and what the relative concentration of each species is.

In Fourier-transform infrared spectroscopy, instead of detecting each individual frequency optical interference techniques are used to scan the entire frequency range rapidly. A single scan contains all frequencies of infrared radiation, and so the time necessary to gather the pertinent data is only a fraction of a second, whereas conventional infrared spectroscopy typically requires half an hour or so to produce a complete spectrum. The individual frequencies are then separated by the mathematical technique of Fourier analysis. The advantage gained by Fourier-transform infrared spectroscopy is obvious: when one is studying rapid chemical reactions, one needs swift and accurate measurements of how the amounts of reactants, reaction intermediates and products are changing over time.

Fourier-transform infrared spectroscopy enables us to follow the kinetics of reactions among chemicals that play an important part in the fracture of glass. We can distinguish among the edge-shared rings (the reactants), the water molecules adsorbed to the rings (the reaction intermediates) and the silicon hydroxide groups (the products). Our work has led to three interesting findings. First, a silicon atom in a strained ring is much more likely to accept electrons than a silicon atom in stress-free glass. Consequently the silicon atom in the strained ring is more likely to adsorb electron-donating molecules such as water, ammonia and methanol. Second, all the chemical





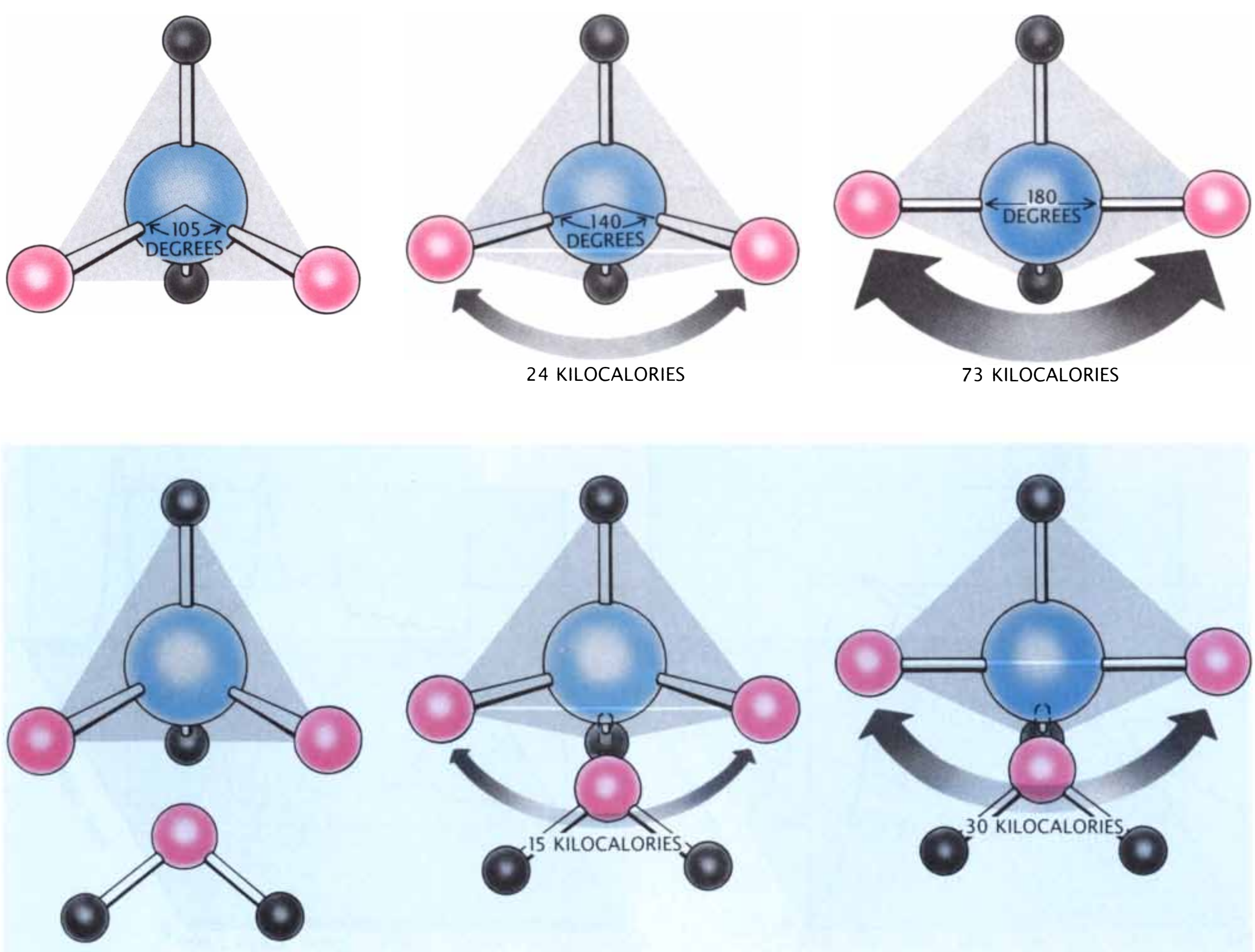
MOLECULAR SIZE of a chemical substance affects its ability to speed the growth of cracks in glass. Water, which has a molecular size of only .26 nanometer (billionth of a meter), causes cracks to grow much faster than methanol (.36 nanometer) does, and aniline (.42 nanometer) has hardly any effect (*top*). In essence, water can readily enter a crack opening (which has a diameter of .4 to .5 nanometer), methanol has difficulty getting in and aniline is so big that it never arrives at the sites of bond rupture (*bottom*).

species that make cracks grow more rapidly in silica glass also dissociatively chemisorb on the strained ring: they cleave one of the distorted silicon-oxygen bonds in the ring. Conversely, chemicals that do not react with the edge-shared rings have no effect on crack growth in silica.

Finally, the rate of ring cleavage when we expose our simple model system to water vapor is more than 100,000 times higher than the rate at which a flat, unstressed surface of silica will react with water. The results provide direct experimental evidence that silicon-oxygen bonds can be broken rapidly by a stress-enhanced reaction with water and other chemicals.

We next sought to determine if the dissociative-chemisorption reaction between water molecules and strained silicon-oxygen bonds is the factor that limits the rate of crack growth. We made direct comparisons between the kinetics of the ring-cleavage reactions in our model compound and the speed of crack growth measured under identical environmental conditions. We found that ammonia is slightly more aggressive than water in rupturing strained silicon-oxygen bonds and also causes cracks in silica glass to grow somewhat faster than water does. Methanol is also more aggressive than water in rupturing strained silicon-oxygen bonds, but it is five orders of magnitude less effective in increasing the rate of crack growth than either water or ammonia. The comparison between reactions of our model compound in methanol and crack-growth measurements clearly indicates that the kinetics of fracture cannot be predicted solely on the basis of the kinetics of the underlying bond-rupture reactions.

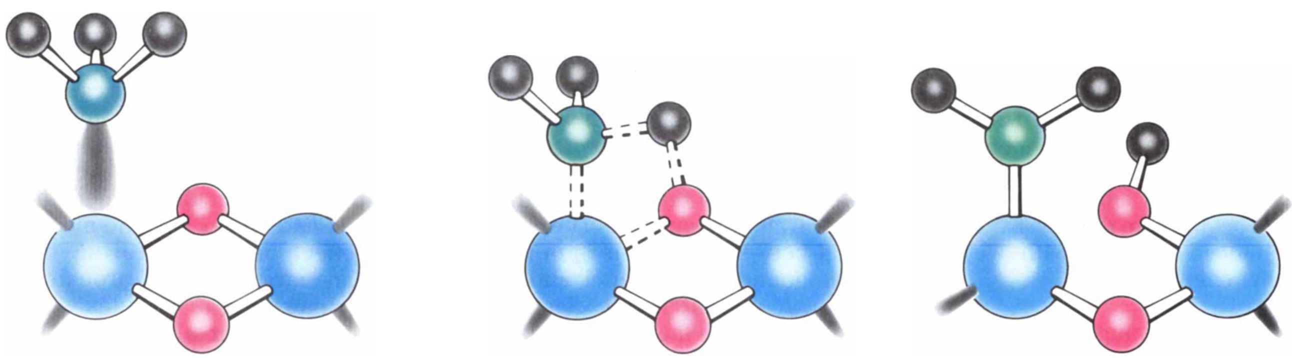
What accounts for the unusual behavior of methanol? As we mentioned above, the size of the attacking molecule also determines the ability of the molecule to promote crack growth. Water and ammonia have nearly identical molecular sizes (.26 nanometer), but methanol is a much larger molecule (.36 nanometer). Smaller molecules such as water or ammonia can readily enter a crack opening (which has a diameter of from .4 to .5 nanometer) and cause reactions that rupture bonds, but larger molecules such as methanol have difficulty getting in. Indeed, we have found that molecules larger than about .4 nanometer have no measurable effect on crack growth; they have an extremely small probability of ever making it to the sites where bond-rupture reactions take place. The size of the reactive molecule can influence crack speeds even when the molecule is smaller than the opening of the crack tip. In effect, the



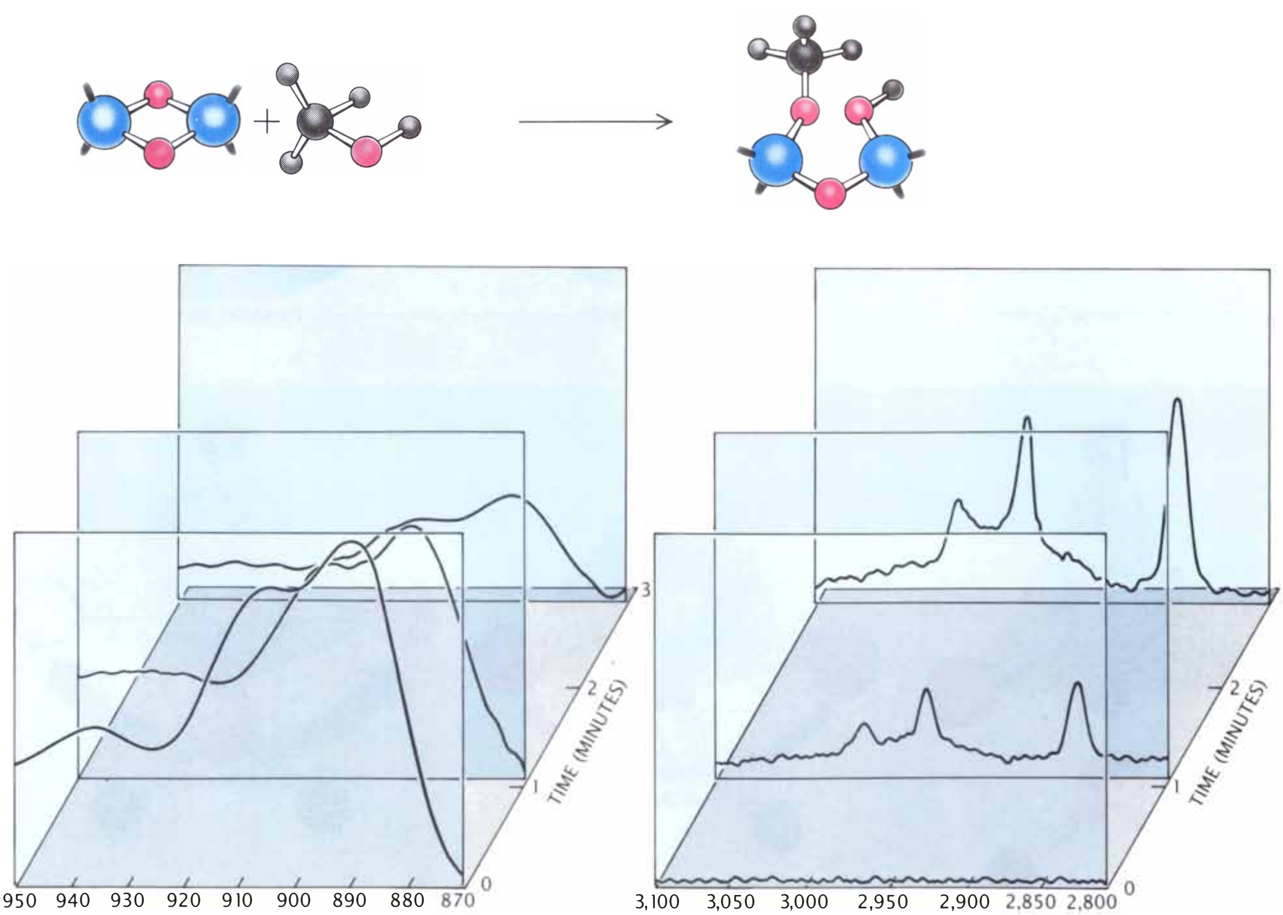
THEORETICAL CALCULATIONS show that water lowers the energy necessary to stress, or distort, the basic tetrahedral unit of silica glass. In the absence of water a total of 73 kilocalories of energy must be expended to bend the angle between two oxygen atoms from 105 to 180 degrees (*top*). In the presence of water the energy is only 30 kilocalories (*bottom*). To reduce the number of electrons that enter into the calculations, two of the oxygen atoms in the tetrahedron were replaced with hydrogen atoms.

region just behind the crack tip acts as a sieve that admits molecules to the tip at a rate determined by the size of the molecule.

A similar effect is observed in zeolite molecular sieves. Zeolites are aluminosilicate crystals containing well-defined cage structures connected by openings of known size. Zeolites are employed to separate different chemicals by their size; the larger molecules are excluded, whereas the smaller ones pass through the openings and adsorb within the cage structures. Using measured molecular diffusion rates in zeolites as our model for the crack-tip opening, we have shown that methanol should diffuse into a crack tip at a rate four orders of magnitude less than the rate of water. By taking into account



EXPERIMENTS confirm the importance of stress in promoting crack growth. The authors worked with small rings consisting of two oxygen and two silicon atoms. Such so-called edge-sharing rings are practical for studying the effects of bond strain, because the bond angles and distances between atoms are greatly distorted from those of normal silica. Here the reaction between an ammonia molecule and an edge-shared ring is shown. Ammonia is adsorbed on a strained silicon site; a dissociative-chemisorption reaction breaks the silicon-oxygen-silicon bonds. Such strained bonds react 100,000 times faster than unstrained ones.



FOURIER-TRANSFORM INFRARED SPECTROSCOPY is utilized to measure the rates at which various chemicals react with edge-shared silicate rings. Shown here is an example of a reaction involving methanol. The rates at which the rings (*left*) disappear and the chemisorbed methanol (*right*) appear are monitored by bathing the samples in infrared radiation. When the frequency of the radiation matches a characteristic vibration of each molecule, a large absorption is recorded (*peaks in curves*).

the rate of molecular diffusion near the crack tip and the rate of dissociative chemical reactions on the strained silicon-oxygen bond, we can now predict to within an order of magnitude the relative rates of crack growth in silica glass exposed to different chemicals.

The results of our atomistic studies of fracture suggest several novel and interesting possibilities for predicting and controlling the strength of glass and other brittle materials. Our discovery of the importance of molecular diffusion near the crack tip indicates that surface coatings might be designed to block the opening of the crack and restrict the passage of small molecules, such as water, that can attack the bonds at the tip and chemically weaken glass. We are currently exploring this possibility by attaching large molecules to the crack walls and measuring their effect on the rate of subsequent crack growth when water is present. We find that we can decrease the crack-growth rate 1,000 times when we first apply such a molecular coating to the crack. If the coating can be applied commercially, it should greatly increase the structural lifetime of glass products.

The same thing should be true for ceramics. Our work on crack growth in other solids leads us to believe that the general conclusions developed for silica can explain the strength behavior of a wide range of brittle materials. The actual crack-tip reactions appear to vary from material to material, however, and the chemistry of each solid must be considered on a case-by-case basis. In complex silicate glasses, for example, the simple bond-rupture mechanism for crack extension can be complicated by the formation of reaction layers that are several micrometers (millionths of a meter) thick. Although the reaction layers do not affect the bond-rupture process directly, they can significantly alter the amount of stress concentrated at the crack tip.

Yet it is important to keep in mind that even though the interplay between surface chemical reactions and mechanical strength can be more complex than the interactions we have described in this article, a thorough understanding of how glass fractures does lay the groundwork for predicting how and when other materials will fracture. It may even provide clues for predicting that most overpowering of fracture events: an earthquake.

H.M.S. *Warrior*

Recently restored after suffering more than a century of neglect, the vessel was the first iron-hulled, armored warship. It deterred the enemy as it sealed its fate by engendering a naval arms race

by Walter Brownlee

Vice Admiral Lord Nelson ensured lasting fame for his flagship *Victory* by commanding from its deck the British Royal Navy in its triumph over the combined French and Spanish fleets at the Battle of Trafalgar in 1805. The warship is now preserved in Portsmouth, England, as a monument to Britain's naval prowess and to the memory of Nelson himself, who died in the battle. Recently another British warship has been restored in the largest and most expensive such project to date to take a permanent berth in Portsmouth near H.M.S. *Victory*. The ship has no battle scars and its decks were not stained with the blood of war heroes. Indeed, it never fired a single shot in anger. Yet it deserves a figurative as well as a literal place of honor beside *Victory*.

The ship is named *Warrior*, and it is the only surviving large Victorian warship. But this is not what makes the ship particularly noteworthy in naval history. *Warrior's* historical impact resulted from its creation: merely by being built it caused the entire pattern of warship design to change throughout the world. For H.M.S. *Warrior* was the first iron-hulled, armored battleship.

Warrior was so superior to any other warship at the time of its construction that its supremacy never had to be challenged in battle. Hence the ship never had the opportunity to distinguish itself in any military confrontation, and no valiant deeds were ever demanded of its crew. As a consequence it was quickly forgotten by the general public a decade after it was launched, forced into a premature retirement as other naval powers adapted and perfected many of the features *Warrior* had introduced.

Ironically, the full implications of *Warrior's* role can be appreciated better today, in a world that has had 40 years' experience with nuclear weapons, than it was during Victorian times. For *Warrior* was a deterrent. The warship emerged in 1860 as an awesome machine of war that rendered obsolete all previous warships and thereby successfully deterred a threatening enemy. Yet like all technology-based deterrents, it could not be uninvented and its emulation by the enemy could not be prevented. It thereby also presented the Victorian world with another concept unfamiliar to their time but familiar in ours: an arms race.

Although various countries fought during the 18th and 19th centuries to establish and protect trade routes across the oceans, before 1840 all essentially played by the same rules: warships were constructed of timber and propelled by wind. Wood was a common, natural and self-replenishing building material. Thousands of years of experience with it had resulted in an industry capable of producing reliable vessels that could sail anywhere in all but the most adverse weather conditions. Provided a landing party could buy, beg, borrow or steal a suitable piece of timber, the repair or alteration of any part of a warship's structure could be accomplished by teams of skilled carpenters, who always formed part of a ship's crew. Wind, of course, was also available worldwide and completely free. Tall masts (typically three) carried a highly complex system of sails, yards and tackles that could be expertly manipulated to propel a warship wherever its captain chose. Only in the case of extreme weather conditions, such as a total absence of wind or a raging hurricane, did this form of propulsion fail old-time sailors.

The basic function of such warships was to deliver enough firepower against an enemy target to destroy it, while ensuring a reasonable degree of safety for the ship's crew. The ships were in essence floating wood castles that protected the crew from enemy gunfire and at the same time provided a firm and steady platform from which cannons could be discharged. The need for massive and stable structures on a fighting ship meant that the shape of its hull had to be rather bulky, which in turn made it ponderous and slow.

A warship's cannons, named for the weight of the shot they fired, came in a variety of sizes. The largest common cannon was a 32-pounder and weighed about 5,000 pounds; smaller sizes were 24-pounders, 18-pounders and 12-pounders. Because the timbers of a wood warship could not withstand the increased weight and the wrenching recoil of guns larger than a 32-pounder, an increase in firepower could be achieved only by distributing as many 32-pounders and smaller cannons as possible on as many decks as possible throughout the ship—the heavier guns on the lower decks and the lighter ones on the upper decks.

To maximize the total destructive firepower of a large number of guns, they were all aimed at the same target and discharged at about the same time—a technique called a broadside. By the 1850's a ship of the line (a large warship capable of holding a forward position in a major battle) carried a total of 130 guns on three decks and could hurl a broadside weight of as much as 1,800 pounds of shot every minute.

Britons in the first half of the 19th century relaxed with pride and satisfaction behind their armed "wood walls." Potential enemies could build similar ships, but Britain was way ahead in numbers and types of warships and had a large pool of trained seamen to man them. Britannia ruled the waves not because she possessed a more advanced technology but by virtue of the quantity of her warships and the quality of her "Jack-tars."

Because the balance of sea power was so clearly in Britain's favor, any innovation that threatened the status quo worried the British Admiralty. As a result the Admiralty kept a close watch on naval developments of potential enemies and then made the minimum response to counteract the move. Yet the admirals could not control the advances of British commercial shipbuilders.

The iron-smelting and steam-engine technologies developed during the Industrial Revolution in Britain made it possible to improve on steamship designs such as the ones originally realized by Robert Fulton and Henry Bell. In 1837 Isambard Kingdom Brunel introduced the first reliable transatlantic paddle steamer, *Great Western*, and eight years later he astounded the world with *Great Britain*, which had an iron hull and a screw propeller.

To the British merchant navy the advantages of iron-hulled, propeller-driven ships were clear: they offered increased speed, reliability and durability. In fact, by the mid-19th century such commercial ships were a common sight. Nonetheless, the Admiralty refused to follow the trend set by the merchant navy beyond commissioning a few small experimental vessels of this type.

At the same time France under Napoleon III had already begun to challenge the Royal Navy's control of the seas, and it was more receptive to the military application of the new naval technology. In 1850 France built *Napoléon*, a large three-decked wood warship fitted with an engine and a propeller. The British Admiralty, which had long regarded France as its "natural enemy," was compelled to react. It responded by converting *Sans Pareil* to propeller power. By 1858 each of the two nations had amassed a total of 32 propeller-driven wood ships of the line.

These developments underscored the limitations of traditional wood hulls. The massive three-deckers sagged under the additional weight of an engine and boilers. Indeed, because the hulls were nearly at the breaking point, a "rippling" broadside technique had to be introduced in which the fusillade typically began at the bow and worked its way down the length of the ship to the stern. A captain who thoughtlessly ordered a

RESTORED *WARRIOR* was towed in June of this year to its permanent berth in Portsmouth, England, where it is open to the public. Its restoration required more than seven years and £6 million. The vessel is the only surviving large Victorian warship. Its construction was a major technological leap in the history of warships: the transition from wood hulls to armored, iron hulls.

traditional broadside from an engine-powered wood warship was likely to see his vessel in drydock for six months while cracked and torn beams and frames were replaced. Nonetheless, since Britain still had the lead in total number of warships, the Admiralty saw no need to rethink the situation.

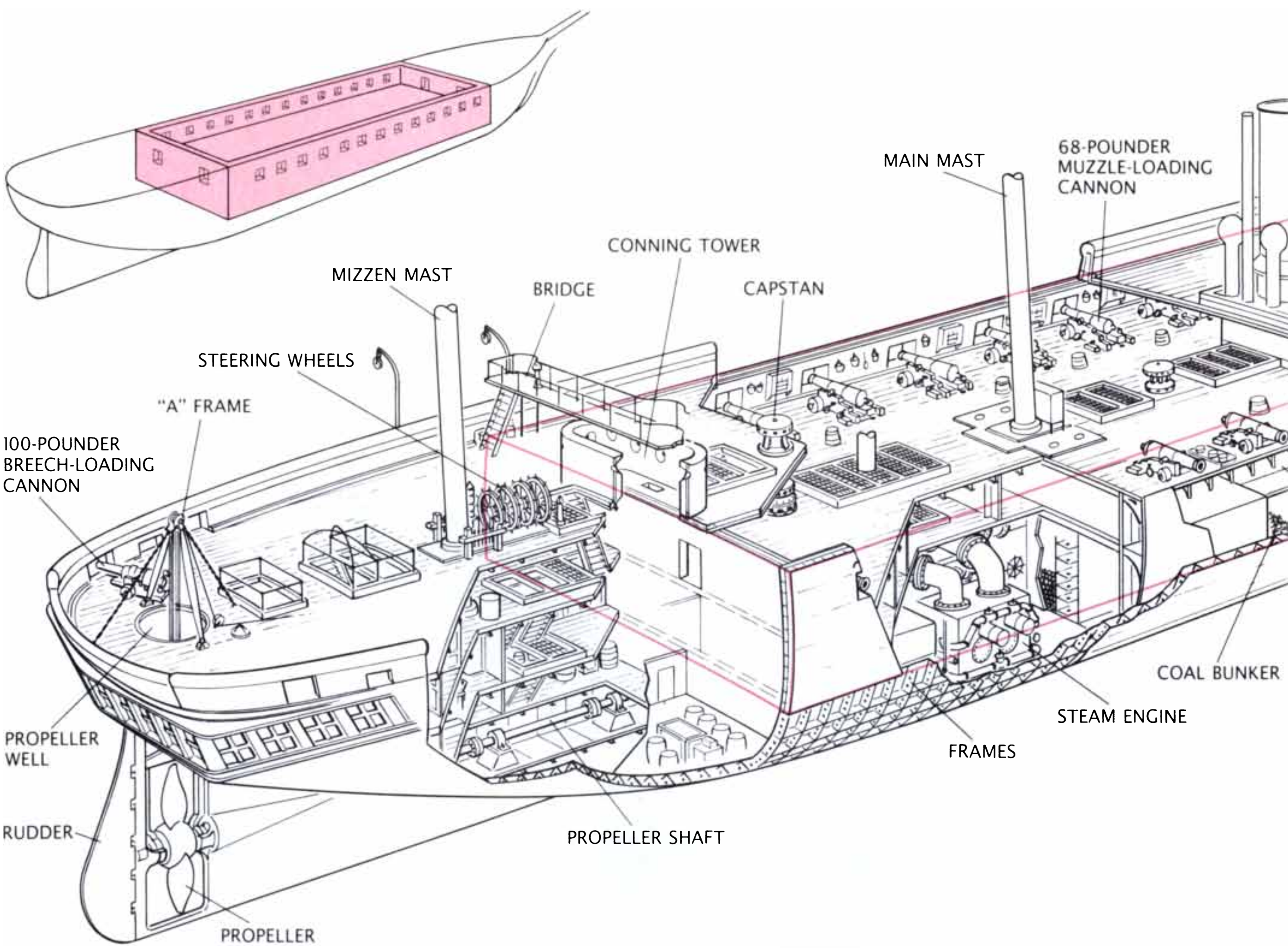
In France, however, the tireless and brilliant ship designer Stanislas Dupey de Lôme had been appointed director of matériel, and he immediately called a halt to the building of wood warships. Dupey de Lôme was aware of the fact that in the Crimean War floating batteries had been armored with iron plates and that the batteries had been able to withstand heavy pounding by shore guns. He recognized that over the years the quality of wrought iron had been greatly improved, a point overlooked by the British Admiralty. His aim was to establish a fleet of iron-hulled, iron-armored and engine-powered warships for the French navy. Yet France's economy was still basically agricultural, and its foundries could not supply enough iron to realize his ambitious plans.

Dupey de Lôme had to settle for a fleet of wood-hulled ships that were merely plated with iron. The first such ship was *Gloire*. The new ship had a wood hull 26 inches thick to which $4^1/_2$-inch wrought-iron plates were bolted; it was primarily a sailing vessel, but it had been equipped with an engine as well. Such a ship would be capable of pounding an entire unarmored wood fleet to pieces, at the same time remaining impervious to returned gunfire.

Work on *Gloire* was begun in March, 1858, and within a few weeks reports of its construction had reached Britain. The news caused an uproar in Parliament. Public reaction was also swift and alarmist, and war rumors spread quickly. The Surveyor to the Admiralty stated on June 28:

"It is not in the interest of Great Britain...to adopt any important change in the construction of a new class of very costly vessels, until such a course is forced upon her by the adoption of a foreign power of formidable ships of a novel character requiring similar ships to cope with them.... This time has now arrived. France has commenced to build frigates of great speed with their sides protected by thick metal plates and this renders it imperative for this country to do the same without a moment's delay."

It was generally assumed that Brit-



CUTAWAY VIEW of *Warrior* reveals holdovers from the preceding century as well as features that were ahead of their time. The ship was outfitted with a figurehead and an elegant stern gallery for its officers. It was also designed to fire traditional broadsides: the simultaneous discharge of cannons lining a warship's sides. Although it was equipped with a steam engine, it relied on its square-rigged sails (not shown) as the main source of propulsion and on the muscles of its crew to weigh anchor, lift equipment and load provisions. On the other hand, the keel, frames, beams and cladding of the hull were all made of wrought iron and its shape was more like a speedy "clipper" than a standard, ponderous ship of the line (a front-line warship). *Warrior* also

ain would simply respond by building a series of similar ships, but First Sea Lord Sir John Packington argued for a different plan. Backed by Chief Constructor Isaac Watts and by many merchant-ship designers and builders, Packington convinced a reluctant Admiralty and Parliament that Britain should take an unprecedented step: the creation of a completely new warship so advanced and powerful that France would have no option other than to retire from the developing naval contest.

The basic design objective was clear and simple, although daunting to realize. The new ship was to be the largest, fastest, most powerfully armed and most heavily armored warship that had ever been built. In spite of the misgivings of many, preliminary plans were quickly drawn up and tenders were called for on April 29, 1859. Because the Royal Navy had little experience in building ships of iron, the construction would have to be in the hands of a merchant shipyard. (The armament, however, was to remain under Admiralty control.) The building contract was won by the Thames Iron Works of Blackwall and the keel of the ship that would be known as *Warrior* was laid on May 25 of the same year.

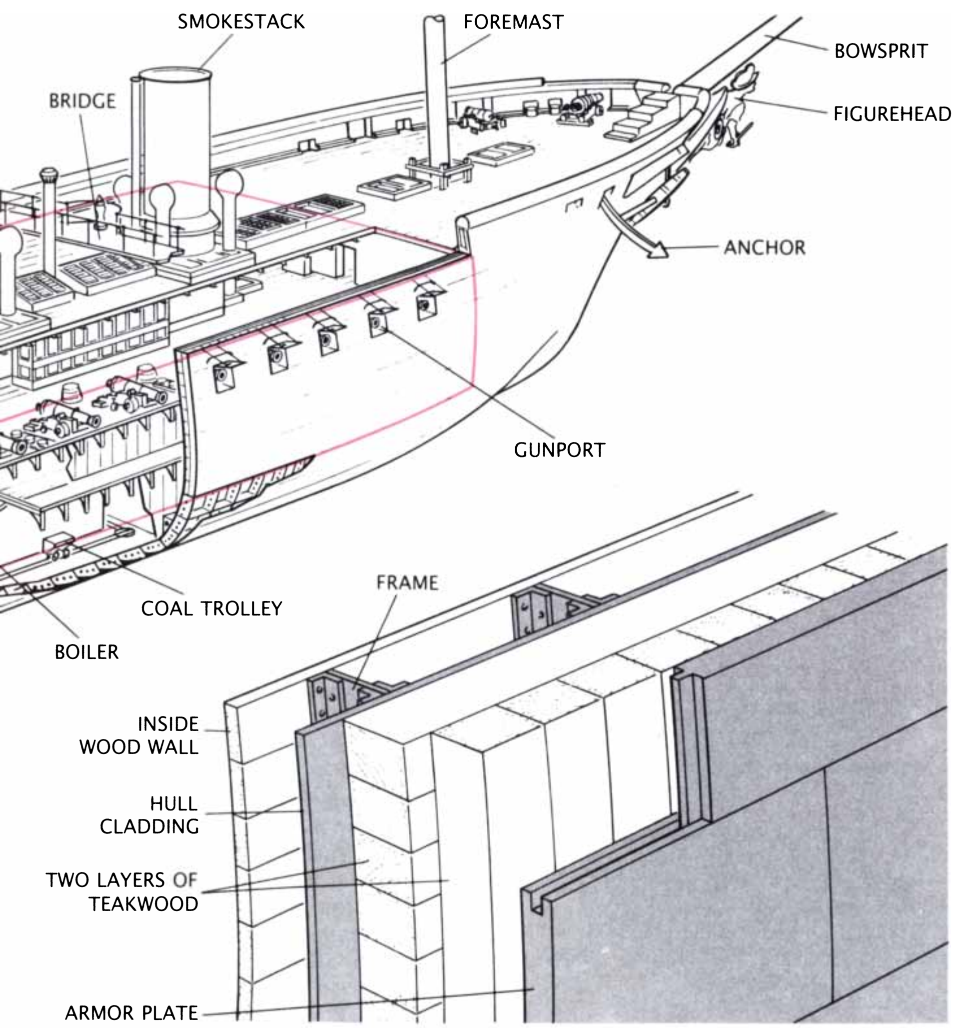
In spite of Packington's forward-looking vision, some aspects of the ship still reflected the practices of the preceding century. *Warrior,* for example, was quaintly outfitted with a huge figurehead and an intricately carved stern gallery. A more important 18th-century holdover was the fact that the ship was designed principally as a stable gun platform, its guns lining the ship's sides in keeping with the old broadside tradition. Moreover, although the ship had a steam engine that powered a propeller, its main form of propulsion was to be the wind. Indeed, its masts and sails were almost identical with the rigging of an 80-gun ship of the line.

The inclusion of two propulsion modes necessitated some special features. In order to provide enough room for the lower square-rigged sails to unfurl, the smokestacks were made collapsible. To lessen the drag caused by the motionless propeller while the ship was under sail alone, the propeller was fixed in a cradle so that it could be disengaged from the propeller shaft and hoisted into the hull [*see top illustration on page 135*]. The propeller had to be lifted by a team of more than 600 crewmen, just as more than 100 men were needed to weigh the ship's anchors (which were the heaviest of the time). The Royal Navy was so convinced of the benefit to be derived from putting its Jack-tars through hardship that it refused to include any steam-operated mechanical aids on the new ship.

On the other hand, many aspects of *Warrior* were clearly ahead of its time. The keel, frames, beams and cladding were all made of wrought iron. Iron bulkheads fitted with watertight doors divided the ship into sealable compartments and also provided additional structural strength. Above the keel a second watertight layer of wrought iron protected the ship against damage from grounding. (*Warrior* was the first Royal Navy warship to have a double bottom.)

The shape of *Warrior*'s hull followed the principles put forward by John Scott Russell in his wave-line theory, which was basically a set of mathematical relations among the length, breadth, depth and cross section of ships with slim, wave-cutting bows and streamlined sterns. The ship was about 425 feet from stern to figurehead and its beam (width) was 58 feet. Its draft (depth below the waterline) was 27 feet and it was designed to displace well over 9,000 tons of water. Today the ship would be regarded as having typical "clipper" lines. With a length-to-breadth ratio (at the waterline) of 6.5, *Warrior* was as speedy as it looked: it could do in excess of 13 knots when powered by sails alone, more than 14 knots when powered by steam alone and an amazing 17.5 knots when powered by both. *Warrior* was capable of overtaking any warship then afloat.

Warrior's armored section, known



had a boxlike armored section (*color, upper left*), known as the citadel, in which most of the ship's heavy cannons were placed. The citadel was sealed off from the rest of the ship by watertight iron bulkheads, so that if the unarmored stern and bow were shattered by enemy gunfire, it could still remain afloat. A more detailed view of the citadel's sides (*lower right*) shows that they consisted of wrought-iron plates $4\frac{1}{2}$ inches thick bolted through 18 inches of teakwood onto the one-inch iron cladding of the hull.

as the citadel, took up the middle 230 feet of the ship and was designed as an integral part of its structure. The citadel was essentially a box whose sides were made up of plates of wrought iron 4¹/₂ inches thick. These plates were bolted through 18 inches of teakwood onto the one-inch iron plates of the hull. No shot from any standard gun of 1860 could pierce the armor. Viewed from the side, it was impossible to distinguish the unarmored ends of the ship from the armored citadel section. The ship was designed so that if the unarmored bow and stern sections were shattered by enemy gunfire, the citadel would remain intact and afloat.

The fore and aft walls of the citadel, made of four-inch wrought-iron plates backed by 12 inches of teakwood, represented another first. The gun decks on warships built before *Warrior* did not have armored walls that spanned the entire width of the ship, leaving sailors manning the guns extremely vulnerable to a raking shot (a shot that enters the bow or stern and traverses the full length of the gun deck).

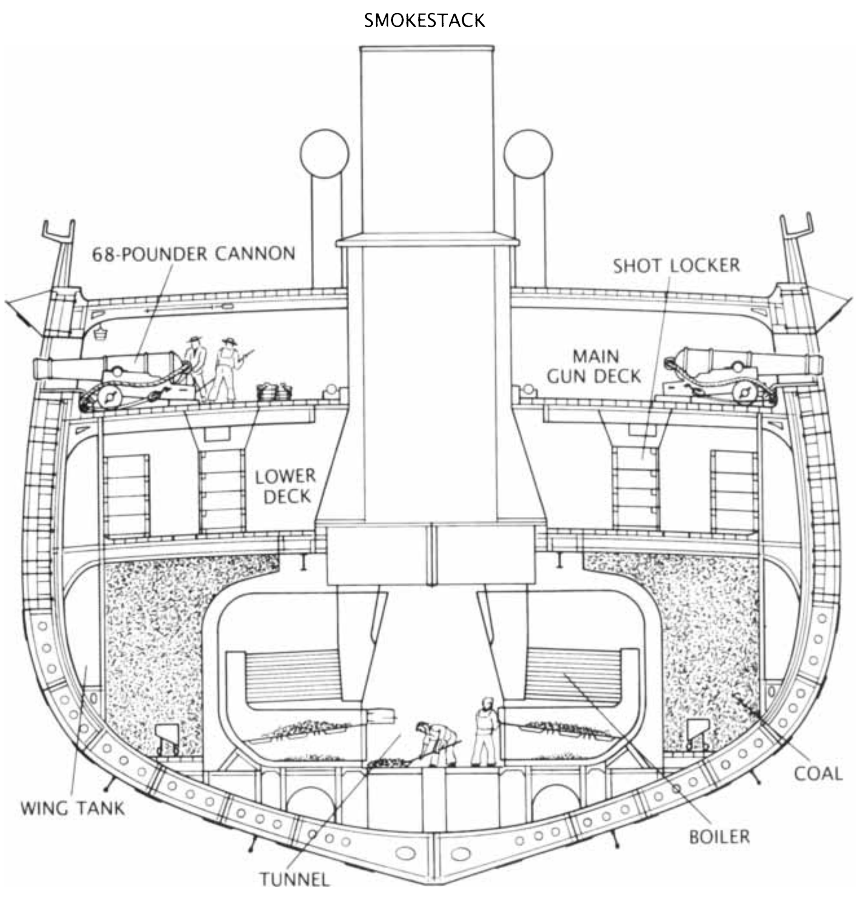
On the top deck a small enclosure made of iron-plated teak—a rudimentary conning tower—also gave officers some protection from small-arms fire (which had killed Nelson) while they oversaw the course of a naval engagement. Small openings in the enclosure allowed them to view the decks and sails, and a large hatch at their feet made it possible to communicate with the helmsmen on the deck below. Although such an arrangement represented a first step toward centralized bridge control, the old-time officers of the Royal Navy objected to it. Perhaps inspired by Nelson's example, they thought it was more appropriate to pace the top deck in full regalia while the battle raged about them.

The engine and the boilers were placed in the belly of the ship, under the armored citadel. The engine, manufactured by John Penn and Sons of Greenwich, had two horizontal cylinders, each with a bore of nine feet four inches, a stroke of four feet and a piston speed of about five miles per hour. Its drive shaft revolved a maximum of 55 times per minute and its nominal horsepower was 1,250. Steam at a maximum pressure of 22 pounds per square inch was fed into the engine from 10 coal-fired boilers, each of which held about 19 tons of water. Some 850 tons of the best Welsh coal filled the ship's bunkers. The iron hull of *Warrior* easily bore the weight of the engine, boilers, coal and drive shaft.

The great load-bearing strength of wrought iron also eliminated the need to spread a large number of small guns throughout the ship. Instead *Warrior* carried fewer but bigger guns, most of them on one deck in the citadel. Its main battery consisted of 26 smooth-bore, muzzle-loading cannons that fired solid balls or shells weighing 68 pounds. One of these 68-pounders, along with its wood undercarriage, weighed more than five tons. The cannon had a theoretical range of three miles, although normal firing range was about 1,000 yards. Ordinary wood warships could carry at most one 68-pounder, and it had to be fired with caution since its weight and recoil could easily strain the hull. *Warrior's* hull soundly supported all 26 cannons, and they could even be discharged in a simultaneous broadside without undue strain.

Warrior was also outfitted with the latest in gun technology: a complement of breech-loading cannons with rifled bores. The cannon's rifling (a spiral groove cast into the bore) caused projectiles to leave the barrel spinning, thereby endowing them with a stable trajectory. Because the projectile and powder charge was inserted through the breech, or rear section, of the cannon, the cannon did not have to be withdrawn from its port for loading. The cannon was designed by Sir William Armstrong and fired conical shot or shells, each of which weighed 100 or more pounds, over a maximum range of five miles. Ten such weapons were placed on the ship soon after its launching. In addition *Warrior* had four smaller rifled-bore, breech-loading guns, giving it an official armament of 40 guns. Not included in the official rating were two 25-pounders, a 12-pounder and a six-pounder.

Any one of *Warrior's* larger cannons would have pierced the 24 inch-



CROSS SECTION AMIDSHIPS reveals the layout of *Warrior's* boilers, coal bunkers and main gun deck in the citadel as well as the structure of its double bottom. *Warrior* was the first Royal Navy vessel to have such protection against damage from grounding.

es of timber armor on a wood warship. From the day *Warrior* took to the seas on December 29, 1860, all other warships of the period were given the nickname "egg shells." It was believed *Warrior* could tackle entire fleets of traditional warships with impunity.

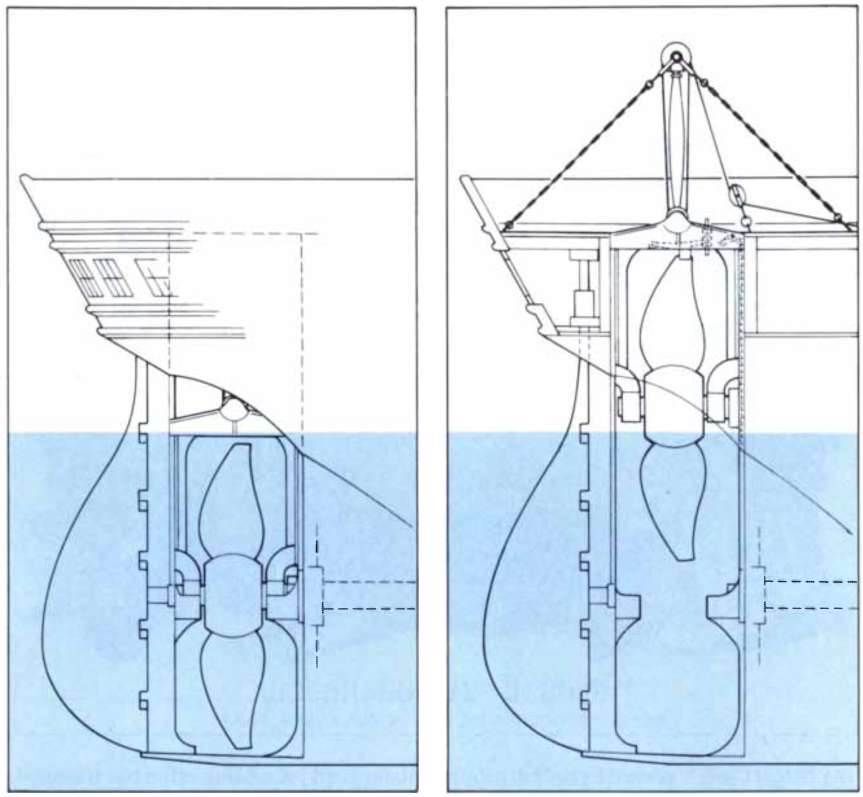
Yet *Warrior* was not immune to criticism at home. Some observers dismissed the ship even before its construction began, because they could not believe an iron warship of more than 9,000 tons' displacement could float. Others attacked the ship simply because they hated its appearance, so different from the colorful high-sided ships to which they were accustomed. *Warrior* was long and thin, was painted a monotonous black and floated low in the water; it soon was dubbed "Black Snake."

The ship's armament also caused some bewilderment. The more informed public was aware that a standard 130-gun ship of the line could hurl a broadside weight of well over 1,600 pounds. Although *Warrior* had the largest guns then available, it hurled a broadside weight of only 1,480 pounds. To satisfy the inquiries of the public, the British government had to explain that a single 68-pound shot had the destructive impact of five 32-pound shots; similarly, a 100-pound shot was equivalent to seven 32-pound shots. By this measure *Warrior* could deliver every minute the equivalent of 3,500 pounds of shot from ordinary guns.

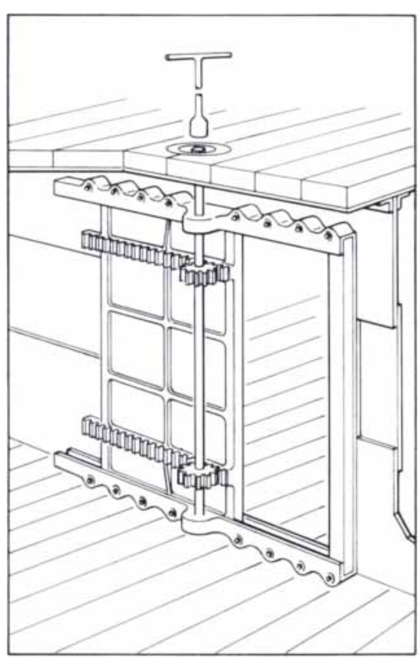
Warrior not only had (in effect) more firepower than two standard wood ships of the line but also cost more than the two ships combined. Whereas the grandest of the wood three-deckers were built for the then staggering sum of £176,000, *Warrior* cost £379,154. In Parliament and in the press there were heated debates over whether such an enormous sum of the taxpayers' money was in fact well spent.

In spite of the criticism, when *Warrior* took up its station in the English Channel, most Britons felt they could breathe a little easier. Napoleon III seemed to have given up ideas of challenging British sea power and turned his attention to conquests in continental Europe. *Warrior* had done the trick: it had deterred a potential enemy.

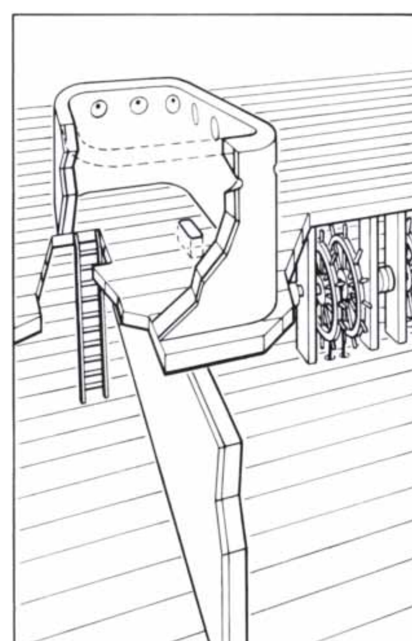
Nonetheless, many within the Admiralty viewed the developing situation with a sense of despair. They recognized that in one fell swoop the Royal Navy's massive wood fleet had



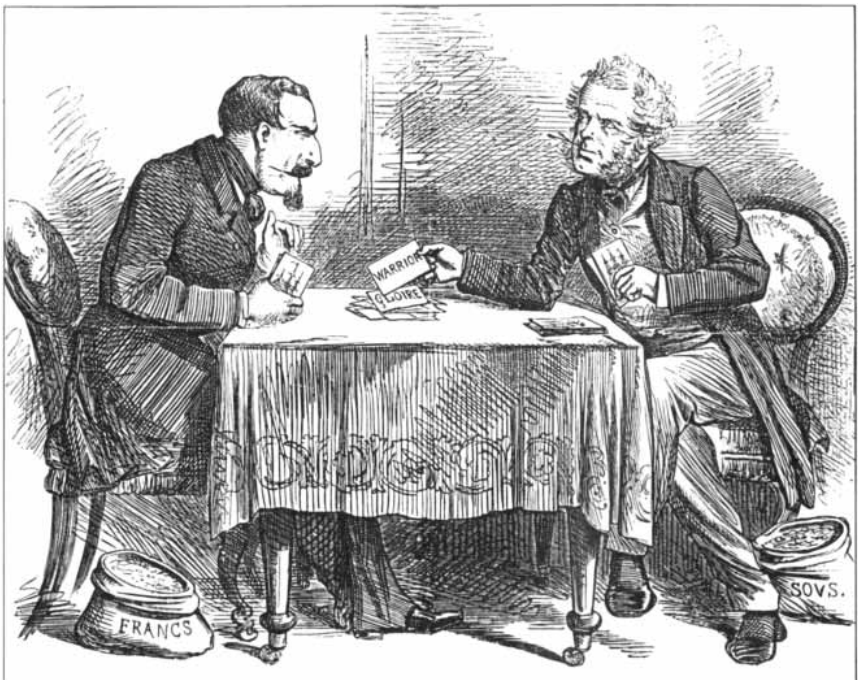
CRADLE (*left*) in which *Warrior's* propeller spun made it possible to disengage the propeller from its drive shaft and retract it into the hull of the ship (*right*) in order to lessen drag when the ship was under sail. More than 600 crewmen had to haul on the lines of a block-and-tackle system on an A-frame to hoist the device, which weighed 35 tons.



WATERTIGHT DOORS were fitted into *Warrior's* iron bulkheads. It is not clear whether the doors were part of the original design, but they certainly were added early in the ship's lifetime. Every door could be opened or closed from the overlying deck by means of a T-shaped bar.



ARMORED CONNING TOWER on the top deck afforded *Warrior's* officers protection from small-arms fire as they oversaw the course of a naval battle through loopholes. A large hatch at their feet allowed them to shout orders to the helmsmen in the citadel on the underlying deck.



"BEGGAR MY NEIGHBOUR."

PAM. "IS NOT YOUR MAJESTY TIRED OF THIS FOOLISH GAME?"

"*WARRIOR* CARD" played by British Prime Minister Lord John Palmerston has trumped the *Gloire* card laid down by Napoleon III of France in a *Punch* cartoon that appeared in 1861. *Gloire* was a wood-hulled warship that had been plated with iron and equipped with steam power. It was regarded as the most capable warship until the launching of *Warrior,* which outdid *Gloire* by far in terms of speed, firepower and armoring. Faced with the prospect of having to engage ships such as *Warrior,* Napoleon chose to avoid entering a naval competition with Britain and turned his attention to the Continent.

been made obsolete. Bolstered by the knowledge that it could be done, other nations would soon build iron-hulled and -armored warships. Moreover, if a nation lacked the wherewithal to build such ships, it could always buy them from Britain, whose commercial shipyards were willing to build them for any nation (except France, of course). If Britain was to maintain control of the oceans, it would have to rebuild the entire Royal Navy in iron.

Warrior single-handedly threw the warship-building industry into turmoil and in so doing paved the way for its own demise. Nations throughout the world revised their military-construction programs to include iron-hulled warships, many of which incorporated improvements on the *Warrior* design. The most obvious improvement was the installation of more powerful guns whose ammunition could pierce *Warrior's* armor. As a countermeasure the armor of subsequent warships was made thicker, but the countermeasure was implemented in vain: the next generation of warships were simply equipped with even larger guns. The placid and conservative Victorians were troubled by these seesawing advances in guns and armor, which today we would easily recognize as symptomatic of an arms race.

Once the pride of Britain, *Warrior* faded from the public eye after a few years. Within five years the ship that had been visited by most of the crowned heads of Europe was virtually forgotten. *Warrior* was so outdated by 1871 that it was no longer classed as a first-line battleship and was relegated to the reserve fleet.

Its subsequent fate was even more ignominious. The ship was offered for sale as scrap metal at the turn of the century, but there were no takers. In 1904 it became part of the Vernon Torpedo Training School at Portsmouth and was renamed *Vernon III.* After another unsuccessful bid to sell the ship as scrap, in 1924 it was turned into a floating jetty at the Pembroke Dock Oil Fuel Depot in South Wales and renamed *Oil Fuel Hulk C77.* Its masts, engine and most of its internal equipment had by then vanished, and a six-inch layer of concrete was spread over its top deck.

By 1979 *Warrior* was the only surviving British capital warship between Nelson's *Victory* and *Belfast,* which took part in the Korean war, yet only a few naval enthusiasts even knew of its existence. *Warrior's* deteriorating condition, however, was noted by H.R.H. The Duke of Edinburgh, Prince Philip, and brought to the attention of the Maritime Trust. In September, 1979, the Royal Navy handed over *Oil Fuel Hulk C77* to the trust. The hulk was then towed to Hartlepool, England, where it was to be restored.

Surprisingly, its wrought-iron hull had enabled it to remain afloat for more than a century without letting a drop of seawater inside! Nevertheless, it soon became clear that the task of restoration would be larger than any previous such project. The *Warrior* Preservation Fund was set up to raise the necessary money, and a restoration team was organized among the local shipyard workers. By 1984 more than 150 people were working on the project.

On June 12 of this year the restored ship left Hartlepool and headed for a permanent home in Portsmouth, where it was opened to the public on July 27. The conservation of the hull and interior and the re-creation of the 19th-century equipment and fittings took more than seven years and an estimated £6 million.

With the masts and rigging once more towering 180 feet above its top deck the reincarnated *Warrior* is again a proud ship. Its lower decks are crowded with huge cannons and ancillary equipment. Mess tables and benches for the crew of 700 are in position along with all the wood cooperage used for eating and drinking. The officers' cabins are resplendent with gold leaf, and scores of brass items reflect the light from the replica candlesticks and lamps. Rifles, pistols, cutlasses and bayonets stand once again in rack after rack. Steering wheels operate the ropes connected to the rudder as they did 127 years ago. The upper-deck capstan, which was used to hoist the anchor, can still accommodate the 120 sailors needed for the job. Replicas of the engine and the boilers have also been installed deep in the bowels of the ship.

Although *Warrior* had revolutionized warship building, for nearly 100 years it survived by sheer good fortune. Now restored and protected, *Warrior* is as much a testament to the technological feats of Victorian Britain as it is a reminder of the vagaries of military technology in any age.

Eddie is one of 11 million reasons why you'll want to watch the Juvenile Diabetes Foundation's "Thanks for Giving" TV Special.

Here are a few more.

Won't you join us?

Join hosts Michael Landon, Gloria Loring and others in this unique program aimed at finding a cure for diabetes, which, with its complications, is the nation's third largest killer disease. For information or to make a donation call 1-800-JDF-CURE.*

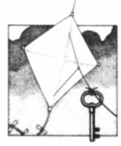
JUVENILE DIABETES FOUNDATION
THANKS
JDF
FOR GIVING
TELEVISION SPECIAL

Tune in Sunday, November 29 on U.S.A. Network.

**Above stars are scheduled to appear.*

THE AMATEUR SCIENTIST

How to capture on film the faint glow emitted when sticky tape is peeled off a surface



by Jearl Walker

When common adhesive tape is peeled off a surface, it glows with a faint blue or blue-white light along the line where it separates from the surface. The glow is generated by numerous miniature sparks that leap between the surface and the tape or along the tape itself. To see the light, adapt your eyes to darkness for at least 10 minutes. Then pull the tape from, say, a glass pane held near your eyes while you look at the line of separation between the tape and the substrate.

What is going on? For the first minute or so after you stuck the tape to the glass, electrons moved across the interface of the two surfaces, drawn to either the glass or the adhesive by electrical forces. The migration, which is called contact charging, left electrically charged patches along the interface. Some patches acquired an abundance of electrons and became negatively charged, whereas others lost electrons and became positively charged.

When you peel away the tape, you stretch, distort and rupture the adhesive material along the line of separation. As the adhesive detaches from the glass and air fills the rupture zone, the patches of charge on the two surfaces separate. When their separation reaches a certain critical value, sparks jump through the air, exchanging charged particles between oppositely charged patches to bring about a degree of electrical neutrality.

Along the way the particles collide with the air's gas molecules, leaving the molecules in excited energy states. The molecules quickly return to the unexcited state by emitting light, some of which is visible. The air's nitrogen molecules, which happen to emit light at the blue end of the visible spectrum, are responsible for most of the glow you see. Additional light may be emitted by the surface of the glass or the tape when they are struck by the particle streams, and there are still other processes that may also generate some light. All the emissions are brief, but the glow appears to be continuous if you keep peeling the tape, thereby creating fresh patches of charge and more discharges.

The separation glow is so faint that it is difficult to photograph by conventional means, but Tom Dickinson and Ed Donaldson of Washington State University have devised a clever way to record the glow on film. Working in a dark room, they peel the tape directly from the film itself or from a thin transparent plate laid on the film. When the film intercepts some of the light emitted by a spark, an image is recorded in the film's emulsion. When the film is developed, it holds a record of the light emitted during the peeling process. Dickinson and Donaldson call such a record an "autograph" of the peeling. The experiments are part of their work on fracture problems involving adhesives, composites and ceramics.

Sometimes the peeling mechanism creates a peculiar pattern of bright and dark stripes. In other cases it produces images of bright, isolated sparks. How the peeling proceeds depends partly on the type of surface to which the tape is attached. A variety of surfaces can be tested, but only transparent ones serve to expose film. A microscope slide works but produces blurred images because it is thick enough to allow the light from a spark to spread before it reaches the film.

Dickinson and Donaldson studied the glow characteristics of 3M Scotch Brand Magic Transparent Tape (No. 810) and 3M Scotch Brand Filament Tape (No. 893). The adhesive on the transparent tape is alkyl acrylate and the backing is cellulose acetate. The adhesive on the filament tape is natural rubber combined with a "tackifying" agent. The sticky face of the tape has more of the agent than the interior does, where glass filaments are bound to the polyester backing.

The film types, all Polaroid brands, were chosen for their variety of surface coatings and sensitivity to light. Type 146 has an ISO rating of 200 and is uncoated. Types 47 and 107C are rated at 3,000 and have a gelatin coating. Type 612 is rated at 20,000; the nature of its coating is proprietary information.

The films were loaded into the opened rear section of Polaroid cameras that accommodate them, with the emulsion side of the film exposed for experimentation. When the room lights were on, the loaded film was protected by an opaque black card. With the lights off and the card removed, tape was pressed onto the film with a finger. After a few seconds it was peeled off and another tape was applied to a different area of the film. After two or three tapes had been peeled, the film was developed in the usual way by bringing it in contact with the developing material. Dickinson and Donaldson wondered if any residue left on the film by a peeled tape might create artifacts in the final pictures. They found no evidence that any residue interfered with the film's development or its sensitivity to light.

Tape peeled directly off the film leaves images, but are they really due to visible light from sparks? Or might they instead be created when the emulsion is struck by particles, acted on by chemical species from the tape or illuminated by ultraviolet light from the sparks? To check for these other possibilities, tape was first peeled off a glass slide placed on the film. The slide was chosen because it blocked particles and chemical species and absorbed ultraviolet light. The images left under the slide were then compared with those produced when the tape was peeled directly off the film. The two sets of images differed only in sharpness (because the slide spread the light), indicating that visible light from the sparks must be the only significant source of the images on the film.

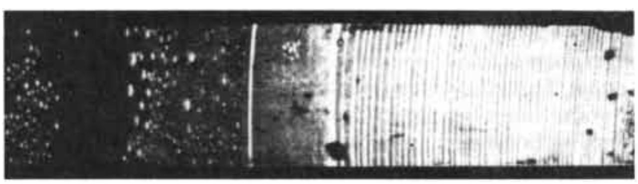
After these early trials, tape was peeled directly off the film, normally a few seconds after being applied. To produce the autographs in the top illustration on the opposite page, the transparent and filament tapes were peeled from left to right off Type 146 film, with the rate of separation accelerating from about one millimeter per second to about five millimeters per second. In both cases the pull on the tape was perpendicular to the plane of the film. Part of the autograph is striped. The rest of the display has many individual images that apparently were produced by rapid, intense bursts of light, making the autograph resemble a gathering of fireflies. Under magnification the structure of some of the individual images suggests that they were produced by sparks flashing parallel to the film surface.

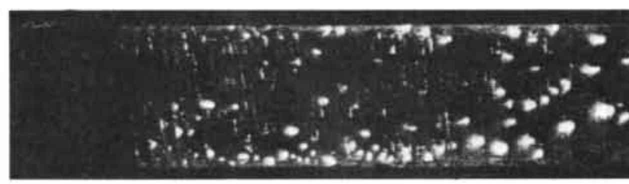
The faster rate of separation produces more pronounced stripes for the transparent tape, whereas for the filament tape a slower separation works best. The stripes result from periodic variations in the speed at which the tape separates from the film. The variations are similar to the "stick and slip" sequence observed when one surface slides over another surface. During the stick phase for the tape the separation is slow, yielding a few feeble sparks that barely expose the film. The sparks leave a dim (but not totally dark) line in the developed autograph. During the slip phase the separation is rapid, generating vigorous, bright sparks that trace a bright line across the film. Both bright and dark regions are lines running across the width of the tape-film interface, because the full width of the tape uniformly underwent either stick or slip.

The sharp features in the pattern must be due to sparks near or on the emulsion, because the light from such sparks had no chance to spread. The more diffuse features probably derived from sparks farther from the emulsion, perhaps along the tape surface after the surface was lifted from the film. Dark circles left in an autograph are caused by air bubbles trapped under the tape when it was applied to the film: charged patches are not created in a bubble area, because the tape is not in contact with the film. Defects in the tacky side of the tape also produce regions that do not spark, thereby leaving unexposed regions on the film.

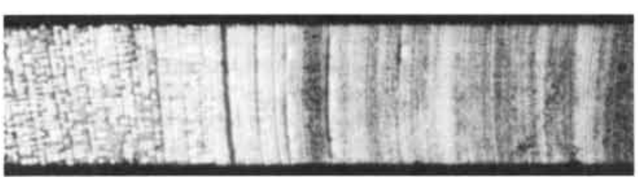
During a stick-and-slip separation the distance between the stripes depends on the properties of the adhesive, the backing and the film surface. For example, the stiffness of the tape partially determines the mechanics of how the adhesive layer on the tape breaks free of the film. Dickinson and Donaldson ran trials with transparent or filament tape stiffened by one or two layers of electrical tape attached to the backing. When the stiff composite was peeled at about the same rate as in previous trials with unstiffened tape, the distance between the stripes doubled to about two millimeters and the intensity of the sparks weakened.

The distance between stripes also depends on the direction in which the tape is pulled. When it is pulled directly upward from the plane of the film, the stripes are closely spaced. When it is pulled along the plane of

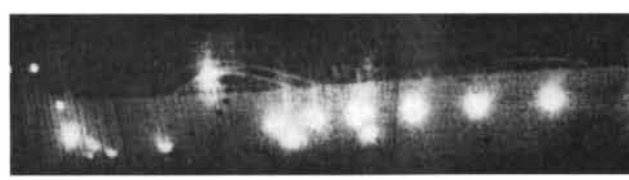




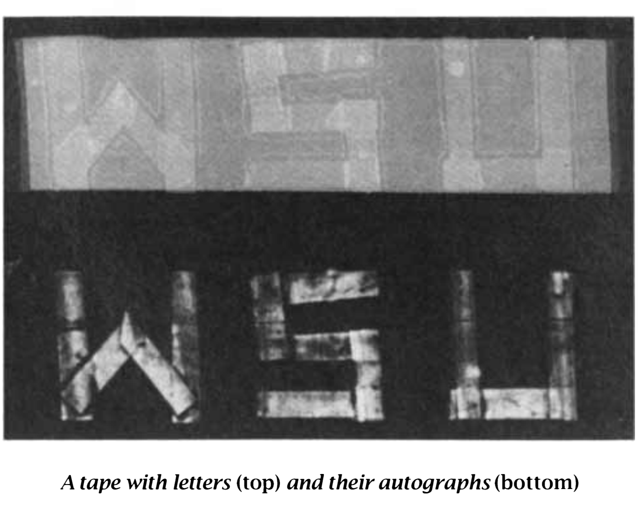
Autographs left by Magic Transparent Tape (left) *and Filament Tape* (right) *that were peeled at an accelerating rate*



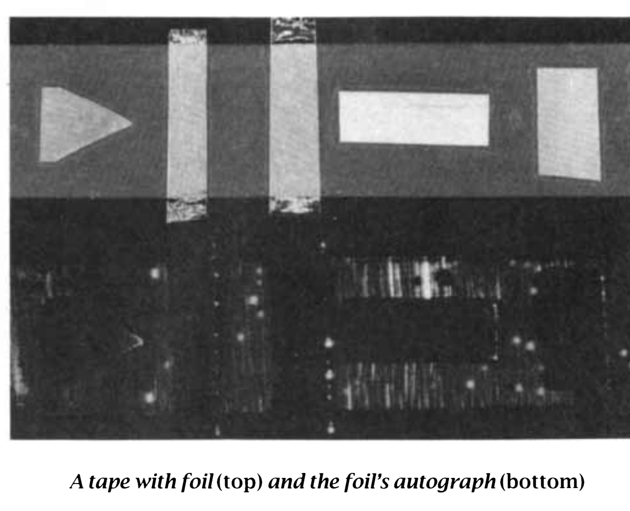
The effect of changing the direction of pull



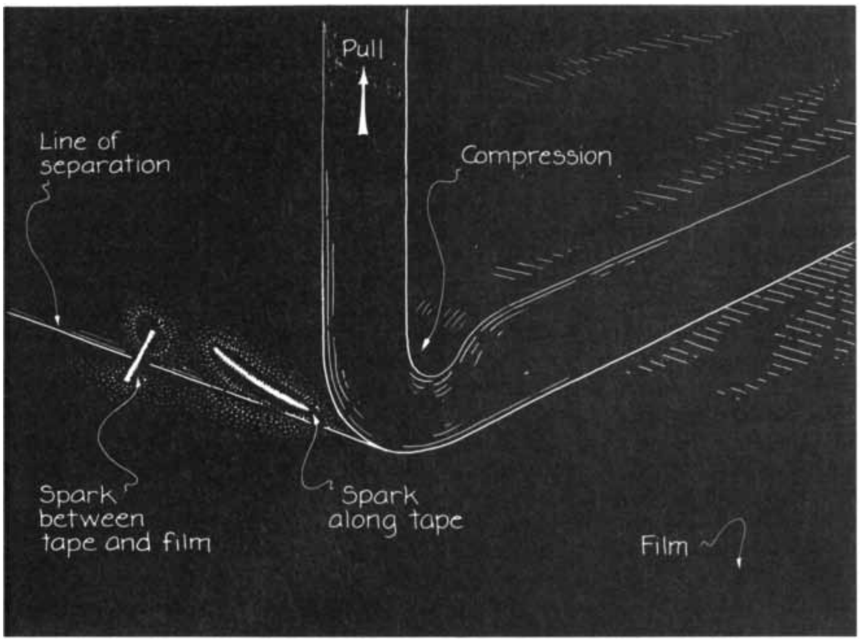
The effect of putting pressure on the bottom half of the tape



A tape with letters (top) *and their autographs* (bottom)



A tape with foil (top) *and the foil's autograph* (bottom)



Discharges from and along a tape that is being peeled

material at an interface with a surface is called adhesive failure. In contrast, the rupture of a single type of material is called cohesive failure. It is normally considerably weaker than adhesive failure in producing charged patches and the consequent sparking. To test this idea, Dickinson and Donaldson found a way to have both types of failure take place during the peeling of a strip of Magic Transparent Tape. They first stuck a layer of Scotch Brand double-sided tape (both sides of which are sticky) to the film. On the tape they attached three letters, *W, S* and *U*, that they had cut from the transparent tape. The sticky side of the letters faced down. Then they stuck a strip of Magic Transparent Tape over the entire length of the double-sided tape.

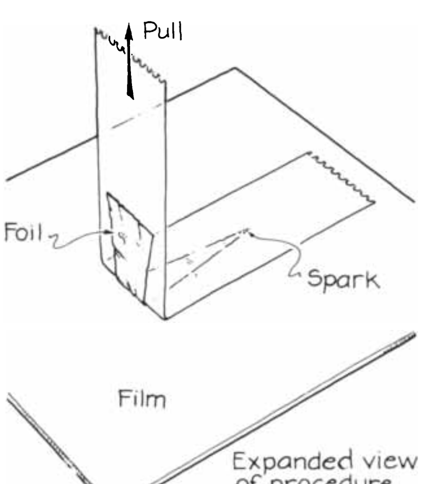
When the strip of Magic tape was peeled away from the double-sided tape, both types of failure took place. Where the strip was pulled directly from the double-sided tape, the Magic tape underwent cohesive failure, because the separating surfaces were identical in composition. Where the Magic tape was pulled from the backing of the tape forming the letters, it underwent adhesive failure, because the separating surfaces were unlike in composition.

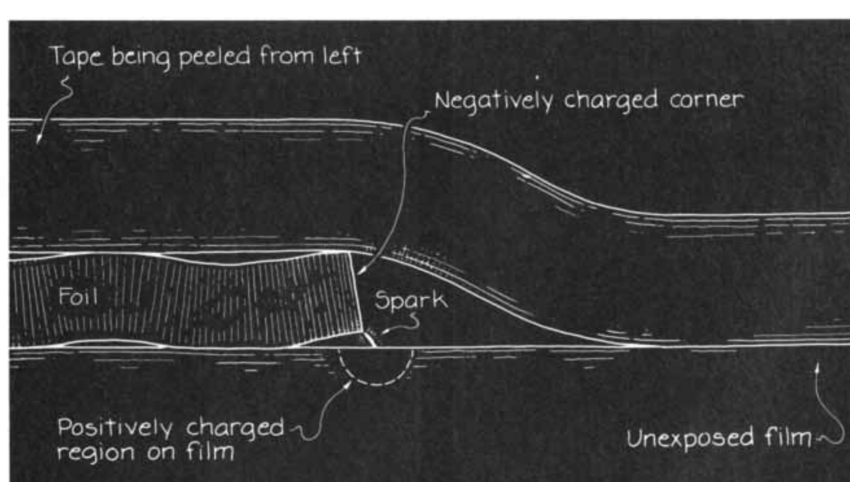
In order to avoid any additional production of light, the double-sided tape was then lifted off the film gently while being wetted with methanol applied with a cotton swab. As the methanol seeped into the crack between the tape and the film, either it eliminated the sparking by providing a conducting path for the electrons or it destroyed the tape's adhesive bond with the film. When the film was de-

the film, the stripes are farther apart. There are also more high-intensity sparks then, presumably indicating that the distortion and rupture of the adhesive must be severe at times, so that highly charged patches are created. To make the autograph in the illustration at the middle left on the preceding page the tape was first pulled along the plane of the film and then pulled directly upward.

The pressure with which a tape is attached to the film influences the extent of sparking. When Dickinson and Donaldson attached one strip of Magic Transparent Tape to a Type 107C film, they pressed particularly firmly on the bottom half of the strip. When the tape was peeled, more and brighter images were produced in the bottom half [*see illustration at middle right on preceding page*]. The influence of pressure can sometimes be seen without extra pressing because of the way the transparent tape lifts off a surface: the edge of the tape flexes just before it lifts off, pressing down on its tacky side. The firmer attachment of the edge prior to separation increases the sparking that takes place in the course of separation, and so the autograph displays a bright outline of the tape's edge.

The rupture of a tape's adhesive





The point discharge between aluminum foil and film veloped, the letters showed up as white on a dark background [*see illustration at bottom left on page 139*]. Apparently the adhesive failure had produced sparking, whereas the cohesive failure had produced little or no sparking. The result was somewhat counterintuitive, because the tape had been more difficult to pull during cohesive failure, incorrectly suggesting that cohesive-failure sparking should be more vigorous.

Dickinson and Donaldson noticed that fingerprints on the sticky side of a tape reduced the sparking when the tape was peeled off film. The oil from the finger reduces the adhesion of the tape, and it may also provide a conducting path, reducing the chance that there will be a discharge through the gas that fills the crack during the peeling. (Uncontaminated regions underwent normal sparking.) Ink deposited by a felt-tip pen on the sticky side of the tape or on the film's emulsion similarly diminished the sparking.

When tape is peeled off film, the distribution of charged patches creates strong electric fields in the crack region. Suppose a metal conductor such as aluminum foil is sandwiched between the tape and the film, so that when the tape is peeled, the foil is pulled up along with it. Does the presence of the foil alter the distribution of charge and subsequent sparking? Dickinson and Donaldson sandwiched several small pieces of foil between Magic tape and film [*see illustration at bottom right on page 139*]. One would expect that when the tape was peeled, sparking would be normal except for where the foil lay—that there the absence of adhesive failure would prevent sparking, leaving a dark replication of the foil's shape on the developed print. Indeed, the print did have such replications, but each had curious bright bands along the right-hand edge (remember that the peeling is from left to right) and sometimes along the top and bottom edges; there was no band along the left-hand edge. The bands were separated from the rest of the autograph by a dark zone where the thickness of the foil had kept the tape from adhering to the film.

What accounts for these bands? When the peeling reaches the left edge of a piece of foil, some of the charges created there are conducted throughout the foil. For example, if the left edge of the foil receives an abundance of electrons, the entire foil becomes negatively charged as the electrons spread over it. Their concentration is greatest along the edges, particularly at sharp irregularities left by the scissors when the foil was cut. Regions of film just below the edges of the foil become positively charged as some of their electrons are driven away by the electrons on the foil. An electric field then exists between the edges and the underlying film. When the field is strong enough, a discharge through the air trapped under the foil brings electrons to the film to neutralize both foil and film. As the peeling continues, more sparks are produced.

If the foil's right edge has been cut to form a point, the electric field is stronger there, because charges congregate more at points than they do along straight edges. The bright sparking that ensues creates a bright band in the replication of the point in the developed film. The peeling need not reach the point to create the band. To show this, Dickinson and Donaldson stopped the peeling seven millimeters short of the point on a triangular section of foil. To prevent further sparking, they then removed the rest of the tape by wetting it with methanol. When the film was developed, a bright band marked the point of the triangle's replication.

The tendency of adhesive tape to emit light when it is peeled off a surface can be judged by peeling it near the antenna of a transistor radio that is tuned to an unused portion of the AM spectrum. (The antenna is probably near the top of the radio.) The discharge brought about by the peeling emits radio noise with an intensity roughly proportional to the brightness of the sparks. Peel the tape a second or so after it has been applied and then try longer delays. Does the noise level vary? Try different surfaces and peeling speeds. Does the radio pick up the noise if it is tuned to the FM spectrum? What happens to the noise if the tape is surrounded by air of high humidity?

You might investigate how tape peels under other circumstances. For example, sprinkle metallic dust, lycopodium powder or fine flour lightly over the sticky side of the tape before applying it to the film. Are the autographs altered? What happens if you sandwich a fine metallic mesh between the tape and the film? You might also like to investigate the peeling characteristics of some nonadhesive materials, such as plastic food wrap. I should be pleased to hear about your findings.

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

COMPUTER RECREATIONS

Simple special effects illustrate the art of converting algorithms into programs

by A. K. Dewdney

Special effects do not always require the technical equipment and financial backing of a major motion-picture studio. Fabricators of fantasy can now generate limited but somewhat startling phenomena on their home computer with the aid of what might be called special-effects programs. Imagine, for example, that writhing worms were suddenly to crawl about vividly on one's display screen. A special-effects program can produce such a sight. What if the display screen were to liquefy unexpectedly? Raindrops falling here and there send circular ripples ever outward. That is also an easily managed special effect. Finally, suppose the screen becomes a window to a universe of stars that explode or stream by. Here one might be on board a spaceship flying down a corridor of stars. A third program produces such an illusion economically.

Economy of means is a feature of such programs, to my mind; if the effect is startling, so much the better if the program producing the effect is startlingly short. What more suitable vehicle could escort readers into the wider universe of ideas and appearances lying in wait behind the display screen? In particular, judging from a small but persistent fraction of my mail, there are computer neophytes who stare disconsolately at the pages of this department, unable to convert the algorithms presented here into working programs. They stand on the sidelines as the parade of Mandelbrot sets, word scramblers, FACEBENDERS, primordial soups and star clusters passes by. They can only guess at what joy there is in seeing an algorithm take on the flesh and blood of programhood.

For those readers, then, I shall do more this month than merely describe how the above special effects can be made. This once I shall explain the nuts and bolts of converting an algorithm into a program. I shall describe how to cover one's screen with worms in both a general algorithmic language and a specific programming language, BASIC. Besides this I shall show how to construct the algorithm itself and give an example of the art of conversion as a way of providing an entrée to past and future columns. Perhaps neophytes themselves will undergo a kind of conversion.

What protohacker can resist the temptation to write a program that will produce some lively, electronic annelids? It deflates expectation only a little to realize that each worm is nothing more than a chain of circles. But to program any special effect (or anything else for that matter), one must temporarily set aside expectation in favor of analysis. How does the worm move? By adding a new circle at the head and removing one from the tail. The resulting effect of motion—that of a whole creature sliding along—is surely an illusion, but it brings the piece off.

The program I call WORMS began to hatch when I first viewed the effect on my new SUN computer. Someone had added this eye-catcher to the basic menu of system and utility programs normally inhabiting a newly delivered machine. This someone had perhaps assumed that the sight of worms writhing on the large SUN display screen would make the buyer think the expense of the machine was justified.

The program actually hatched when I wondered how the effect of a crawling worm was realized. A close inspection of the screen made it more or less immediately obvious that the entire sense of motion was achieved by drawing a new circle at the head of the worm and erasing an old one from the tail. "Aha," I said to myself, "easy!" I briefly visualized a very short program until I realized that a program that would erase a circle previously drawn must first remember where the circle is. Since every circle in a worm must eventually be erased, every circle would have to be remembered.

One of the simplest data structures for remembering a large number of similar items is the array. I therefore thought next about how I would store the circles in an array. Whatever language one is writing in, the graphic command to draw a circle (or to erase one by drawing it black) normally uses the two coordinate points of the circle's center as well as its radius. The radius of the circles constituting my worm would be constant but centers would change from one circle to the next. It was therefore necessary to store only center coordinates. Because they were different numbers, it was convenient to store them not in one array but in two. I decided to call them *xcirc* and *ycirc*. Thus *xcirc*(*i*) would be the *x* coordinate of the *i*th circle and *ycirc*(*i*) would be its *y* coordinate.

The resulting arrays posed the major programming challenge of WORMS; somehow the program would have to add new circles to the array even as it deleted old ones. Like a worm, the arrays themselves would constitute a chain of circles. But the arrays would not move across the screen. Instead a special variable called a pointer would indicate which array position currently held the creature's tail-end circle. In both mathematics and computing there is a marvelous power in the simple act of naming. Calling the pointer variable *tail* immediately gave me a key statement in WORMS:

$$tail \leftarrow tail + 1$$

In other words, with each major cycle of the program's operation the tail pointer would move one position along in the arrays.

If, for example, *tail* were currently equal to 7, the last circle in the worm would have its center coordinates at *xcirc*(7) and *ycirc*(7). When the assignment shown above is carried out, the new value of *tail* will be 8. Meanwhile a new head circle must be drawn and its coordinates put in the arrays *xcirc* and *ycirc*. The logical place to put these coordinates is in the position currently occupied by the coordinates of the old tail: position 7.

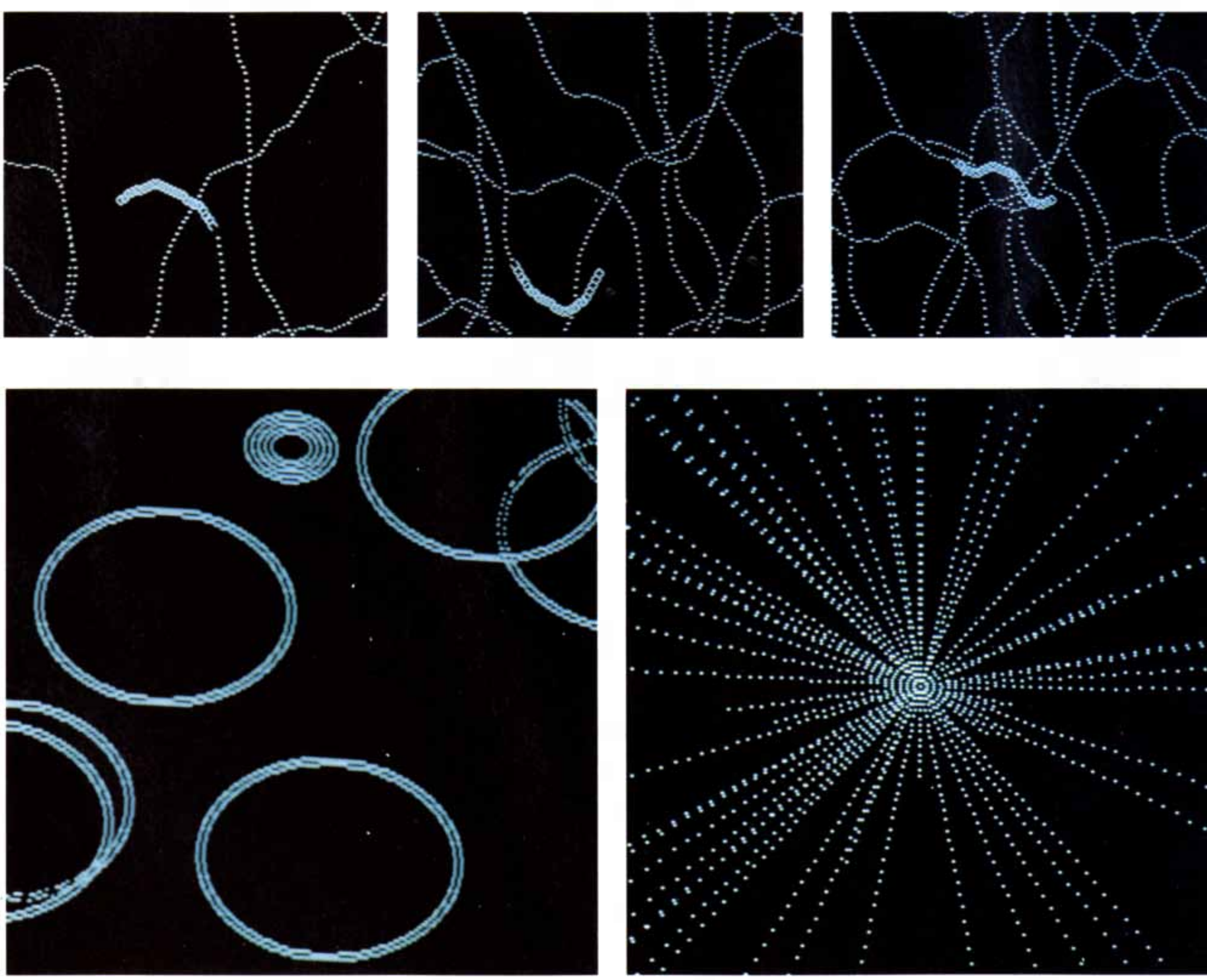
At this point I had not the slightest idea how the program would decide where to draw new head circles on the screen; this was a completely separate issue that I would deal with later. A frequent cause of difficulty in designing programs lies in the confusion that results when different computational demands become mingled in the mind. Divide et impera.

Before even beginning to sketch an algorithm for WORMS on paper I had the idea of creating the two arrays that would contain all the circles currently appearing in a given worm. The pointer *tail* would move along each array, replacing an old tail-circle coordinate by a new head-circle coordinate. What would happen, however, when *tail* reached the end of the array? It would jump back to the beginning, of course, using modular arithmetic. If the worm were composed of 10 circles, for instance, and the old value of *tail* were 9, the new value of *tail* would be 0. Here is what the algorithm looked like at this point:

```
loop
  tail ← tail + 1 (modulo something)
  get new head coordinates
  insert in xcirc(tail) and ycirc(tail)
  draw new head
  erase tail
end loop
```

The real usefulness of the algorithmic approach to the design and specification of computer programs lies in the wonderful vagueness and flexibility of algorithmic language. The designer may say anything he or she wants at any point in the algorithm, as long as the statement has a potential meaning that can be fleshed out later. For example, I had put off the question of just how the program would arrive at new head coordinates by writing "get new head coordinates." This would probably become several lines in a more detailed algorithm. Even at such a coarse level of detail I found that the algorithm immediately led to a clearer vision of just how the program should be structured. I even found a mistake: how would the program know how to "erase tail" if it did not know where it had been? Farther up in the loop the tail coordinates had just been overwritten with the new head coordinates. Erasure of the old tail would have to be done earlier in the loop, namely before the coordinate replacement was carried out. I amended the algorithm accordingly. Beginners may want to try to fix the algorithm themselves before going on and seeing what I did.

It was now time for a second pass through the algorithm. What kind of loop would I embed the entire thing in? The question immersed me right away in the problem of terminating the worm. The simplest scheme I could think of involves a fixed number of iterations in a for-loop in which



A worm that leaves tracks **(top),** ***falling rain*** **(bottom left)** ***and a stellar explosion*** **(bottom right)**

some index mindlessly counts out several thousand iterations during which the worm would wriggle about the screen. But the preferred mode for ending the worm's career is to hit a key. Hence a repeat loop seemed in order. The program would repeat the instructions inside the loop until a human being hit a designated key.

I also had to decide what to substitute for "something" in the first instruction of the loop. The modulus would have to be the size of the array. This led directly to the question of how long to make the worm. I arbitrarily decided that 25 circles would yield a fairly active-looking worm.

The algorithm for WORMS now looked like this (where "mod" stands for "modulo"):

```
repeat
   tail ← tail + 1 (mod 25)
   erase tail
   get new head coordinates
   insert in xcirc(tail) and ycirc(tail)
   draw new head
until keystroke
```

Continuing to refine the algorithm, I replaced "erase tail" by something a little closer to the command I intended to use ultimately. Since I wanted to draw a black circle where formerly a light one (on my otherwise dark screen) had been, I inserted the statement "black circle at *circ*(*tail*)," which means that a black circle with its center at *xcirc*(*tail*) and *ycirc*(*tail*) should be drawn.

I now had to face how to "get new head coordinates." By what mechanism might the program decide where to draw the next head circle? It obviously had to be close to the preceding one and somewhere in "front" of it. Furthermore, in order for the twists and turns of the worm to look natural, the new direction taken should not be too different from the old one. Here the solution lay in a single variable called *dir*. Short for direction, it would contain the current heading of the worm in degrees. At each iteration of the main loop the program would increase or decrease *dir* randomly by 10 degrees.

It was time now to decide how close the circles should be and, indeed, how big to make them. Tentatively, a radius of 4 seemed reasonable. Since the worm had to look connected, the circles would have to overlap. I replaced "get new head coordinates" by

```
change ← random
if change < .5
   then dir ← dir + 10
   else dir ← dir − 10
x ← xcirc(tail − 1)
y ← ycirc(tail − 1)
newx ← x + 4 • cosine(dir)
newy ← y + 4 • sine(dir)
```

The algorithm fragment selects a random number called *change*. Most programming systems provide random numbers between 0 and 1. Since a random number in this range will be less than .5 roughly half of the time, the variable *dir* will be incremented half of the time. This will be true only in the long run, of course. In the short run *dir* will be increased or decreased unpredictably by 10 degrees. The coordinates *x* and *y* are simply the coordinates of the previously drawn head. The last two statements are just high school computations that employ the elementary trigonometric functions cosine and sine to get *x* and *y* increments scaled to size 4. Coordinates of the new head emerge.

It took only two more steps for me to refine the algorithm. Putting the new head coordinates in *xcirc*(*tail*) and *ycirc*(*tail*) was now easy—it was a simple assignment statement. Drawing the new head involved a pseudoplot command of the form used earlier to erase the tail.

Finally, I had to initialize the arrays and variables at the beginning of the algorithm. Here I set the pointer *tail* to 1 and *xcirc*(1) and *ycirc*(1) to 100 each. The remaining entries of these arrays were assumed to be automatically set to 0. The direction variable *dir* was initialized to 0. The worm would thus begin at the point (100,100) and head in the 0-degree direction. The point (100,100) is a kind of shorthand for the actual center of one's screen. This may vary from machine to machine.

One problem remained: What if the worm crawled off the screen? I decided to incorporate a wraparound feature in the program. An *x* or *y* coordinate that wandered outside the range of allowable screen values would automatically be converted into a coordinate on the opposite side of the screen. Assuming a hypothetical screen on which coordinates range from 0 through 199, the following computation converts each coordinate outside this range into one that lies inside it:

```
newx ← newx mod 200
newy ← newy mod 200
```

The modulus function is available in virtually every programming language. It divides the number to be converted (*newx*) by the modulus (200) and computes the remainder. The variable to which the remainder is assigned (*newx*) now carries the value of the modular computation.

So much for WORMS. Or should I have called it WORM? The algorithm provides for just one. To add S meaningfully, readers may employ one set of arrays for each worm that is to appear. In any event, the algorithm I have so far specified appears in the illustration on the opposite page. An equivalent BASIC program appears alongside it. It would be most upsetting to this teacher to learn that any reader had merely typed the BASIC program into his or her machine without absorbing the lessons carefully spelled out here.

A line-by-line comparison of the algorithm and the program discloses a fairly close similarity between them. Indeed, two major blocks of instructions within the algorithm survived the translation into BASIC (IBM PC BASIC version 3.0) with very little change. Major differences revolve around implementation features of BASIC's key facility on the IBM PC, and also on the absence of the repeat loop from this version of BASIC.

In the first category, KEY(1) (the function key F1 on the IBM PC keyboard) was selected as the trigger for turning off the program once a user had become tired of watching the worm wriggle about on the screen. In line 70 the program turns the key on, signaling the underlying BASIC system to watch for signals emanating from the F1 key. In line 90 the KEYOFF command removes the BASIC options display at the bottom of the screen to make way for the worms. The main instruction involving the F1 key, however, occurs at line 60. Here the system is instructed to jump to statement 290 if the user presses F1. At line 290 the screen is cleared and execution stops at the END statement.

In the second category, the absence of a repeat loop made it necessary to construct an equivalent form of execution control by means of the GOTO command at line 280 of the program. This causes execution to jump back repeatedly to line 110 until someone pushes the F1 key. The position of the resulting loop is signaled by two comments, one at line 100 and the other at line 270. These comments were inserted for the sake of indicating the intended loop structure to readers.

Other differences also came about from writing WORMS in BASIC. Statement 10 sets the dimensions of the arrays to 25 entries each. Statement 80 sets up the high-resolution screen for the IBM PC. Here the points are close together. In line 110 a new variable *wastail* takes the value of *tail* just before the latter variable is decremented. Hence *wastail* in the program can replace *tail* – 1 in the algorithm where *x* and *y* are computed. The main reason for using *wastail*, however, is that *tail*, an array index, cannot have the value zero in this version of BASIC. Consequently in statement 120 the 1 is added after the modulus is taken, not before. In the algorithm I was merely being sloppy, imagining that the index *tail* ran from 0 through 24. The actual circle-drawing command of the IBM PC BASIC occurs at line 130. The center of the circle is at XCIRC(TAIL) and YCIRC(TAIL); the radius is 4 and the color is 0 (black).

In line 150 the variable DIR is incremented not by 10 degrees but by the roughly equivalent angle of .1745 radian. (The IBM PC BASIC does not use degrees.) In lines 200 and 210 the variables NEWX and NEWY are limited to sizes 600 and 200 respectively. This forms a box just inside the IBM PC display screen. The modulus facility in BASIC reduces positive numbers quite nicely but fails to work on negative numbers. Thus when a worm crawls off the screen in a negative direction, the modulus is taken by adding the appropriate screen dimensions, as in lines 220 and 230. A worm that crawls off one edge of the box will reappear on the other side. At line 290 the SCREEN 0 command resets the screen for any ensuing textual displays.

WORMS lovers can enhance their program by inserting a RANDOMIZE command before the loop begins, preventing the worm from retracing its path. They can also add provisions to restart the program, if they want to. They can also add a little trail of dots as we have done in the exemplar worm displayed in the top of the illustration on page 143.

The other special effects I mentioned above might now conceivably be programmed, even by neophytes, if the following hints suffice.

The program I call RAINDROPS selects a random point on the display screen and draws a succession of circles of ever increasing radii about it. Then the program selects another point and repeats the circle-drawing procedure. RAINDROPS continues in this way until the rainmaker depresses a key. Thus there are two loops, an outer loop in which the point is selected and an inner loop in which the surrounding circles are drawn. RAINDROPS converts two random numbers between 0 and 1 into screen coordinates by multiplying them by the respective dimensions of the screen. The resulting point (*x*,*y*) becomes the center for a succession of circles drawn by the inner loop. Let an index *k* run from 1 through 25 while white (or light-colored) circles are drawn. The centers are at (*x*,*y*) and the radii take on the values 4*k*. It is simpler to leave all the circles so drawn on the screen. But it is more realistic to have just one circle or a few circles spreading outward. To obtain just one, RAINDROPS must erase (redraw in black) the circle with radius 4(*k* – 1) after the circle with radius 4*k* is drawn.

STARBURST displays a stellar explosion. The experience can also be seen as the trip down a vast corridor of stars that pass one's armchair spaceship at warp speed. The stars move by small increments along lines of perspective toward the view-

```
xcirc(1) <— 100
ycirc(1) <— 100
dir <— 0
tail <— 1
repeat
    tail <— (tail + 1) mod 25
    black circle at circ(tail)
    change <— random
    if change < .5
      then dir <— dir + 10
      else dir <— dir - 10
    x <— xcirc(tail - 1)
    y <— ycirc(tail - 1)
    newx <— x + 4 · cosine (dir)
    newy <— y + 4 · sine (dir)
    newx <— (newx + 100) mod 200 – 100
    newy <— (newy + 100) mod 200 – 100
    xcirc(tail) <— newx
    ycirc(tail) <— newy
    white circle at (newx, newy)
until any key
```

```
10  DIM XCIRC(25), YCIRC(25)
20  XCIRX(1) = 100
30  YCIRC(1) = 100
40  DIR = 0
50  TAIL = 1
60  ON KEY(1) GOSUB 290
70  KEY(1) ON
80  SCREEN 2
90  KEY OFF
100 'repeat
110   WASTAIL = TAIL
120   TAIL = (TAIL MOD 25) + 1
130   CIRCLE (XCIRC(TAIL), YCIRC(TAIL)),4,0
140   CHANGE = RND
150   IF CHANGE < .5 THEN DIR = DIR + .1745 ELSE DIR = DIR - .1745
160   X = XCIRC(WASTAIL)
170   Y = YCIRC(WASTAIL)
180   NEWX = X + 4 * COS(DIR)
190   NEWY = Y + 4 * SIN(DIR)
200   NEWX = NEWX MOD 600
210   NEWY = NEWY MOD 200
220   IF NEWX < 0 THEN NEWX = NEWX + 600
230   IF NEWY < 0 THEN NEWY = NEWY + 200
240   XCIRC(TAIL) = NEWX
250   YCIRC(TAIL) = NEWY
260   CIRCLE(NEWX, NEWY),4,1
270 'until key pressed
280 GOTO 110
290 SCREEN 0
300 END
```

Converting WORMS from an algorithm into BASIC

er. The increments themselves increase as the stars approach. Nearer objects appear to move faster than distant ones. STARBURST is more complicated than RAINDROPS, however. It must keep track of all the stars currently appearing. Readers may use an array called *stars* to this end—more correctly two arrays, as in WORMS—to store the x and y coordinates of individual stars. The z coordinates are implicit in the apparent motion of the stars.

When a star first appears at the origin (100,100), motion increments are at first very small. The increments get steadily larger as stars approach the observer (or vice versa), increasing by an amount inversely proportional to z, its supposed distance from the observer. Programmers may plan on, say, 25 increments per star before it flashes past. The distance z to a given star may thus be taken as $25 - k$, where k is the number of iterations in the star's history.

The grand loop for STARBURST involves the creation of a new star, the destruction of an old one and an increment for every star in the arrays.

The resourcefulness of our readers never fails to amaze me. Having implied in the August column on word webs that anyone without a computer need not concern herself or himself with certain problems, I find those very people knocking on the door, solutions in hand.

In particular, I posed the general question of converting one n-letter word into another by way of a sequence of intermediate words. Only one letter at a time may be changed. The result is a word ladder. In the case of $n = 2$, for example, all two-letter words (in the dictionary I provided) are connected by word ladders. I left open the question of whether all n-letter words are connected for higher values of n. It proved hardly necessary to use computers to explore the interconnections.

Many readers offered *gnu* to settle the three-letter case. For $n = 5$ a number of others offered *xylem*. These forays depended only on the insight that a word that cannot be connected to any others settles the issue; then others cannot be connected to it. According to George A. Miller of San Francisco, there are 1,217 English words to which *horse* (my own example) cannot be connected. Miller has a program that determines all possible word ladders. It can be ordered from him at 2426 Bush Street, San Francisco, Calif. 94115.

A highly interesting variation of word ladders is suggested by Paul L. Blass of Florham Park, N.J. In addition to substituting single letters, let the words crawl sideways like worms, shortening or lengthening by one letter at each end.

By employing one or two slippery intermediaries (such as *orad*), Joseph M. Erhardt of Richmond, Va., discovered a way to transform *evil* into *good.* Dennis Farr wrote from Framingham, Mass., to display a ladder stretching from *solid* to *plane.* He did some *flips* and a *slide* on the way.

The Martian word problem dealt with webs of a different kind: Given a dictionary that lists allowed substitutions of one string for another, transform an initial word into a final one. The problem was solved by Alfred V. Perthou, III, of Seattle, Wash., and Edward J. Groth of Scottsdale, Ariz. The puzzle falls apart once it is realized that the few dictionary entries amount to simple rules that allow one to shuttle symbols about.

Some of the best fun arose from the computational word-bender created by Ron Hardin and featured in the August column. Readers may recall the strange verse entitled "Tweeze Denied Beef Worker Isthmus," a tortuous sound-alike of the familiar Christmas poem by Clement Clarke Moore. The art was already old in the hands of computerless humans. Ronald C. Read, a mathematician and gamester at the University of Waterloo in Ontario, told me of a French (kind of) book called *Mots d'Heures: Gousses, Rames,* or "Mother Goose Rhymes," which contains such perversions as "Lit-elle messe, Moffette."

In English there is literature of this ilk. Two readers sent in the charming "Ladle Rat Rotten Hut." For such, a computer is hardly necessary. D. H. Wood of Montreal, Quebec, rendered the carol "It Came upon the Midnight Clear" with new lyrics: "Eat cane a pond am I'd knight glare." Finally, Richard Tilden wrote from Somerville, Mass., to remind me that Walt Kelly, the beloved American cartoonist, had constructed a number of pieces including "Deck the Halls with Boston Charlie" and "My Bonnie Lice Soda Devotion."

I am subscribing forthwith, after much needless delay, to the magazine *Word Ways.* I have heard so much about all the art in the field of word webs from the editor, A. Ross Eckler, that I cannot wait any longer. Readers can subscribe by writing to Eckler at Spring Valley Road, Morristown, N.J. 07960.

STATEMENT OF OWNERSHIP, MANAGEMENT AND CIRCULATION (required by 39 U.S.C. 3685). *1A.* Title of publication: Scientific American. *1B.* Publication number: 00368733. *2.* Date of filing: September 24, 1987. *3.* Frequency of issue: monthly. *3A.* Number of issues published annually: 12. *3B.* Annual subscription price: U.S. and its possessions, 1 year $24; all other countries, 1 year $33. *4.* Complete mailing address of known office of publication: 415 Madison Avenue, New York, N.Y. 10017. *5.* Complete mailing address of the headquarters or general business offices of the publisher: 415 Madison Avenue, New York, N.Y. 10017. *6.* Full names and complete mailing address of publisher, editor, and managing editor: Publisher, Harry Myers, 415 Madison Avenue, New York, N.Y. 10017. Editor, Jonathan Piel, 415 Madison Avenue, New York, N.Y. 10017. Managing Editor, Armand Schwab, Jr., 415 Madison Avenue, New York, N.Y. 10017. *7.* Owner: Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Holtzbrinck Publishing Holdings Limited Partnership, 521 Fifth Avenue, New York, N.Y. 10175; Handelsblatt GmbH, Kasernenstrasse 67, Postfach 1102, 4000 Duesseldorf, Federal Republic of Germany. *8.* Known bondholders, mortgagees, and other security holders owning or holding 1 percent or more of total amount of bonds, mortgages or other securities: none. *9.* For completion by nonprofit organizations authorized to mail at special rates (Section 423.12, DMM only). The purpose, function, and nonprofit status of this organization and the exempt status for Federal income tax purposes: Not applicable to this organization. *10.* Extent and nature of circulation: *A.* Total number of copies (net press run): average number of copies each issue during preceding 12 months, 758,689; actual number of copies of single issue published nearest to filing date, 790,506. *B.* Paid and/or requested circulation: *1.* Sales through dealers and carriers, street vendors and counter sales: average number of copies each issue during preceding 12 months, 131,061; actual number of copies of single issue published nearest to filing date, 148,645. *2.* Mail subscription (paid and/or requested): average number of copies each issue during preceding 12 months, 483,258; actual number of copies of single issue published nearest to filing date, 484,661. *C.* Total paid and/or requested circulation (sum of *10B1* and *110B2*): average number of copies each issue during preceding 12 months, 614,319; actual number of copies of single issue published nearest to filing date, 633,306. *D.* Free distribution by mail, carrier or other means, samples, complimentary, and other free copies: average number of copies each issue during preceding 12 months, 4,825; actual number of copies of single issue published nearest to filing date, 3,089. *E.* Total distribution (sum of *C* and *D*): average number of copies each issue during preceding 12 months, 619,144; actual number of copies of single issue published nearest to filing date, 636,395. *F.* Copies not distributed: *1.* Office use, leftover, unaccounted, spoiled after printing: average number of copies each issue during preceding 12 months, 8,289; actual number of copies of single issue published nearest to filing date, 14,389. *2.* Return from news agents: average number of copies each issue during preceding 12 months, 131,256; actual number of copies of single issue published nearest to filing date, 139,722. *G.* Total (sum of *E, F1* and *2*—should equal net press run shown in *A*): average number of copies each issue during preceding 12 months, 758,689; actual number of copies of single issue published nearest to filing date, 790,506. *11.* I certify that the statements made by me above are correct and complete. (Signed) Harry Myers, Publisher.

BOOKS

A vacation trip for young readers around the world of science

by Philip and Phylis Morrison

Physical Sciences

EXPLORING THE NIGHT SKY: THE EQUINOX ASTRONOMY GUIDE FOR BEGINNERS, by Terence Dickinson. Principal illustrations by John Bianchi. Camden House Publishing Ltd., distributed by Firefly Books, 3520 Pharmacy Avenue, Unit 1-C, Scarborough, Ontario, M1W 2T8 ($9.95).

The starry sky has been caught well in these full-page paintings; they evoke—particularly for the beginner—the sense of pattern and the subtle contrast better than the camera can. One among the plates offers the sky surrounding bright belted Orion, viewed across a snow-covered stretch of pine woods. Some six dozen stars are painted in, along with a key map, names, time and direction in which to look. So far this is familiar, if unusually convincing; every star guide has a version of such an image, the myriad fire folk breathtaking in the velvet night. The next page is unique, yet exactly what tyros need. Here is the same stretch of sky. Now the foreground is a city scene, and the stars are city stars in the perpetual twilight of the city sky: belted Orion and scarcely a dozen more. This is unusually realistic help, fulfilling the book's claim that it is an "astronomy guide for beginners."

The first two sections of this mainly pictorial book complement the sky plates very well: they enhance the astronomical imagination rather than the grasp of the unaided eye. A cosmic voyage is set out in 10 paintings of scenes at growing scale. The first step shows a view of the moon's surface from close by, bright crescent earth as background. The last is a cluster of galaxies. On the way we visit many planets; we even gaze at the grand disk of the Milky Way from one in the Magellanic Cloud. The second section treats yet more "alien vistas," the stuff of speculative headlines, from other planets to quasars. Here a little hyperbole enters, although not much: icicles on Pluto are hard to credit, and black holes are given enthusiastic play. A table with a fine diagram of nearby stars drawn in color and to relative size is perhaps the most abstract page of this admirable work from Canada. It is a beginner's bargain that is hard to beat, well suited for readers in the middle grades on up (and their elders too).

THE GEYSERS OF YELLOWSTONE, by T. Scott Bryan. Revised edition. Colorado Associated University Press (paperbound, $7.95).

Old Faithful is so worn a souvenir of casual tourism that its wonder is now somewhat obscured. Never did it jet faithfully "every hour, on the hour"; its actual rhythms have been carefully measured. Nearly 50,000 events yielded a mean interval of 64.9 minutes, between observed extremes of 33 and 120 minutes. This small bargain volume is a comprehensive field guide to the 300-plus active geysers that play within our oldest national park. It covers the subject spout by spout from the little fellows like one-foot Lime Kiln Spring up to Giant, whose rare big burst of steamy water reaches past a height of 200 feet. The book is a tribute to and a tool for the geyser gazers, the amateurs who watch fondly night and day over the entire striking assembly of geysers (more than half of all the active ones in the world are there in Yellowstone National Park) with patience and enthusiasm.

Like all subjects admired by naturalists, geysers have names, characteristics, locations. All the known geysers are mapped, disputed over, judiciously listed and compared in this volume. The water in Sawmill's pool spins during an eruption. For Jelly, dormant intervals are long and brown algae grow thick in the pool. These inanimate fountains form an ecology, interacting, shifting, even competing. The lucky observer can be first to report such changes. The intricate plumbing of crack and crevice through which cold groundwater finds its way to hot rock and back, and the unstable consequences of boiling flow under the variation of pressure with depth, rule the geyser phenomenon.

T. Scott Bryan makes clear how the subtler chemistry of the basin rock is decisive as well. In Yellowstone the dozen geyser basins lie in recent volcanic fields of silica-rich rhyolite. The silica dissolves and redeposits as opal-like stuff to seal the unseen plumbing and build the visible cones and pools that house the geysers. Underground changes invisibly control the flow: Daisy's Thief, a new jet that emerged in 1942, spouts moments before Daisy, siphoning off just one of her regular eruptions. Earthquakes induce more extensive changes over the years.

The lore is outlined here in genuine fondness for the geyser-gazing art. It is also recounted in defense of these somewhat fragile organisms of water, steam and brittle rock against attack by boot, spade and willful blocking. Although the book focuses primarily on Yellowstone, there is an unusual up-to-date compilation of geysers the world around (geysers are to be found in half a dozen of our states). Any visitor to Yellowstone for whom geysers carry any appeal at all needs the guide. Good readers far from Wyoming will also enjoy the narrative; it is a fresh example of what can be made out of the specificity of the natural world.

ICEBERGS AND GLACIERS, by Seymour Simon. William Morrow and Company, Inc. ($13).

In a score of big brilliant pages, ice blue and cloud white, this skillful author has assembled a fascinating display of ice at geologic scale, supplemented by a brief, clear text written for readers in the lower grades. It is systematic enough to convince (a single snowflake shown large is presented as the beginning of it all) and dramatic enough to catch eye and mind in the striking results of time and snow. The roped parties of geologists picking their way among the enormous crevasses of the wide Juneau Ice Field, the blued cliffs of the thick tabular bergs frowning high above the open ocean, the bright red Coast Guard patrol cutter at sea warily close to a jagged berg, and the false-color satellite image of Iceland

and its ice only sample what is here. These photographs are large and glowing enough to compete with the fleeting images of television. The motion of glaciers is made clear, and the mechanics are suggested well enough, although one figure to support the story of the old surveys of glacier motion would have been helpful.

The work glaciers do on land is not forgotten: rock flour, scratches, sheeplike rocks, moraines and other traces are shown. The book ends with a page or two to remind the young reader of the thick ice that once was here and will come again.

SCIENCE FARE: AN ILLUSTRATED GUIDE AND CATALOG OF TOYS, BOOKS, AND ACTIVITIES FOR KIDS, by Wendy Saul, with Alan R. Newman. Harper & Row, Publishers (paperbound, $14.95).

This ambitious book is for parents, teachers, librarians and leaders of youth groups of any kind, people who may feel themselves far from science or its teaching. It centers on informal education in science, out of school, although many teachers will put it to good use.

The text is in two parts; the shorter first portion reviews how children might learn science. It treats how to get started, reading about and within science, approaches to learning, questions and answers, visits, working with others and science fairs. All those matters are complex, subjective, even divisive, and so the position taken matters; not everyone will accept the premises, even though Saul writes fairly and openly.

A few remarks seem so striking and sensible that they may serve to calibrate the author's point of view. They stand in a chapter that addresses three worries: mess, unrealistic expectations of grand success and safety. On messes: "Hands-on science tends to be *messy*," although palliatives are offered. On lack of realism: Headline adult science is Cape Canaveral and giant dinosaur bones. A child cannot enter that high circle but can go far anyhow. On safety: "Although coating a coin with mercury may be fun, the thrill of the shine is *definitely* not worth the danger of messing with a dangerous chemical." In contrast, "rockets can be dangerous, but I believe the value an interested and reliable kid can realize from an intense interest in model rocketry is worth the dangers inherent in the activity." If that sympathetic and courageous judgment does not suit you, the book will have only minor worth; if you find the view plausible, the book is a rich resource.

The larger part of the text is an evaluative listing of toys, kits, posters, books and other things you can buy to help children into science. Since the estimates flow in part from the viewpoint of the authors (aided by expert consultants), the two portions are connected, although concreteness widens the appeal of the lists. Those lists extend from the main sciences to include computers, electronics and optical devices. Even if the work were not evaluative, it would be fascinating just to see what is out there and where to get it. A worm farm, a big magnifier, a batch of geologic data from your own state, astronaut ice cream, the welcome given children by the organized network of amateur radio—maybe the book is for kids after all! Hundreds of good books and magazines are knowingly annotated. Perhaps you recall fondly your old chemistry set with its rows of chemicals ranged in a gleaming metal box; alas, "there is nothing comparable on the market today." This book is a contemporary resource that fulfills the same description.

A SKELETON IN THE DARKROOM: STORIES OF SERENDIPITY IN SCIENCE, by Gilbert Shapiro. Harper & Row, Publishers ($13.95).

The darkroom was Wilhelm Roentgen's laboratory, where he worked alone in the winter evenings, seeking the faint glow from the cathode rays he could make in his clumsy discharge tube. It was an all-glass tube, and he would run the voltage up so that the rays Professor Lenard had just reported could penetrate the thickish wall. The glow they would make in air was very faint; he sought real darkness so that he might see it, or better yet, see the glow from the little fluorescent screen he had prepared. But even while he was setting up in the dark at low voltage, there came a faint glow from the end of the laboratory table. "Damn! The cover must have fallen off the coil." Only when he had struck a match could he see that it was in fact the screen on the table that had glowed; the glow was green as it should be, but Roentgen had mistrusted that judgment, for he was a little color-blind. Within a few impatiently awaited nights the skeleton itself had appeared: the bones of Roentgen's living hand shadowed clearly on the screen. Within weeks X rays were made worldwide; the world of physics had entered its 20th century, only a few years early.

Six other major discoveries in science are also well and briefly rendered as tales of surprise, based on carefully selected secondary sources. They begin with Hans Christian Oersted's wire and compass needle in 1820 and end in 1979, when the Alvarezes, father and son, and their nuclear-chemist colleagues, Frank Asaro and Helen Michel, sought and found the iridium fingerprint at the top of the Cretaceous rocks that somehow spelled the end of the dinosaurs. Between them we read of penicillin, of the faint radiation from the big bang, of the J/Psi meson (a double name because it was a double discovery, made at two places simultaneously) and of the bit of scruff on the output chart of a new radio telescope. That scruff signaled in 1967 the first pulsar; graduate student Jocelyn Bell remembered that the same mark had been seen before at the same place in the sky.

None of these tales is particularly new (granted, the J/Psi story is told much more intimately than is usual). Yet the sense of adventure is so well brought out in each of the narratives that they constitute an *Arabian Nights* of the laboratory. The book can be enjoyed by any good reader; the more science you know, the richer the tales become. Their sources are given in full.

MUSICAL SOUND: AN INTRODUCTION TO THE PHYSICS OF MUSIC, by Michael J. Moravcsik. Paragon House Publishers ($24.95).

A man of taste in physics, in music and in popular exposition of science, Michael Moravcsik, a nuclear theorist at the University of Oregon, has written an unusually readable text. With a genuine minimum of mathematics the book treats the physics of waves, propagation and perception, musical time and musical scales, as well as the making, hearing and recording of music. The aim throughout is an honest mapping of everyday understanding into the more general language and crisper ideas of physics.

Is it the inverse-square law you want to grasp? Most texts might draw concentric spheres swelling with radius and leave it at that. This author shows waves sent down a tunnel and out into a closed shed as well: different decay in idealized circumstances. Kinetic energy and potential energy? A vibrating string is drawn in 18 successive positions and the energy relationships are tabulated.

Analysis according to M. Fourier is given a step-by-step account; there is a good supply of curves and analogies but there are no formulas. The idea of musical scales is explained by means of explicit arithmetic, the intervals rising through the simple rational fractions to yield the diatonic major. The task is repeated starting from another note, and the needed arithmetic complicates itself by short division before your eyes. Logarithms are avoided, although the rules of decibel appear earlier in an account of loudness. The tempered scale enters as welcome compromise in a world that wants to use the fixed pitches of winds or clavier.

A logical flow from energy source through primary vibrator to resonators unifies the treatment of the musical instruments. The Donald Duck sound of the helium-filled singer is evidence that resonance of the oral cavities is important to voice quality, even if the primary vibration frequencies of the vocal cords cannot change on inhalation of the gas. Room acoustics and recording are similarly discussed, with a sense of what is important; technicalities are kept to a severe minimum. This book should serve well any serious musical readers able to work out a graph, from those who are old enough for junior high to the nonscience college students for whom it was designed. Its apt concreteness keeps it quite free of the glib generalizing that mars many simplified treatments of a piece of real physics.

Mathematics and Image

When Sheep Cannot Sleep: The Counting Book, by Satoshi Kitamura. Farrar Straus Giroux ($9.95).

One night Woolly, a serious-looking young sheep with small horns, was too restless to sleep. He got up to walk around the meadow in the lovely calm night. He found plenty to do: chase a butterfly, watch ladybugs, talk with squirrels. He even found some apples on a tree and climbed a handy ladder to reach them. There were fireflies; later a squadron of big lights zipped across the sky, looking just like UFO's. Frightened, he ran off across a bed of red tulips to a large house, full of doors and windows. One room had colored pencils, and so he made pictures (a profoundly mathematical set of them); the kitchen held peas that he cooked and ate when he felt hungry. He went to bed in the quiet house, thinking as he rested of all his distant dear woolly friends and relatives. Finally Woolly fell sound asleep and filled the air with *Z*'s.

Everything that Woolly saw can be counted, although no digits appear at all until the index page, which lists the sights in order: from a single butterfly up through 22 *Z*'s. Very young counters will find these tallies hard to resist, although in the endpapers and on the dust jacket Woolly encounters the starry sky: too many! This is a funny and somehow a joyful book both for its arithmetic, deeper than mere memorization, and for the sentence or two of text on each page.

A Grain of Rice, written and illustrated by Helena Clare Pittman. Hastings House, Publishers ($12.89).

A handsome young peasant, Pong Lo, had the temerity—his clever head seemed at risk—to ask the Emperor for the hand of the Princess Chang Wu. The gentle Princess was clearly won by Pong Lo and persuaded her father to engage him to help clean the Imperial Storeroom. He rose by sheer merit and charm to cook for the Emperor himself. The Princess grew so radiant that soon her father assembled the young noblemen so that from among them she might choose a husband.

What happened next? Of course poor Chang Wu languished, fell ill and came near death. She was saved only by a potion from Pong Lo, brought to her by her worried father with the promise that if the medicine worked, the young man might ask for any gift. His remedy cured swiftly; he chose, of course, the hand of the Princess. Impossible! But an alternative, a single grain of rice, was granted freely, including a whimsical little stipulation that the amount must double every day for 100 days. Even by the 12th day the affair was still ceremonial: 2,048 grains came to Pong Lo's door in a box covered with silk, to join the precious little jade and ivory packages of the week before. By the 40th day...elephants on parade. The abacus beads had clicked and clacked: 2 to the 39th power means some 50 tons of good rice. The grave advice of the Imperial Exponentiator was compelling. Soon the wedding of the Princess to the richest man in all the land was celebrated with a grand feast, organized by the gifted Prince, who out of delicacy served not one grain of rice.

The story of the wheat on the chessboard has been amusingly Sinified (its ambience now not gaming but dining) with evocative drawings; any young reader (with calculator handy) will enjoy the tale and can lengthen it.

Dots, Spots, Speckles, and Stripes, by Tana Hoban. Greenwillow Books ($11.75).

The 100 eyes of the peacock's tail, the colorful mottling of strawberries, kiwi, lilies and the child's freckles; the field of sunflowers aligned and backlit, the big Dalmatian, the cow, the ladybug, the zoo tiger and the confetti spilled on the street... Was that next picture, a smiling masquer adorned with both spots and stripes, made at the same little carnival? In strong contrast and delicate ripple, this photographer once more sets before us a diverse feast made one, images suited for all who are beginning to ponder the pages of books. The 29 full pages without words celebrate dappled things; they appear here more in sweetness than as Gerard Manley Hopkins saw them, "counter, original, spare, strange."

Opt: An Illusionary Tale, by Arline and Joseph Baum. Viking Penguin, Inc. ($11.95).

Across the 20-some picture pages of this droll little book any young reader (even those who still prefer being read to) can follow the stylish story of a visit to a kingdom where King and Queen strongly resemble face cards and children fill the colorful costumes of the court, from jester to Princess. There are strange pets, along with a cat and a friendly dragon (he made a perfect guest). The few lines of text on each page are even in rhyme.

Pleasant enough, all of that an agreeable icing. The clever magician-artist pair who made this book in fact use it to present to young readers the full standard set of optical illusions. Cleverly woven into the images are lines that are the same length but look longer or shorter in their context, colors that similarly are belied by their surroundings, false nonparallels, hidden figure-ground image pairs and visually concealed faces; words that cannot be read except at glancing view, shifting gray spots and aftercolors, impossible perspectives, and drawings of hollow boxes—or are they solid blocks? A few afterpages explicitly point out and rationalize in simplest terms all those perceptual effects (and a few more). As a reader you are offered advice that will help you make and test your own illusions and are encouraged to do so. The book is a small tour de force of cheerful, if rather mannered, exposition.

Prelude to Mathematics, by W. W. Sawyer. Dover Publications, Inc. (paperbound, $4.50).

The little classic by a Toronto professor remains one of the best books through which a high school student can come to see that mathematics is neither the set tasks of the school curriculum nor the professional elaboration of calculations. The book was first written in the 1950's; a few final pages were added about 10 years ago. It consists of two parts, different in content but akin in tone.

The first quarter of the book, although it offers very interesting examples, is really a look at just what a mathematician is and how one grows. The wish to explore for oneself, to seek pattern, to generalize and unify is given vivid meaning by the examples, which are not always elementary. (A fine footnote explains the ubiquity of the Laplacian operator in applications; it is the simplest differencing operator that treats all directions and all points in space on an equal footing.) Neither mathematics seen as art nor mathematics seen as utility gives an adequate model of this subject with its "enormous bulk," although the theory of the artist "would do less harm."

The remaining text treats a number of topics selected for their novelty and interest, often with quickly grasped and yet striking results (not all the text is easy going). The topics include non-Euclidean geometries and a little of their physics, matrices, projective geometry and its applications within mathematics itself, finite arithmetic, some algebraic logic and the theory of groups. The decades have not made these topics less interesting; groups are perhaps even of greater interest today than they were a mere 10 years ago. True, there is no hint of complexity or computability or fractals or similar computer-related topics. It is still a useful exercise to try to describe geometric matters over the telephone to an angel who lacks earthly experience.

A few pages explain the Galois group of the equation for the three cube roots of 2. It is the same as that for the movements of an equilateral triangle. Its structure is compound, and through it we begin to glimpse the nature of the nonprime simple groups, "interesting and somewhat rare," lately noticed for the remarkable intricacy of their kind, the smallest having 60 elements, the next 168. Today the list of such interesting simple groups extends to gigantic proportions. Professor Sawyer has the knack of opening topics up, not of closing them off; explorers, notice.

Technology

The History of Invention: From Stone Axes to Silicon Chips, by Trevor I. Williams. Facts On File, Inc. ($35). **The Genius of China: 3,000 Years of Science, Discovery, and Invention,** by Robert Temple. Simon and Schuster ($19.95).

The shelves of every good library for readers in science are weighed down by two big and wonderfully illustrated multivolume works; both are comprehensive references and mighty good reading at the same time. One of these is the eight-volume general history of technology, completed in 1984 after three decades of work under managing editor Trevor Williams. Its senior editor is Charles Singer. The other work is the 15 volumes (and counting) on the history of science and civilization in China, produced during approximately the same years by Joseph Needham of the University of Cambridge and his friends. The pair of books reviewed here are up-to-date, authoritative single-volume abridgments. Gone are the leisurely and learned apparatus and full documentation, the glorious footnotes and obscure sources; what has been gained, besides affordable price and small bulk, is colored illustrations in plenty and a concise narrative text that opens the once academic material to many more readers, above all those of high school age. This is a fine year for the reference bookshelves.

Dr. Williams has organized his narrative on the same lines as his big book. We are led along the continuous thread of the history of technology from ancient times on into the medieval world, the Renaissance and the Industrial Revolution. An expanded account treats the 20th century. For each period—the charts and illustrations are integral with the text—we find a readable survey of all relevant fields, say agriculture, domestic arts and crafts, industry and building, transport and communications, even the arts of war. Of course, the detail of treatment is adjusted to the nature of the period.

Early pages show and comment on Olduvai Gorge and the key graph of world population over time; one of the final images is a shuttle launch. The 300 large pages in between span the entire domain of what we know. Here is a rendering of the famous Roman grain mill near Arles in the south of France, in the fourth century the most powerful installation in the world. Its 30 horsepower came from 16 big waterwheels fed one after another by an aqueduct led down the steep hillside. A few pages later a similar drawing describes the remarkable copper mine at Rio Tinto in Spain, where Roman engineers arranged a set of eight wheels that one after another raised water a total of 100 feet to keep the mine works from flooding; human beings supplied the pumping power.

The second book is by a scholar whose own previous books are excellent credentials for the study of the history of science and technology. Dr. Needham himself introduces Robert Temple's volume as a "brilliant distillation." It sets out the whole of Chinese technology topic by topic, from agriculture through astronomy, engineering, mathematics, acoustics and weapons, to name but half its chapters. The book shares Needham's constant and delighted search for unexpected priorities in Chinese technology. Again a single amazing sample will have to suffice. Seeking salt from brine wells, the Chinese craftsmen-engineers of the first few centuries A.D. drilled to depths of 4,000 or 5,000 feet, anticipating the oilmen of Pennsylvania by 15 centuries and more. There were entire fields, hundreds of deep wells, in China by the year 1000; the gas they yielded as a natural by-product was burned to evaporate the brine. The mile-long cables that held the drill bit were pieced out of strips of bamboo, flexible enough to wind on drums, 10 times as strong as hemp rope. Bamboo piping lined the borehole and carried off the gas. Month after month the hard cast-iron bit, deep below, would pound up and down a couple of feet at a stroke, steadily powered by a team of six men on the surface rhythmically jumping on and off a long lever board. The tall bamboo derricks are in use in our times; there are clear photographs.

Not everything we read is right, even in such definitive treatises, a lesson itself worth learning. Claims that empty eggshells can be made to float in the air as hot-air balloons are flawed, if authentic; it simply will not work. Williams remarks in an aside that the verdant limestone sinkhole holding the wonderful radio dish at Arecibo is an extinct volcano. What counts is how well these fine books bring the best of our learning to a new audience.

Life on a Fishing Boat: A Sketchbook, by Huck Scarry. Prentice-Hall, Inc. ($14.95). **Cleared for Take-off,** by Meredith Hooper. Illustrations by Tony Talifero. Angus & Robertson Publishers, U.S. distributor Salem House Publishers ($14.95).

These two books are closely observed and vividly reported narrations of how expert people work for a living with some very special tools. The tools are of course described; our attention is on their knowing use, not on how they were made. Authors and artists in each case went along to watch hour by hour; these reports stay close to what happens.

The two accounts are akin, yet quite different. The tools in the first book are fishing boats of several kinds. The smallest is a bright red wood diesel trawler—a tugboat for fishnets—out all day raking its net along the bottom for half a ton of good fish off the Brittany coast. It is not a tool for one person's hand, although the captain does own it. The fiberglass wheelhouse holds radar, depth and fish-finding sonar and a Decca radio navigator. The captain depends on the expert help of his two sons, as once his father depended on him. Many large, detailed pencil sketches show us the boat and its family, the nets, the instruments, the people, the fish in variety (flat, round, whiskery), the dockside market, the fishmongers... It is easy for us to imagine ourselves aboard on these particular voyages, sharing a little of the life of a hunter at sea. We join two or three other French boats, out for Norway lobster or silvery sardines, then on to the cold, windy North Sea, turning back as shore radio warns of Force 8 winds to come. A few less intimate pages extend to North American shores, where the fisherman's world looks little different.

The second book is also about a journey of a day or two, but one seven miles high and halfway around the world (and back). We are aboard a Boeing 747-238B, one owned by Qantas, called VH-EBQ, delivered in Seattle in December, 1979, and veteran at this flight's end of 2,090 landings. It carries along "its own biography and its own vital statistics" in a big book. There is a first-aid kit of vital spare parts too, along with oxygen, knife sharpeners and a silver caviar bowl, but no spare engine. This big plane resides at the Qantas Sydney Jet Base, where it was fitted the week before our flight with four new Rolls-Royce engines; for now they must be used only at the same thrust rated for the older engines. The engineers have been retrained, the new spare parts are stored out there along the route, but the final certification is not yet in hand. The new engines will save high-cost fuel anyway (on each leg of the journey the big jets drink 100 tons), although they can save much more once they are allowed to operate as designed. It is this kind of lifelike detail the book celebrates; a heavy jet is a complex tool tied even more than any fishing boat to intricate techniques, legal obligations, constant negotiations and many skills, organized internationally for smooth function worldwide.

This trip takes us from Sydney to Singapore, Bahrein, London and back, all in some 46 hours one weekend; we leave, but the jet is off again, late Monday afternoon. The sense of participation is strong; the reader is by turns with the intense flight crew up front, the big, busy cabin crew, the caterers, mechanics, dispatchers, controllers, couriers. You are one of an overlarge family of 400 intimate strangers, restlessly sleeping in your cramped seat, washing by turns, waking repeatedly to eat in place, watching sunsets or movies, maybe sniffing the outside air at the nightbound stops across Asia. The maximum space between passenger headrests is always close to 34 inches in economy class; "humans come in odd shapes and sizes but airlines do their best to make them all fit a standard space."

The good drawings show the plane in detail, the airports from runway to passenger queues to kitchen; there are telling photographs, plans and maps. This is the closest thing to a long international flight, both tedious and exhilarating (roomier in first class), to be found in print.

Bridging the Golden Gate, by Kathy Pelta. Lerner Publications Company ($10.95).

The great bridge opened just 50 years ago; a quarter of a million celebrators during last summer's anniversary flattened the span's six-foot camber. This excellent small book (for anyone who can read well) recounts how the popular wonder came to be. In 1916 the proposal for a suspension bridge to span the rough waters of so many ferry rides was put forward by a newspaperman-engineer. He recognized that it would be the longest bridge ever built. The county supervisors on both sides ordered a study; soundings were carried out after the war, in 1920. Feisty engineer Joseph Strauss headed a firm that had built 400 unremarkable bridges worldwide; the daring idea appealed to him mightily, all the more so as a widespread opposition grew on safety, economic and aesthetic grounds. It took until 1928 to win the go-ahead in court. Strauss won the competition for chief engineer. His first design of 1921 is shown here; no beauty, it was a pair of towers with cantilever arms joined by a cable portion. (The present beautiful design was by an associate.)

In the depression year of 1930 the people voted a bond issue for the bridge by a count of three to one. Work began in 1933; the forecast was three million days of abundant labor. Construction was full of novelty; many stages are shown in photographs. The building of the Golden Gate saw the first use of concrete trucks that stir the mixture as they go and the first use of safety nets in bridge construction. For more than three years there was not one fatality on the job; lead poisoning from the paint was the chief worry. In the end 11 men died, 10 of them when scaffolding gave way as forms were being removed during the completion of the roadway. The bridge was finished a few months late, a million dollars under the $35-million budget and well below the grim expectations of one life lost for every $1 million of construction.

The first day was for pedestrians. "Be quiet," a man called to the strolling crowd. "Listen to the bridge." The beautiful soaring structure was a giant wind harp. (It had a touch of the Tacoma Narrows disease, oscillating in strong winds; after the Tacoma accident the Gate bridge was treated by cross bracing.) The billionth car traversed the famed bridge in 1985; all the loans had been paid off through tolls by 1971. Strauss was once asked how long the bridge would last; he "replied calmly, 'Forever!'"

Wheels at Work: Building and Experimenting with Models of Machines, by Bernie Zubrowsky. Illustrated by Roy Doty. William Morrow and Company, Inc. (paperbound, $5.95).

What a contraption! It is clearly drawn, in operation by the fourth- to sixth-grade boy-girl partners who built it. Waterpower drives it. The flow siphons down from a bucket of water on a high stool. The jet is directed against a waterwheel at floor level that consists of plastic cups edging a pair of plastic plates held on ball-point-pen tube bearings. A wire shaft mounted on milk cartons supports the assembly. Masking tape fastens and seals everything. A loop of ribbon couples the driving wheel to another wheel on the tabletop above. That wheel's rim repeatedly carries four pipe-cleaner loops through a soapy solution held in a shoe storage box. A household fan blows across the film-covered loops to loose a storm of rainbow spheres: bubbles made wholesale by machine. Safe and well tested until it really worked, the design is just what we count on from this savvy engineer at the Boston Children's Museum.

The book evolves with easy text and many drawings to that complicated wonder (and a little past it to windmills and rubber-band boats), through simpler pulleys, gears, windlasses and the like. Each device lies within the same genre of materials and style. These tasks are scaled to small teams of children, good friends, club members or students. The projects call for very little money; they do require adult tolerance for sand or water spilled on the floor or in some close-by yard or driveway. The devices are fine examples of engineering design; they can be realized in materials available in household and supermarket: cat-litter trays and coat-hanger wire.

Archaeology and Folklore

Digging to the Past: Excavations in Ancient Lands, by W. John Hackwell. Charles Scribner's Sons ($13.95).

The beautifully painted amphorae of the opening page give way at once to an archaeologist carefully sketching on his plane table; this science is no treasure hunt but a study of the past by methods that inescapably destroy part of their context. A responsible dig is a controlled one, but it is never beyond surprises.

In full-page paintings from life and a clear firsthand text, this artist-archaeologist recounts his experiences at a tell in Jordan. The survey may lay out 10,000 squares, each about the size of a living room. A supervisor, helped by one or two assistants and some local labor, will excavate each square chosen. Perhaps 20 squares can be opened in one hot summer season, working daily from first light to sunburned noon. The tools are familiar: brush, pick, hoe, trowel, sieve. It may take a full day of touchy work to uncover without crumbling them only five or 10 ancient bricks. The sifters will come on thousands of shards; broken pottery is everywhere. Each find is labeled and entered in the computer file. Sift meticulously; just one shard with scratched writing "can be more valuable than any other find on the tell" during 10 years of excavation. In every square a meter-wide strip at north and east is saved from excavation; the smoothed vertical walls of these "balks" graph the past in layers to be read anywhere in the dig.

You can get lucky and find tools, coins, figurines, seals, striking mosaics, walls built of big boulders. Graves that contain offerings and even human skeletons are treasured finds. It is no wonder that most of the diggers are hopeful volunteers, usually undergraduates working for credit. The evening meal of falafel, flatbread filled with salad and beans, is satisfying after a hard day, while someone on a more fortunate square recounts an exciting little find. At season's end the dig is legally inspected by the hosts; then heavy equipment can come to haul away all the unwanted silt. You go home after a farewell feast of herbed roast lamb for all; sometimes the "young Bedouin boy" who was hired to gather stones is met 10 years later as a full-fledged archaeologist, representing his national department of antiquities as director of another dig. One day the long report will come out and the museums will be enriched: more of the human past.

Folktales of India, edited by Brenda E. F. Beck, Peter J. Claus, Praphulladatta Goswami and Jawaharlal Handoo. The University of Chicago Press ($29.95).

Ninety-nine folktales from all over India, each only a few pages long, are set down in English translation. They are unadorned; a map, a set of indexes and a valuable foreword and introduction are what you get. The stories were collected by a dozen or more cited scholars. For each story there is an introduction; a note recounts the source in detail. Here are stories told by 13-year-old girls, by famous professional storytellers, by a garage mechanic or a village housewife or a Brahman professor during a university tea break. This is the stuff itself.

Of course, India is so big, so complex and so long devoted to the narrative and the parable that in these stories Sanskrit myth can be heard amidst today's wit. Who can know which came first? Those who wrote down the classics long ago had certainly heard the storytellers in the vernacular. "Myths, by and large, divinize the human; folktales humanize the divine." Folktale is of the earth and says little about other lives than those lived here, although gods and ogres do appear. These tales speak intimately; not every youngster is ready for them. (High school kids know all the themes.)

They yarn about family life, courtship, newlyweds, parents and children, sisters and brothers, domestic strife and about virtue and knowledge and beginnings. A tale from Kashmir, taped by a professional folklorist in 1971 from the telling by a housewife 84 years old, has become so widely known that it has entered the theater. Listen for its overtones and connections:

A saintly beggar appears at a devout home where husband and wife have seven girls but no son. He promises them a son, whom they may raise for 12 years only; at that time the boy must be returned. The bargain is struck. They love and admire young Akanandun as he grows, takes the sacred thread, becomes a diligent schoolboy. But the beggar-saint returns, takes a knife, dismembers the boy when he comes home and even demands that the parents eat the child's boiled flesh. The distraught mother is calmly asked to taste the dish to judge its salting. The very stars hide in the face of such shame and disgust. All at once the boy Akanandun himself enters cheerfully to share the meal. He grows up as he should, the lifelong joy of the family. The storyteller commented that it requires a heart of steel to tell this particular story.

Living Things

Wild Boars, by Darrel Nicholson, photographs by Craig Blacklock. Carolrhoda Books, Inc. ($12.95). **Morning Glories,** by Sylvia A. Johnson, photographs by Yuko Sato. Lerner Publications Company ($10.95). **Discovering Flies,** by Christopher O'Toole, illustrations by Wendy Meadway. The Bookwright Press, 387 Park Avenue South, New York, N.Y. 10016 ($10.40). **Sparrows,** adapted by Katherine Pohl. Raintree Publishers Inc. ($14.65).

Colorful and interesting species abound worldwide, and naturalists everywhere study them fondly. Each life cycle is recorded, nowadays often by admirable closeup photography. The role of each species in its context is to some degree known and often easily described in simple language. There are young readers everywhere too, and publishers to provide them with easy text and fine images of natural history. Once a book has been well received somewhere, it is no difficult task to translate it, adapting the text as needed, and to republish it for other readers in another tongue. The photographs can usually remain unchanged. Two universals meet in these pages: the natural world and the interests of young readers. Admirable.

All the same, the meeting is at a place of least resistance. The simple text and striking pictures vary from form to form in wild variety. Yet the concept is single, the pool of resources plentiful. The result is a tide, many good books of the same kind; on the whole the differences in quality are not large compared with the disparate appeal of this or that kind of plant or animal to any particular reader. The four books discussed are members of a long series; with them one might also list six dozen related, more or less similar works on other living forms. The same phenomenon can be found on television, where nature film also abounds worldwide, for both young and adult viewers. These books need to be good to compete with those images moving on the video screen; they are, and they do.

A few words for the individual little volumes; like animals, no two books are alike. The wild boars are American: blends of introduced European boars and feral domestic pigs. These animals resemble their wild Eurasian forebears. They are found in half a dozen states, fierce, adaptable, hardy, like the ones shown in snowy Minnesota woodlands. The morning glories are Japanese; the seedlings are found first underground, and all the major parts of the plants are followed through time until at last new seeds appear.

Some flies are of course almost everywhere. Here are commonplace greenbottles and exotics as well; one West Indian wood-boring fly is a full two inches long. The fly text is aimed at early readers; its unusual images are taken from footage shot by a number of English filmmakers. The sparrows were recorded by a Japanese photographer; they are the European tree sparrow, somewhat different from the usual American varieties, and indeed from our now common English sparrow as well. They all mate, nest, fight, preen, chirp and flutter in a most sparrowlike way.

Remember, a dozen or so books like each one of these are out there.

DINOSAURS PAST AND PRESENT: VOLUME 1, edited by Sylvia J. Czerkas and Everett C. Olson. Natural History Museum of Los Angeles County in association with University of Washington Press ($35).

These big pages offer dramatic pictures, old and new. On the dark forest floor a protobird, still rather reptilian, seizes a large dragonfly as the big, shadowy feet of a sauropod swing past. What happens when a ponderous herbivore seeks aquatic refuge from a little herd of fierce allosaur predators? Why, they follow it into the water to tear at the fleeing giant while we watch from the depths nearby. Across a dry cypress swamp trot a herd of big-beaked reptiles. Their broad crests bear a pattern in red and white as showy as some ceremonial New Guinea mask. All these are the work of contemporary artists. Way back in 1897 the artist Charles R. Knight painted a dramatic scene of energetic combat, two dryptosaurs leaping and rolling in mortal combat "like a pair of half-ton fighting cocks." Neither the lumbering file in mid-distance nor ritual combat with toothy tyrants need limit dinosaur images today.

The book reports a traveling museum exhibit of representations of the dinosaurs, from those of Benjamin Waterhouse Hawkins at the time of the Exhibition of 1851 to the scrupulous and imaginative work of today's painters. In addition to these images there are half a dozen papers by paleontologists and artists (one or two can claim both professions) that address, often quite personally, the problem of restoration of those fellow creatures long since gone. A piece by Robert T. Bakker strengthens his vision of the old reptiles as warm-blooded, fast-moving, diverse. Size estimates based on trackways (laid down by both the earliest reptiles and today's mammals), prey and predator counts as well as bone structure all concur; those heavy shafts of bone bespeak adaptation to extraordinary stress, not a clumsy slowness. The chief theme of the other papers is how the synthesis that the artists seek can best include the detailed analytic data of the expert. No reader will go away without admiration for the devotion and originality of the artists' effort. There is even a "dinosauroid" sculpture, an expert thought experiment about what saurians might have yielded along the *sapiens* path had they and not the little mammals survived. On this planet no test seems possible.

A companion volume is due in the spring, and seven big museums in turn will host the show through 1988. A serious work for adult readers, this first volume will challenge and delight dinosaur enthusiasts from the middle grades up. There is a lot to learn at many levels: fresh images, comparisons, striking graphs, entire parades of tracks and plenty of informed argument.

WHEAT: THE GOLDEN HARVEST, by Dorothy Hinshaw Patent, photographs by William Muñoz. Dodd, Mead & Company ($12.95).

Every year the average American eats at least 100 pounds of wheat products, more than he or she eats of sugar, fats, beef or pork. And so the field of ripe wheat stretches golden to the mountain edge, every single head heavy with the tiny grass seeds that will nourish us. Those are shown up close too, as are the first green leaves of winter wheat peeking through the snow. Other pages compare soft white spring wheat with hard red winter grains. Bread is watched in all its stages, from the bowlful of ingredients—the wet dough indeed looks unpromising—to the plump, browned, fragrant freshly baked loaf. We learn a little of elevators and mills, of grain drills and ammonia tanks, of white flour and bulgur and bran and wheat germ and of the giant combines that now harvest the crop across the plains.

This Montana writer-photographer pair manage to make vivid both the broad farms with their big machines and what we carry home from the supermarket. One remark helps to illuminate the history of our times: the acre of ripe wheat that 12 men with a dozen horses, mules or oxen would cut and thresh in three days is now gathered up in six minutes as the combine rolls, one person at the controls.

EGGS: NATURE'S PERFECT PACKAGE, by Robert Burton. Photographs by Jane Burton and Kim Taylor. Facts On File Publications ($22.95).

"The photographs in this book represent three years of dedication to eggs," to the animals that produced them and to those that hatched from them. The take is dramatically visible: big vivid pictures, which are particularly helpful when they show sequences of events, enrich the well-written nontechnical text. The emphasis is on a thoughtful comparative account, evolutionary but not at all molecular or even genetic; the entire chromosome dance draws only a column of explanation. We read of and see the budding of hydras and the fusion of protozoa, but it is the widespread occurrence of the egg and the related seed, and the sexual reproduction they generally represent, that sound the motif.

The hard-shell hen's egg, outwardly inanimate, and "the animated, mewling kitten crawling towards its mother's teats" make a strong cognitive contrast. Even the biologists were misdirected by those externalities, and doubted that a mammal might ever have laid eggs. The platypus was a wonder. But in fact the egg and its development are remarkably alike across the entire span of multicellular life, from the tiny brine shrimp to the great gray slug, the rainbow trout, the common frog, birds (of course) and the newborn kitten. All of those are shown along some developmental sequence, particularly the melodrama of birth or hatching. Yet the egg is a powerful single theme with variations. There are three egg architectures. The egg of the mammal is tiny, fed and guarded by the mother's body; the bird and the reptile supply plenty of yolk and a big container, but a tiny embryo within enters on the same dance. The sea urchin and the brine shrimp make little maternal investment; there are plenty of minute unfurnished eggs, and the small larvae must now emerge quickly to set off on their own.

This is good reading for anyone interested in biology, made easier by the striking illustrations; the language is clear and unlabored, with only a little of the Hellenic jargon of the textbooks. The complete absence of explicit indications of scale is a problem presented by many, but of course not by all, of the images. Action is everywhere in image and text, since the topic of this book is life in swift change. The American lobster, for instance, is a unique mother. No other parent helps as clearly with egg hatching; when the time is ripe, she beats her swimmerets, on which her eggs are glued, "so violently that. . .the embryos are shaken out."

THE NATURALIST'S YEAR: 24 OUTDOOR EXPLORATIONS, by Scott Camazine, with original illustrations by Cynthia

Camazine and the author. John Wiley & Sons, Inc. (paperbound, $14.95).

The versatile author is a physician in Ithaca, N.Y., where he plainly adorns the Cornell tradition of keen amateur naturalists who transcend mere field-guide identification. The 24 activities described with affection and mastery in this attractive and personal volume are divided equally among the seasons of the year.

The cold winter winds carry the dry flakes of snow, to deposit them behind obstacles. A little aerodynamics helps to make sense of the snow cornices of rolling country and the avalanches of open snowy lands. Snow fences upwind form lee drifts, keeping clear the roadway beyond; they work best when they are not solid walls but half-open pickets. A trial with your own test obstacles (perhaps fireplace logs) to follow what drifts they induce fills the outdoor hours more happily than shoveling snowpiles.

Spring will bring bad weather as well as good. Study the creatures in the home on a bad day. Flies, silverfish, clothes moths, yes, and cockroaches—"even I find them somewhat loathsome"—all repay study. A key is given to begin identification of the insects you find; only a few flies are easily identified, even under the stereomicroscope. These pages give a nice sense of the problems of taxonomy, at and beyond the level of the field guides. Cluster flies deserve even more attention; their exquisite response to temperature sets many nice problems. In summer, seek out and examine the milkweed and the birdfoot trefoil flowers. These blossoms, common in the northeastern U.S., are as cleverly contrived to extort proper pollination from their insect visitors as any of the orchids celebrated by Charles Darwin. They too use such clever devices as baited slots, springs and clips. Test the mechanisms yourself with a toothbrush. For autumn, turn biochemist. Here is an account of paper chromatography you can carry out on the brightly colored leaves; household equipment, such as the blender and coffee filters, is needed; so are reagents such as vinegar and solvents such as acetone and methyl alcohol. They are found among the paint thinners at your hardware store.

Scott Camazine has mapped some fascinating trails clearly open to high school students, or to younger ones with some help from older friends. His book is a reliable guide. It is worth a line to expose an old and disagreeable stereotype: the artist Cynthia Camazine is the author's talented and patient stepmother.

A Shovelful of Earth, by Lorus J. Milne and Margery Milne. Illustrated by Margaret La Farge. Henry Holt and Company, Inc. ($12.95).

Call it what you will, soil or dirt or ground covers most of the lands, only feet thick but continent-wide. Even the biologists somewhat neglect it; soil science is mostly a product of the 20th century. These two biologist-writers point out a domain open to fresh exploration at modest effort almost everywhere.

Dig in the forest if you can; between the tree roots the ground is soft and moist. Hands and a shovel are plenty of tools; mark off with string and nails four square feet, somewhat less than a tenth of a milliacre. First is the litter layer, where the parts of plants can be recognized by eye. Then comes dark brown topsoil, organic but without recognizable plant parts. Below that is the *B* horizon, the paler subsoil, grittier, denser, with scattered pebbles, not very organic. The pebbles grow in number with depth, until finally there is bedrock, too deep to reach, it may be, although disclosed at road cuts. In each layer there is a specialized and often abundant life.

The litter holds the sparrow hunting (listen for its scratching among the dry leaves) for the beetles, snails and grubs that crawl or scurry there. Earthworms dominate the topsoil; there are also springtails, fungus networks, roundworms, ants and yellow jackets. Deep down lie the strange, lonely cicadas, programmed for resurrection.

The drawings make this darkened life vivid, and the text tells of the ways of counting and collecting it. You will enjoy some finds, but not all the methods are easy enough for the amateur to use. Soils differ; there are brief chapters on what is found under the evergreen forest, under land on which rain falls every day, under dry deserts and under frozen ground. How a single sand grain could be shown to have a history as a distinct particle for a couple of hundred thousand years, in desert dune, glacial ice and tidal flat, is well explained; for the electron-microscopic images that were the signs of that past you would need to look up the reference given. This is a fresh and important 100-page contribution, a first-rate small book for grade school biologists (or older beginners).

BIBLIOGRAPHY

Readers interested in further exploration of the subjects covered by the articles in this issue may find the following lists of publications helpful.

TECHNOLOGY IN SERVICES

DEINDUSTRIALIZATION AND THE SHIFT TO SERVICES. Ronald E. Kutscher and Valerie A. Personick in *Monthly Labor Review,* Vol. 109, No. 6, pages 3-13; June, 1986.

A COMPARISON OF EMBODIED TECHNICAL CHANGE IN SERVICES AND MANUFACTURING INDUSTRY. Richard A. Barras in *Applied Economics,* Vol. 18, No. 9, pages 941-958; September, 1986.

SERVICE SECTOR PRODUCTIVITY. John W. Kendrick in *Business Economics,* Vol. 22, No. 12, pages 18-24; April, 1987.

THE MOON'S ANCIENT MAGNETISM

CONVECTION IN THE MOON AND THE EXISTENCE OF A LUNAR CORE. S. K. Runcorn in *Proceedings of the Royal Society of London, Series A,* Vol. 296, No. 1446, pages 270-284; 1967.

PRIMEVAL DISPLACEMENTS OF THE LUNAR POLE. S. K. Runcorn in *Physics of the Earth and Planetary Interiors,* Vol. 29, No. 2, pages 135-147; August, 1982.

LUNAR MAGNETISM, POLAR DISPLACEMENTS AND PRIMEVAL SATELLITES IN THE EARTH-MOON SYSTEM. S. K. Runcorn in *Nature,* Vol. 304, No. 5927, pages 589-596; August 18, 1983.

THE PRIMEVAL AXIS OF ROTATION OF THE MOON. S. K. Runcorn in *Philosophical Transactions of the Royal Society of London, Series A,* Vol. 313, No. 1524, pages 77-83; November 27, 1984.

HOW ANIMAL CELLS MOVE

THE LOCOMOTION OF FIBROBLASTS IN CULTURE. III. MOVEMENTS OF PARTICLES ON THE DORSAL SURFACE OF THE LEADING LAMELLA. M. Abercrombie, Joan E. M. Heaysman and Susan M. Pegrum in *Experimental Cell Research,* Vol. 62, Nos. 2-3, pages 389-398; October, 1970.

REDISTRIBUTION AND PINOCYTOSIS OF LYMPHOCYTE SURFACE IMMUNOGLOBULIN MOLECULES INDUCED BY ANTI-IMMUNOGLOBULIN ANTIBODY. Roger B. Taylor, W. Philip H. Duffus, Martin C. Raff and Stefanello de Petris in *Nature New Biology,* Vol. 233, No. 42, pages 225-229; October 20, 1971.

THE CRAWLING MOVEMENT OF METAZOAN CELLS. M. Abercrombie in *Proceedings of the Royal Society of London, Series B,* Vol. 207, No. 1167, pages 129-258; February 29, 1980.

MEMBRANE RECYCLING BY COATED VESICLES. Barbara M. F. Pearse and Mark S. Bretscher in *Annual Review of Biochemistry,* Vol. 50, pages 85-101; 1981.

ENDOCYTOSIS: RELATION TO CAPPING AND CELL LOCOMOTION. Mark S. Bretscher in *Science,* Vol. 224, No. 4650, pages 681-686; May 18, 1984.

COSMIC STRINGS

COSMOLOGICAL FLUCTUATIONS PRODUCED NEAR A SINGULARITY. Ya. B. Zeldovich in *Monthly Notices of the Royal Astronomical Society,* Vol. 192, No. 3, pages 663-667; September, 1980.

EVOLUTION OF COSMIC STRINGS. A. Albrecht and N. Turok in *Physical Review Letters,* Vol. 54, No. 6, pages 1868-1871; April 22, 1985.

COSMIC STRINGS AND DOMAIN WALLS. Alexander Vilenkin in *Physics Reports,* Vol. 121, No. 5, pages 263-315; May, 1985.

SUPERCONDUCTING COSMIC STRINGS. E. M. Chudnovsky, G. B. Field, D. N. Spergel and A. Vilenkin in *Physical Review D,* Vol. 34, No. 4, pages 944-950; August 15, 1986.

COLLECTIVE COMPUTATION IN NEURONLIKE CIRCUITS

COMPUTING WITH NEURAL NETWORKS. Edited by John Denker. American Institute of Physics, 1986.

PARALLEL DISTRIBUTED PROCESSING: EXPLORATIONS IN THE MICROSTRUCTURE OF COGNITION. Edited by David E. Rumelhart, James L. McClelland and the PDP Research Group. The MIT Press, 1986.

COMPUTING WITH NEURAL CIRCUITS: A MODEL. J. J. Hopfield and D. W. Tank in *Science,* Vol. 233, No. 4764, pages 625-633; August 8, 1986.

COURTSHIP IN UNISEXUAL LIZARDS: A MODEL FOR BRAIN EVOLUTION

FEMALE OF THE SPECIES: SEX AND SURVIVAL IN THE ANIMAL KINGDOM. Bettyann Kevles. Harvard University Press, 1986.

EVOLUTION OF MECHANISMS CONTROLLING MATING BEHAVIOR. David Crews and Michael C. Moore in *Science,* Vol. 231, No. 4734, pages 121-125; January 10, 1986.

BEHAVIORAL FACILITATION OF REPRODUCTION IN SEXUAL AND UNISEXUAL WHIPTAIL LIZARDS. David Crews in *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 83, No. 24, pages 9547-9550; December, 1986.

THE PSYCHOLOGY OF REPRODUCTIVE BEHAVIOR: AN EVOLUTIONARY PERSPECTIVE. David Crews. Prentice-Hall, Inc., 1987.

THE FRACTURING OF GLASS

A MOLECULAR INTERPRETATION OF STRESS CORROSION IN SILICA. Terry A. Michalske and Stephen W. Freiman in *Nature,* Vol. 295, No. 5849, pages 511-512; February 11, 1982.

SLOW FRACTURE MODEL BASED ON STRAINED SILICATE STRUCTURES. T. A. Michalske and B. C. Bunker in *Journal of Applied Physics,* Vol. 56, No. 10, pages 2686-2693; November 15, 1984.

CLOSURE AND REPROPAGATION OF HEALED CRACKS IN SILICATE GLASS. Terry A. Michalske and E. R. Fuller, Jr., in *Journal of the American Ceramic Society,* Vol. 68, No. 11, pages 586-590; November, 1985.

H.M.S. *WARRIOR*

THE BLACK BATTLEFLEET. G. A. Ballard. The Nautical Publishing Co., 1980.

WARRIOR: THE FIRST MODERN BATTLESHIP. Walter W. Brownlee. Cambridge University Press, 1985.

THE AMATEUR SCIENTIST

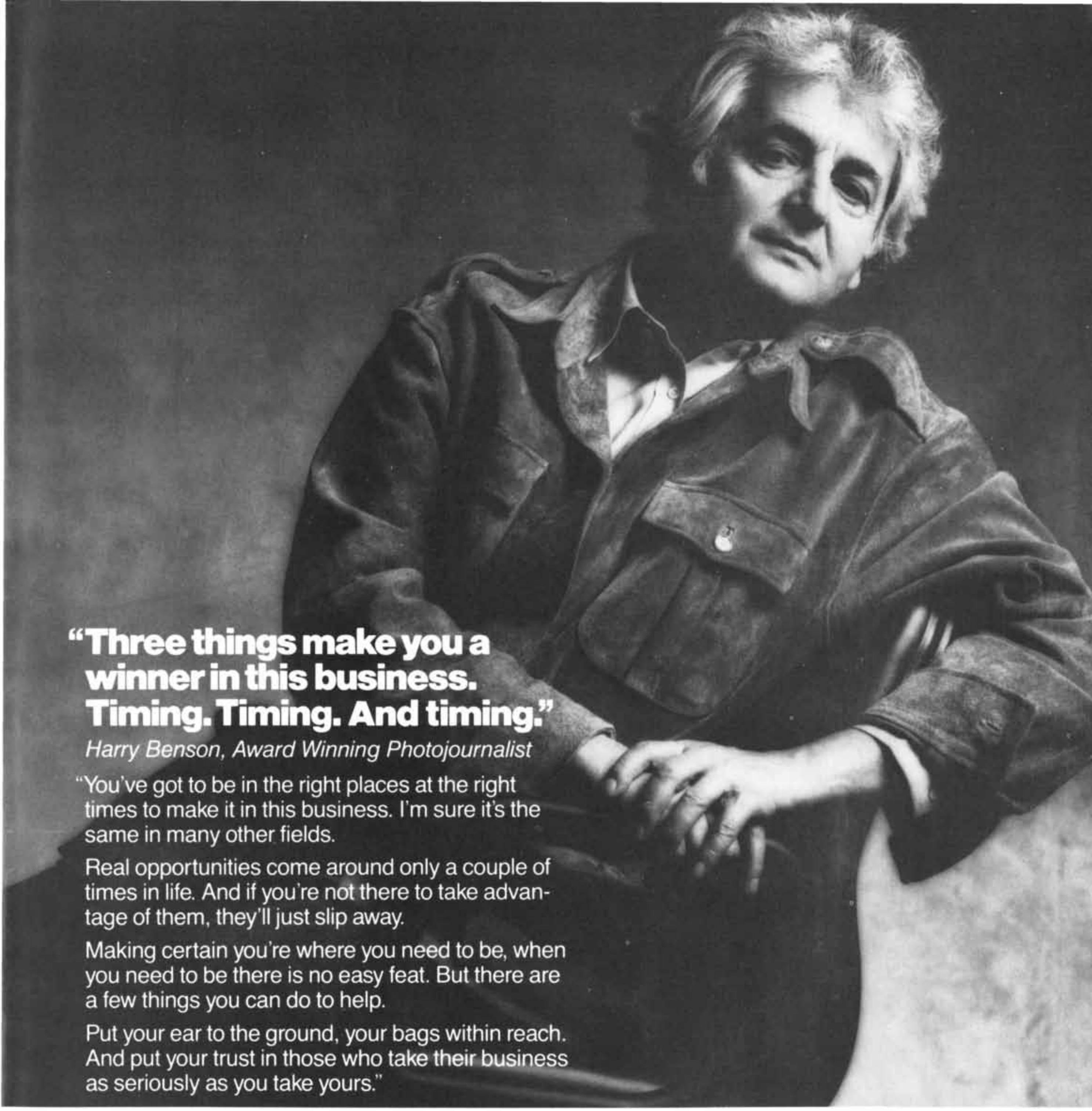
FRACTO-EMISSION FROM POLYMERS, CRYSTALS, AND INTERFACES. J. T. Dickinson and L. C. Jensen in *Proceedings of the Society of Photo-Optical Instrumentation Engineers,* Vol. 743, pages 68-75; January, 1987.

AUTOGRAPHS FROM PEELING PRESSURE SENSITIVE ADHESIVES: DIRECT RECORDING OF FRACTURE-INDUCED PHOTON EMISSION. J. T. Dickinson and E. E. Donaldson in *The Journal of Adhesion,* Vol. 24, Nos. 2-4, pages 199-220; December, 1987.

COMPUTER RECREATIONS

MOTS D'HEURES: GOUSSES, RAMES. THE D'ANTIN MANUSCRIPT. Luis d'Antin van Rooten. Penguin Books, 1980.

The Annual Index, heretofore published in December, will appear in the January issue.

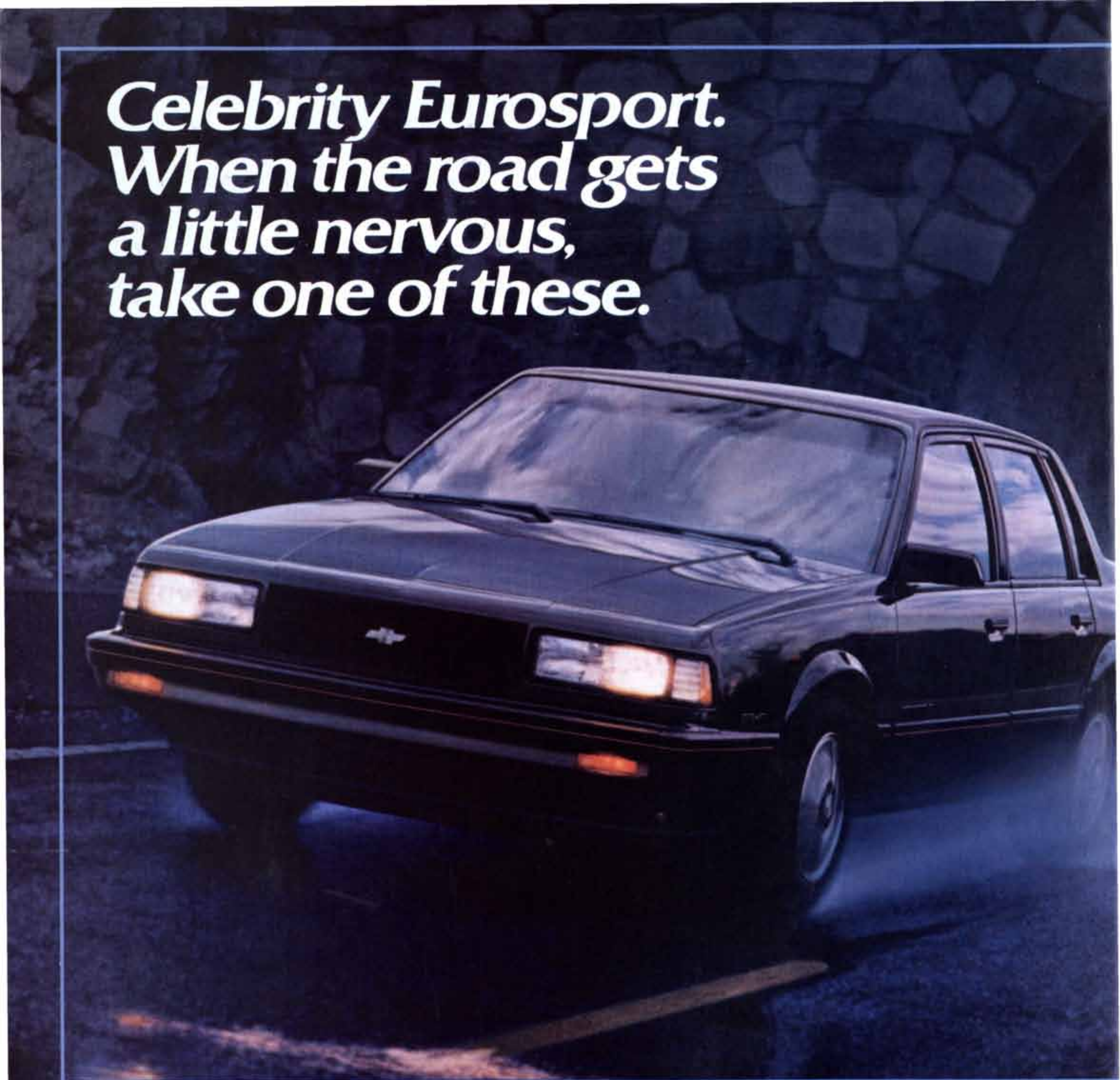


No one is more serious about your international travel needs than Lufthansa. Our outstanding on-time arrival record will tell you that. So will our meticulously maintained fleet and our friendly, courteous service. That's why in 161 cities around the world, when people need to be in the right place at the right time, they fly Lufthansa.

People expect the world of us.

Lufthansa
German Airlines

Lufthansa is a participant in United's Mileage Plus program and Delta's Frequent Flyer program. See your Travel Agent.



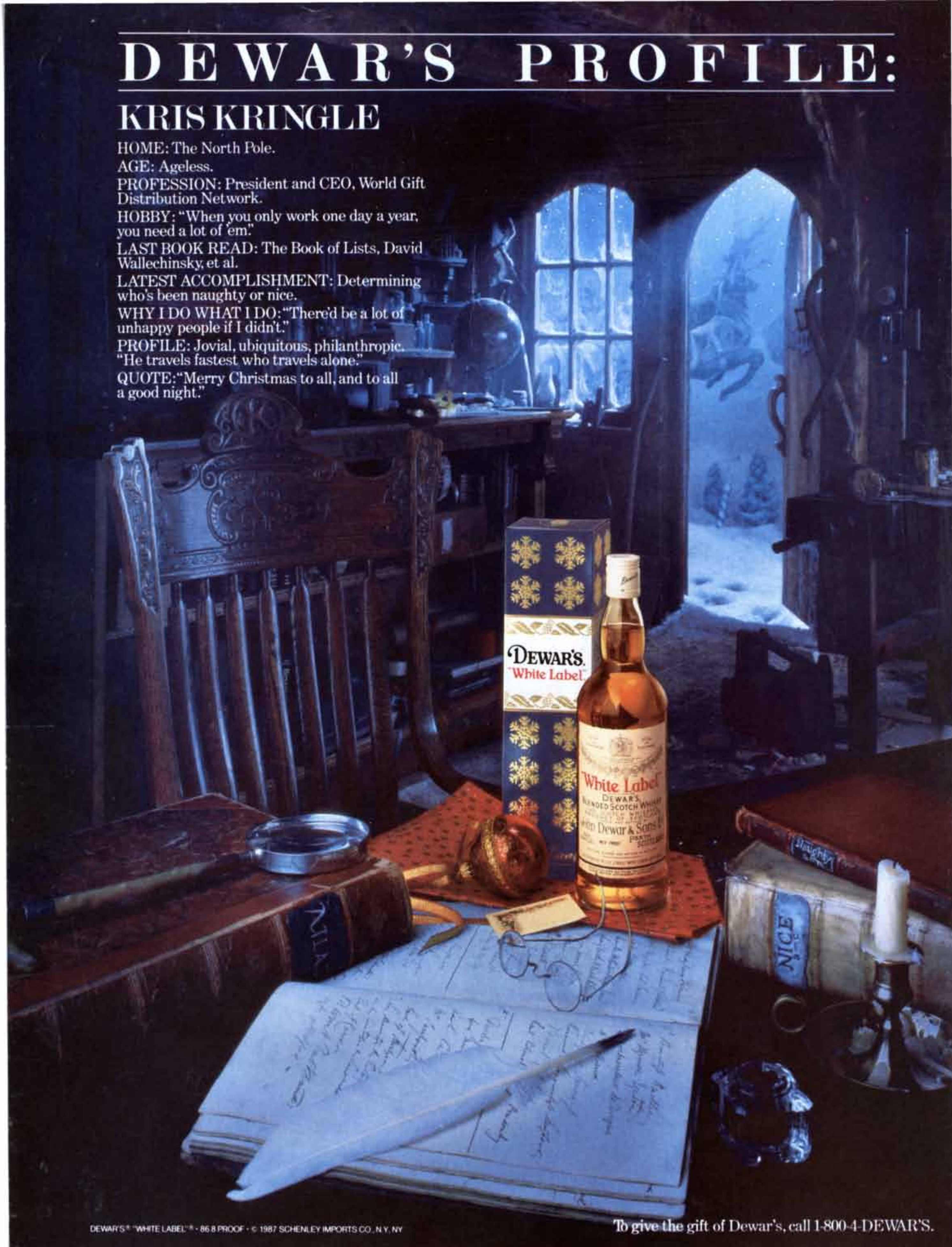
There are two reasons why Celebrity Eurosport was designed as a road car: To make dry, twisty back roads more exciting. And to make wet, twisty, dark back roads...somewhat less so.

Oddly enough, it's on the latter that you might appreciate Eurosport's performance most.

With large stabilizer bars designed to do precisely that. Stabilize. With performance-tuned struts and shocks to keep more of your tires on the road for maximum adhesion. And with available P195/70R-14 Eagle GT + 4 all-season radials all-around, that's a lot of adhesion.

Of course, we don't say that driving a Celebrity Eurosport will make you actually *want* to venture out in inclement weather. But it just might make you feel a bit more secure when you have to.


Let's get it together buckle up.
GM 6/60
QUALITY COMMITMENT PLAN
Front-drive traction and fast-ratio rack-and-pinion steering.
Heavy-duty front-and-rear anti-sway bars for road stability.
Performance-tuned struts and shocks.
Available P195/70R-14 Goodyear Eagle GT + 4 all-season radials.
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


DEWAR'S PROFILE:
KRIS KRINGLE
HOME: The North Pole.
AGE: Ageless.
PROFESSION: President and CEO, World Gift Distribution Network.
HOBBY: "When you only work one day a year, you need a lot of 'em."
LAST BOOK READ: The Book of Lists, David Wallechinsky, et al.
LATEST ACCOMPLISHMENT: Determining who's been naughty or nice.
WHY I DO WHAT I DO: "There'd be a lot of unhappy people if I didn't."
PROFILE: Jovial, ubiquitous, philanthropic. "He travels fastest who travels alone."
QUOTE: "Merry Christmas to all, and to all a good night."
DEWAR'S
"White Label"
DEWAR'S® "WHITE LABEL"® - 86.8 PROOF - © 1987 SCHENLEY IMPORTS CO., N.Y., NY
To give the gift of Dewar's, call 1-800-4-DEWAR'S.