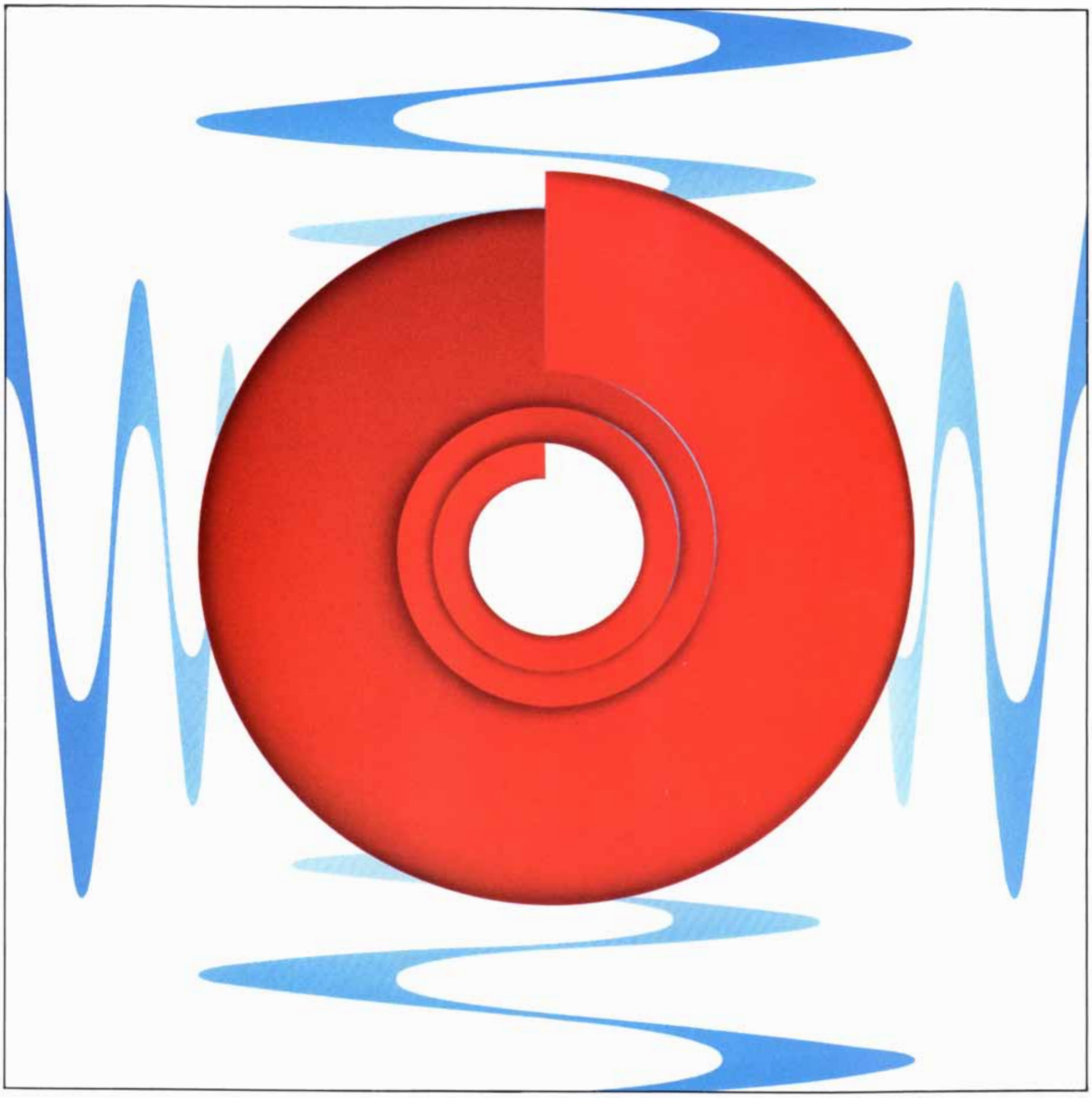
# SCIENTIFIC AMERICAN

MAY 1988
$2.50

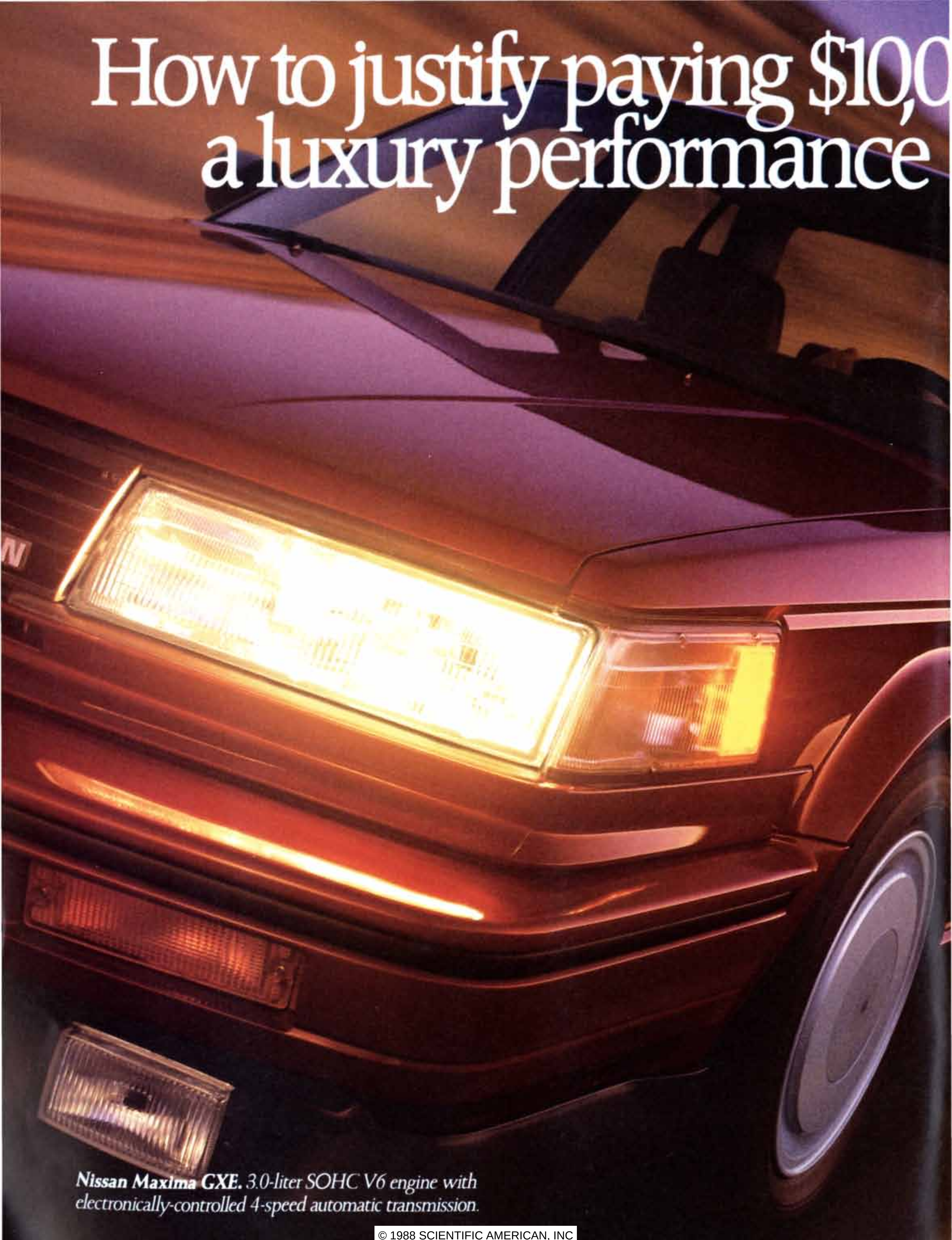
*Exploiting the body's own cancer remedy.*

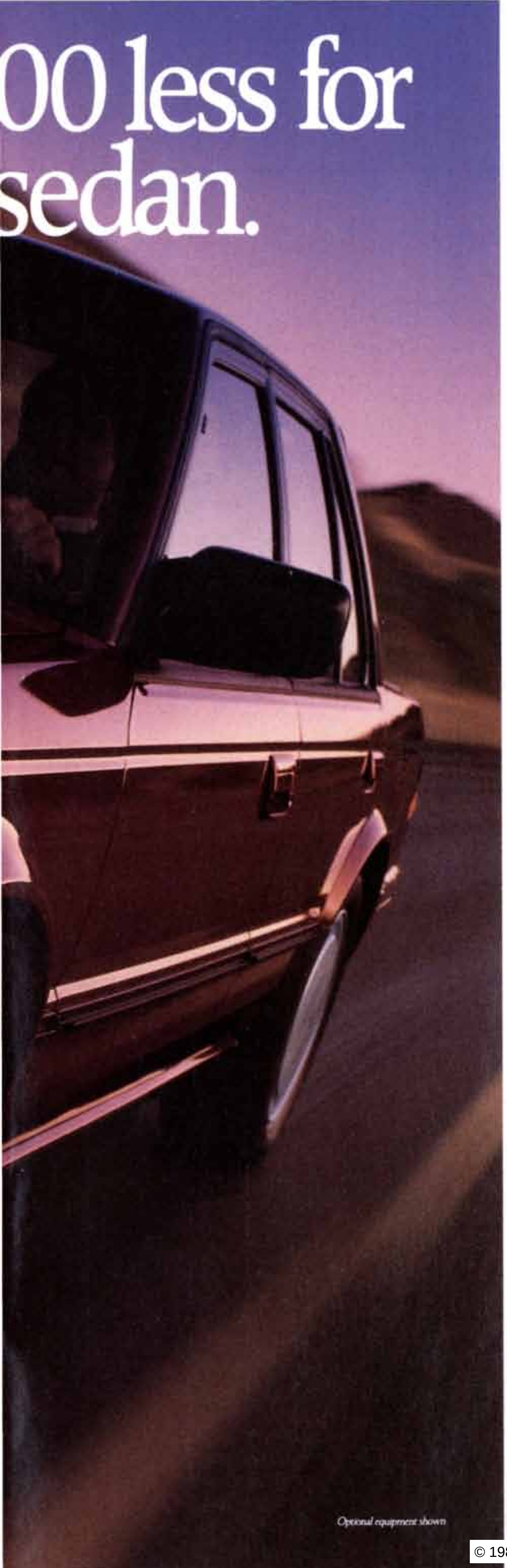
*Hazards of indoor air pollution.*

*Is the earth's magnetic field decaying?*



***Squeezed light*** *could enable physicists to evade quantum uncertainty in fundamental measurements.*


How to justify paying $10,0
a luxury performance
Nissan Maxima GXE. 3.0-liter SOHC V6 engine with
electronically-controlled 4-speed automatic transmission.




For some people, status means having the privilege of spending $27,000 for a European car.

Nissan has engineered a car for people who find status in something else: Just being smart.

The new Nissan® Maxima® GXE. A car *Motor Trend Magazine* called "The best performance sedan you can buy for under $25,000."* Which we take as an enormous compliment, since our car costs far less than that.

But as flattered as we are, we're not surprised. Because at the heart of the Maxima is the same basic fuel-injected V6 engine you'll find at the heart of our famous Z-Car.

And thanks to our revolutionary Sonar Suspension,™** the Maxima GXE not only monitors your driving behavior, it actually "reads" the road surface and automatically adjusts itself for improved comfort and handling.

Naturally, the Maxima also performs well in other areas. It has power windows, locks, mirrors, and antenna. As well as air-conditioning and a six-speaker electronically-tuned AM stereo/FM stereo radio and cassette deck.

Look at it this way. You could either buy a pricey, European luxury sedan. Or, for the same money, you could buy yourself a Nissan Maxima. And then go on a pricey European vacation.

NISSAN®

**Built for the Human Race.™**

*Based on manufacturer's suggested retail price. **Optional, but definitely worth it.*

BELT YOURSELF

SCIENTIFIC AMERICAN
May 1988 Volume 258 Number 5

## DEPARTMENTS

THE COVER depicts electric and magnetic fields making up a beam of "squeezed" light (see "Squeezed Light," by Richart E. Slusher and Bernard Yurke, page 50). Vertical waves represent the magnetic field, horizontal waves the electric field, the spiral their union. Because of quantum uncertainty the fields have no definite value at any time: they are swaths rather than thin lines. They are "squeezed": parts of each are more definite (thinner) than other parts.

## THE ILLUSTRATIONS

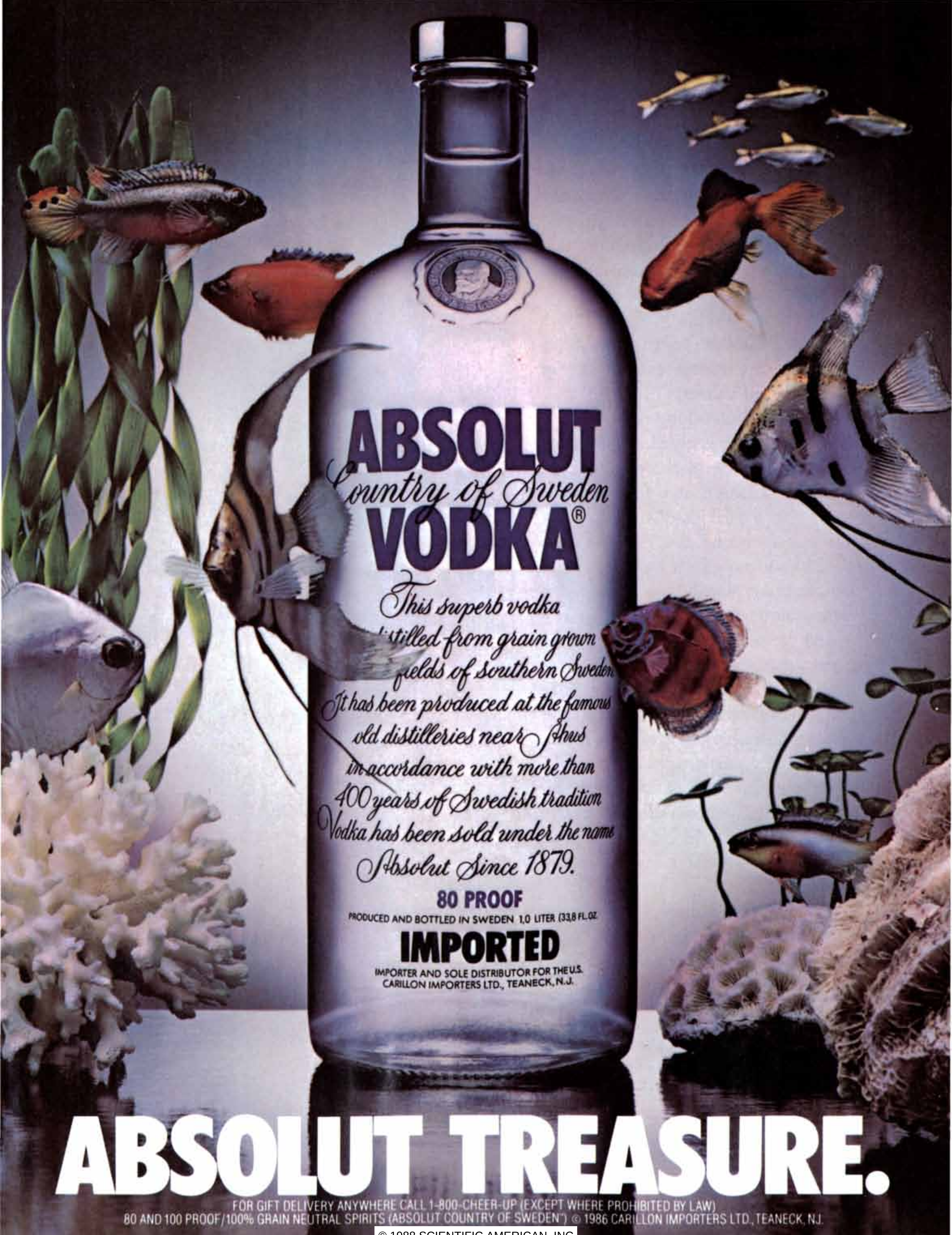
Cover painting by George V. Kelvin, Science Graphics

| Page | Source | Page | Source |
|---|---|---|---|
| 43 | David Scharf | 77-83 | Ian Worpole |
| 44-48 | George Retseck | 85 | M. Prenzel, Animals Animals |
| 51-56 | George V. Kelvin, Science Graphics | 86-89 | Patricia J. Wynne |
| 59 | Lloyd J. Old | 90 | David C. Fritts, Animals Animals (*top*); Mervyn Griffiths (*bottom*) |
| 60 | Helen Coley Nauts, Cancer Research Institute | 91 | Patricia J. Wynne |
| 69 | Carol Donner (*top*), Lloyd J. Old (*bottom*) | 93 | Quesada/Burke |
| 70 | Andrew Christie | 94 | Thomas C. Moore (*top*); Wolfram Görlitz, University of Würzburg (*bottom*) |
| 71 | Carol Donner (*top*); Michael A. Gimbrone, Jr., Harvard Medical School (*bottom*) | 95-97 | Thomas C. Moore |
| 72 | Murray Greenfield | 99 | National Park Service |
| 73 | Carol Donner | 100-103 | Tom Prentiss |
| 74 | Edward A. Havell and Robert J. North, Trudeau Institute, Inc. (*top*); Barbara D. Williamson, Memorial Sloan-Kettering Cancer Center (*bottom*) | 104 | The Champlain Society |
| 75 | Murray Greenfield | 107-112 | Joe Lertola |
| | | 115 | Charles A. Knight |
| | | 116-117 | Michael Goodman |
| | | 118 | David Wiseman |
| | | 119 | Stanley S. Miller |
| | | 120 | David Wiseman (*top*), Bridge Software (*bottom*) |
| | | 121 | Laurie Grace, courtesy of Lascaux Graphics |

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017. Copyright © 1988 by Scientific American, Inc. All rights reserved. Printed in the U.S.A. No part of this issue may be reproduced by any mechanical, photographic or electronic process, or in the form of a phonographic recording, nor may it be stored in a retrieval system, transmitted or otherwise copied for public or private use without written permission of the publisher. 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.

**SCIENTIFIC AMERICAN**

*Established 1845*


ABSOLUT
Country of Sweden
VODKA®
This superb vodka
distilled from grain grown
fields of southern Sweden
It has been produced at the famous
old distilleries near Åhus
in accordance with more than
400 years of Swedish tradition
Vodka has been sold under the name
Absolut Since 1879.
80 PROOF
PRODUCED AND BOTTLED IN SWEDEN 1,0 LITER (33,8 FL.OZ.
IMPORTED
IMPORTER AND SOLE DISTRIBUTOR FOR THE U.S.
CARILLON IMPORTERS LTD., TEANECK, N.J.
ABSOLUT TREASURE.
FOR GIFT DELIVERY ANYWHERE CALL 1-800-CHEER-UP (EXCEPT WHERE PROHIBITED BY LAW)
80 AND 100 PROOF/100% GRAIN NEUTRAL SPIRITS (ABSOLUT COUNTRY OF SWEDEN™) © 1986 CARILLON IMPORTERS LTD., TEANECK, N.J.

# LETTERS

To the Editors:

In "The Moon's Ancient Magnetism" [SCIENTIFIC AMERICAN, December, 1987], S. K. Runcorn describes the painstaking research he and others have done on the relict magnetism of lunar rock samples. He explains how the results, together with data from the Apollo global magnetic surveys, indicate that the ancient moon may have possessed a powerful dipole field.

One remarkable finding is that the polar axis of this ancient field, and by inference the axis of lunar rotation as well, changed position dramatically with respect to the lunar surface as many as three times. Runcorn goes on to propose a possible explanation for the moon's polar wanderings (or, more correctly, its surface wanderings). In doing so, I think his compass becomes unreliable, so to speak.

He proposes that lunar satellites in gradually decaying equatorial orbits fragmented and impacted on the moon's equatorial zone on three distinct occasions between 4.2 and 3.85 billion years ago. On each occasion this bombardment is said to have redistributed enough surface mass to destabilize the moon's rotation, leading to polar wandering.

Now, it is well known that lunar satellites, both natural and artificial, would impact from a slowly decaying orbit at a velocity no greater than 1.8 kilometers per second—much less than the cosmic velocities of 15 or more kilometers per second usually specified for the impacting objects said to have formed the basins.

Are we really expected to believe that an object striking the surface at a velocity of 1.8 kilometers per second and an angle of five degrees or so is capable of creating a great lunar basin such as Imbrium, Serenitatis or even Crisium? It is my belief that a lunar satellite impacting on the equatorial zone would actually add mass to it, thereby improving the stability of the rotational axis. And what about the legendary "mascons"? Analyses of the orbital paths of lunar probes have indicated the presence of significant mass concentrations beneath many of the Maria—Imbrium, Serenitatis, Crisium, Nectaris, Humorum and Orientale—in direct opposition to the theory being proposed.

Personally, I make no apology for supporting the theory that the moon of today is largely the product of forces from within rather than from without. I also acknowledge that I am in a decided minority. The impact theory is championed far and wide, but its cause and the cause of planetary science in general can hardly be furthered by the kind of woolly reasoning contained in this theory of ancient lunar satellites.

ROBIN W. STEVENS

Tonbridge, Kent
England

To the Editors:

S. K. Runcorn's enjoyable article "The Moon's Ancient Magnetism" describes the basis for his imaginative interpretation of lunar magnetic anomalies in terms of impacting lunar satellites, lunar polar wander and a reversing lunar dynamo. The reader is able to judge for himself how far the data require this particular interpretation. The assertion on the cover of the issue [December, 1987] that "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," however, should not pass without comment.

Runcorn was generous enough to mention work done at the University of California, Santa Barbara, by Stanley M. Cisowski, J. R. Dunn and me. It has focused on the magnetism of the returned samples of lunar rock, with the goal of understanding the mechanism by which the magnetism was acquired and the strength of the ambient fields in which the magnetization took place. Unfortunately, much of the initial work was done before recent developments in paleomagnetic techniques and has not been repeated, and so the analysis remains somewhat unsatisfactory. Moreover, as Runcorn pointed out, standard methods of determining the intensity of the magnetizing fields do not work well with lunar samples. Nevertheless, it is very clear that reports of magnetization requiring fields with a strength as much as twice that of the earth's field are rare indeed and also controversial. As we have discussed elsewhere, values an order of magnitude smaller would be a more realistic upper limit on the field strength during the proposed lunar magnetic era between about 3.9 and 3.6 billion years ago. Hence "may" is the operative word in the cover caption.

No matter how difficult interpreting lunar paleomagnetism proves to be, one cannot help feeling some awe at a record preserved for almost four billion years. At present it may be wise to regard the moon's magnetism as a fascinating enigma. Such a conservative approach does not have the appeal of Runcorn's fascinating discussion. It would be a pity, however, to misinterpret the priceless message that has come to us over the ages.

M. FULLER

Department of Geological Sciences
University of California
Santa Barbara

To the Editors:

Clearly if volcanic rather than impact processes account for the circular craters and basins of the moon, as Mr. Stevens believes, then my interpretation of lunar paleomagnetism is moonshine. Harold Urey used to preface persuasive accounts of events in the early solar system disarmingly: "Now, I was not there at the time but..." Finding diagnostic evidence in this field is challenging.

In "The Geologic History of the Moon" (U.S. Geological Survey Professional Paper No. 1348, 1987) Don E. Wilhelms of the U.S. Geological Survey demonstrates conclusively the fundamental importance of impact processes. Only recently has a piece of evidence testing my theory been identified: the great multi-ring basins have slight asymmetries—an elliptical shape and ejecta concentrated in a "butterfly" or "bow tie" pattern—similar to those produced in hypervelocity impact experiments or observed in missile craters when the projectile's approach angle is less than from five to 10 degrees. The velocity vectors of bodies that formed the multi-ring basins are found to parallel the paleoequators of corresponding age. This is strong evidence that the basins were formed by satellites in decaying equatorial orbits, even though, as Stevens implies, much remains to be learned about the physics of these huge impacts.

The congruence of the paleomagnetic data, the evidence from mechanics, and Euler's principle (that a body tends to rotate about its axis of maximum moment of inertia) all strongly support the theory I outlined. Like Dr. Fuller, I am awestruck by the extreme age of the fossil magnetism in lunar rocks. But I think one should not be dumbfounded. Supreme Court justice Oliver Wendell Holmes wrote: "The law is an experiment, as all life is an experiment." He might as well have been writing about scientific theories, which are experiments to see how far observations, often apparently conflicting, can be understood rather than left as "enigmas." One risks being refuted, but that is how science advances.

I and my colleagues at the University of Newcastle upon Tyne, David W. Collinson and Alan Stephenson, agree with Fuller that more laboratory experiments to determine the strength of the field that magnetized the lunar rocks are urgently required if we are to understand the rise and eventual decay of a core dynamo in the moon. We agree that the exact calibration of the method developed by Fuller and Cisowski can be debated. Their view that the lunar field never reached an intensity as high as what we measured using the Thellier-Thellier method, however, does not undermine a conclusion with which we would all agree: considering the smallness of the moon's iron core compared with the earth's, it did indeed show remarkably powerful dynamo action between three and four billion years ago.

S. K. RUNCORN


CIVILIZATION
HAS NEVER
ADVANCED
QUITE SO
RAPIDLY.

# 50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

MAY, 1938: "On a flat, dry plain 18 miles north of Roswell, New Mexico, rises a 60-foot tower of steel that has roused more curiosity, and has probably had a greater influence on the future of the world, than any other feature of all New Mexico's arresting landscape. From this tower, at irregular intervals, a Massachusetts physicist and his assistants send roaring into the skies certain gleaming, cigar-shaped projectiles of metal, powered by gasoline and liquid oxygen and landed by parachutes. The physicist is Dr. Robert Hutchings Goddard."

"Sand-blown and desolate ruins of the Lost City of Nevada will soon rub elbows with the fishes. The continued rise of the water behind Boulder Dam will shortly form a great lake that will cover the area of the ancient city completely. The Lost City was probably an active center of Indian habitation about A.D. 800."

"In Laurel, Mississippi, in the heart of one of the South's principal sweet-potato-growing sections, there is being manufactured a product that bids fair from present indications to form the basis of a flourishing new southern industry—the production of sweet-potato starch. Long existing only in the minds of chemists, sweet-potato starch of a purity, color, and quality equal to that of the finest starches is now an actuality: hundreds of thousands of pounds are being produced commercially."

"How far can one read a newspaper by the world's most powerful light? A group of technicians sought the answer to this question recently, when the 2,000,000,000-candlepower beacon atop the Colgate-Palmolive-Peet Building in Chicago was turned into the world's largest reading lamp for 90 minutes. Flying at 7,000 feet over Chicago, passengers on board a special United Air Lines' *Mainliner* were able to read a newspaper by the light of the huge airway beacon at a distance of 27 miles."

"There was a time, not so long ago, when automobile engine bearings had to be taken up after about 15,000 miles of service. Today the motorist seldom gives a thought to bearings. He expects fine performance for an indefinite number of miles and gets it because the automobile is the beneficiary of progress in the field of abrasives. The slow substitution for natural abrasives of artificially created materials that can be controlled as to nature and performance have made abrasives the handmaiden of mechanical precision."

SCIENTIFIC AMERICAN

MAY, 1888: "During the late violent storms in the Channel the waves washed through a high and hard sand bank near the Isle of St. Malo, France, nearly four meters thick, laying bare a portion of an ancient forest that was already passing into the condition of coal."

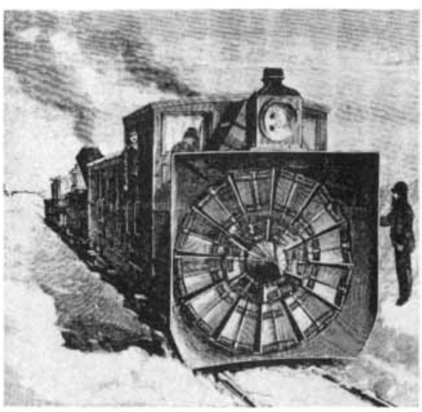
"A modified form of the phonograph has been perfected by Mr. Emile Berliner, of Washington, D.C., who styles his new instrument the 'gramophone.' One of the distinguishing features of this invention is that the indentations of the transmitting diaphragm are made upon a flat plate instead of a cylinder, as in the Edison and Bell phonographs."

"Mr. H. Sullivan Thomas, who has been lecturing on the mosquito before the literary society of Madras, is ungallant enough to say that it is only the female that does the biting."

"It has been ascertained that of the millions of bison that once roamed on the prairies of the West only seventy-five or a hundred remain, and these are located in the extreme southwestern portion of Texas. An expedition is soon to start for Texas to round up buffalo there. The leading purpose is to perpetuate a species of animal that is thoroughly typical of American animal life, one of the controlling ideas of the trip being to kill none of the animals while corralling them or after their capture."

"By recent advances in the study of celestial physics, especially with the spectroscope, we are enabled to know something of the actual condition and relative ages of some of the stars. We may assume that the age of stars, other things being equal, will depend upon their temperature, and that their temperatures are higher in proportion as their spectra are richer in violet rays. The majority of the stars that are visible to the naked eye are white or bluish, and therefore at a high temperature; but many are yellow or orange, like our sun, showing that they have passed their youth, while others are from dark orange to dark red, showing that their sidereal evolution is far advanced."

"By far the most satisfactory method yet devised of removing snow from railway tracks is found in the rotary steam snow shovel. The machine, with the engines and boiler, is mounted on an eight-wheeled car. It has a weight of some fifty tons, or about that of an ordinary locomotive, and is pushed into the snow by one or two locomotives. The operation of the machine is described as being a marvelous sight. Such is the tremendous centrifugal force of the wheel that the snow is discharged as a great stream or cloud and hurled to a distance of from one hundred to three hundred feet from the track."



*The new rotary steam snow shovel, shown idle* (left) *and at work* (right)

LANCE NEIBAUER AND JOHN MARCONI. NEW ACHIEVERS IN SCIENCE.

*we never stop asking*

Lance Neibauer. Aircraft designer. His award-winning Lancair kitplanes are the nation's top-selling homebuilts. Now using the new HP-28S, the only calculator that can do symbolic algebra and calculus, retrieve and combine graphs instantly. HP Solve lets him enter his own formulas. More than 1500 functions and 32K bytes of RAM make it the world's most powerful scientific calculator. John Marconi. Air pollution control engineer. Project leader for study on cancer-causing air pollutants. Needs to solve statistical and financial problems. Perfect job for the new HP-27S. It's the only scientific calculator with business functions. It also offers easy-to-follow menus and optional solutions books. Hewlett-Packard's new range of calculators is built for your success. For the name of your nearest Hewlett-Packard dealer, call 1-800-752-0900, Ext. 142C

# Innovation takes a calculator packed with new ideas.

NEW ACHIEVERS: THE HP-28S AND HP-27S SCIENTIFIC CALCULATORS.

HEWLETT PACKARD

# THE AUTHORS

ANTHONY V. NERO, JR. ("Controlling Indoor Air Pollution"), is principal investigator and coleader of the building-ventilation and indoor-air-quality program at the Lawrence Berkeley Laboratory (LBL) of the University of California, Berkeley. After completing a doctorate in physics at Stanford University in 1971 and postdoctoral work at the California Institute of Technology, he was made assistant professor of physics at Princeton University. He has been at the LBL since 1975. Nero, who spent a year at the U.S. Arms Control and Disarmament Agency and has published a guidebook to nuclear reactors, says that over the years he has become increasingly interested in how society makes decisions on technological, environmental and health issues.

RICHART E. SLUSHER and BERNARD YURKE ("Squeezed Light") are experimental physicists at the AT&T Bell Laboratories. Slusher earned his Ph.D. at the University of California, Berkeley, in 1965. He has studied interactions between laser light and many forms of matter: semiconductors, solid helium, tokomak plasmas and beams of atoms. He is now head of the department of solid-state and quantum physics at the Bell Laboratories, where he is seeking to understand the fundamental limits of optical nonlinearities in materials. Yurke got his Ph.D. from Cornell University in 1983; he went to the Bell Laboratories to develop his many ideas for generating and detecting squeezed light. These ideas matched Slusher's experimental resources and a fruitful collaboration was born. Yurke's current experiments are aimed at generating squeezed microwave radiation at ultralow temperatures.

LLOYD J. OLD ("Tumor Necrosis Factor") holds the William E. Snee Chair of Cancer Immunology at the Memorial Sloan-Kettering Cancer Center in New York. He has worked there in a variety of capacities since 1958, the year he received his M.D. at the University of California School of Medicine. From 1973 to 1983 he served as vice-president and associate director of scientific development at the Sloan-Kettering Institute for Cancer Research. His early experience as a fellow at Sloan-Kettering convinced him to dedicate himself to cancer research rather than clinical practice. For the past 20 years he has focused much of his attention on a search for cancer-specific antigens. The work has led to a detailed study of antibody responses to tumor cells in patients who have melanoma and to the development of a melanoma-vaccine program. Old is also medical director of the Cancer Research Institute, an organization that supports fellowships and grants in the area of cancer immunology.

KENNETH A. HOFFMAN ("Ancient Magnetic Reversals: Clues to the Geodynamo") is professor of physics at the California Polytechnic State University in San Luis Obispo and an editor of the *Journal of Geophysical Research* (Solid Earth and Planets). He got his bachelor's degree in physics in 1966 and his Ph.D. in geophysics in 1973, both from the University of California, Berkeley. He then did postdoctoral work at the University of Minnesota before going to California Polytechnic in 1974. Hoffman, who has spent much of the past decade studying the process of geomagnetic reversal, has been an author or coauthor of the past three quadrennial reports on geomagnetic reversal submitted by the U.S. national committee of the Union of Geodesy and Geophysics to the international organization. He is currently conducting paleomagnetic investigations on Molokai in Hawaii and on Réunion Island in the Indian Ocean.

MERVYN GRIFFITHS ("The Platypus") retired in 1975 as a principal research officer in the Division of Wildlife at the Commonwealth Scientific and Industrial Research Organization (CSIRO) in Australia. He now studies milk secretion by monotremes and marsupials as an honorary research fellow at the CSIRO and as a visiting fellow at the John Curtin School of Medical Research of the Australian National University. He got his bachelor's, master's and doctoral degrees from the University of Sydney between 1937 and 1959; during that period he also won a scholarship that allowed him to study physiology at other institutions, and he served for four years in the Australian air force. He went to the CSIRO in 1959 and it was there that he became particularly interested in the evolution of marsupials and egg-laying mammals.

JOCHEN FRICKE ("Aerogels") is professor of experimental physics at the University of Würzburg in West Germany. He received a Ph.D. in physics in 1967 from the Technical University of Munich. He then spent two years as a postdoctoral research fellow at the University of Pittsburgh. He has been editor of the German physics magazine *Physik in unserer Zeit* since 1970. Fricke's research interests now center on the optimization of thermal superinsulations, the properties of highly porous solids and questions of scaling.

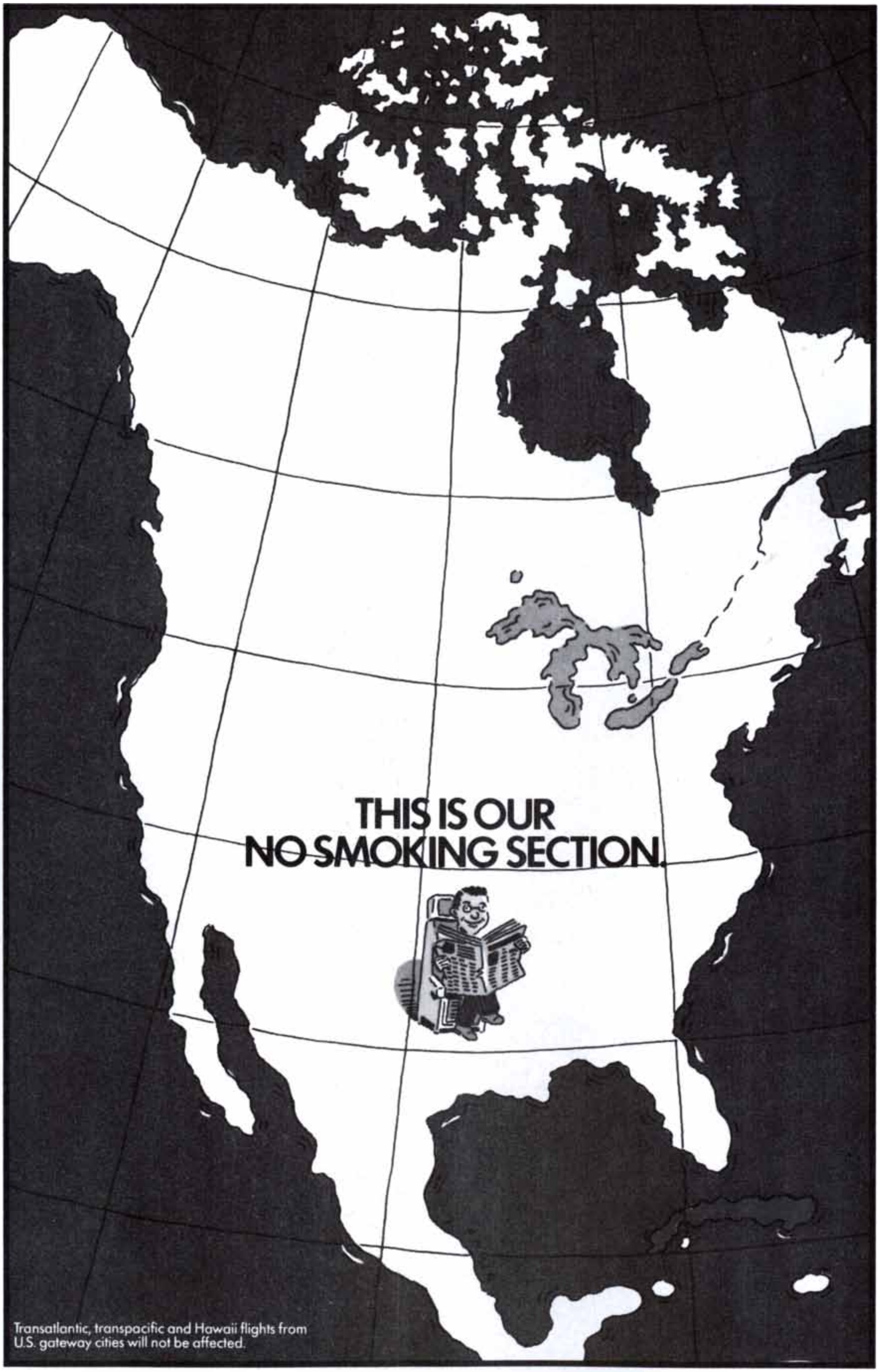
FRANCIS P. MCMANAMON and JAMES W. BRADLEY ("The Indian Neck Ossuary") worked together on the National Park Service's excavations of the ossuary and other sites on the Cape Cod National Seashore. McManamon, who is chief of the Park Service's Archaeological Assistance Division, got his Ph.D. from the State University of New York at Binghamton in 1984. He has a long-standing interest in the prehistory of northeastern North America, archaeological techniques, and historic preservation, and he recently began to study the archaeology of the Pacific coast. Bradley is director of preservation planning for the Massachusetts Historical Commission. He holds a doctorate from Syracuse University and has written extensively about early contact between native Americans and Europeans and the resulting processes of cultural change. Bradley's book *Evolution of the Onondaga Iroquois; Accommodating Change: 1500–1655* was recently published by the Syracuse University Press.

LARRY ABBOTT ("The Mystery of the Cosmological Constant") is associate professor of physics at Brandeis University. He earned his Ph.D. in physics from Brandeis in 1977. He joined the faculty there in 1979 after working at the Stanford Linear Accelerator Center and at CERN, the European laboratory for particle physics in Geneva. Abbott writes: "Over the past several years I have been interested in the application of new ideas in particle physics to cosmology and have worked on inflationary cosmology, dark matter and the large-scale structure of the universe. At the same time I have become intrigued by the problem of the cosmological constant. More recently I have studied the implications of the observation of neutrinos from supernova 1987A and have become interested in the physics of neural networks."

Northwest was the first to stop smoking on all flights in North America. We are committed to giving you the service you want. Unlike other airlines who prohibit smoking on flights under 2 hours, we want you to have a smoke-free environment on all flights in North America.

So if you don't smoke and don't like it when others do, we're behind you all the way. No butts about it.

NORTHWEST




**LOOK TO US NORTHWEST AIRLINES**

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

# SCIENCE AND THE CITIZEN

## The Violent Yanomamö

*A new study rekindles a debate over the roots of warfare*



In a benign environment isolated from modern civilization, what kind of society is likely to develop? Not long ago optimists could point to the "gentle Tasaday," inhabitants of a Philippine rain forest who were supposedly so peaceful that their language lacked a word for anger. Unfortunately the Tasaday, "discovered" in 1971, recently have been reported to be a hoax. Some anthropologists now cite a grimmer model: the Yanomamö, a tribe whose 15,000 members inhabit some 200 villages in Brazil and Venezuela, deep in the Amazonian jungle.

The Yanomamö are known largely through the writings of Napoleon A. Chagnon of the University of California at Santa Barbara, who has studied them for 24 years. His 1968 book *Yanomamö: The Fierce People* described a potentially idyllic world, offering more than enough game and edible flora to feed all the villagers. Nevertheless, the polygamous males who dominated the villages fought viciously and often; they seemed to revel in ferocity. The book and subsequent reports on the Yanomamö by Chagnon have been hailed by scientists—some now known as sociobiologists—seeking a genetic basis for social behaviors. But critics accused Chagnon of ignoring environmental factors that may have spurred the Yanomamö to fight and so of implying that war is inevitable.

Chagnon, who has spent some 50 months living with the Yanomamö, has rekindled the controversy with a report in *Science*. He documents the violence anew: 30 percent of the males die violently and almost half of those who are age 25 or older have participated in a killing. Chagnon suggests that a brutal, Darwinian logic underlies the killing. Through violence a Yanomamö male seems to enhance his reproductive success and that of his kin: he becomes "fitter."

Killings usually result from protracted and apparently senseless feuds between villages. Yet these feuds, Chagnon notes, always stem originally from competition for "reproductive resources," or women. Men from one village kidnap women from another village, or two men in the same village fight over a woman both of them desire. In the latter case, if one man is killed, the village often "fissions": close relatives of the victim leave to establish their own village. The victim's male relatives, armed with bows and arrows, then raid the original village to exact vengeance. The resulting feud of "blood revenge" may last for years.

Blood vengeance makes sense in terms of kin-selection theory, according to Chagnon. A man who quickly avenges an assault on his kin, he explains, may develop a reputation that deters attacks on them and therefore on his village (the two are often synonymous). Such a man is highly esteemed by both men and women in the village. Indeed, Chagnon has recently determined—he says the finding was "startling" even to him—that killers have on the average two and a half times as many wives and three times as many children as nonkillers. Moreover, warriors apparently do not run an inordinately high risk of dying in a raid. Raiders usually sneak up on a village, kill the first one or two people they encounter and quickly flee, Chagnon reports. Both the attacking warriors and those in the attacked village are likely to survive such a raid.

Chagnon says successful Yanomamö warriors are not necessarily more genetically "predisposed" toward aggression than others; they may simply be better at learning how to use violence to achieve success in their community. (Chagnon says he has been falsely accused of claiming that there is a "warfare gene.") But he suggests the procreative drive that underlies the Yanomamö's violence

*Yanomamö killers have twice as many wives and three times as many children as nonkillers*

*YANOMAMÖ WARRIORS holding bows and arrows prepare for a raid on another village. Prior to the raid each man stands before the other warriors and declares what harm he intends to inflict on the enemy. The photograph is by Napoleon A. Chagnon.*

is indeed innate and common to all humanity. "One thing you can never get enough of," he says, "is sex." Closer observation of other tribal societies, he proposes, would show that the Yanomamö culture is "not atypical." Indeed, he implies that the link between reproductive success and violence in early societies may explain why "military achievements are valued and associated with high esteem...in many other cultures, including our own."

A persistent critic of Chagnon's work is Marvin Harris of the University of Florida. Studies other than Chagnon's, Harris says, indicate that the Yanomamö are subject to environmental stress: as their population density increases, they must expend greater energy obtaining food, and meat in particular. He argues that feuds are an adaptation to this stress, serving to control the population and thereby to prevent food shortages. Competition for material resources, and not the procreative urge, underlies all warfare, primitive and modern, according to Harris. "This proposal that all humans strive to have as many babies as possible, which lies at the heart of the sociobiology paradigm, is flawed," he says. "People are more interested in having TV's than children."

R. Brian Ferguson of Rutgers University says "the intensity of fighting described by Chagnon is very unusual. Even close to the Yanomamö there are other tribes that engage in no warfare whatsoever." The broad range of behavior in primitive societies, contends Ferguson, shows that human nature is much more "plastic" and subject to environmental influences than Chagnon suggests.

But Ralph L. Holloway of Columbia University maintains that a strictly environmental view of warfare is naive, given the growing body of research linking behavior to genetics. "Most anthropologists want this view that there is no end to human possibility," observes Holloway, himself an anthropologist. "But you can't just say it's all environmental variance."

Edward O. Wilson of Harvard University, considered the father of sociobiology, agrees. Unlike Chagnon, who shies from the issue, Wilson directly addresses the possibility that there may be a genetic component to human violence. "I'm really curious about why people pussyfoot around the human aggression element," he says. "Humanity has been wading in blood for as long as it's been around. If we have a strong biological predisposition toward violence, we can't just wish it away." Chagnon's discovery of a "powerful, potentially selective" link between aggression and reproductive success, Wilson suggests, may represent an important step toward understanding and possibly eliminating warfare. *—John Horgan*

## Star Games

*Can computer simulations test the untestable?*

Construction has just begun on a research facility that is crucial to the future of the Strategic Defense Initiative. The setting of the so-called National Test Facility matches the scope of the project's ambition: located in the shadow of the Rockies, within Falcon Air Force Station near Colorado Springs, it will house a staff of almost 2,000 within 325,000 square feet of office space. The facility will be the hub of a nationwide computer network, called the National Test Bed. The Martin Marietta Corporation has won a five-year, $508-million contract to build and operate the network; the total cost may reach $1 billion. Groundbreaking officially took place in March, but research is already under way in temporary quarters. The goal? Simulating the response of an SDI command, control and communications ($C^3$) system to a nuclear attack.

Maintaining an effective command, control and communications system for a multilayered defense consisting of hundreds of ground- and space-based weapons is seen as a formidable task even by SDI champions. The system would have to analyze data from myriad sensors, allow military commanders to select an appropriate response, track thousands of warheads and distinguish them from thousands of decoys, assign weapons to targets and monitor kills, all within minutes.

Such a system could never be tested realistically. Planners will therefore rely heavily on simulation, and the National Test Bed will be one of the most powerful simulation systems ever built. Its heart will consist of two Cray supercomputers and two large IBM machines; it will boast numerous lesser computers and experimental parallel processors. Hooked up to major computing facilities across the country, the National Test Facility will model attack scenarios and defensive responses. It will also simulate various disruptions of its own $C^3$ and the effects of nuclear bursts on sensors.

At present only a small part of a large attack can be modeled, and with quite low fidelity, according to Col. Thomas L. Leib, the National Test Bed program director. Virginia P. Kobler, chief of battle-management $C^3$ at the U.S. Strategic Defense Command, says that with more sophisticated simulations the test bed will probe the effectiveness of the distributed architecture of the $C^3$ operation. Such a system can in principle be built so that failure at one or more points does not bring it down.

How much confidence can one have in such simulations? Aircraft flight simulators, for example, are realistic because there has been extensive experience with how real aircraft behave; there is no parallel experience with multiple nuclear explosions in space, Richard L. Garwin of the IBM Thomas J. Watson Research Center points out.

Often the hardest problem for a defensive system is to define precisely the characteristics of the offensive weaponry. According to David Weiss of the Software Productivity Consortium in Reston, Va., who studied missile-defense software for the Office of Technology Assessment, any change in a weapon requires extensive reprogramming of major parts of the entire simulation. John Shore of the Entropic Processing Corporation in Washington raises the difficulty of unforeseen effects: "What they are unlikely to have enough information about is nature.... You never know that you are simulating at the level of detail that matters in the real world." Some critics see another kind of vulnerability. "I fear it will be a large black box and people will fiddle with inputs until the output is right," Garwin says.

Gerold Yonas, former chief scientist at the Strategic Defense Initiative Organization, says simulation at the National Test Bed will increase U.S. officials' confidence enough to make possible a decision on deployment; moreover, the credibility of the simulation will be crucial because "the Soviets have to feel confident the system will succeed."

"We understand the threat," Leib says; simulation of the defense will evolve in the course of time. Leib points out that the Strategic Defense Initiative Organization has a "Red team" that will try to defeat defenses; in addition the newly established Strategic Defense System Op-


There are some places where gambling is perfectly acceptable.


# It's not just how fast you go.

**GTI 16V** The GTI 16V goes from 0 to 50 faster than a Porsche 944.*

We hope you're unimpressed.

For our GTI is engineered for those who seek out not merely the fast lane.

But the smart lane.

So, besides a highly refined 16-valve, 123-horsepower engine, you'll find the technology to use it wisely.

Handling is so supple, directional control so sure, Car and Driver wrote:

"Driven hard, the GTI comes into its own. The steering is precise and buttery smooth." Talk about grace under pressure.

*Based on comparison of manufacturer performance data.

erational Test Organization will independently scrutinize simulations. "We will do all we can to validate data and test systems," he says. "Data like these are better than no data at all." —*Tim Beardsley*

## Embarrassment of Riches

*Foreign-born students fill U.S. science programs*

Where does the U.S. find its future technical talent? Increasingly the source lies outside our borders. More than a fourth of the science Ph.D.'s and fully 60 percent of the engineering Ph.D.'s awarded in 1986 went to noncitizens, according to figures compiled by the National Science Foundation; a report published in January by the National Research Council found that about two-thirds of postdoctoral appointees in engineering are foreign citizens. Such figures have been increasing since the 1970's, and many observers view them as a reflection of the dismal state of U.S. science education.

About 60 percent of the foreign graduate engineering students stay in the U.S. after getting their degree, according to the NRC report. In 1982 noncitizens constituted 3.5 percent of the engineering labor force, and the number quite likely will grow. In March the Senate passed a bill that would increase by about 20 percent the number of residence visas allotted to people with special skills.

The influx is in some ways a boon. Foreign-born students are typically their countries' best and brightest, and many of them have already completed their undergraduate education abroad. Because those who stay on in the U.S. after receiving their degree "provide definitely needed supplements to our labor force...their absence would lead to curtailment of important programs," the NRC report states.

By the same token, the foreign talent is keen competition for U.S. citizens. Frank E. Lord, chairman of the manpower committee of the Institute of Electrical and Electronics Engineers' U.S. Activities Board, charges that the many engineering graduates from other countries who are eager to stay in the U.S. allow "exploitative employers" to pay lower salaries than they would offer otherwise. The NRC study found that foreign-born engineers earn as much as their U.S.-born colleagues but acknowledged that their availability might lower engineers' salaries generally. Within universities, other observers fear, the foreign competition may deter American students from becoming graduate students or junior faculty members.

In response many universities, including the University of Illinois and the University of California, have limited the number of noncitizens entering some programs. Nevertheless, the presence of foreign-born faculty and students is shaping American higher education. Poor English ability among teaching faculty is widely seen as a problem. Foreign students often come from a more theoretical and less practical background than their U.S.-born counterparts, according to Dorothy S. Zinberg of Harvard University, who worked on the NRC study.

Then there are strategic concerns. Foreign nationals and immigrants with relatives abroad often have difficulty obtaining security clearance. "If we had to mobilize for defense, we would really have a shortfall of clearable people," Lord maintains. A survey of R&D directors in 20 major high-technology companies indicates that they are increasingly dependent on foreign talent.

What to do? "We don't have a problem with too many foreigners—it's too few Americans," Zinberg says. The number of U.S. students earning doctorates in science and engineering fell by 16 percent between 1971 and 1985, whereas the number of foreign citizens studying those subjects rose by 40 percent. Native-born students, in particular engineering students, simply do not favor graduate school. The attractive salaries available to undergraduate engineering majors in industry are a major reason, according to the NRC report.

A more fundamental problem may be the state of science education in high schools. For example, 30 percent of U.S. high schools do not offer a physics course, according to Bassam Z. Shakhashiri, head of science and engineering education at the National Science Foundation. It is in high school, Shakhashiri says, that the curiosity that produces a professional scientist or engineer first buds—and too often dies. In a recent study of students in 17 countries, the International Association for the Evaluation of Educational Achievement found that U.S. 14-year-olds ranked 14th in their knowledge of basic science, tied with Singapore and Thailand; only Hong Kong and the Philippines fared worse. (Hungary was first and Japan second.) U.S. students were also among the worst at age 18.

The NSF is expanding continuing education for teachers, developing teaching materials and forging new partnerships between industry and schools. It is also trying to reverse a recent drop in enrollment of women and U.S. minority groups in engineering programs. The NSF's education budget increased by 40 percent this year, to $139 million, and the Administration has asked for a further increase of 12 percent for 1989. Most of the money is going to precollege programs. Shakhashiri is not complacent: he points out that in real terms the NSF's current spending for science education is only one-third of what it was in the 1960's. —*T.M.B.*

# PHYSICAL SCIENCES

## Shear Magic

*Jupiter's great red spot is conjured up in laboratories*

In 1664 the English scientist Robert Hooke peered through a telescope at Jupiter and discerned an elliptical red spot on its southern hemisphere. Although the spot has waxed and waned over the years, it is still there, bulging within one of the bright, latitudinal "zones" that encircle the planet.

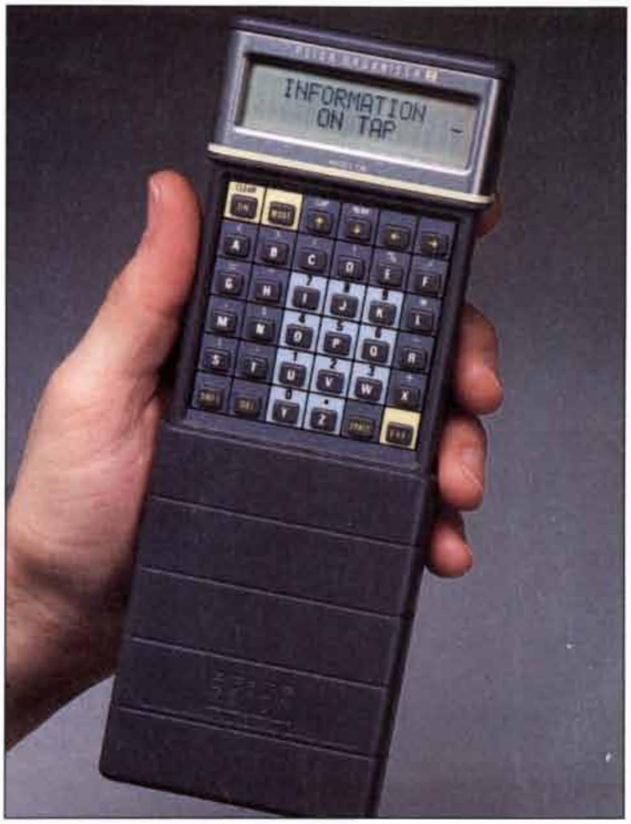
Astronomers at one time conjectured that the red spot, whose diameter is more than twice that of the earth, is somehow related to a protuberance or a hollow on the planet's surface. That notion was rejected after Voyager interplanetary probes in the late 1970's showed Jupiter to be almost entirely fluid, except perhaps for a small solid core. Voyager images also revealed that the spot is a vortex spinning counterclockwise and that the zone encompassing it is extremely turbulent. Winds blow west at the zone's northern border and east at its southern border; they swirl violently in between. How does the spot persist amid such chaos?

Two simulations described in *Nature,* one done on a computer and the other in a water-filled vat, suggest

# THE ORGANISER

## A 3 INCH BY 6 INCH PERSONAL COMPUTER

Imagine being able to have every address, every telephone number, every appointment, every fact and figure you need at your fingertips. The Psion Organiser II makes it possible with a powerful "personal computer" so compact it fits in the palm of your hand. *"The smallest serious computer ever..."* That's what *PC Magazine* said about the *Organiser II.* It represents the most significant downsizing in a serious computing instrument since the introduction of the laptop computer. Pull down the protective sliding cover, touch "On" and the brilliant 2 line 16 character display comes to life. Adjust the brightness control and you're ready to go. Easy to use, the Organiser II accesses all functions from the "menu" which is automatically displayed when you switch on. With 32K of built-in software and powerful programming facilities, this 9-oz. palm-size computer replaces your address book, calendar, appointment diary, calculator, and all those little reminders you keep scribbling to yourself and stuffing in your pocket or purse.


### PHONE, ADDRESS AND NOTE BOOK:


Type in a name, address, home phone number, office number, birthday, anniversary or any other personal or business data, press one key and it's stored. To find it again, enter any part of the information from the record stored and *in less than a second all of the information for that record is automatically displayed on the screen.* Lines can be up to 254 characters and will scroll horizontally so you can read the entire line.

### APPOINTMENT DIARY:



The built-in diary program lets you record in half-hour increments, all business appointments, social engagements and reminders through the year 2000. It can turn itself on and buzz you up to an hour before an appointment. It also has 8 separate alarms. You can easily keep track of daily expenses.

### CALCULATOR:


Your calculations are displayed on the screen as they are entered, so they can be checked and changed if you want to make several similar calculations. It has 12 digit accuracy, ten memories and full mathematical and scientific functions.

### REMOVABLE DATAPAKS:

*Just like a desk-top computer* the Organiser II has the equivalent of two disk drives — tiny solid-state memory units called Datapaks. These perform the same function as floppy disks and keep information safe even when removed. With the *Model XP*, you can have up to *288K* of useable storage on line . . . the *largest* of any hand-held computer. While the *Model CM*, gives you up to *136K* of useable storage.

### PROGRAMMABLE COMPUTER:

Organiser II has an extremely powerful, built in, easy to use, programming language, enabling you to write both very simple and highly sophisticated programs. In addition, you can run "off-the-shelf" software by plugging in optional program packs.

### COMMUNICATIONS:


The Organiser II provides the unique ability to communicate. Through the RS232 interface, it can easily be connected to printers, other computers and telecommunications devices. *Over 150,000 Organiser II's have been sold in Europe* in the past year and a substantial amount of software is available. Through an exclusive arrangement, *The Lifestyle Resource*, is pleased to offer it to you in the U.S. Comes with a catalog of hardware and software available, and comprehensive, user-friendly manual. Operates on one nine-volt battery (supplied).

**FOR THE FASTEST SERVICE ON CREDIT CARD ORDERS CALL TOLL-FREE 800-872-5200**

If you prefer, mail check or credit card number with expiration date and authorized signature. Include $12.95 with each order for UPS and Insurance, along with each item number. Mail to:

**THE LIFESTYLE RESOURCE, DEPT. SFAE18; 921 EASTWIND DR., SUITE 114 • WESTERVILLE, OH 43081**

Model CM 8K Organiser II ............. $179.95 #2030;
Model XP 32K Organiser II ............. $249.95 #2040;
Additional 64K Plug-in Datapaks ........ $89.95 #2050;
Optional AC Adapter .................. $19.95 #2060.

**NO RISK 30-DAY RETURN PRIVILEGE**

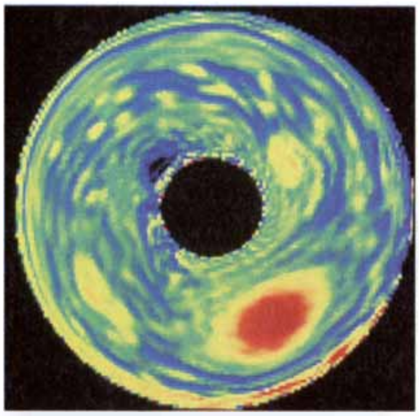
that the spot is a by-product of the powerful shearing forces in its zone. "The attractiveness of these models," says Andrew P. Ingersoll of the California Institute of Technology, "is that they are based on very simple assumptions." Earlier simulations, he notes, were more complicated and less successful at mimicking the red spot.

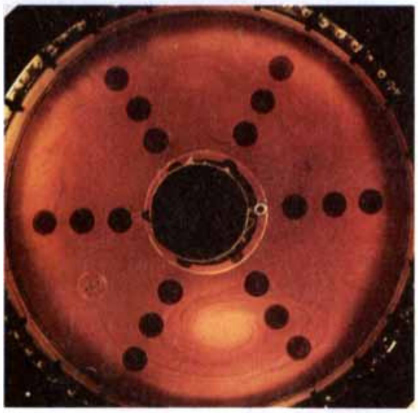
Philip S. Marcus of the University of California at Berkeley developed a computer model in which the zone is viewed from Jupiter's south pole. The zone is represented as a circular band with wind blowing counterclockwise on its outer (northern) border and clockwise on its inner (southern) border. He found that as the strength of the winds increases, small vortexes swirling counterclockwise emerge in the region of shear between them.

These vortexes tend to merge into progressively larger ones. Eventually one large, elliptical vortex develops, driven like "a paddle wheel" by the winds to the north and south of it. "It all looks random and crazy, then something coherent forms," Marcus says. He notes that as the Voyagers flew by Jupiter they filmed vortexes forming and merging into larger ones, including the red spot.

Marcus convinced colleagues at the University of Texas at Austin to test his model with a physical simulation. Joël Sommeria, Steven D. Meyers and Harry L. Swinney constructed a circular vat one meter wide with an inner ring of inlet valves and an outer ring of outlet valves. The workers spin the vat to simulate the rotation of Jupiter while pumping water into and out of the valves. The net effect is that water toward the edge of the vat flows counterclockwise and water near the center flows clockwise. Water in the turbulent shear zone behaves just as Marcus predicted it would: small vortexes merge into larger ones, and finally one large, elliptical vortex develops.

The successful simulation underscores a further mystery. Marcus points out that large vortexes and the shear zones that create them are common in the Jovian atmosphere, although no other vortex is as large and durable as the red spot. What gives rise to the powerful latitudinal winds, blowing around the planet in opposite directions, that create the shear zones and give the planet its distinctive banded appearance? Marcus says the winds are evidently linked to the planet's rapid rotation (it revolves once every 10 hours), but no one has been able to determine exactly how. —*J.H.*





*JUPITER'S GREAT RED SPOT was simulated on a computer at the University of California at Berkeley* (top) *and in a water-filled vat at the University of Texas.*

## Dark Solitons

*Physicists generate durable pulses of darkness*

"I hate to do science on the dark side," Daniel R. Grischkowsky remarks with a somewhat diabolic grin. That said, he continues telling a visitor how he and three colleagues at the IBM Thomas J. Watson Research Center have attempted to make "dark-pulse solitons."

To understand dark-pulse solitons one must first understand bright-pulse solitons. These are pulses of light that travel long distances without dispersing, or broadening. (Soliton is a generic term for a wave that retains its shape indefinitely.) Optical pulses ordinarily disperse, Grischkowsky explains, because they consist of photons of slightly different frequencies, which travel at slightly different speeds.

Workers can prevent a pulse from dispersing, he says, by exploiting an optical property of glass fibers. As the intensity of light passing through a fiber increases, its velocity decreases. At low intensities this "nonlinear" effect is insignificant, but at high intensities it can somehow compress a pulse just enough to make it a bright-pulse soliton. "Nobody really understands why it works," Grischkowsky notes.

Theorists have speculated that a similar technique might generate dark-pulse solitons: gaps in a beam of light that remain dark as they travel through a fiber, indefinitely resisting the encroachment of surrounding photons. The nonlinear Schrödinger equation, which predicts how waves behave in any nonlinear system, provided a mathematical basis for the speculation.

Last year Dieter Krökel, Naomi J. Halas, Giampiero Giuliani and Grischkowsky finished building a device they thought might create dark-pulse solitons. Called the ultrafast light-controlled optical fiber modulator, it can generate and measure pulses—bright and dark—lasting for less than a picosecond, or a trillionth of a second. With the device the investigators generated a .3-picosecond dark pulse embedded in a beam of laser light and sent it through a 10-meter-long fiber. When the power of the laser beam was low, the dark pulse quickly dispersed. But when the workers increased the power to what Grischkowsky calls the "truth-and-beauty regime," the single dark pulse split into two pulses, which then passed unchanged through the length of the fiber.

The nonlinear Schrödinger equation predicts just this "solution," Grischkowsky says: a single dark pulse begets a pair of solitons. Nevertheless, in order to verify that the pairs of dark pulses are solitons, the group would like to observe them

propagate without change through a much longer fiber. The workers' report in *Physical Review Letters* consequently calls the results "preliminary." (Workers from the Free University of Brussels in Belgium and the French National Center for Scientific Research who described a similar experiment last year in the journal *Optics Communications* also hedged their conclusion.)

Bright-pulse solitons have obvious applications: in long-distance optical communications, for example, they could reduce the need for "repeaters" that periodically regenerate the data-carrying pulses. (Other repeaters would still be needed to counter dimming due to absorption.) Could data be represented by dark-pulse solitons? "Other groups can get enormously eloquent about applications," Grischkowsky says. "That's not our goal. We just want to study nonlinear propagation." —*J.H.*

## Ultraviolet Verdict

*A global decline in atmospheric ozone appears to be real*

The ozone-depletion jury is finally in, and the verdict is not a welcome one. In a study done for the World Meteorological Association and the United Nations Environment Program an international panel of scientists has concluded that the ozone layer in the earth's atmosphere is being depleted globally. If the depletion continues, as it is expected to, exposure to ultraviolet radiation from the sun, which is associated with skin cancer in human beings and harm to plants and fisheries, will almost certainly increase.

More than a decade ago F. Sherwood Rowland and Mario J. Molina of the University of California at Irvine put forward a troubling hypothesis: chlorofluorocarbons (CFC's), released into the atmosphere during the manufacture of insulating materials and in the form of refrigerants and aerosol propellants, catalyze the destruction of ozone in the stratosphere. The hypothesis was suddenly revived two years ago when data gathered by the National Aeronautics and Space Administration's *Nimbus 7* satellite, combined with measurements from ground-based and airborne detectors, revealed dramatic declines in ozone concentration over Antarctica during the southern spring. Later studies suggested that the depletion could extend to other parts of the globe, but those claims were considered controversial.

The Ozone Trends Panel was created to settle the controversy. It has concluded from ground-based measurements that between 1969 and 1986 the level of ozone, averaged over the year, decreased by from 1.7 to 3 percent in much of the Northern Hemisphere after natural trends were taken into account. Winter levels fell the most: over the region between 53 and 64 degrees north latitude, which includes northern Canada, Alaska, Siberia and northern Europe, they declined by as much as 6.2 percent. A reanalysis of data gathered by *Nimbus 7* between 1978 and 1985 indicated that the atmosphere above the regions inhabited by most of the world's population experienced a total ozone decrease of about 2.5 percent.

Earlier doubts about the global extent of the depletion were rooted in the fact that the calibration of satellite instruments tends to drift. In order to resolve the issue the panel estimated the drift in the *Nimbus 7* instruments by comparing the satellite data with readings from nearby ground stations. Independent measurements carried out with different types of instruments also confirmed the downward trend in ozone, which was most pronounced at an altitude of about 40 kilometers.

The panel notes that the evidence "strongly indicates" that chlorine from CFC's is "primarily responsible" for the ozone "hole" in Antarctica, where reactions on ice crystals in so-called polar stratospheric clouds seem to play a role. It is not clear exactly how much of the global decrease is due to CFC's and other trace gases, which can also deplete ozone. Nevertheless, according to Robert T. Watson of NASA, the panel's chairman, "everything we do seems to be suggestive that CFC's are responsible for a loss of ozone."

Because ultraviolet radiation not only is shielded by ozone but also stimulates its production, part of the recent, global decrease detected by *Nimbus 7* could be the result of a slight decrease in the sun's ultraviolet output in the course of the natural 11-year solar cycle. But such effects are expected to be independent of season or latitude, whereas models of the effects of CFC's and trace gases predict larger decreases at high latitudes and in winter, which is the observed pattern. The models cannot account for the severity of the winter depletion, however. Measurements as the sun brightens over the next few years will be needed to separate the trace-gas and solar effects, according to Richard S. Stolarski of NASA, the panel's vice-chairman [see "The Antarctic Ozone Hole," by Richard S. Stolarski; SCIENTIFIC AMERICAN, January].

The U.S. Senate in March ratified the Montreal protocol, which could lead to a 50 percent decrease in developed countries' CFC production by 1999. In view of the new results, Watson recommends that officials "look long and hard at whether the Montreal protocol is adequate." Even if the convention is fully implemented, Rowland says, the level of chlorine from CFC's in the stratosphere, currently about three parts per billion, is likely to increase to at least six or seven parts per billion before stabilizing. —*T.M.B.*

# TECHNOLOGY

## Built for Speed

*Semiconductor devices become still smaller and faster*

Investigators at the Siemens Research and Technology Laboratories in Princeton, N.J., and at Cornell University have recently made transistors that are capable of switching at a rate of 113 billion times per second. Among devices with possible near-term applications, the speed approaches a record.

To increase the speed of their device, a modulation-doped field-effect transistor (MODFET), Allen Lepore of Cornell, Erhard Kohn of Siemens and their fellow workers resorted to two strategies. First they shortened the distance the electrons would have to travel. In a field-effect transistor, charge carriers (electrons or positive holes) move from one electrode, the source, to another, the drain. As they do so they pass a gate electrode, which regulates their flow as its electric potential is varied. By employing a technique known as electron-beam lithography, the workers were able to shape a gate measuring only .1 micrometer along one of its crucial dimensions. In commercially available

# MACSYMA

## automates symbolic mathematics. And yields enormous improvements in productivity, accuracy and modeling power.

**MACSYMA** combines exact solutions, symbolic approximations, and numerical methods into a powerful automated approach to scientific and engineering computing. Major benefits include:

- **Improved Productivity:** For many types of computations MACSYMA can increase your productivity by 10 to 100 times. ***It is that revolutionary.***
- **Increased Accuracy:** Manual computational errors are virtually eliminated. Use exact or approximate symbolic solutions in place of less accurate numerical ones.
- **Enhanced Mathematical Power:** You can dare to perform automated computations which you would not believe practical using traditional methods.

***Wide Range of Capabilities***

**MACSYMA** offers the widest range of capabilities for combined symbolic and numerical mathematics of all commercially available software.

- Algebra and Trigonometry
- Calculus and Differential Equations
- Symbolic Approximation Methods
- Numerical Analysis
- Graphics

***Available on Many Computer Systems***

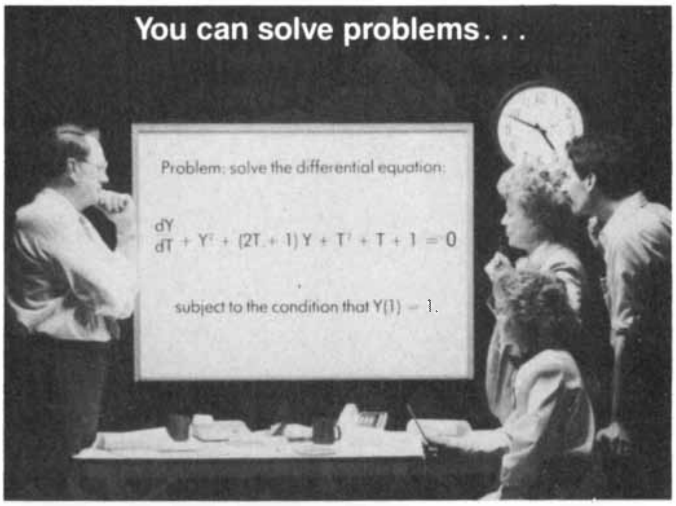
- Apollo
- Masscomp
- SUN-2 and SUN-3
- Symbolics 3600™ series
- Vax family

For an information kit about all the ways MACSYMA can work for you, just call

**1-800-MACSYMA.**

In Mass.,(617) 621-7770.

**Computer-Aided Mathematics Group**
**Dept. M-SA4**
**Symbolics, Inc.**
**Eleven Cambridge Center**
**Cambridge, MA 02142**




Symbolically. . .

(C1) DEPENDS(Y,T)$

(C2) DIFF (Y,T) + Y ^2 + (2*T + 1)*Y + T ^2 + T + 1;

(D2) $\frac{dY}{dT} + Y^2 + (2T + 1)Y + T^2 + T + 1$

(C3) SOLN:ODE(D2,Y,T);

(D3) $Y = -\frac{\%C\ T\ \%E^T - T - 1}{\%C\ \%E^T - 1}$

(C4) SOLVE(SUBST([Y = 1, T = 1],D3),%C),NUMER;

(D4) [%C = 0.5518192]

(C5) SPECIFIC_SOLN:SUBST(D4,SOLN);

(D5) $Y = -\frac{0.5518192\ T\ \%E^T - T - 1}{0.5518192\ \%E^T - 1}$

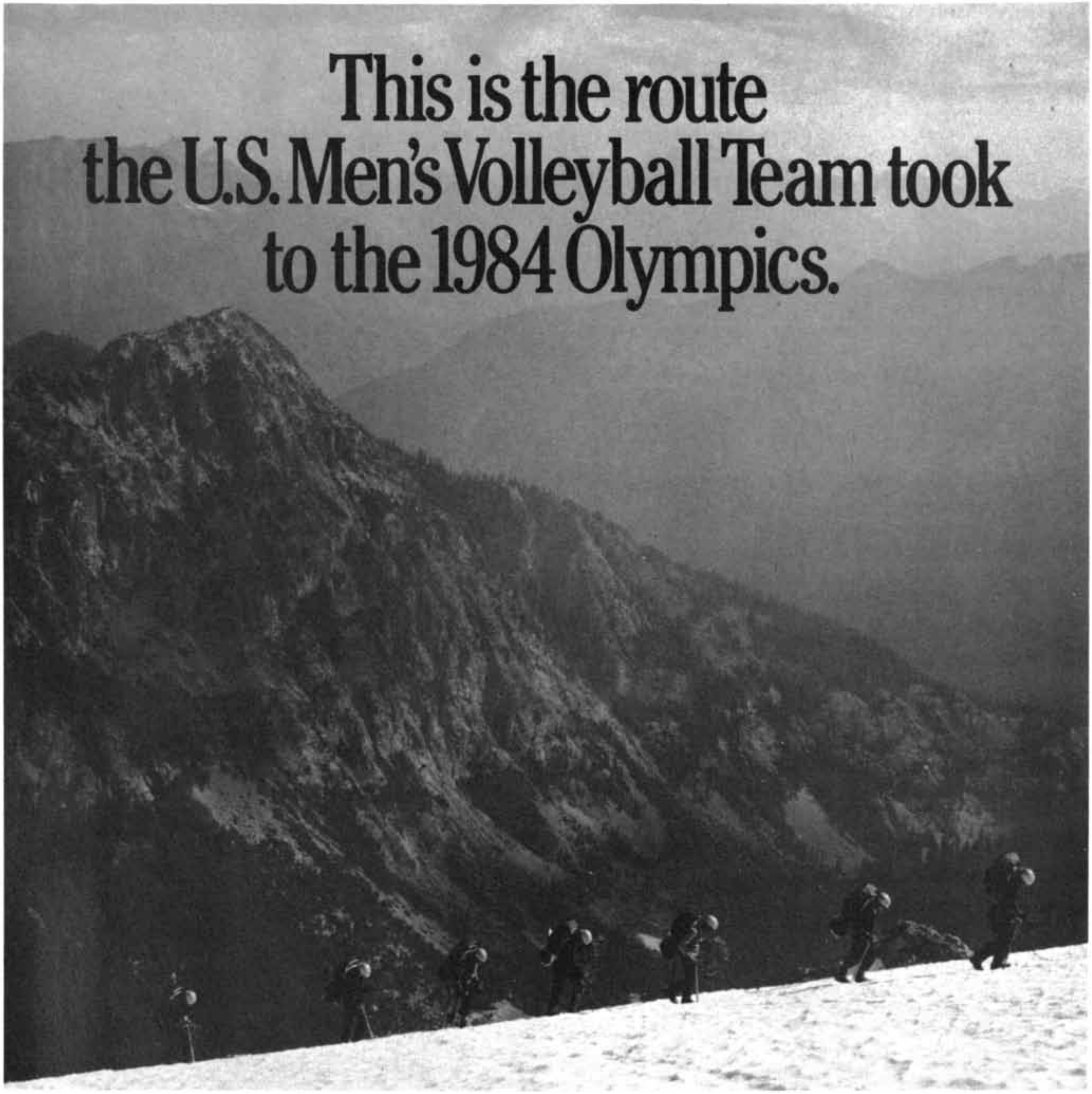
and Numerically.

```
(C6)  FORTRAN(D5)$
          Y = – (0.5518192*T*EXP(T) – T – 1)
     1          /(0.5518192*EXP(T) – 1)
```

## MACSYMA

The most comprehensive software for mathematical computing.

symbolics™




They had to wind their way through waist-high snowdrifts to get there.

You see, before the U.S. Men's Volleyball Team went on to win the gold medal at the 1984 Olympics, they went on an Outward Bound winter adventure. Previous to that trip, they were ranked thirteenth in the world. *Thirteenth.*

Now, we can't take complete credit for this startling turnaround. But it is dramatic proof that Outward Bound helps bring out the best in its participants and encourages them to work with and rely on others in ways they never thought possible. Furthermore, it's also proof that our programs not only build character but inspire self-confidence and leadership.

But don't start thinking that Outward Bound is for die-hard athletes. It isn't. Nor is it some "survival of the fittest" competition.

Outward Bound is, however, a series of exciting and challenging outdoor adventures held in some of the nation's most breath-taking settings.

There are adventures designed to accommodate everyone from teenagers to CEOs to grandmothers. There are also special programs for troubled youth, substance abusers, the handicapped—even the blind.

For more information, call us at 1 (800) 243-8520 today.

Now, we won't presume to think that when you leave Outward Bound you'll go on to win an Olympic gold medal yourself. But, then again, after discovering a new sense of pride and self-reliance, who knows?

**Outward Bound®**

**As rewarding as all outdoors.**

**1 (800) 243-8520**

(Conn. residents call (203) 661-0797)

Outward Bound is a nonprofit, nondiscriminatory organization. Outward Bound National Office, 384 Field Point Rd., Greenwich, CT 06830.

high-speed devices the same dimension is 10 times as large.

In addition to reducing the distance charge carriers have to travel, the investigators increased their speed through the material. They created a heterostructure, a sandwich of semiconducting materials in which the crystal lattice continues uninterrupted across the junction between layers of differing composition. Near the junction the electron energy levels favor higher speeds than are possible in specimens of pure silicon or other material. The workers built the heterostructure by a method called atomic planar doping.

First a layer of aluminum gallium arsenide was grown by molecular-beam epitaxy on a surface of pure gallium arsenide. Then, when the layer was some 10 atoms thick, about one in every 100 of the aluminum or gallium atoms in a single atomic plane was replaced by a silicon atom. Succeeding atomic layers were again pure aluminum gallium arsenide. The layer containing silicon atoms provides a surfeit of electrons that carry charge between two electrodes attached to the top of the semiconductor sandwich. These electrons move as a two-dimensional "gas" in the adjacent layers of gallium arsenide; there, under the influence of the nearby junction, they can travel faster than they could in either of the component materials alone.

Other workers have investigated materials that offer still greater advances in speed. April S. Brown and Umesh K. Mishra of the Hughes Research Laboratories in Malibu, Calif., have built experimental heterostructure devices that employ a junction between layers of aluminum indium arsenide and gallium indium arsenide. According to Mishra, the devices have switched 125 billion times per second, even though the gate is longer than it is in Lepore and Kohn's device.

New structures and materials, such as gallium arsenide, may soon replace silicon for applications such as high-speed signal processing and optoelectronics. For many applications in computing, where a number of components must be put on a single chip, silicon will remain superior, says George A. Sai-Halasz of the IBM Thomas J. Watson Research Center. He has built circuits utilizing silicon field-effect transistors that switch 75 billion times per second, or 10 times as fast as the fastest devices being used today. Miniaturization accounts for the speed: the gates are as little as

# 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


.07 micrometer, or 700 angstrom units, across—probably the smallest ever built. In order to achieve useful high-speed performance, however, the devices must be cooled in liquid nitrogen. *—T.M.B.*

## Have You Heard DAT?

*New recordings strain the music industry*

A clamor as discordant as a concerto by John Cage continues over digital audio tape (DAT) recorders, which will enable consumers to capture any sounds, including those on compact discs, in recordings that preserve virtually all the clarity of the original. Some people say this capability threatens companies' profits and musicians' property rights; others maintain it will strengthen the music industry. DAT hardware manufacturers meanwhile have remained reluctant to market the recorders.

The machines store sound on tape as a series of discrete numerical values. Whereas high-quality analog tapes produce recordings with a signal-to-noise ratio of about 68 decibels, the ratio can reach 96 decibels for a DAT, according to one manufacturer, Philips Gloeilampenfabriek N.V. in Eindhoven, the Netherlands.

U.S. consumers were first able to buy digital audio recordings in March, 1983, when the Sony Corporation and Philips introduced compact discs. The format received scant attention from most major U.S. recording companies until demand began to skyrocket, recalls Marc Finer, director of Communication Research, a consulting firm in Pittsburgh. The recording industry here has historically adopted new technologies grudgingly, adds Christopher M. Greenleaf, president of Classic Masters, a small studio in Brooklyn, N.Y.

The first U.S. compact-disc replicating facility—a $20-million investment—finally opened in September, 1984. The money was well spent: last year consumers bought more than 100 million compact discs, possibly topping record sales for the first time and making the digital discs "the fastest-growing music medium in the last 25 years," Finer says.

But even as the U.S. recording industry invested in compact-disc technology, a convention of some 80 companies, mostly from Japan, were standardizing specifications for DAT technology. They released the standards in 1986 and introduced the recorders last year. Worried about its compact-disc investment, the recording industry went into battle.

The Recording Industry Association of America (RIAA) says DAT technology will only add to the "some $1.5 billion annually" the industry loses in sales when consumers tape music with conventional analog recorders. It also argues that since musicians lose royalty payments when consumers tape instead of buying, professional artists will suffer.

Congress took up the theme last year in two bills requiring DAT recorders to incorporate a safeguard that stops unauthorized copying. A close inspection by the National Bureau of Standards found the proposed system is technically inadequate. Although the legislation has been quelled, the RIAA continues to look for other technical fixes.

Yet DAT recorders still have not flooded the U.S. "It is unlikely that there will be significant sales of DAT recorders unless the recording industry's demands are satisfied," says Jan D. Timmer, chief executive officer of Philips Consumer Electronics, a DAT hardware manufacturer. Observers nonetheless predict that the recording industry will embrace DAT technology eventually. Since consumers like digital recordings, Finer predicts that DAT's will begin to replace analog cassettes just as compact discs are replacing vinyl records. Small recording studios, such as Classic Masters, have already welcomed DAT technology. "I haven't been in one studio in the last six months that didn't have a DAT recorder," Greenleaf remarks.

Whatever the technology's fate, DAT is not likely to be the final movement of this concerto. Industry sources predict that home compact-disc recorders will appear within a few more years. *—Elizabeth Corcoran*

## ISDN: Betting Billions

*Who will buy the powerful communications fabric?*

Within the next two years James F. Mathieu will try to convince his managers at Lockheed Missiles and Space Co., Inc., to invest in a new communications technology called the integrated-services digital network (ISDN). Telephone companies herald ISDN as a powerful tool that puts any executive or clerk in com-

# 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 0036-8733/88. $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. 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 __________

mand of a spectrum of voice, data and video information and communication. But Mathieu, who supervises telecommunications at Lockheed, is not ready to bet money on a McLuhanesque millennium. Instead he hopes to tell his managers simply that the technology will transmit voice and data more cheaply and more quickly.

Telephone companies around the world are pumping billions of dollars into ISDN, even though the technology is still looking for a market. They are hoping that since information is seen as the "strategic weapon" of the 1990's, ISDN will provide the way to deploy it.

To a user, ISDN looks like the all-purpose wall plug that accepts telephones, computers, facsimile machines and video displays. Instead of separate lines for each device, a user will need only one ISDN line for sending and receiving all types of information. Each line consists of several channels; voice and data are transmitted together, while operational commands are sent separately. This arrangement speeds data transmission severalfold and grants users more control over information.

Some two dozen year-long ISDN trials are under way in the U.S. At the U.S. National Bank of Oregon an ISDN-based "call manager" posts the number of an incoming call on a computer screen as the telephone rings. The caller's name and financial records can also be displayed automatically. Lockheed has used its prototype ISDN equipment to process data, exchange computer-aided design files and even remotely watch parking lots.

Mark N. Ricca, a vice-president of the Eastern Management Group in Parsippany, N.J., a research firm specializing in telecommunications, expects that by 1992 U.S. carriers will glean a mere $500 million from ISDN traffic—roughly half a percent of last year's revenues for domestic transmission of voice and data. In about 1996, if ISDN gains widespread acceptance, Ricca predicts that revenues could jump to $17 billion in the U.S. and to $39 billion worldwide.

But acceptance depends on several factors. The first is compatibility. In the early 1980's international standards groups sketched out how information should be formatted and how fast it should be transmitted. Telephone companies in Japan, Europe, Canada and the U.S. separately began to fill in the details, building a network of Babel. The National Bureau of Standards has recently initiated efforts to reconcile the proposed U.S. systems.

Pricing strategies also need to be formulated. Whereas companies look toward a return on investment and the bottom line, many national regulatory agencies are anxious to see the service widely used. In Japan the monthly rates proposed by the Nippon Telegraph and Telephone Corporation were reportedly five times as high as those suggested by the ministry. The protracted negotiations between the two delayed the initial commercial offering of ISDN by more than a month. U.S. telephone companies are signing up their first ISDN customers contract by contract. Eventually ISDN charges are likely to be between 125 and 175 percent of the cost of a standard telephone line, possibly saving money for users who currently pay for separate data and voice lines.

The greatest challenge that telephone companies face will be proving ISDN is not only a cheaper tool but also a different one. "The U.S. trials have not shown any enticing applications to encourage users to sign up for ISDN," observes Mary A. Johnston, a senior researcher at the BBN Communications Corporation in Cambridge, Mass. She adds: "The companies that can bring new applications to market are the ones most likely to make money." —*E.C.*

# BIOLOGICAL SCIENCES

## Patented Remedy

*A piece of an AIDS-virus protein holds out promise for a vaccine*

There has been considerable pessimism recently among researchers working on a vaccine for AIDS. For one thing, the AIDS virus is so genetically variable that it has seemed unlikely a single vaccine would ever be able to counter all its many strains. A small segment of a recently patented protein, however, shows promise in laboratory experiments that it may ultimately form the basis of a broadly effective vaccine.

The protein is called GP120, because it is a glycoprotein that has a molecular weight of 120,000 daltons (a hydrogen atom has a weight of one dalton). GP120 forms part of the outer envelope of HIV (the AIDS virus); it appears there as a set of "studs" projecting from the virus's roughly spherical surface. Those studs are the part of the virus that first interacts with human cells in HIV infection. In 1984 Myron Essex and Tun-Hou Lee of Harvard University discovered that GP120 evokes the largest immune response of any protein in the AIDS virus.

In February, Harvard was granted a patent for GP120 on behalf of the two investigators. The university in turn has given exclusive rights to commercialize the protein to the Cambridge Bioscience Corporation, a biotechnology company that funded some of the work in Essex' laboratory. The patent may be lucrative, because most efforts to develop an AIDS vaccine (and also some therapeutic approaches) are based on parts of GP120 or on the whole protein.

For the moment that potential remains largely theoretical, because no AIDS vaccine is anywhere near commercial introduction. Work by David D. Ho of the University of California at Los Angeles School of Medicine and his co-workers, however, may suggest a route toward a practical vaccine. Ho has found that antibodies to a segment of GP120 that is relatively uniform from strain to strain of the virus have the capacity to prevent infection by the AIDS virus.

GP120 displays several types of segments: some are highly variable across viral strains, others are highly conserved from one strain to the other and still others are intermediate. Ho synthesized a highly conserved segment of the protein, designated C21E, and injected it into the bloodstream of rabbits, where antibodies to it were made. Those antibodies were found to be capable of blocking infection of human cells by three widely varying strains of HIV.

The antibodies do not keep GP-120 from interacting with the surface of the human cells. Therefore Ho concludes that they must interfere with some event that takes place after the virus binds to the cell surface. The results are of great significance because they identify a portion of GP120 that is capable of evoking neutralizing antibodies. In contrast, antibodies to the whole protein (which are present in AIDS patients) have a much lesser capacity to stem infection by a variety of viral strains.

In a recent report in *Science* Ho

# Even its reviews are beautiful.

On Boston's famed Beacon Hill.

## The 1988 Buick Regal.

**Motor Trend**
"...View one parked next to last year's car — or drive them back-to-back — and you will realize how far the stretch really was. Nice job."

**Car and Driver**
"The interior design has a fresh, modern look...but the inspiration is clearly Buick."

**The New York Times**
"Handling is just great, and more than that can't be said. On the highway, the Regal is quiet and...well, regal. On the twisties, it is pure fun, flat and glued to the road."

**The New York Times**
"Surprisingly, in a car that is 8.4 inches shorter than the '87 Regal, there is a lot of room in the back seat. Not surprisingly, all of the seats are comfortable."

**Car and Driver**
"...Regal's suspension is now fully independent. The front end takes the kinks out of Mother Earth with a pair of struts and coil springs. At each rear corner, road ripples are absorbed by a strut located by a trailing link and two lateral links."

Buckle up and see your Buick dealer today.

BUICK
USA
OFFICIAL CAR OF THE 1988 U.S. OLYMPIC TEAM

GM

The Great American Road belongs to Buick.

maintains that "the C21E domain may be an important site to consider when designing an AIDS vaccine." He adds that a "small, well-defined [protein segment] may be a more effective vaccine than a large, all-inclusive [protein] if it can specifically direct the immune response against a biologically important domain that is conserved in all strains of HIV." —*John Benditt*

## Fetal Fossils

*The embryonic forms of ancient reptiles are revealed*

The fossil record offers remarkably few glimpses of early reptiles, including dinosaurs, as they appeared while in the womb or sealed in eggs. Until recently the only reported fossils of whole embryos represented ichthyosaurs, dolphinlike reptiles that vanished about 100 million years ago. Now well-preserved embryos of two other extinct reptiles, from widely disparate places and times, have come to light.

One fossil, found in shale from southern Switzerland, is a two-inch specimen of a small aquatic reptile from 230 million years ago, in the Triassic period. Called *Neusticosaurus*, the reptile grew to about 12 inches in length and propelled itself with a long, thin tail and paddlelike limbs through the shallow, warm waters that once covered southern Europe.

*EMBRYO of* Neusticosaurus, *an aquatic reptile that lived some 230 million years ago, is imprinted on a piece of shale.*

The fossil is the smallest of the 800 *Neusticosaurus* specimens in the collection of the Paleontological Institute and Museum of the University of Zurich. The very abundance of the remains may explain why the embryo fossil, found in 1928, was overlooked for so long, according to P. Martin Sander, a worker at the university who describes the fossil in *Science*. The specimen's small size, disproportionately large head and unformed bones all indicate that it was unborn. The clinching evidence is the fossil's fetal position: virtually all other neusticosaur fossils are preserved in a belly-down, "road kill" position, Sander points out.

*Neusticosaurus* is thought to have evolved from egg-laying terrestrial reptiles. Had it developed the ability to bear its young live in the water, like the ichthyosaurs? Or did it crawl back onto the land to lay eggs, like the modern sea turtle (also descended from terrestrial creatures)? The fossil was found in a marine deposit and there was no trace of an eggshell nearby. These facts suggest that *Neusticosaurus* was viviparous, but the lack of any pregnant specimens suggests otherwise. Resolution of the issue, Sander says, will require more findings.

A fossil embryo described in *Nature* by John R. Horner of the Museum of the Rockies and David D. Weishampel of Johns Hopkins University, in contrast, was found in an egg. Horner discovered the fossil four years ago at a site in Montana littered with many eggs and adult bones. He and Weishampel say the fossils represent a previously unknown herbivorous dinosaur from the Cretaceous period, some 75 million years ago. The investigators named the dinosaur *Orodromeus*, or "mountain runner." The adults were about eight feet long and walked erect on long, slender hind legs.

Horner CAT-scanned the eggs and found that one contained an intact embryo (embryos in dinosaur eggs are usually highly fragmented). He laid the embryo bare by painstakingly chipping away the shell. The fact that the bones were "very well formed," he says, indicates *Orodromeus* hatchlings were relatively well developed and independent at birth, as are many modern reptiles.

Horner has found embryo-containing eggs of other dinosaurs, including the hadrosaur, a duckbilled, bipedal herbivore that grew to about 20 feet in length. His preliminary analysis suggests that its hatchlings were undeveloped and required substantial nurturing after birth, "more like birds," he says. —*J.H.*

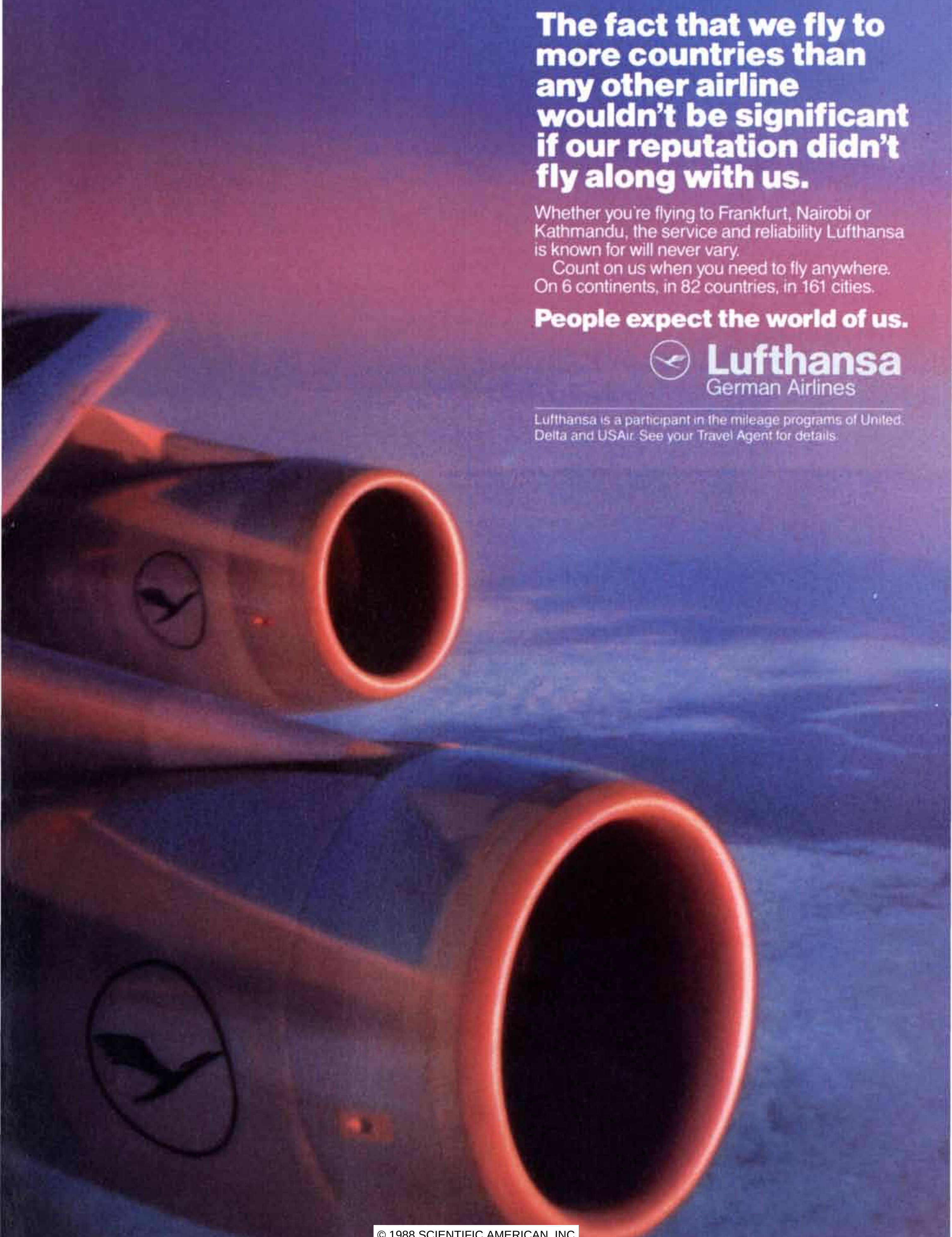
## Malaria Vaccines

*Seeking an Achilles' heel on an elusive target*

Malaria vaccines made with bioengineering techniques that became available only in the past decade offer new hope in the fight against this tropical scourge. The vaccines have successfully immunized some human volunteers, and although fully preventive vaccines remain a distant prospect, the work is the first significant advance since the 1940's, when DDT and the drug chloroquine were hailed as miracle weapons. Malaria-bearing mosquitoes and the malaria parasite, however, developed resistance to those chemicals. Today malaria afflicts between 200 and 400 million people and kills two million every year.

The parasite, *Plasmodium*, adopts several guises during its sojourn in human hosts, making it an elusive quarry for the body's immune defenses. Malaria sporozoites injected into a human by a mosquito quickly invade liver cells. Eight to 10 days later thousands of merozoites burst out of each cell and invade red blood cells, where they multiply further, destroying the cells and causing fever. Some merozoites become male or female gametocytes. A mosquito ingests the gametocytes, which reproduce sexually in its body. Eventually a new generation of sporozoites migrates to the mosquito's salivary glands, and the cycle begins again.

At each phase the parasite possesses a distinct protein coat. An effective vaccine, then, must induce antibodies to at least several of these. Unfortunately the parasite cannot be grown in culture to yield vaccines in the conventional way. Bioengineering offers a solution to these quandaries. Last year in *Nature* Victor Nussenzweig and his colleagues at the New York University Medical Center described the first human tests of a vaccine made from a synthetic peptide, or protein fragment. "We took off-the-shelf amino acids and built a chain of 12 amino acids found on a sporozoite protein," Nussenzweig said. The vaccine immunized one of


The fact that we fly to more countries than any other airline wouldn't be significant if our reputation didn't fly along with us.
Whether you're flying to Frankfurt, Nairobi or Kathmandu, the service and reliability Lufthansa is known for will never vary.
Count on us when you need to fly anywhere. On 6 continents, in 82 countries, in 161 cities.
People expect the world of us.
Lufthansa
German Airlines
Lufthansa is a participant in the mileage programs of United, Delta and USAir. See your Travel Agent for details.

the three vaccinated subjects in the test and delayed infection in the remaining two. Another team, at the Walter Reed Army Medical Center, reported similar success for tests of a vaccine produced by bacteria that had the gene for the peptide inserted into their DNA.

Investigators are also developing vaccines against later stages. The N.Y.U. vaccine has incorporated peptides that stimulate *T* cells to produce gamma interferon, which attacks the liver phase. A group led by Manuel E. Patarroyo of the National University of Colombia in Bogotá reported in *Nature* that synthetic-peptide vaccines had immunized three out of nine vaccinated volunteers against merozoites.

Louis H. Miller of the National Institute of Allergy and Infectious Diseases cautions that Patarroyo's vaccine includes a peptide in which two of the amino acids are known to be different in one of the four sequenced strains of the parasite. The vaccine might not protect against these variants. Such pitfalls underscore the need to identify peptides that are vulnerable to immunological attack and yet are not prone to variation. Miller says, "We have to find an Achilles' heel." —*June Kinoshita*

## Cause and Affect

*How does lithium counter both mania and depression?*

Forty years ago the Australian psychiatrist J. F. J. Cade injected guinea pigs with lithium salts and, inspired by the animals' subsequent lethargy, went on to discover the therapeutic effects of lithium in treating manic depression. Yet neither Cade nor anyone since has been able to explain how lithium works.

Perhaps the most curious aspect of lithium's action is its apparent ability to stabilize both poles of the disorder: the frenzy of mental and physical hyperactivity and the paralyzing depression that ensues. Now an Israeli group says it may have found a common biochemical ground from which the element exerts its dual effect. That ground is a class of molecules called G proteins, which mediate intracellular messengers for signals received by surface receptors.

Writing in *Nature,* Sofia Avissar, Gabriel Schreiber, Abraham Danon and R. H. Belmaker of Ben Gurion University of the Negev show that in rat brain tissue lithium blocks the activation of G proteins by neurotransmitters binding to both adrenergic and cholinergic receptors. At clinically relevant concentrations of lithium the G proteins in effect "ignore" signals to which they ordinarily react. Although it has long been known that lithium cripples a separate messenger system linked to cholinergic receptors, the Ben Gurion group provides the first demonstration of an adrenergic connection.

The group's findings also constitute the first evidence of how lithium acting at a single site could modulate both adrenergic and cholinergic responses. That notion is appealing because it jibes with what is probably the most widely accepted theory of manic depression. Put forth in 1972 by David S. Janowsky, then at the Vanderbilt University School of Medicine, and his colleagues, the theory maintains that adrenergic-cholinergic imbalances cause the mood swings characteristic of the disorder.

The authors do not attempt to explain why, considering the ubiquity of G proteins, lithium's influence is limited to the central nervous system. "If that's its mode of action, it really should affect every cell in the body," says Alan H. Drummond, a pharmacologist at British Bio-technology Ltd. in Oxford, England. Yet at therapeutic doses the side effects of lithium are negligible.

William R. Sherman, a biochemist at Washington University's Clinical Department of Psychiatry, points out that the problem of "CNS selectivity" plagues most of the mechanisms that have been proposed to describe lithium's efficacy. And no direct evidence in human beings substantiates any of them. He notes that studies of lithium and manic depression, like all studies of drug action in psychiatry, are disarmed by their reliance on subjective information. In spite of the fact that thousands of Americans lead normal lives because of lithium therapy, Sherman says "it's almost impossible to get real evidence of what's going on." —*Karen Wright*

# OVERVIEW

## Heartening News

*New ways to treat and prevent heart disease are saving lives*

More than one million people will suffer heart attacks in the U.S. this year and about a third of the attacks will prove to be fatal. These grim statistics nonetheless have a positive side: a decade ago 100,000 more people were dying from heart attacks every year. The improvement has come as a result of significant advances in the treatment and prevention of heart disease. Surgical procedures such as heart transplants and bypass operations have become established treatments for the severest cases. Heart attacks can now be arrested while they are in progress by means of new classes of drugs and nonsurgical procedures. People can significantly lessen the chance of suffering a heart attack to begin with either by taking special medication or by modifying their behavior. These advances in cardiology have given physicians the ability to intervene decisively and effectively at every stage of heart disease.

The major cause of heart disease is atherosclerosis: the buildup of cholesterol and other organic matter in deposits called plaques along the inner wall of blood vessels. Plaques narrow the channel through which blood flows, creating areas of constriction, or stenosis, that lead to ischemia: inadequate oxygen supply to a tissue. Heart attacks, or myocardial infarctions, strike when the blood supply to the heart-muscle tissue (myocardium) is severely reduced or stopped, resulting in tissue damage or death. In the vast majority of cases the reduction in the blood supply can be traced to the formation of a clot (thrombus) in the plaque of a coronary artery, which supplies blood to the myocardium. An occlusive clot is the difference between chronic coronary-artery disease (such as ischemia) and an acute heart attack.

For patients who are suffering from advanced heart disease invasive procedures are usually called for. In the most extreme cases this might entail actual replacement of the heart. The American Heart Association reports that 1,024 such transplants were done in the U.S. in 1986, more than 10 times the number done in 1982. Two decades ago a transplant recipient had less than a 20 percent chance of surviving one year; today the one-year survival rate is more that 80 percent. Such success rates prompted

# THE LIFESTYLE RESOURCE™

*At The Lifestyle Resource we give you all the facts and details necessary to make an informed purchase. Your satisfaction is our primary concern. If your purchase doesn't meet your expectations, return it in original condition within 30 days for prompt refund.*

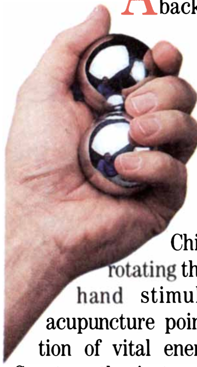
## EXPERIENCE NATURAL STATES...THE VIDEO

At last—the real thing in relaxation videos. No more inane fishtanks or fireplaces: Miramar's "Natural States" cassette presents a 45 minute trip through the Pacific Northwest, from California's Big Sur to Washington's Cascade Mountains. This is NOT a travelogue — not a word is spoken as the unique "steadicam" camera glides through rainforests, past verdant waterfalls, over and around mist-shrouded peaks. This unusual technique, mated to superb aerial shots, produces something very nearly like an out-of-body experience. The incredibly crisp footage is enhanced by an award winning score that was specially composed "note-to-frame," rather than just layered on in post-production. This is the first nature/music video to "go gold" in sales, and for good reason: its breathtaking beauty will bring tranquility to the worst case of stress and burnout. Just slip "Natural States" into your VCR; in 45 minutes you'll be human again! Hi-fi Stereo with Dolby Surround plays mono on regular machines. **$29.95** #2010.

## WORRY-FREE WATERING

Water grows things best when it arrives at regular intervals, a fact that sometimes gets by old Mother Nature. Now you can give your lawn or garden the advantage of timely watering. The new Rainmatic 2000 *automatically turns water on and off* up to four times a day, seven days a week, a different schedule every day. Lets you set any length watering interval. Easily programmed to set up your own customized watering schedule or use built-in program. One-touch manual override. Great for sprinklers and drip/trickle systems. Operates all season on four alkaline D-cells (not included). Weather-protected electronic circuitry. LCD readout with digital clock. Locks on to faucet. Only **$49.95** #1080 with one-year warranty and instruction manual.

## TO YOUR HEALTH



Ancient mandarins dating back 800 years believed that these Chinese Exercise Balls induced well-being of the body and serenity of spirit. These treasured gifts were given to President Reagan and his wife while visiting the Peoples Republic of China. The Chinese say that rotating the balls in the palm of each hand stimulates the fingers and acupuncture points, and improves circulation of vital energy throughout the body. Sports enthusiasts, musicians, computer users and health-conscious people everywhere consider them great muscle conditioners. Arthritis sufferers feel a decided benefit from this gentle but challenging excercise. Very effective for relaxation and meditation, Chinese Exercise Balls emit a distantly mysterious chime as you turn them. Beautifully hand crafted, 45mm. hollow polished chrome balls are perfectly weighted and fit comfortably into the average man's or woman's hand. In exquisite silk brocade box. **$29.95** #1700.

## ZONE OF CALM

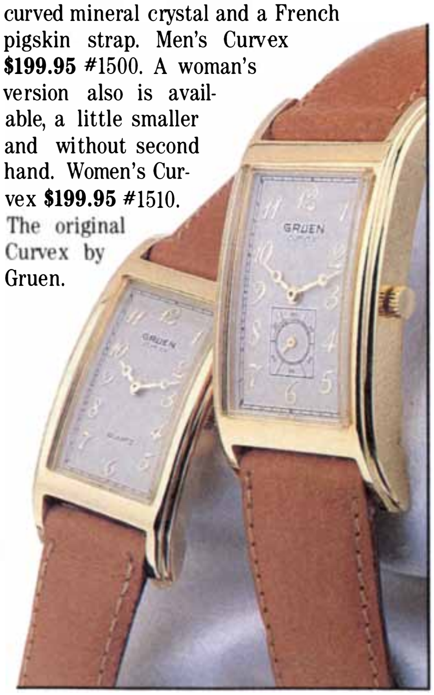
Noise produces stress in the human organism. Today the volume of civilization nearly everywhere seriously interferes with our abilities to relax, read, sleep, concentrate or otherwise function at optimum efficiency. But you needn't be victimized by noise pollution. The Marsona Sound Conditioner electrically synthesizes a variety of pleasing natural sounds that help mask out and reduce the annoyance of unwanted noise. You can simulate the sounds of ocean surf, summer rain, mountain waterfalls. You control not only the volume but also wave pattern, wave or rain rhythm, the seeming nearness or distance of the source. The Marsona is the finest instrument of its kind on the market, solid-state electronics designed for continuous use. UL-listed. **$139.95** #1860.

## 38,000 HOUR LIGHTBULB!

The problem with bulb savers is they reduce light output by almost 50%. But the patented, UL listed Power Mizer has solved that problem. These little buttons are activated by a state-of-the-art microcontrol switch that limits voltage to the bulb to 85%, thereby prolonging the life of the filament. Light bulbs using Power Mizer lose so little perceptible light that a foot-candle meter is needed to calculate the light loss. Under normal conditions, the Power Mizer will extend the life of a regular one-way 100W bulb from 750 hours to 38,000 hours or about four years if used 24 hours everyday! Many major hotels, corporations and universities use these types of devices in hard-to-reach places. They pay for themselves in light bulbs, labor and energy savings in no time. Set of 6 Power Mizers **$19.95** #2000; 12 for **$29.95** #1990.

## UPDATED CLASSIC TIMEPIECE

The Gruen Curvex, one of the most famous watches of all time, was designed in 1935 with an exclusive curved case and movement that conformed to the shape of the wrist. If you're lucky enough to have inherited one, you know how much they're worth. Now the Gruen watch company of Switzerland, founded in 1874, is once again offering the Curvex, at a fraction of the collector's item price. These precision watches are made from the original molds, and are extensively hand crafted. The only change since 1935 has been the addition of a modern Swiss quartz movement, accurate to within a minute per year. The case is 10 micron heavy 22 karat gold plate, fitted with a curved mineral crystal and a French pigskin strap. Men's Curvex **$199.95** #1500. A woman's version also is available, a little smaller and without second hand. Women's Curvex **$199.95** #1510.




The original Curvex by Gruen.

# THE LIFESTYLE RESOURCE™

## ALLERGY AND COLD BUSTER


Now a major scientific breakthrough — The Viralizer® System—gives you relief from cold, sinus and allergy symptoms. It's the newest development of a concept pioneered at the Pasteur Institute in Paris. The cause of the common cold is the Rhinovirus family which lives and multiplies in the nose and throat, but cannot thrive in temperatures over 110°F. The Viralizer is designed to deliver a gentle, controlled heat which penetrates the nose and throat, creating a hostile environment for cold viruses. After a pleasant heat treatment, the Viralizer dispenses either of two mild, over-the-counter, medicated sprays. Vira-Spray I is an analgesic, anti-bacterial spray. Vira-Spray II is a decongestant that provides temporary relief of nasal congestion due to colds, hayfever, sinusitis or allergies. These therapeutic sprays further discourage the stubborn cold and sinusitis germs so you're less likely to be re-infected or spread your cold to others. The Viralizer can produce effective relief by using it for only 3 or 4 minutes, several times a day. Proven in clinical tests 90% effective in eliminating the symptoms of upper respiratory infection in 24 hours or less, the Viralizer works without pills. Viralizer, is safe for children and adults, has been tested and recommended by doctors. The complete Viralizer® System includes 1 electric Viralizer with Vira-Spray I and Vira-Spray II plus a 3-pak refill of medicated sprays. **$39.95** #1690.

## ROWING FOR FITNESS

Rowing tones and firms your whole body and provides aerobic conditioning, as well. The Vitamaster Stowaway works all major and most minor muscle groups in your back, stomach, arms, shoulders and legs. Weighing only 28 lbs., it requires only 12″ x 15″ x 34″ when in operation and collapses after use for compact storage.

Actually takes up less room than a suitcase. The Stowaway is constructed of heavy tubular steel and has a full-size 54″ bed. The contour seat is padded for comfort and mounted on nylon rollers which glide on chrome rails. Single rowing arm with hydraulic tension resistance has 12 adjustable tension settings—from light to heavy—to provide a varied workout. Molded plastic pedals with adjustable footstraps rotate individually to simulate rowing action. Non-mar footpads protect floor. Comes assembled. Save yourself an expensive gym membership. For the one-time cost of **$119.95** #1970.

## YOUR SMILE IS BEAUTIFUL

*Hand brushing's motion misses plaque.*

*The counter-rotational movement cleans plaque.*


But you're only fooling yourself if you think it's proof of healthy teeth and gums. The unseen enemy is a natural mouth substance deposited on the teeth — plaque. Unchecked, plaque develops into rock-hard tartar behind which decay breaks down the tooth enamel and invades the soft, sensitive substructure. You have to remove the plaque before it hardens to prevent the process. Until recently there was no easy solution to this problem, but now there's Interplak, a precision-engineered, university-tested, three-speed plaque removal instrument. Its 10 independently driven bristle tufts rotate around their own axes: each tuft spins in the direction opposite to the spin of its neighbor at 4,200 rpm. The bristles are soft and rounded, adhering to dentistry recommendations, so that virtually every particle of plaque — *98% of it* — is removed gently and without irritation. Interplak even reaches between teeth and around orthodontic braces. Children think it's fun, so they're eager to brush. Interplak is cordless and easily stores in a free-standing or wall-mounted recharger/holder. Individual color-coded brush heads are available. The Interplak Home Plaque Removal Instrument with charger and two brush heads. **$99.95** # 1110. Two additional brush heads **$29.95** # 1111.

## WHEN YOU'RE IN A HURRY

*When you're running short of time, ask one of our phone operators about UPS Blue Label delivery. In most cases, we can get your shipment to you in three or four working days.*

**CALL TOLL-FREE 800-872-5200**

---

**THE LIFESTYLE RESOURCE™** — CREDIT CARD ORDERS CALL TOLL-FREE 24 HOURS **800-872-5200**

DEPT. SFAE08; 921 EASTWIND DR. SUITE 114: WESTERVILLE, OH 43081

| | ITEM NO. | QTY. | DESCRIPTION | ITEM PRICE | TOTAL PRICE |
|---|---|---|---|---|---|
| SEND TO (PLEASE PRINT) | | | | | |
| ADDRESS | | | | | |
| CITY STATE ZIP | | | | | |
| ☐ CHECK OR MONEY ORDER ☐ MASTERCARD ☐ VISA ☐ AMEX SIGNATURE | | | | | |
| ACCT. # EXP. DATE | | | Shipping Charge covers UPS, handling and insurance for guaranteed delivery. | SUB TOTAL | |
| | | | | SHIPPING (see table at left) | |
| | | | | TOTAL | |

**ORDER WITH CONFIDENCE**
- *We ship most orders within 48 hours.*
- *Credit Card orders billed only upon shipment.*
- *No risk 30-day return privilege.*

UPS 2ND DAY AIR

| | | |
|---|---|---|
| Up to $20 | $ 3.95 | |
| $20.01 to $30 | $ 4.95 | |
| $30.01 to $40 | $ 5.95 | |
| $40.01 to $50 | $ 6.95 | |
| $50.01 to $60 | $ 7.95 | |
| $60.01 to $70 | $ 8.95 | |
| $70.01 to $100 | $10.95 | |
| Over $100 | $12.95 | |

UPS Second Day available for an additional $7.50 per order.

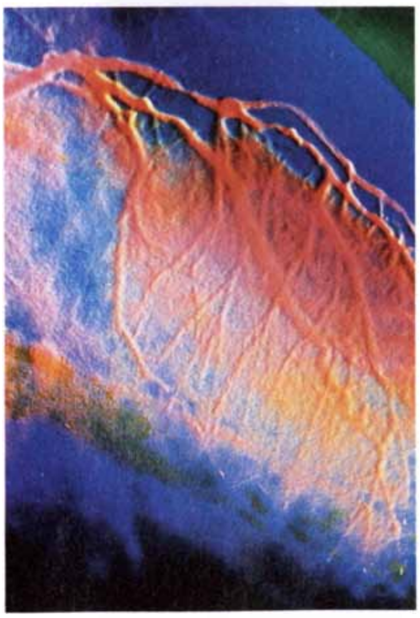
THE MEDIA DEVELOPMENT GROUP

©1988 G27 ADVERTISEMENT

John Speer Schroeder and Sharon Hunt of the Stanford University Medical Center to write in the *Journal of the American Medical Association* that transplants have become "a valid form of advanced therapy for end-stage heart disease." Schroeder and Hunt project a need for 15,000 transplants per year. Yet according to Gerald M. Lawrie of the Baylor College of Medicine, no more than 2,000 donor hearts are potentially available.

Some patients waiting for a suitable donor heart might be able meanwhile to rely on mechanical replacements for the heart or "assist devices," which merely take over the work of a malfunctioning ventricle. Indeed, a group led by David J. Farrar and J. Donald Hill of the Pacific Presbyterian Medical Center in San Francisco concludes in an article in the *New England Journal of Medicine* that using mechanical hearts as "bridges to transplant" is both safe and effective. If artificial hearts ever are permanently implanted, they are likely to be self-contained, electrical devices whose power can be delivered by induction across the skin. Eugene Passamani of the National Heart, Lung, and Blood Institute expects clinical trials of such electric hearts "in a year or two," with a view toward making them generally available in the late 1990's.

Another form of cardiac surgery that has become routine is the coronary-artery bypass graft. In this procedure blood vessels from elsewhere in the body are grafted onto the diseased coronary artery so that the blood can be rerouted around the obstruction. An estimated 230,000 such operations were done in 1985. Bypass grafts, like transplants, are highly traumatic (the patient's chest must remain open for several hours) and require a prolonged hospital stay and recovery period.



*CONSTRICTED ARTERY* (topmost vessel), *the result of a buildup of plaque, reduces the blood supply to the heart muscle.*

For that reason cardiologists are turning to much less invasive procedures to increase the blood flow to the heart. The preeminent method is technically known as percutaneous transluminal coronary angioplasty. It involves inserting a catheter, or thin flexible tube, into an artery through a skin incision (percutaneously). The catheter is manipulated so that it travels along the inside of the vessel (transluminally) to the stenotic lesion of a coronary artery. Inflating a small balloon at the end of the catheter causes distension of the artery (angioplasty) and also compresses the plaque against the artery wall, thus enlarging the opening. More than 130,000 people underwent the procedure in 1986 (almost three times the 1984 number), and some 80 percent of them showed a lasting increase in blood flow.

Nevertheless, angioplasty has certain limitations, one of which is that it is extremely difficult to clear total occlusions of coronary arteries. For such cases, investigators are developing methods for delivering a laser beam through a flexible optic fiber in a catheter. The beam can bore a hole through the blockage, creating an opening in which a conventional angioplasty balloon can be inflated.

In spite of these new cardiac-care techniques, damage to the myocardium is sometimes unavoidable, resulting in arrhythmia: disturbances in the heart's rhythm. Yet Ary L. Goldberger of the Harvard Medical School and his colleagues, who have analyzed the heart's rhythm in terms of fractal dynamics, suggest that analysis of electrocardiographic data may one day provide ample warning of life-threatening arrhythmias.

Perhaps the greatest progress in the treatment of heart-attack victims has been in the introduction of new thrombolytic, or clot-dissolving, drugs. They include streptokinase, urokinase and tissue plasminogen activator (TPA), a natural anticlotting agent. A genetically engineered form of TPA has recently been approved by the U.S. Food and Drug Administration. Indeed, intravenous injection of TPA is often enough to halt a heart attack within minutes and save heart tissue from damage.

In this connection even ordinary aspirin appears to have a potential role. According to a preliminary report that was published in the *New England Journal of Medicine,* aspirin can cut the risk of a heart attack by almost half even for those who have no overt signs of cardiovascular disease, apparently by working to reduce the incidence of clot formation.

Although thrombolytic drugs account for a significant reduction in mortality from heart attacks, the most sensible strategy in combating heart attacks is to prevent cardiovascular disease in the first place. Studies have shown that certain readily identifiable physiological features characterize those individuals who are at high risk of developing the disease. Two such factors are hypertension, or high blood pressure, and high cholesterol levels in the blood.

As a result medications that counter hypertension are being prescribed as part of a preventive approach to heart disease. The substances include so-called beta-blockers and ACE inhibitors, which lower the heart rate and the output of blood from the heart. Drugs such as cholesteramine and nicotinic acid are also being prescribed, since they lower cholesterol levels in the blood. More intriguing is the possibility of discovering drugs for the control of low-density lipoproteins (LDL), which are the main carriers of cholesterol in the blood. Because high levels of LDL appear to promote the deposit of cholesterol on artery walls, substances that interfere with LDL metabolism open a new avenue for intervention. By administering such drugs a group led by Thomas Carew of the University of California at San Diego School of Medicine has been able to reduce the extent of atherosclerosis in rabbits by 50 percent.

Ultimately, however, the most effective measures are those that depend on an individual's own initiative. Both the risk factors associated with lifestyles (such as smoking, a diet excessive in saturated fat and salt, and lack of exercise) and the decision to seek emergency cardiac care are under an individual's control. According to American Heart Association estimates, more than 350,000 heart-attack victims die every year before reaching a hospital—and this in spite of the fact that an average of three hours elapses between the initial appearance of symptoms and death. —*Gregory Greenwell*

TOYOTA SUPRA
NERVE CENTER.
Analog instruments monitor every rev. 8-way adjustable driver's Sport Seat and optional glove-soft leather interior add even more luxury to performance.
3000
TOYOTA
TWIN CAM 24 VALVE
turbo
POWER STRUCTURE.
Good-bye crowd. Turbo-charged 3.0-liter twin cam 24-valve engine with inter-cooler generates 230 hp at 5600 rpm.
IT DRAWS A CROWD, AND LEAVES IT BEHIND.
It's an imposing presence. An eye-magnet. But the 1988 Toyota Supra Turbo is not simply for show. The low-slung aerodynamics separate it from the crowd, even at standstill. But the gap widens quickly when its electrifying turbocharged potency is released: 0 to 60 in a mere 6.7 seconds.* And control is precise: race-bred double wishbone fully independent suspension; ventilated 4-wheel power disc brakes; optional Anti-Lock Braking System (A.B.S.). Toyota Supra—#1 in its segment in customer satisfaction.** The crowd can look but they can't touch it.
Get More From Life... Buckle Up!
TOYOTA
*United States Auto Club Certified Performance Figures.
**J.D. Power & Associates 1987 Customer Satisfaction with Product Quality and Dealer Service (for 1986 vehicles).
© 1987 Toyota Motor Sales, U.S.A., Inc.
TOYOTA QUALITY
WHO COULD ASK FOR ANYTHING MORE!

SCIENTIFIC AMERICAN May 1988 Volume 258 Number 5

# Controlling Indoor Air Pollution

*Airborne combustion products, toxic chemicals and radioactivity are more abundant indoors than outdoors. Should indoor air be regulated? If so, how? Putting risks in perspective helps to answer both questions*

by Anthony V. Nero, Jr.

Because emissions from factories, power plants, waste dumps and automobiles can harm people as well as the biosphere in general, many countries have established extensive systems for identifying and controlling such pollution. Yet the greatest exposures to airborne combustion products, volatile toxic chemicals and radioactivity typically occur not outdoors but inside residences, offices and other nonindustrial buildings—settings that traditionally have been neglected by pollution-control agencies. Indeed, the health risks incurred merely by breathing the air at home can substantially exceed the general limits on risk that regulatory agencies impose in controlling pollutants in outdoor air or drinking water.

On the other hand, the risks from inhaling indoor pollutants are typically less than those associated with many voluntary activities that are regulated only marginally, if at all. By choosing to smoke cigarettes a person greatly increases the likelihood of later suffering from heart disease and lung cancer. Yet smoking is controlled indirectly in the U.S., where manufacturers are required only to print warnings on cigarette packages and in advertising.

The risks posed by indoor pollutants are in fact comparable in magnitude to those associated with exposure to chemicals or radiation in industrial settings. Living in certain houses and working at certain jobs are also similar in that both involve assuming some risk from exposure to pollutants in exchange for some personal benefit (a home or a salary). Yet the approach for regulating occupational exposures, which is based primarily on setting concentration limits for each pollutant, would be extremely difficult to apply in controlling the quality of indoor air. The causes of indoor air pollution are so diverse and the concentrations are so variable that continual monitoring would be required in virtually all buildings—more than 80 million in the U.S. alone. Moreover, governments may hesitate to intrude into private houses in the name of pollution control, preferring merely to warn people of the dangers and let them decide for themselves whether or not to take action.

Although building codes could be rewritten to reduce the chance of having high pollutant levels in new structures, the incidence of buildings with indoor concentrations five, 10 or even 100 times the average justifies an active effort to find and fix such buildings now. The effort may affect hundreds of thousands of houses in the U.S., requiring as large an environmental program as has ever been undertaken. The scope of such a program, the variability in structures and pollutant concentrations and the balancing of personal choice, risk and benefit make construction of an effective and sensible overall strategy for the control of indoor air quality a daunting challenge to the research and regulatory community.

The pollutants found in indoor air are similar to those found outdoors and in some instances actually come from outdoor sources. Yet the pollutants measured in the highest concentrations indoors are those that arise from within buildings or their substructures. They reach such levels simply because they are emitted into a small volume—the indoor atmosphere—from which they cannot easily escape.

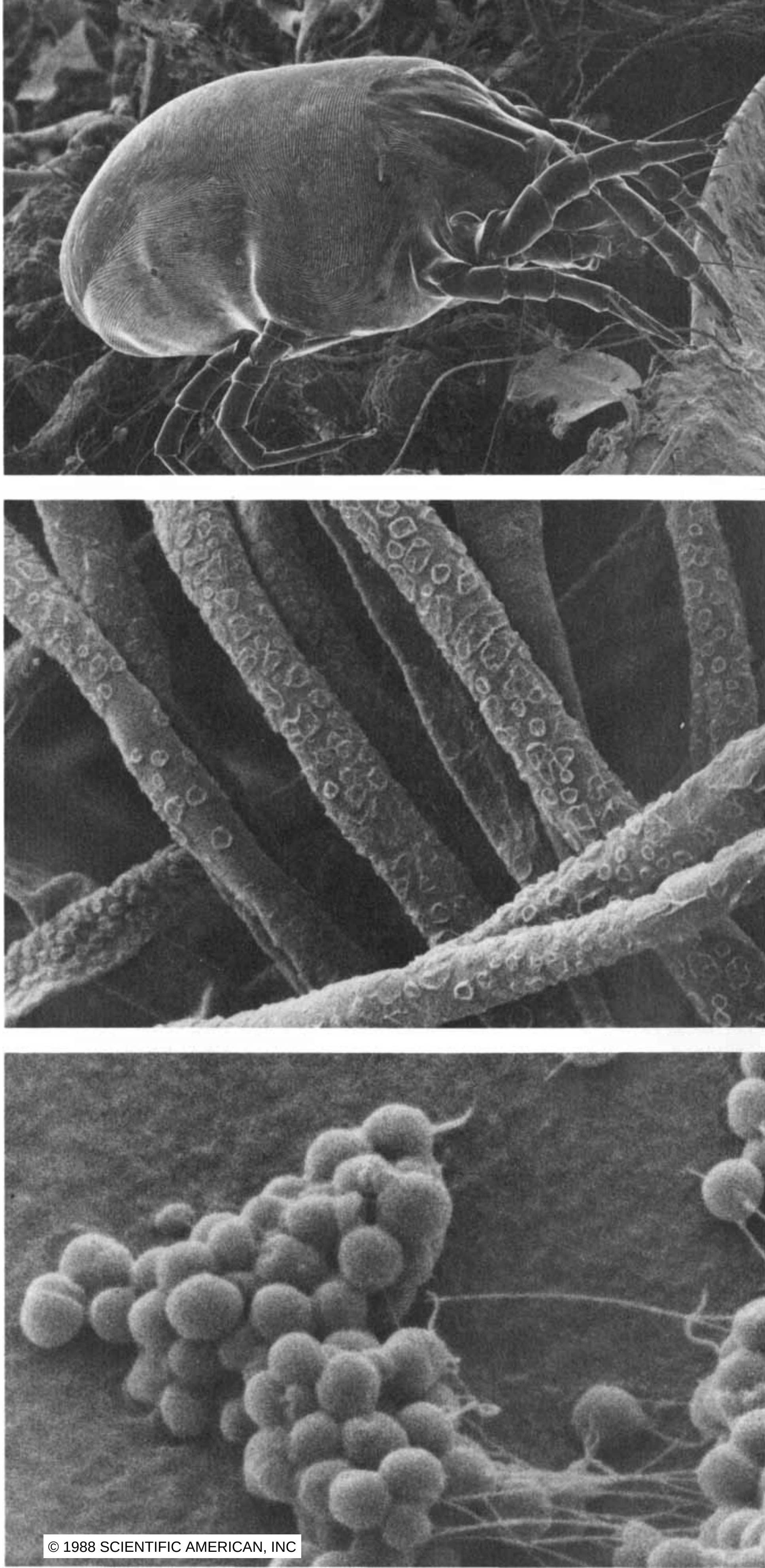
One of the most familiar indoor pollutants is cigarette smoke, which is composed mainly of organic aerosols, or tiny airborne particles. Heating and cooking appliances that burn natural gas, kerosene, oil and wood (or peat, coal and dung, as is the case in some countries) also emit varying quantities of respirable particles along with carbon oxides, nitrogen oxides and trace organic chemicals. Less familiar indoor pollutants include methylene chloride, formaldehyde and a vast range of more complex organic chemicals that are given off by building materials, furniture, cleaning fluids, pesticides, paints and paint strippers. Respirable fibers of asbestos can be released indoors from insulating material in older buildings. Indoor air pollutants even include the products of living organisms or the organisms themselves, such as bacteria, fungi and house mites. Perhaps most disconcerting for the occupants of many houses is the fact that the ground on which the houses sit is a major source of radon: a radioactive gas that is generated naturally as a product of the nuclear decay of radium, an element present in trace quantities throughout the earth's crust.

These classes of pollutants, which

are almost always present in indoor air, vary over a wide range of concentrations among residential and office buildings. The variability depends primarily on the rate at which each pollutant is emitted into the indoor atmosphere. A home with vented heating and cooking appliances (and without smokers), for example, will ordinarily not have high concentrations of combustion products in the air. Similarly, measurements of radon concentrations done in 100 U.S. houses in the late 1970's showed that differences in the entry rate of the gas accounted for most of the variability in the concentrations among the houses.

It is now understood that radon enters houses in the air that is drawn from the underlying soil by small differences in air pressure between indoors and outdoors. These pressure differentials (amounting to no more than .0001 atmosphere) arise in part from the "stack" effect: the tendency for air to rise whenever it is warmer than the surrounding atmosphere. (The stack effect is the working principle of a chimney.) Hence whenever it is warmer indoors than outdoors, air tends to flow in through the lower part and out through the upper part of a building's shell (its substructure, walls and roof). It does this by "infiltrating" through various openings, particularly around windows and doors and penetrations for pipes and wiring. Winds blowing against a building can also produce similar pressure differences—and thus airflows—across the building's shell, although the airflows exhibit a pattern substantially different from the patterns produced by the stack effect. Only a small part of the infiltrating air actually comes from the soil, but this is what carries the radon into houses. The entry rate of the gas therefore depends on the permeability of the soil to airflow, along with other geologic, meteorologic and structural factors.

The bulk of the infiltrating air actually comes from the outdoor atmosphere and accounts for the general ventilation of small structures such as houses. (In contrast, large structures such as office buildings usual-



**MICROORGANISMS found in ordinary household dust, such as mites (*top*), fungi (*middle*) and bacteria (*bottom*), are classified as indoor air pollutants because they can cause allergic reactions and other illnesses, as more familiar pollutants do.**

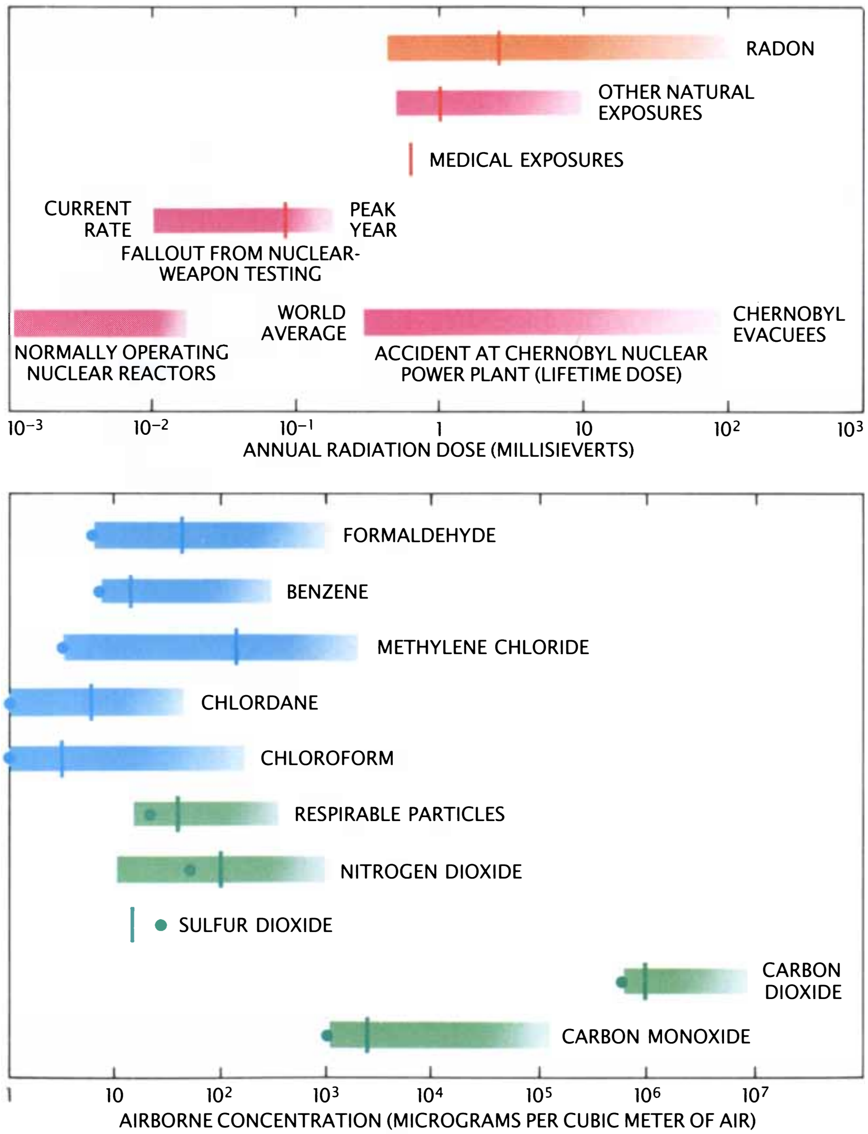
ly rely on fans, air conditioners and ducts to provide ventilation.) A surprising amount of outdoor air manages to leak into a house by means of infiltration. In most houses outdoor air replaces the volume of indoor air every one to two hours.

Such an exchange of air represents a constant energy "drain," since energy has to be expended to heat or cool the outdoor air as it replaces the indoor air. Consequently in the mid-1970's energy-conscious homeowners in the U.S. and elsewhere began to reduce infiltration by weather stripping, plugging openings in the building's shell and caulking around windows and doors. Studies indicate that these measures reduce ventilation rates by roughly 10 to 30 percent, depending on how carefully the measures are implemented. While the modest decreases result in useful energy savings, the resulting changes in the concentrations of indoor pollutants are small compared with the ten- or hundredfold difference in concentrations observed between one house and another.

There is in addition a third factor that determines indoor pollutant concentrations: the rate at which a particular pollutant reacts with other airborne species or interior surfaces. Nitrogen dioxide, for example, is found to be removed from indoor air as much by such reactions as by ventilation. The chemical form and concentrations of radon's decay products (isotopes of polonium, lead and bismuth) also depend on the amount of airborne particles and the pattern of air movement in a particular building, influencing the radiation dose the products ultimately impart to the lungs when they are inhaled. Many other potentially important aspects of indoor-air chemistry remain virtually unexplored.

Taken together, the variability in entry rates, ventilation rates and reaction rates is the cause of an impressive range of concentrations for most indoor pollutants. No better example can be given than that of radon. In the U.S. concentrations in single-family houses vary over four orders of magnitude (factors of 10)—from a few becquerels per cubic meter of air to more than 10,000, with an average level of about 50 becquerels per cubic meter. (One becquerel is equal to one radioactive decay per second. Another common unit of measurement is the picocurie per liter, which is equal to 37 becquerels per cubic meter.) The average indoor level represents a radiation dose about three times larger than the dose most people get from X rays and other medical procedures in the course of their lifetime. Those exposed to higher levels receive proportionately higher doses. Indeed, hundreds of thousands of Americans living in houses that have high radon levels receive as large an exposure of radiation yearly as those people living in the vicinity of the Chernobyl nuclear power plant did in 1986, when one of its reactors exploded and released radioactive material into the environment.

The wide range of pollutant types and concentrations entails a correspondingly wide range of health risks. Cigarette smoke, asbestos fibers, the decay products of radon, formaldehyde and many other organic chemicals are demonstrated or potential carcinogens. Most of these pollutants can also lead to chronic or acute diseases, such as respiratory infections and allergic responses, as can combustion products in general and a variety of indoor bacteria and fungi. Extremely high levels of carbon monoxide—a combustion product—can even result in immediate death. Yet only in a relatively few cases, such as acute allergic reactions or carbon monoxide poisoning, is there a clear-cut relation between a given exposure to an indoor pollutant and an associated health effect. More often than not a given instance

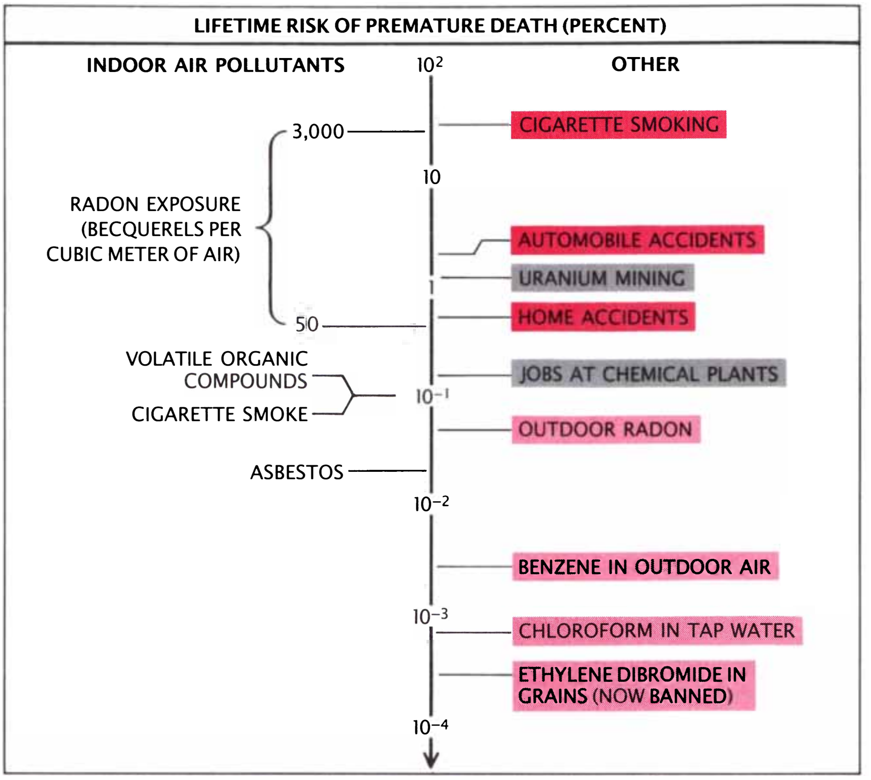



**INDOOR AIR-POLLUTION LEVELS (*bars*), measured in terms of radiation exposure due to radon (*orange*), concentrations of organic chemicals (*blue*) and concentrations of combustion products (*green*), span orders of magnitude (factors of 10). Furthermore, the average indoor levels (*lines*) are generally higher than the average outdoor levels (*dots*). Exposures from other sources of radiation (*red*) are shown for comparison.**

of respiratory disease or cancer cannot be directly attributed to a specific cause, environmental or otherwise.

Instead scientists study the occurrence of pollution-related diseases in heavily exposed groups (sometimes human beings but more often animals) in order to obtain probabilistic relations between exposures to pollutants and the chance that the diseases will appear in the general population. This approach provides the basis for estimating, albeit usually with substantial uncertainty, the risk of cancer and other diseases associated with the lower pollutant concentrations generally found in water and air, both indoors and outdoors.

In quantifying such health risks one must also consider the risk of suffering nonfatal diseases. Because these illnesses can occur with higher frequency, the overall risk they pose may be judged to be as important as, say, the risk of cancer. Yet it is difficult to treat all illnesses—fatal or otherwise—on a common basis; this would require an equivalence between days sick and days of life lost. It may also be inappropriate to do so, since acute illnesses such as allergic responses are immediately apparent to the sufferer, not merely hypothetical risks, as is the case with cancer.



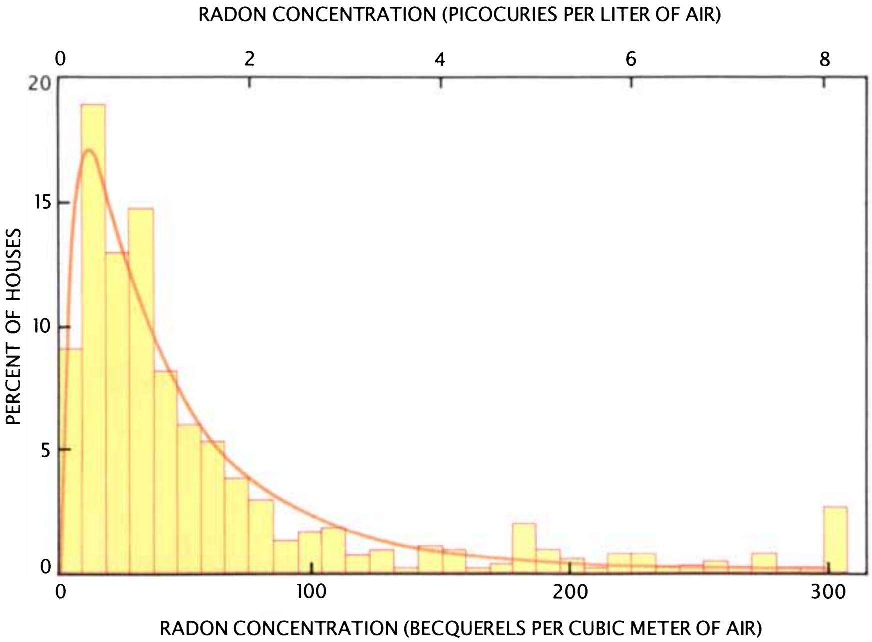

**ESTIMATED PROBABILITY of suffering a fatal disease is substantially higher for exposure to indoor air pollutants (*black*) than for exposure to the pollutants in outdoor air, drinking water and food (*pale red*). The risk of death from exposure to indoor pollutants, however, is no more than that from certain voluntary activities (*red*), such as smoking, and occupational hazards (*gray*), such as those faced in mining uranium.**

In order to estimate the risks posed by indoor radon, for example, results from epidemiologic studies of underground miners who were exposed to high concentrations of radon's decay products are extrapolated to the lower exposures characteristic of a typical house. Based on these estimates, the average indoor concentration of radon in the U.S. corresponds to a chance of contracting lung cancer of about one in 250, or .4 percent, which would account for approximately 10,000 lung-cancer deaths per year in the country. Although the risk estimates for indoor radon exposure do have a degree of uncertainty, it is much less than for estimates of risk from other environmental-pollutant exposures, such as those from toxic chemicals, which are usually based on extrapolating doses and responses a thousandfold. In fact, no extrapolation is needed to estimate the health risks in houses with exceedingly high levels of radon. For example, people who have lived for 20 years in houses that have 1,000 becquerels of radon per cubic meter (which number in the tens of thousands in the U.S.) face an additional 2 or 3 percent chance of contracting lung cancer.

These figures are nothing short of remarkable. Pollutants in the outdoor environment are regulated so that the estimated risks of premature death from exposure to them are usually less than .001 percent. Indeed, just about the only environmental risks—at least for fatal disease—that are comparable to indoor radon arise from other indoor pollutants. Although the estimate is highly tentative, the risk of cancer arising from exposure to a wide range of organic chemicals in the indoor environment can be said to be about .1 percent. The risk of premature death due to typical exposures to asbestos is thought to be about .02 percent, with most of this risk resulting from indoor exposures. Both of these estimates require more than an order of magnitude greater extrapolation from epidemiologic or animal studies than the estimates of the risk associated with radon exposure. Finally, the risk of lung cancer attributed to breathing smoke from other people's cigarettes is estimated to be about .1 percent.

Yet the level of risk posed by pollutants at concentrations found indoors is either in the same range as or lower than other risks that are accepted in exchange for some personal benefit. These risks include diseases caused by occupational exposures to toxic chemicals and work-related accidents as well as automobile accidents, which people are willing to accept—within reasonable bounds—in order to earn a salary or to have the convenience of personal transportation. The risk of death from automobile accidents, for example, averages about 2 percent in the U.S. People also seem willing to accept a .5 percent risk of dying in a fall or fire at home in return for the comfort of living indoors. Many are even ready to accept the 30 percent risk of premature death associated with smoking—a risk that is rivaled only by exposure to the highest indoor radon levels known—for the sake of personal pleasure.

How, then, can one approach the problem of indoor air pollution, which entails risks that exceed common "environmental" risks but not the risks people tacitly accept when driving, smoking or simply living




**DISTRIBUTION OF RADON CONCENTRATIONS suggests that about 2 percent of the houses in the U.S. (numbering some one million) have concentrations greater than or equal to 300 becquerels per cubic meter of air, which is five times the average level.**

in houses? Any overarching strategy for controlling the risks associated with indoor-pollutant exposures requires three basic, interdependent elements: a system of advisory or regulatory standards that determines the overall attack on the problem, a methodology for identifying the situations of greatest concern and a framework for selecting control techniques suited to each situation.

The underlying system of standards can take fundamentally different forms, depending on the objectives. One objective might be to control the average exposure of the entire population; a contrasting objective is to avoid extreme levels, thereby limiting individuals' risk of disease, fatal or otherwise. In any case, the objectives can be achieved by formulating standards that either control the factors affecting pollutant concentrations or establish limits on the concentrations themselves.

Actually both types of standard are applied in the control of outdoor pollution. The release of pollutants from automobiles and power plants, for example, has been controlled by standards that modify such factors as combustion processes. Concentration limits in turn are embodied in standards for outdoor air and water, which are meant to protect the population at large to a higher degree than individuals who are exposed to the same pollutants in the workplace.

Concentration limits for outdoor air usually apply to a large environmental region—an "air basin," such as that of Los Angeles—so that exposures are more or less consistently limited throughout a large population. Because conditions vary greatly from one building to another, a comparable approach to controlling indoor pollution would entail dealing with some 80 million air-quality-control "districts" in the U.S.—one for each building. As a result concentration limits for indoor pollutants are most effectively aimed at avoiding excessive individual exposure rather than controlling the average exposure of a population.

The fact that the indoor environment has significant inherent risks has to be recognized and near-term attention focused on the exceptional situations—including truly high levels of radon, organic chemicals or combustion products, as well as the occasional excessive levels of flaking asbestos and even house mites. This approach contrasts sharply with suggestions to limit formaldehyde concentrations to zero (within the limits of instrumental sensitivity) or to limit indoor radon concentrations to outdoor levels (about 10 becquerels per cubic meter of air). These proposals do not appear to recognize that the risks associated with average exposures to formaldehyde and radon are solidly within the range of risks that are normally accepted.

Nevertheless, average exposures to indoor pollutants can also be gradually lowered as a long-term goal. This can be achieved by establishing standards that regulate the factors affecting indoor-pollutant concentrations. For this reason it is important to identify how source, ventilation and structural characteristics affect the concentrations. Such knowledge can be applied in the formulation of specific criteria for the design, fabrication and utilization of new buildings and furnishings that will ensure acceptable indoor air quality in the great majority of cases.

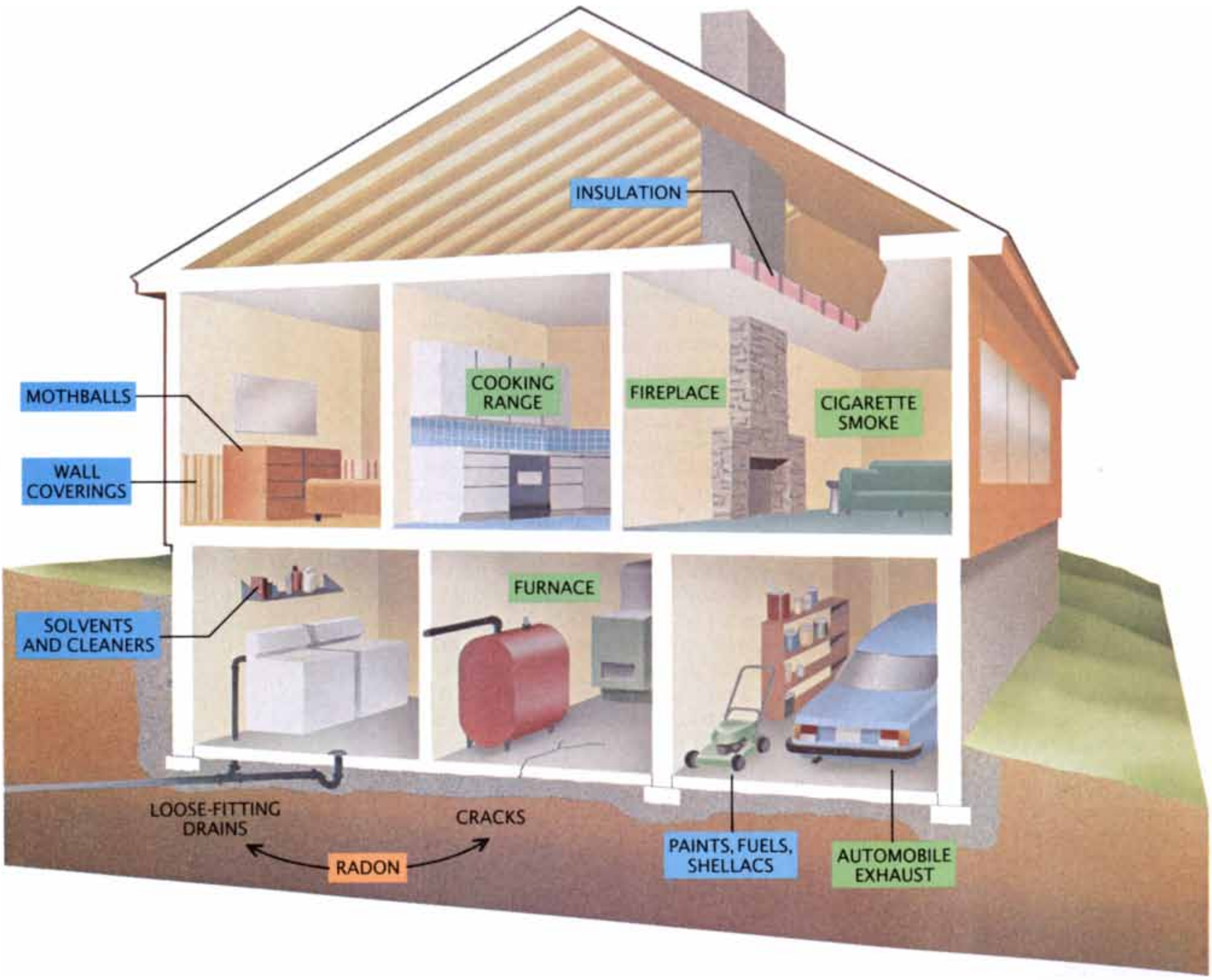
Studying the behavior of indoor pollutants also helps in the development of a methodology by which to identify buildings that have or are likely to have excessive pollutant levels. For example, knowledge that certain materials or appliances are often associated with high levels of organic pollutants or combustion emissions can lead to monitoring in buildings containing those products. Similarly, analysis of the general geologic, architectural and meteorologic factors affecting indoor radon concentrations might serve to identify regions where excessive levels are likely to occur.

Once a building has been identified as needing remedial action (whether it is already constructed or still in the planning stages), control techniques can be implemented. A number of control techniques have already been developed, corresponding roughly in both design and effectiveness to the fundamental factors affecting indoor concentrations. Given that the emission or entry rate is the primary determinant of concentration levels, measures to reduce pollutant sources provide the primary control—assuming that ventilation rates are in the normal range. Emissions of formaldehyde and other volatile organic substances, for instance, can be lessened by changing the way particle board, adhesives and other products are manufactured. Good burner and exhaust designs in heaters and cooking appliances can decrease the concentration of combustion products. Entry of radon from the ground can be diminished markedly by means of simple systems of pipes and fans that draw air from (or blow air on) the soil or gravel immediately under the substructure of a house.

If it is determined that neither infiltration nor opening windows is sufficient to ensure an adequate ventilation rate, mechanical systems can be installed. In large buildings these systems can be quite complex, designed to meet performance criteria that are often incorporated into building codes. In houses with little infiltration much simpler systems can be employed, such as a single exhaust fan. (To save energy the system can be designed so that exhaust and intake airstreams exchange heat.) Although infiltration and mechanical ventilation can provide a basic level of protection, they cannot be relied on to reduce pollutant concentrations substantially: the required increase in ventilation rate would ordinarily be more difficult and more costly to achieve than eliminating or reducing the source of the pollutants.

An alternative means of control might be to physically clean the air of airborne gases or particles. This approach suffers some of the same limitations as ventilation, namely that marked reductions in pollutant concentrations would require high rates of processing the air. In fact, many of the less expensive and popular tabletop air cleaners provide only very low and generally inadequate cleaning rates. Of more concern is the fact that better devices may not even reduce the overall exposure. For example, in drastically reducing the total concentration of radon's decay products, filter systems and electrostatic precipitators (which remove particles suspended in air by imparting an electric charge to them) increase the fraction of the decay products that are not attached to airborne particles. Unfortunately the free decay products appear to cause the greater radiation dose to the lung. Hence in spite of the fact that air cleaning may reduce the total concentration of radon's decay products, it does not necessarily reduce the exposure to radiation.

Air cleaning may find a role in control of biologic particles, such as bacteria, fungi or residue from house mites, that do not behave like chemical or radioactive pollutants. In particular these organisms actually multiply given the right conditions, requiring a different perspective on control. The most effective approach




**SOURCES OF AIRBORNE POLLUTANTS in a typical house are myriad. Combustion products (*green*) are traced to cigarette smoke, heating and cooking appliances and perhaps automobile exhaust. Organic chemicals (*blue*) are given off by substances in paints, plywood, solvents and adhesives. Radon (*orange*) is drawn into a house by small differences between outside and inside air pressure; the gas seeps through cracks in the house's foundation and through openings around loose drainage pipes.**

in their case would be to combine air cleaning or increased ventilation and a reduction in indoor humidity.

Current efforts to control indoor air quality have in fact followed the two basic approaches: setting specific concentration limits and modifying the design and manufacture of buildings and their contents. The Netherlands has adopted standards limiting indoor formaldehyde concentrations to 120 micrograms per cubic meter, and Canada has set a limit of 150 becquerels per cubic meter for indoor radon in uranium-mining communities. West Germany and the U.S., on the other hand, have set standards that limit formaldehyde emission from wood products, such as plywood, and have considered regulating the emissions from unvented fossil-fuel heating appliances. Policymakers, however, have come to acknowledge their naiveté in adopting concentration limits as the primary basis for control: emission and ventilation standards have been found to be more workable.

Attempts have also been made in recent years to include optional concentration limits for identified classes of pollutants as part of the ventilation standards incorporated into building codes. This illustrates a dangerous tendency to expand each of the approaches in an overall control strategy to include the other, leading to an overt or implicit confusion of objectives. A self-consistent and comprehensive strategy would rely on the factors that determine ventilation and emission rates as "handles" for keeping concentrations within (implicitly) acceptable ranges in most new buildings and adopt explicit concentration limits mainly as criteria for reducing truly excessive levels found in existing structures.

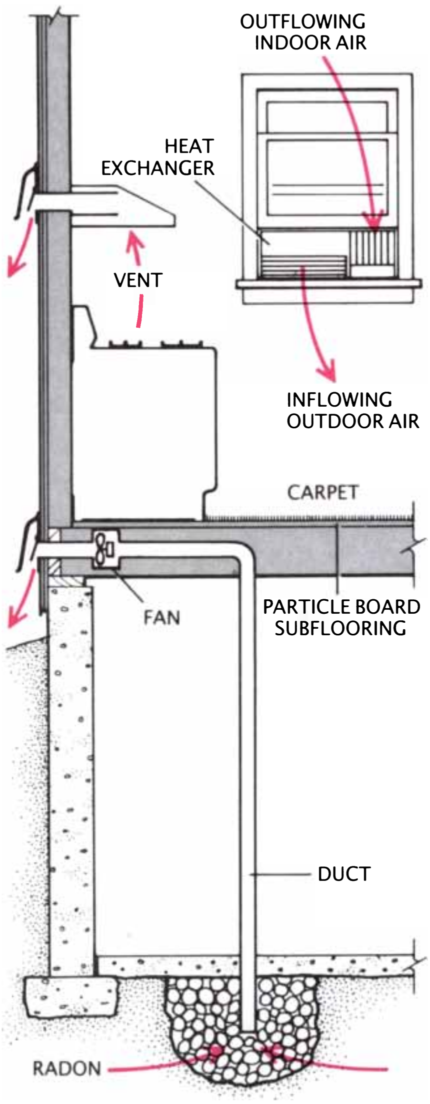
In spite of the confusion in objectives, there at least is some consensus on where the major responsibilities lie for implementing control strategies. This consensus is important, because many actors are directly involved. They include national and local government agencies, professional organizations, manufacturers of building materials and household appliances, builders and contractors. For that matter, anyone who is active in issues of health, environment, housing, energy and consumer products as well as in related areas such as demography, meteorology and geography should necessarily be involved.

Yet the success or failure of a program to control indoor air quality ultimately hinges on the behavior of the owners and occupants of buildings. At present occupants are often not aware of potential health problems caused by the way they use certain appliances or substances—or for that matter of the effect they have on others by smoking cigarettes in enclosed spaces. Even building managers may not know for what activities the building was originally designed or how the ventilation equipment in the building is meant to be operated.

For this reason some local governments are considering requiring every manager of a building to conform to the assumptions made in its design. In this way a chain of responsibility from the engineer, architect and builder to the owner and occupants can be maintained. An important element in such a system might be a supporting document associated with a building, much like a deed, that describes the design assumptions relevant to indoor air quality and records changes in the building's occupancy or ventilation equipment. In the case of office buildings the document could list a building's smoking areas and indicate the extent to which furnishings that emit organic chemicals can be used.

A simpler document for private houses could maintain a record of radon measurements and of any remedial techniques and maintenance procedures that may be needed. It would be similar to the document required in many areas that certifies a house for sale is free of termites. Indeed, in areas where high radon concentrations have been reported real-estate transactions are already including this information. Such a document would be one of the more visible manifestations of a comprehensive control strategy.

The control of indoor air pollution goes beyond the conceptual basis of current regulatory structures, which were built for controlling exposures to pollutants in industrial or outdoor settings. The fact is that the health risks posed by indoor pollutants have to be considered in their own right, and the objectives and approaches for a rational pollution-control strategy must be thought out anew. Such rethinking depends on science to evaluate the health risks, elucidate the correlation between pollutant concentrations and the factors that influence them, develop methods for measuring pollutant concentrations and determine the effectiveness of control techniques. The resulting science and policy of indoor air quality might even change how we think about other pollutant exposures, leading to a more realistic perspective on environmental risks in general.




**CONTROL MEASURES seek to lower concentrations of pollutants by lowering the rate at which they enter the indoor atmosphere or by increasing the rate at which outdoor air replaces indoor air. The amount of formaldehyde released from particle board and carpets can be reduced by changing the resins and binders used in their manufacture. Venting a cooking range usually results in a marked drop in the concentration of combustion products. The rate at which radon enters a house can be lowered by reducing the air pressure under the building. This can often be done by means of a single fan and a duct. Finally, heat exchangers can improve a house's general ventilation rate without significantly increasing energy costs, since they heat or cool inflowing outdoor air with outflowing indoor air.**

# Squeezed Light

*Quantum noise, or fluctuation, in light waves limits the precision of measurements. A solution is to make part of a wave quieter by making another part noisier*

by Richard E. Slusher and Bernard Yurke

Silent light. Can a beam of light ever be truly "quiet," free from the random fluctuations called noise? Or are physicists destined forever to make observations and measurements with imperfect, noisy light? After all, the quantum theory states that any light must be accompanied by a certain minimum amount of fluctuation, which limits the fundamental precision of observations carried out with light beams. Nevertheless, there seems to be a ray of hope. Several groups of investigators, including our own, have found that it is sometimes possible to "squeeze" the noise in a light beam: to redistribute it so that parts of the light wave are less noisy than before, although other parts of the wave become noisier. The resulting wave can then be used in highly precise measurements.

A beam of light consists of an oscillating electromagnetic field. In the world view of classical physics, the oscillations of the field can be pictured as a smooth wave, whose shape can be described with absolute certainty. According to the quantum-mechanical uncertainty principle, however, such certainty is unattainable; the best one can say is that the wave's shape fits within a particular "envelope" of uncertainty. That uncertainty is manifested as noise: small, random fluctuations in the electromagnetic field.

Indeed, according to quantum mechanics there must be some amount of noise even in "darkness"—when no other light is present. In classical physics the wave representing darkness would be flat, with no undulations (in some sense, it would not really be a wave at all). In quantum mechanics, on the other hand, one can say only that the wave is flat to within some small degree of uncertainty; within the envelope of uncertainty the wave fluctuates randomly. In practical terms, this means that even in a vacuum, with no external light sources, there must still be small fluctuations in the electromagnetic field.

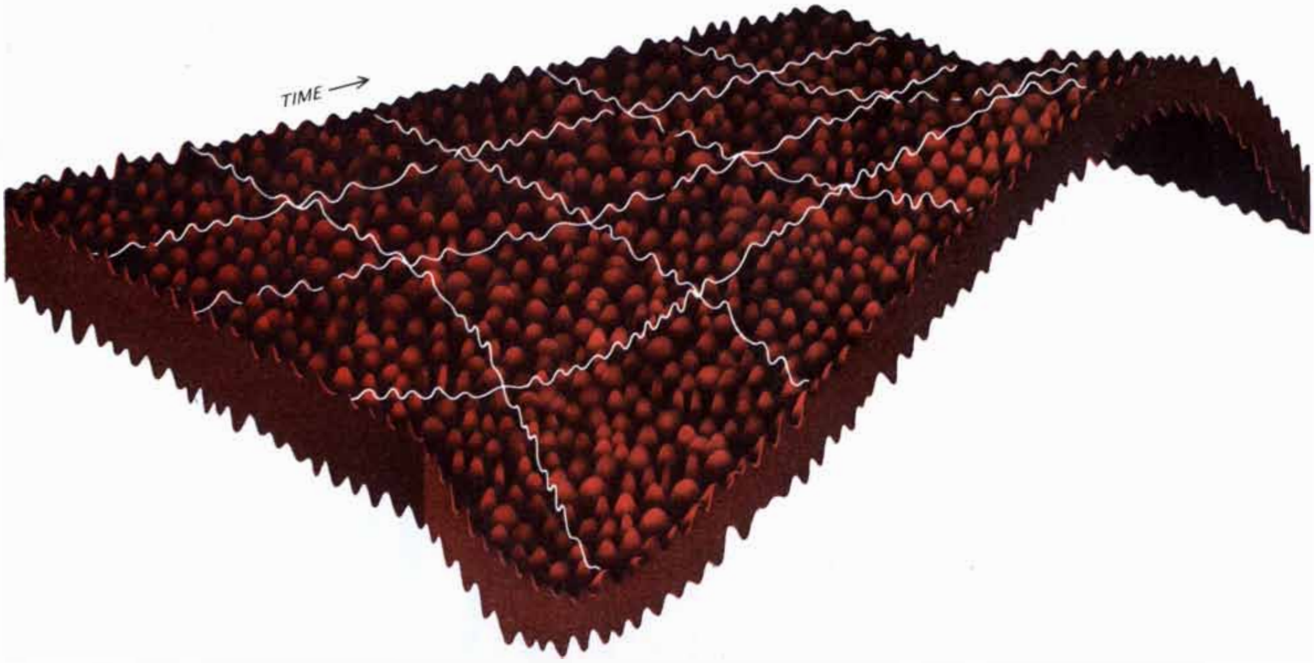
In a sense it is really these vacuum fluctuations that underlie the fluctuations within ordinary beams of light. The wave representing a light beam can be seen as consisting of the fuzzy, irregular vacuum fluctuations superposed on the smooth wave described by classical physics. One could say that it is interference with the vacuum fluctuations that causes ordinary light waves to be noisy.

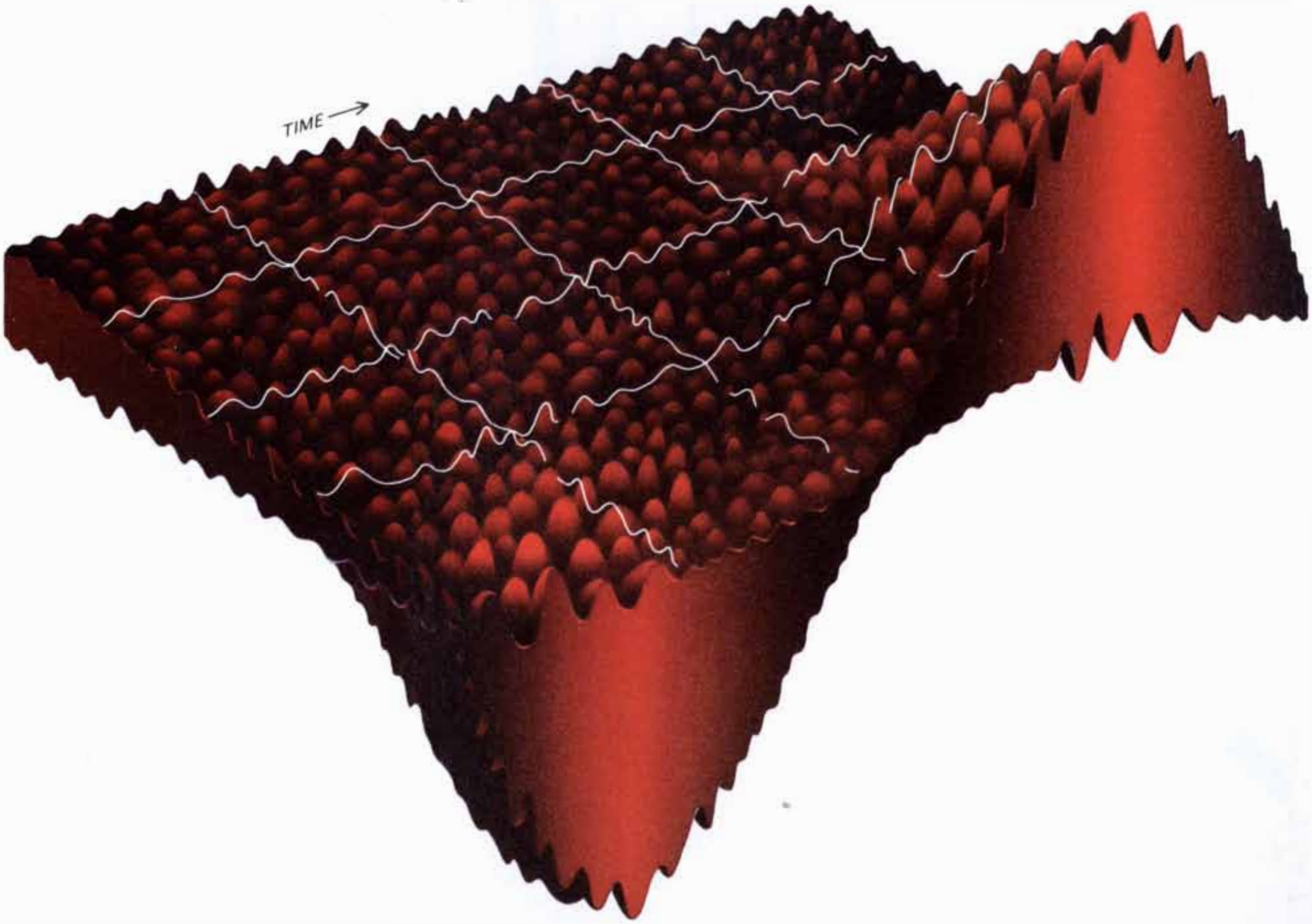
This noise fundamentally limits the degree of precision possible in measurements involving light. It limits, for example, the precision of interferometers that gauge the interference pattern arising between beams of light that have been reflected off massive objects and then combined; such interferometers, which detect very small changes in the relative positions of the massive reflectors, are components of devices designed to detect gravitational waves. Noise also limits the precision of spectroscopy, in which the frequency and intensity of the radiation emitted by atoms or molecules yield information about their properties. Eventually quantum noise will also limit the power of such technologies as optical computing and optical communications.

Squeezed light provides a way around some of these limits. The noise that bedevils such measurements can be squeezed from some parts of a light beam or, more directly, from the quantum-mechanical vacuum itself. The end product of squeezing the vacuum is a propagating beam that, in a special sense, contains less fluctuation than does an absolute vacuum: a beam of "light" that is darker than dark. By making observations with a beam of squeezed light and a detector that looks only at the squeezed part of the light wave and not at the noisy part, one can make measurements with more precision than seems possible under the constraints of the uncertainty principle. (In fact, the uncertainty principle is not actually violated; the necessary uncertainties are merely redistributed in time.)

How does the experimenter go about squeezing light? A somewhat extended physical analogy is an aid to understanding the mechanisms involved. The oscillations of the electromagnetic field are analogous in some ways to the motions of a child's playground swing. Like the motions of a light wave, the oscillation of the swing has a particular frequency (the number of cycles, or round trips, the swing makes in a given amount of time) and amplitude (the height to which the swing rises in each cycle). The relative motion of two children on swings, like that of two light waves, is expressed as their relative phase: the children are said to be swinging in phase if in every cycle they both reach the top of their motion at the same time. They are out of phase, to a greater or lesser degree, if they reach the top of their motion at different times.

Suppose a child in a playground chooses to pump a swing in a slightly unconventional way. Instead of leaning back and then forward once per cycle (the conventional method of pumping), he decides to stand on the seat or squat down on it at different times during the swing's cycle. If he squats as the swing approaches the top of its motion and stands when the swing gets closest to the ground (thereby doing work against centrifugal force), he will be adding energy to the swing's motion and amplifying it:



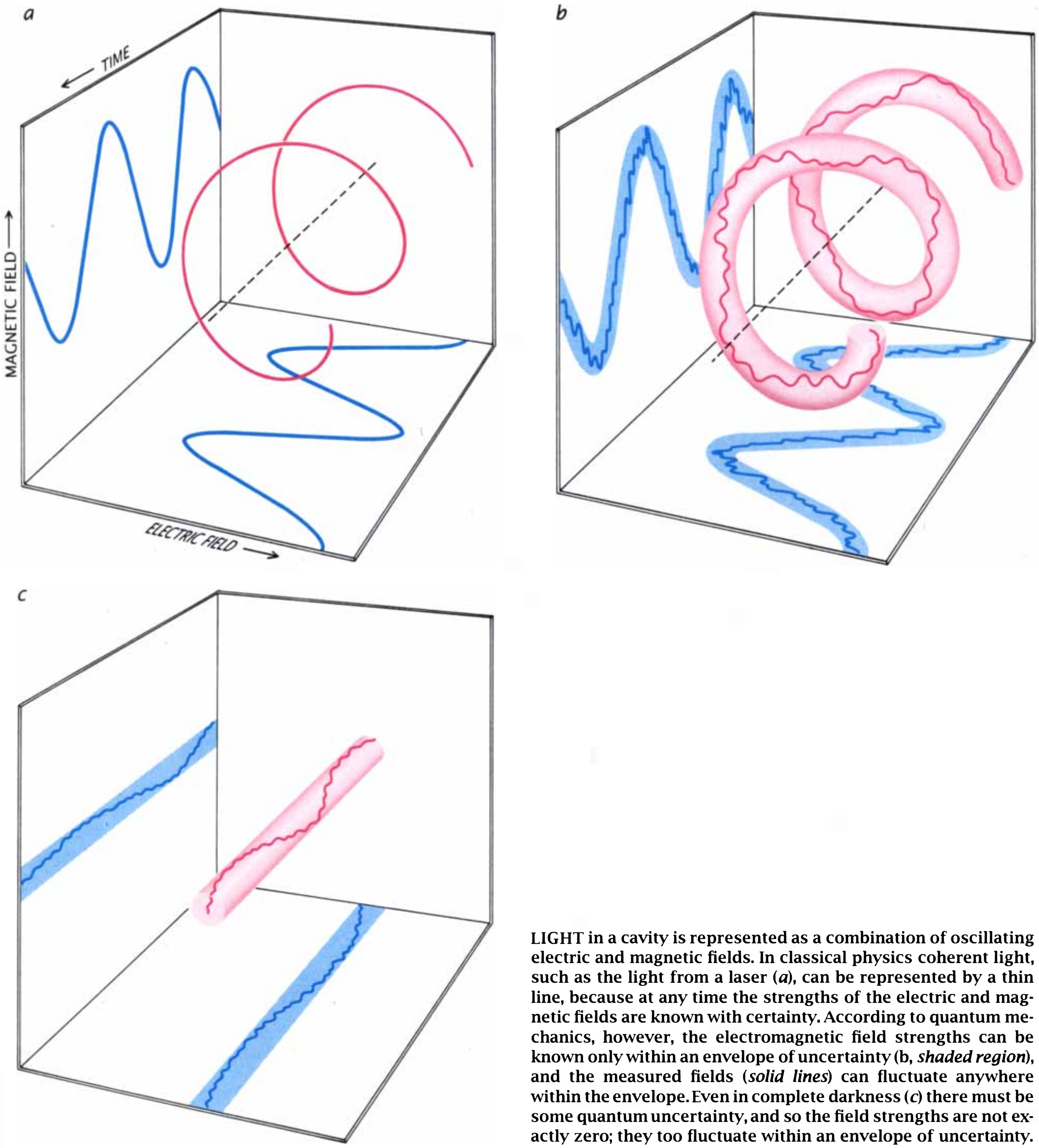

QUANTUM NOISE—the random fluctuation required by the laws of quantum mechanics—undermines the precision of light waves. The two shapes depicted here represent light waves. The noise inherent in the waves is reflected in their thickness and irregularity; if there were no noise, light waves would be represented by smooth, two-dimensional sheets. In ordinary laser light (*top*) the noise is uniform throughout the wave. Uniform noise is also present in the darkness outside the laser beam (represented here by the flat part of the irregular shape). In "squeezed" light (*bottom*) noise has been manipulated so that part of the wave is less noisy (thinner and smoother in this depiction) than unsqueezed light, although other parts are noisier.

the swing will go higher in every cycle. On the other hand, if the child squats when the swing is nearest the ground and stands as it rises, he will deamplify the swing's motion and it will not go as high in each subsequent cycle.

For the sake of our analogy, the child's pumping motion has two important characteristics. The first is that the child stands and squats twice for every round trip made by the swing. Put in more technical terms, the frequency of the pumping action is twice the frequency of the swing's motion. The second important characteristic is that the swing's motion can be amplified or deamplified, depending on the relative phase of the standing and squatting (that is, depending on whether the child stands or squats at the top of the cycle).

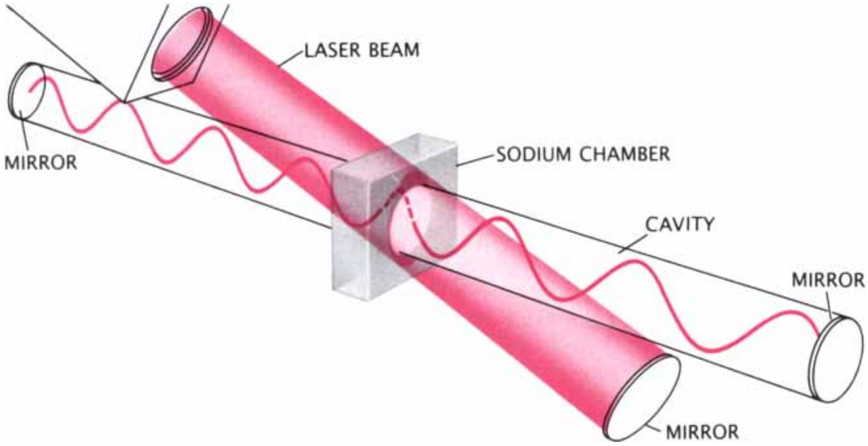
A similar kind of pumping can be applied to a light wave confined in an elongated cavity that has mirrors at both ends. Suppose the length of the cavity is an exact multiple of the wavelength of the light, so that a whole number of wavelengths fit within the cavity. As it is reflected from each end of the cavity the confined wave interferes with itself and resonates as a so-called standing wave, much as a sound wave resonates within an organ pipe.

Now suppose it were possible to move the mirror at one end of the cavity back and forth (and thereby change the length of the cavity) at ex-




**LIGHT in a cavity is represented as a combination of oscillating electric and magnetic fields. In classical physics coherent light, such as the light from a laser (*a*), can be represented by a thin line, because at any time the strengths of the electric and magnetic fields are known with certainty. According to quantum mechanics, however, the electromagnetic field strengths can be known only within an envelope of uncertainty (b, *shaded region*), and the measured fields (*solid lines*) can fluctuate anywhere within the envelope. Even in complete darkness (*c*) there must be some quantum uncertainty, and so the field strengths are not exactly zero; they too fluctuate within an envelope of uncertainty.**

actly twice the frequency of the light wave. Changing the length of the cavity in such a regular way would add or subtract energy from the light wave as it bounced off the mirror. Like the child's pumping motion, the mirror's vibration would either amplify or deamplify the light wave, depending on the relative phase of the mirror's vibration and the light wave's oscillation. If the mirror's motion has the right phase in relation to the light wave's oscillation, the wave is amplified: the oscillating electromagnetic field grows stronger. If the mirror's motion has the complementary phase, the wave is deamplified and the oscillating electromagnetic field grows weaker.




**APPARATUS for squeezing light is centered on a cavity that has mirrors at both ends. The cavity's length is an exact multiple of the wavelength of the light to be squeezed, so that the light resonates within the cavity, forming a set of standing waves. A chamber of sodium gas is placed within the cavity (*center*), and a laser beam is directed through the gas. The laser beam is reflected back on itself by a mirror (*lower right*) to form another standing wave (not shown) within the sodium chamber. This wave causes rapid variations in the optical properties of the sodium, and hence rapid variations in the time light takes to travel the length of the cavity. As a result some of the waves resonating within the cavity are amplified, whereas others are damped. The effect is to increase the total amount of noise in some parts of the light and decrease the noise in others.**

To return to our analogy, now imagine an entire playground full of children, all on swings. They are all swinging at the same frequency, but they are not in phase: at any given moment some children may be near the top of the cycle while others may be close to the ground. Suppose a teacher comes into the playground and shouts "Stand...squat...stand...squat..." into a megaphone at exactly twice the frequency of the children's swinging. Some of the children will be at or near the bottom of the cycle every time the teacher shouts "Stand," and others will be at or near the top. Those who stand near the bottom of the cycle will amplify the motion of their swings, while those who stand near the top will deamplify the motion of theirs.

Eventually the children who stand near the bottom of the cycle will be swinging with much greater amplitude than before, and those who stand near the top of the cycle will hardly be swinging at all. Then the playground will contain a number of children who are swinging very high—and are swinging almost exactly in phase with one another and with the orders given by the teacher—and a number of other children who are essentially not swinging at all, even though they continue to stand and squat.

It turns out that the vacuum noise in a cavity (the noise that is present even in darkness) bears a close resemblance to the playground full of children swinging randomly. The vacuum noise can be thought of as consisting of many waves, all having the same frequency but randomly varying amplitudes and phases. When all these waves are added together, they produce the fuzzy, uncertain wave of the vacuum noise.

If the many waves that make up the vacuum noise are analogous to the many children on swings, what component of the cavity is analogous to the teacher? This role is played by the vibrating mirror. Remember that the mirror amplifies all waves that have a certain phase with respect to its vibration and deamplifies all waves that have the complementary phase. The mirror will therefore amplify some components of the vacuum noise and damp out other components. Eventually the noise inside the cavity will consist of a number of waves of relatively high amplitude that are nearly in phase with one another (and with the mirror's vibration) and a number of other waves that have extremely low amplitudes and are out of phase with the high-amplitude waves. This combination of waves is called a squeezed state: the vacuum noise has been squeezed out of some phases and into others.

Squeezed vacuum noise is fundamentally different from ordinary noise. Ordinary noise is a constant fuzz of uncertainty; in the squeezed state, on the other hand, the amount of noise increases and diminishes periodically. To see why that is so, consider just one of the component waves and compare it with a single child on a swing. In one cycle of the swing the child goes from a position near the ground to a position high above the ground in front of the swing set, back down to a position near the ground, up to a position high above the ground behind the swing set, and back down again to a position near the ground. The wave in the cavity represents oscillations of the electromagnetic field. Like the position of the child in relation to the swing set, the strength of the electromagnetic field during one wave period goes from zero up to some maximum value, back to zero, down to some maximum negative value and back up to zero again.

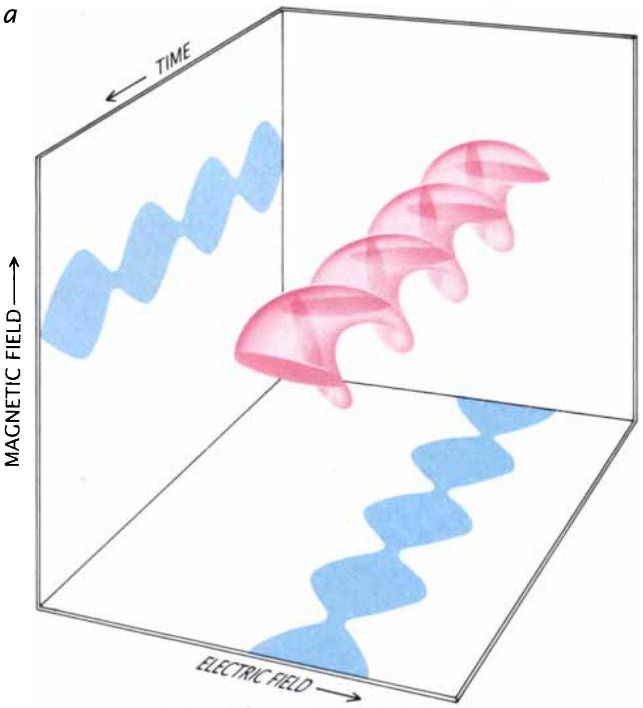
In every cycle of the wave, then, there are moments during which the strength of the field passes through zero. In the squeezed state, the waves that make up the noise are all approximately in phase, and so they all pass through zero at about the same time. This means that the field strength in the cavity (the sum of the fields represented by all the waves) will periodically be very low—just as if there were almost no electromagnetic noise at all. Between these moments the field strength will be very high as all the waves simultaneously reach their maximum positive or negative amplitude. The squeezed state thus consists of moments of what seem to be very low noise alternating with moments of what seem to be very high noise. Put another way, the squeezed vacuum is represented by a wave made up of alternating regions of low-amplitude fluctuations and high-amplitude fluctuations [*see illustration on next page*]. The envelope of uncertainty within which the field fluctuates has been pinched together in the former regions and made wider in the latter.

Our analogy seems to make impossible demands of the "teacher": the vibrating mirror. In fact one cannot vibrate a mirror at the necessary high frequencies. Nevertheless, we and our colleagues at the AT&T Bell Laboratories have used a mechanism that has the same effect as the hypothetical mirror to generate the first squeezed light ever observed.

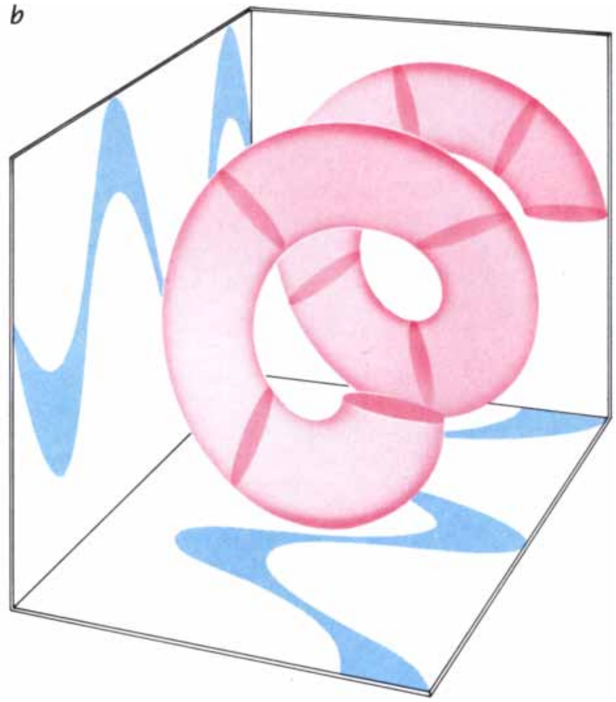
The role of the hypothetical mirror was to change the length of the cavity periodically. In our experiment we did not vary the actual length of our cavity. Instead we varied what is called the effective optical length of the cavity. We did so by means of a chamber filled with gaseous sodium, which we placed in the cavity.

The speed of light in sodium gas is lower than the speed of light in a vacuum, because the sodium atoms temporarily borrow some of the light's energy as it passes through the gas. Thus a beam of light takes longer to traverse a path through the sodium gas than to traverse an equivalent path through the vacuum, and so the path through the sodium is in a sense optically "longer." One can increase the speed of light in sodium gas, however, by "exciting" the gas with a laser of the appropriate frequency. Excited sodium atoms are less likely to borrow energy from a light wave, and so light is able to travel more

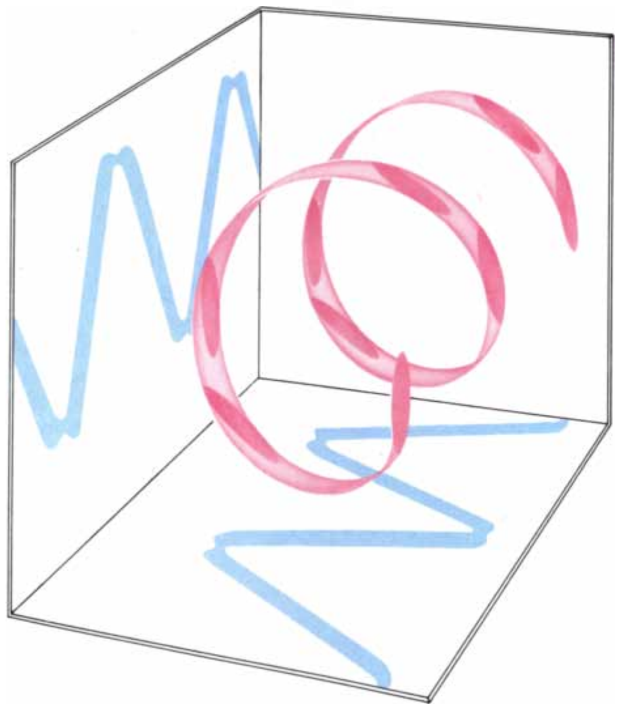
*a*




*b*



*c*



**SQUEEZED LIGHT can take several forms. In each of them the envelope of uncertainty is pinched close together in some parts of the light wave and made wider in other parts. In the squeezed vacuum (*a*) brief periods in which random fluctuations are very small (represented by a narrow envelope of uncertainty) alternate with periods in which the fluctuations can be quite large. In one form of squeezed laser light, called phase-squeezed light (*b*), the envelope of uncertainty is wide when the field strength passes through its maximum value but narrow when the field strength passes through zero. Hence the wave's amplitude is uncertain, but its phase—its timing—is relatively certain, because it is relatively certain when the wave will go through the start of its cycle. In amplitude-squeezed light (*c*), on the other hand, the amplitude is relatively certain but the phase fluctuates widely.**

quickly through excited sodium than through unexcited sodium. A path through excited sodium gas is therefore optically shorter than a path through unexcited sodium. A laser could thus make it possible for us to cause rapid changes in the optical length of the sodium chamber within our cavity, and hence in the effective optical length of the cavity.

To squeeze the vacuum fluctuations in our cavity, then, we aimed a laser beam at an oblique angle through the sodium chamber. A mirror reflected the laser beam back on itself, forming a standing wave. The standing wave periodically excited the atoms in the sodium chamber. We had arranged the apparatus in such a way that the frequency with which the standing wave excited the sodium atoms was exactly twice the frequency of the vacuum fluctuations we wanted to squeeze.

The effect was precisely the same as if we had moved one of the resonant cavity's end mirrors back and forth at the right frequency: the periodic change in the optical length of the cavity squeezed the vacuum fluctuations. The mirror at one end of the cavity was only partially silvered, so that about 2 percent of the light impinging on it emerged from the cavity as a beam of squeezed light. The mirror at the other end was as close to totally reflective as possible, to prevent unsqueezed vacuum fluctuations from entering the cavity.

The technique we used to squeeze the vacuum fluctuations is an example of a more general technique known as four-wave mixing or optical phase conjugation [see "Applications of Optical Phase Conjugation," by David M. Pepper; SCIENTIFIC AMERICAN, January, 1986]. The first investigators to suggest that four-wave mixing could squeeze light were Horace Yuen and Jeffrey H. Shapiro of the Massachusetts Institute of Technology. Although we used sodium gas as the active medium in our apparatus, four-wave mixing can be based on any of a number of materials called nonlinear optical materials: gaseous, liquid or solid materials whose optical properties vary depending on the amount of light passing through them. Examples include glass fibers, organic polymers and semiconductors such as gallium arsenide.

Since our demonstration other investigators have also squeezed the vacuum fluctuations. Ling-An Wu, H. Jeffrey Kimble, John Hall and Huifa Wu of the University of Texas at Austin have provided the most spectacular demonstration to date. They used a crystal of lithium niobate as their active medium and pumped it with a green laser at the right frequency to squeeze fluctuations in the near-infrared part of the spectrum. They have recently been able to reduce the noise in one part of the wave by as much as 70 percent. The corresponding noise increase in the other part of the wave was no larger than could be expected from the uncertainty principle. Their ability to reduce the noise even further was limited only by imperfections in the optical components of their apparatus.

In another interesting result, Robert M. Shelby, Marc D. Levenson, S. H. Perlmutter, Ralph G. DeVoe and Dan F. Walls of the IBM Almaden Research Laboratories have been able to squeeze fluctuations in an optical fiber rather than in a conventional resonant cavity. They were only able to achieve a noise reduction of about 13 percent, because of the scattering and absorption of light by the glass of which the fiber was made.

Once squeezed light has been generated, how can it be detected? How can an experimenter determine that he has indeed squeezed the vacuum fluctuations? There are two major difficulties involved in designing a detector for squeezed light. First, the detector itself has to operate without adding noise to the delicate signal. Second, a squeezed-light detector must be able to look at only a part of the light emerging from the cavity. Remember that the squeezed vacuum consists of an alternating sequence of high-amplitude fluctuations and low-amplitude fluctuations. Conventional photodetectors would register the high-amplitude fluctuations, obscuring the brief moments when the fluctuations were smaller than the vacuum noise.

Our detector is too complex to describe in detail here, but we shall give a brief account of the principles by which it operates. A key element in the detector is a beam of high-amplitude laser light called the local oscillator. The local-oscillator beam is actually a portion of the laser beam that "pumps" the sodium chamber, and so it has the same frequency as the squeezed light that emerges from the cavity. The squeezed light itself, when it emerges from the cavity, is divided into two beams by a half-silvered mirror. The local-oscillator beam is directed toward the same half-silvered mirror in such a way that it will also be split into two beams, each of which combines with one of the squeezed-light beams. The result is two separate beams, each made up of a squeezed-light beam combined with a beam from the local oscillator. Each of the combined beams then impinges on a photodetector, generating an electronic signal.

The photodetectors measure only the amplitude of any impinging light wave; that is, they tell only how high the "peak" of the wave is. The local-oscillator wave has a much higher amplitude than the squeezed-light wave, and so the shape of the combined wave—including the location of its peak—is determined largely by the local-oscillator wave. Because the peak is the only part of the combined wave the detector looks at, the only part of the squeezed-light wave that can have any effect on what the photodetector registers is the one that roughly coincides with the peak of the local-oscillator wave. The function of the local-oscillator wave, then, is to determine which part of the squeezed wave is able to register on the photodetector: in effect, to turn the photodetector on for a specific segment of the cycle.

After the combined waves have impinged on the photodetectors, the signals from the two photodetectors are electronically combined in a way that subtracts out the amplitude of the local oscillator. This procedure also cancels any fluctuations there may be in the local-oscillator beam. What is left is the tiny signal due to the part of the squeezed wave that coincided with the peak of the local-oscillator wave. By changing the relative phase of the squeezed wave and the local-oscillator wave (which can be done by causing the local-oscillator beam to follow a slightly longer route on its way to the half-silvered mirror), the experimenter can examine different parts of the squeezed wave and thereby build up a picture of the entire wave of squeezed light.

Up to now we have discussed the squeezing of the vacuum fluctuations but said little about how the fluctuations in light beams, such as laser beams, can be squeezed. The techniques involved are similar to those by which the vacuum fluctuations are squeezed. The end product is quite different, however, because of the nature of the uncertainty that accounts for the fluctuations.

The fluctuations inherent in laser light embody two kinds of uncertainty. For one, they represent uncertainty in amplitude: the amplitude of a wave of coherent light fluctuates randomly within an envelope of uncertainty. They also represent a subtler kind of uncertainty: the phase of the wave also varies within an envelope of uncertainty. Put another way, one cannot predict with absolute accuracy either exactly what the maximum strength of the electromagnetic field will be or when the field strength will pass through zero. (One can describe the phase of a wave by specifying when it passes through the "start" of its cycle, at zero.)

In squeezing laser light it is possible to reduce either one of these uncertainties, but only at the expense of increasing the other. The reason is that, just as in squeezing the vacuum fluctuations, whenever the envelope of uncertainty is pinched together in one part of the wave, it must be made wider in other parts.

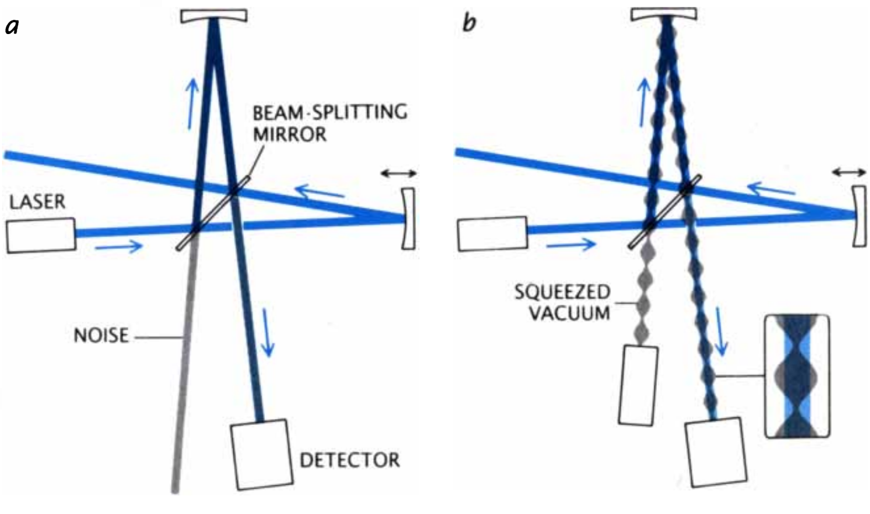
For example, one can squeeze the envelope so that it is very narrow near the peak of the wave. Then one would know the wave's amplitude with great certainty. In order to do so, however, one would have to broaden the envelope in the region where the wave passes through zero, thereby increasing the uncertainty of the wave's phase. Conversely, one could squeeze the envelope of uncertainty near where the wave passes through zero, thus decreasing its uncertainty in phase, but only at the expense of making the envelope wider near the wave's peak and so increasing uncertainty about its amplitude. Either kind of squeezed laser light can be useful, depending on whether one wants to measure an effect that depends on phase or one that depends on amplitude.

Most applications for squeezed light that have been developed to date rely on squeezed vacuum fluctuations. One example involves interferometers that are being developed to measure the gravitational waves thought to be caused by such catastrophic events as the gravitational collapse of a star that triggers the explosion of a supernova [see "Gravitational Wave Observatories," by Andrew D. Jeffries, Peter R. Saulson, Robert E. Spero and Michael E. Zucker; SCIENTIFIC AMERICAN, June, 1987]. In one such interferometer light from a laser is split into two beams. Each beam is sent down a separate arm of the interferometer, where it is reflected back toward the center of the device by a massive mirror. The beams are recombined, and changes in their interference pattern are used to gauge changes in the positions of the massive mirrors.

According to an analysis by Carlton M. Caves of the California Institute of Technology, the precision with which such devices can be operated has already reached the limits set by the presence of vacuum fluctuations in the device. By aiming a propagating beam of squeezed vacuum fluctuations into the device and thereby quieting the vacuum fluctuations during part of each wave cycle, it should be possible in principle to achieve an increase in accuracy by a factor of about a million. Practical considerations, such as the imperfection of optical components, sharply limit the gain in precision that can actually be achieved, but with present-day equipment it may still be possible to increase the accuracy of interferometers by as much as a factor of 10. Several squeezed-light interferometers are now being built to demonstrate this increased sensitivity.

Optical ring gyroscopes provide another area of application. Such devices detect and quantify rotational motion by gauging changes in the interference patterns of laser beams sent in opposite directions around a fiber-optic ring [see "Optical Gyroscopes," by Dana Z. Anderson; SCIENTIFIC AMERICAN, April, 1986]. One interesting experiment involving optical ring gyroscopes would be to measure the distortion of space due to the earth's rotation that is predicted by the general theory of relativity. The effect should be very small; the experiment would require measuring a quantity about a billionth as large as the earth's rate of rotation. Nevertheless, the experiment might be possible if the accuracy of ring gyroscopes could be increased by a factor of 10; squeezed light could provide that increase in sensitivity.

In some fields, then, physicists have already achieved such a high degree of accuracy that in order to make any further progress they must somehow evade the limits to precision that seem to be demanded by the uncertainty principle. In these fields squeezed light, in which uncertainty is redistributed, can be put to work immediately. In the coming years we can expect that experimental accuracy in an increasing number of other studies will also reach this point. One can even imagine that sometime in the future such practical disciplines as optical communications and optical computing will rely on devices so finely tuned that they too will have to turn to squeezed light for greater fidelity and precision. This new kind of light is certain to help physicists see more clearly for many years to come.




**INTERFEROMETER designed to detect gravitational waves could be made more precise with the aid of squeezed light. In current designs (*a*) a laser beam (*left*) is split by a half-silvered mirror (*center*), bounced off massive reflectors (*top and right*) and recombined at the mirror, where the separated beams interfere with each other. Part of the combined beam strikes a detector. When one of the reflectors moves under the influence of a gravitational wave, the interference pattern of the recombined beams changes, which changes the amount of light entering the detector. The interferometer's precision is limited by noise, which enters the system at the beam-splitting mirror and obscures the beam that reaches the detector. A solution (*b*) is to aim a propagating beam of squeezed vacuum fluctuations at the mirror. Then the noise in the output beam would diminish periodically; a special detector measures only quiet parts of the beam.**

# Tumor Necrosis Factor

*First identified because of its anticancer activity, the factor is now recognized to be one of a family of proteins that orchestrate the body's remarkably complex response to injury and infection*

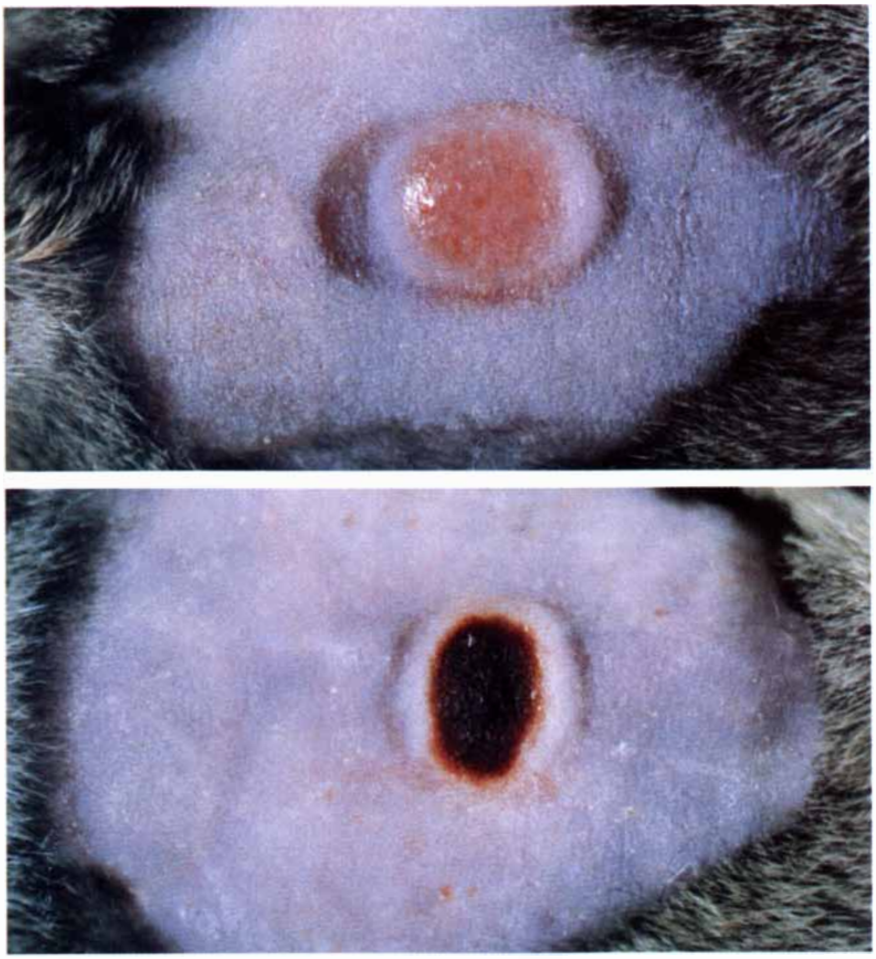
by Lloyd J. Old

Rare events, properly interpreted, have been the source of much progress in science. The spontaneous regression of cancer is a case in point. Before the turn of the century a few astute physicians observed that shrinkage of malignant tumors in patients sometimes coincided with the development of bacterial infections. They postulated that infectious agents or their products might somehow fight cancer.

This notion, and the later data that supported it, prompted decades of search for a mechanism that could lead from infection to cancer regression. Some evidence suggested that the bacteria did not kill tumors directly but instead strengthened the activity of forces in the body that are capable of restraining cancer. In pursuit of this idea, my colleagues and I at the Memorial Sloan-Kettering Cancer Center some 15 years ago discovered a small polypeptide, or protein, that is produced by the body in the course of bacterial infections and that kills tumors in mice. We and others are now in the early stages of testing the substance, which we named tumor necrosis factor, as an anticancer treatment for human beings.

Although the factor was initially discovered because of its cancer-killing activity, efforts to elucidate its functions further have revealed that it is also a central regulator of inflammation and immunity, the intertwined processes that limit and repair injuries and fight infection. It is one of a family of so-called cytokines: polypeptide mediators that transmit signals from one cell to another. Together with other substances, cytokines constitute the molecular language of inflammation and immunity and form a complex interacting and overlapping network of signals that orchestrate the body's defensive reactions. These potent and sometimes toxic proteins can elicit, enhance or inhibit one another's effects.

Like tumor necrosis factor, certain other cytokines, such as interferon (a general term for several structurally related molecules), are known to have anticancer activity and are also showing some promise as cancer therapies. A gradually deepening understanding of the individual and combined effects of the cytokines is leading to treatments for other conditions as well. For example, interferon has been shown in human trials to



HEMORRHAGIC NECROSIS of a cancerous tumor in a mouse occurs soon after the animal is injected with endotoxin, a component of gram-negative bacteria. Whereas the cancer in an untreated animal thrives (*top*), the tumor in the treated animal bleeds into itself (hence the black color) and dies (*bottom*). It is now believed that this effect of endotoxin is not a direct one. Instead the endotoxin causes certain cells in the body to secrete tumor necrosis factor (TNF), which then acts as an agent of tumor destruction.

**WILLIAM B. COLEY stimulated much of the research that led to the discovery of tumor necrosis factor. Near the turn of this century he began to treat cancer patients with a killed-bacteria vaccine and observed that their tumors sometimes regressed. Later work showed that the vaccine did not kill cancers in the test tube, a finding that provoked a search for latent anticancer forces aroused in the body by bacterial products. The search resulted in the identification of TNF in the 1970's.**

control certain viral infections; other cytokines that stimulate the production of infection-fighting blood cells are also being tested in the clinic. It seems likely that cytokines themselves, or substances that induce their release, will eventually be administered to strengthen the body's ability to fight a wide range of diseases. Conversely, inhibitory factors that restrain the polypeptide mediators in instances where they are toxic are also likely to become valuable therapeutic agents.

The story of the discovery of tumor necrosis factor properly begins with William B. Coley, a surgeon at Memorial Hospital in New York from 1892 to 1931. In the late 19th century he and a few other physicians had some success in treating cancer patients by infecting them with live bacteria. There were, however, serious problems with this approach. Infection could not be induced in some patients. Moreover, in the pre-antibiotic era the difficulty of controlling the infections that did result was cause for concern. Coley therefore developed vaccines of killed bacteria, which came to be known as Coley's toxins. These reproduced many of the symptoms of bacterial infection, such as fever and chills, but they could be administered without fear of producing an actual infection. Tumors in some patients treated with the toxins diminished or disappeared, but (as was the case with infection by live bacteria) the results were inconsistent. In the end radiation therapy and chemotherapy essentially supplanted Coley's approach.

Interest in the potential value of microbes as treatments for cancer might have died with Coley, but his daughter Helen Coley Nauts of the Cancer Research Institute dedicated herself to making physicians aware of his results. Interest also lived on in the laboratory. There investigators confirmed that a range of infectious agents and their products had anticancer effects in animals. In particular, they demonstrated that the injection of live or killed strains of gram-negative bacteria could cause hemorrhagic necrosis of mouse tumors: the tumors bled into themselves, turned black and dried up.

In work that was important to the later discovery of tumor necrosis factor, Murray J. Shear and his colleagues at the National Cancer Institute in 1943 identified and purified the active component of the gram-negative bacteria, determining that it was a complex fat-and-sugar compound now called lipopolysaccharide (LPS). Subsequent work showed that the compound is a constituent of the bacteria's outer wall and that it has both beneficial and harmful effects. In addition to causing hemorrhagic necrosis of tumors, LPS increases an animal's resistance to new bacterial infections and to lethal doses of X rays. In minute quantities the substance also causes fever, which in moderation may well help to combat infection; in greater amounts LPS can lead to shock and death. (The toxic properties of LPS account for its other name, endotoxin.)

In the late 1950's Baruj Benacerraf, then at the New York University Medical Center, and I studied another bacterial agent that later would also play a role in the discovery of tumor necrosis factor: bacillus Calmette-Guérin, or BCG. This microbe, an attenuated form of the organism causing tuberculosis, induces a self-limiting infection in mice and makes them more resistant to subsequent bacterial infection and to tumor growth.

These animal studies and others demonstrated that bacterial products could indeed lead to the destruction of cancers, but the findings by no means explained how they did so. Test-tube studies provided a hint: neither LPS nor BCG inhibited or killed tumor cells directly. Surely the microbial action was indirect and was mediated by something in the host. It was at this stage that Elizabeth A. Carswell, Robert L. Kassel, Barbara D. Williamson and I discovered tumor necrosis factor.

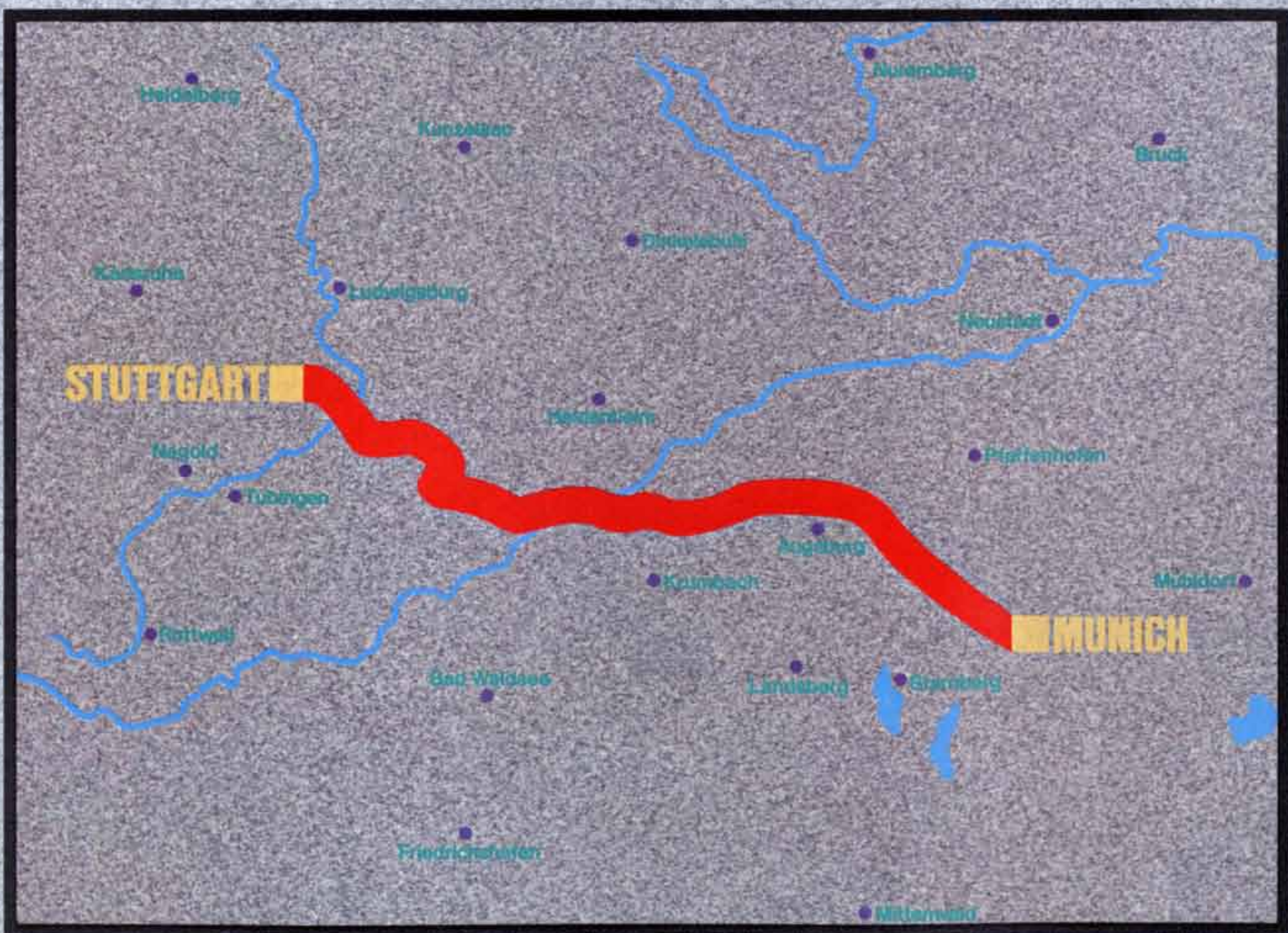
While searching for substances produced by the body that would restrain the growth of cancer and yet leave normal cells unharmed, we had noted that blood drawn from normal mice inhibited the growth of leukemic cells in other mice while exerting no apparent effect on their healthy tissues. Now we wanted to find ways to increase the level of the putative cancer-inhibiting factor. Since LPS and BCG render mice more resistant to tumor growth, we injected these substances into healthy mice, drew blood and examined the effect of the blood on tumors in other mice.

In one set of tests we injected animals with BCG and, some days later, injected them with LPS. Blood from these mice induced hemorrhagic necrosis in tumors of other animals and was highly toxic to cancer cells in the test tube. After ruling out the possibility that residual BCG and residual LPS (a more difficult possibility to exclude) in the blood might have been responsible for the tumor damage in the treated mice, we were left with one probable explanation for our findings: the animals injected with BCG and LPS had produced large amounts of an antitumor factor, and it was this factor that was responsible for the cancer killing. In view of the striking damage the substance did to mouse tumors, we called it tumor necrosis factor. The factor has subsequently been found to reproduce—and to be a mediator of—many of the effects of BCG and LPS.

What was the source of tumor necrosis factor? For several reasons we assumed it came mainly from activated macrophages, or cells that engulf and degrade bacteria, dead cells and other debris in the body. In particular, we knew that both BCG and LPS activate macrophages and increase their number in the body, and we knew that when macrophages stimulated by LPS or BCG are mixed with tumor cells, the tumor cells are killed.

Once we discovered tumor necrosis factor in mice, we quickly showed

# THE TOUGHEST STRETCH FOR MOST GERMAN TOURING SEDANS ISN'T FROM STUTTGART TO MUNICH.



# THE CURVES THAT GIVE GERMAN PERFORMANCE COUPES THE MOST TROUBLE AREN'T ON A MOUNTAIN ROAD NEAR FREIBURG.

# THEY'RE ON YOUR SPINAL COLUMN.

Negotiating the twists and turns of the Black Forest comes naturally to most German sports coupes.

Negotiating the twists and turns of the driver, alas, does not.

Happily, the turbocharged XR4Ti coupe was engineered to be the master of both terrains.

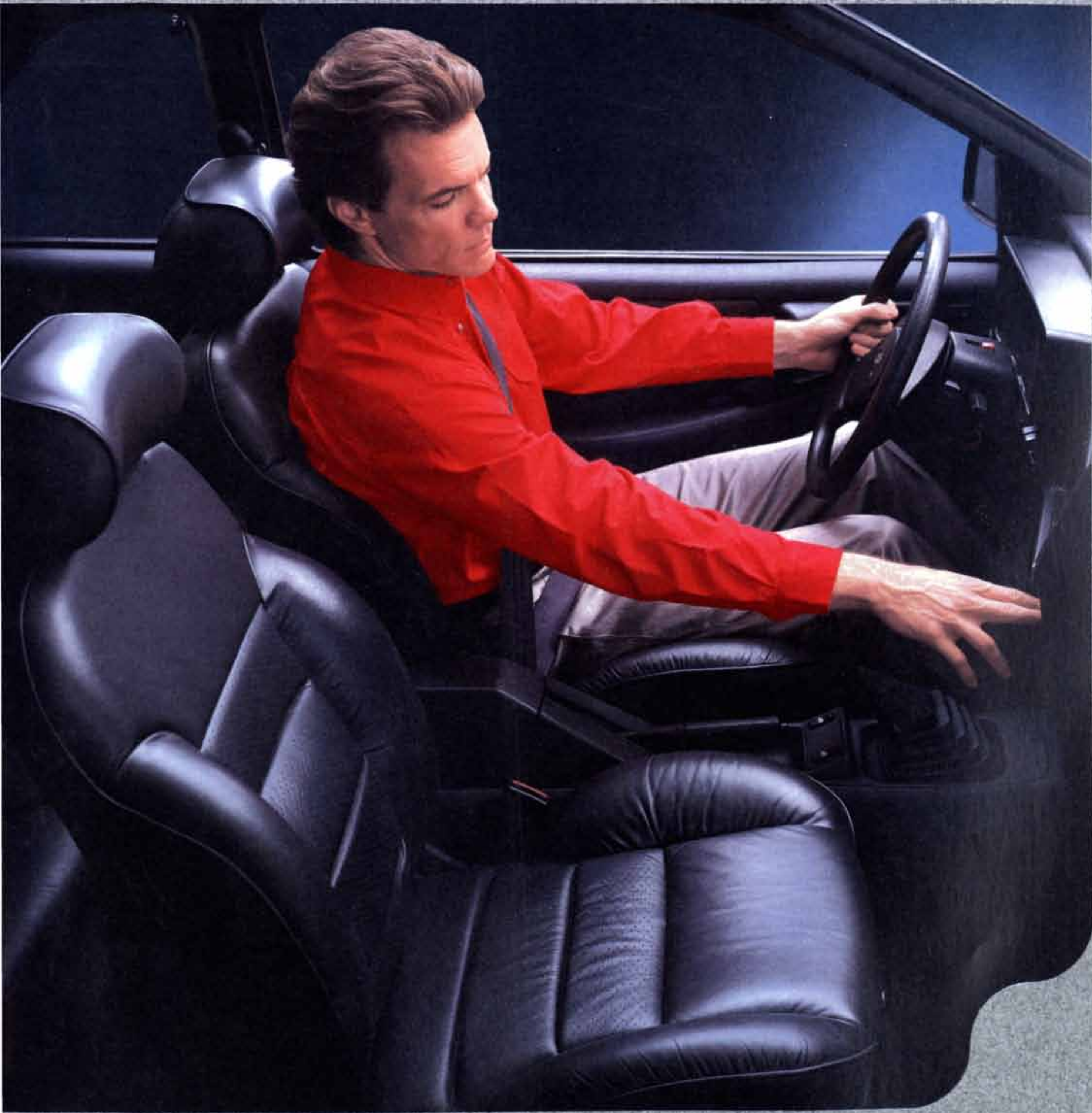
On the road, XR4Ti sprints from zero to 60 in just 7.8 seconds. Credit goes to a fuel-injected, overhead-cam 4-cylinder engine that uses a Garrett turbocharger to deliver more power than many engines with twice the displacement and double the cylinders.

Add to this XR4Ti's fully independent suspension and low-profile Pirelli P6s, and the result is a vehicle with the taut feel and spirited reflexes you expect from a true European performer.

What you may not expect is XR4Ti's devotion to driver comfort and ease of movement. With anatomically designed front seats that help keep you firmly in place. Lumbar supports that adjust to your dimensions. And a clean, no-nonsense presentation of analog gauges, controls, and power conveniences. All carefully positioned to maximize your command over XR4Ti's 175 horsepower.

The XR4Ti coupe from Germany. Performance that sends a chill up your spine. Without putting an ache in your back.

IMPORTED FROM GERMANY FOR SELECT LINCOLN-MERCURY DEALERS.



X R 4 T i

# SCORPIO AND XR4Ti. GERMAN PERFORMANCE YOU CAN BE COMFORTABLE WITH.

# GET YOUR FREE TOUR GUIDE TO SOME OF THE MOST COMFORTABLE PLACES IN GERMANY.

Read about the surprisingly comfortable accommodations offered by these remarkable German automobiles when you select either the Scorpio Tour Guide or the XR4Ti Tour Guide. Both full-color brochures include an interesting driving review of Scorpio or XR4Ti performing on the roads for which they were originally designed. And you'll receive other important information about the car of your choice.

BUCKLE UP–TOGETHER WE CAN SAVE LIVES.

LINCOLN-MERCURY DIVISION Ford

## SCORPIO

- four-door touring sedan
- seats five
- 2.9-liter V-6 engine
- four-wheel disc brakes
- Anti-lock Brake System (ABS)

## XR4Ti

- two-door sports coupe
- seats five
- 2.3-liter turbocharged four-cylinder engine

Lincoln-Mercury reserves the right to change product specifications at any time without incurring obligation.

## Your Tour Guide includes:

- a driving review of the selected vehicle
- large interior and exterior photographs of the featured car
- information detailing standard features and options
- engineering and ergonomic analysis by former World Champion racing driver Jackie Stewart

Complete the attached postcard and mail it today, or call our toll-free number: 1-800-822-9282. If the card is missing, request the Scorpio or XR4Ti Tour Guide by calling the toll-free number or writing to:

Scorpio/XR4Ti Tour Guide
P.O. Box 7041
Lincoln Park, MI 48146

SXRC07

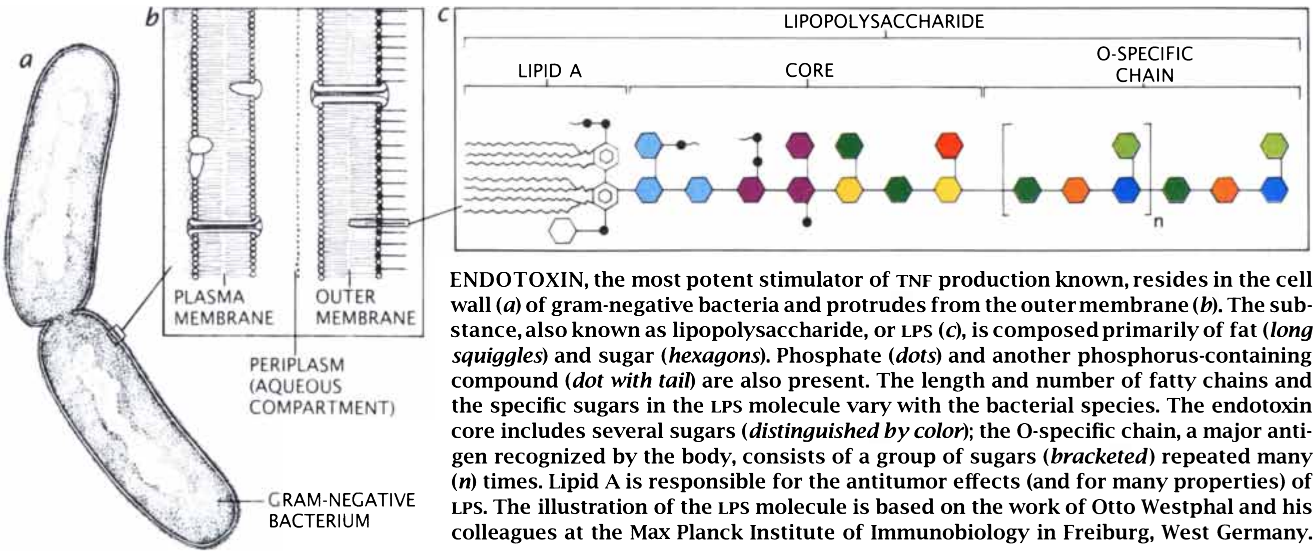
NO POSTAGE NECESSARY IF MAILED IN THE UNITED STATES

**BUSINESS REPLY MAIL**

FIRST CLASS PERMIT NO. 317 LINCOLN PARK, MI

POSTAGE WILL BE PAID BY ADDRESSEE

SCORPIO/XR4Ti TOUR GUIDE
P.O. Box 7041
Lincoln Park, MI 48146-990

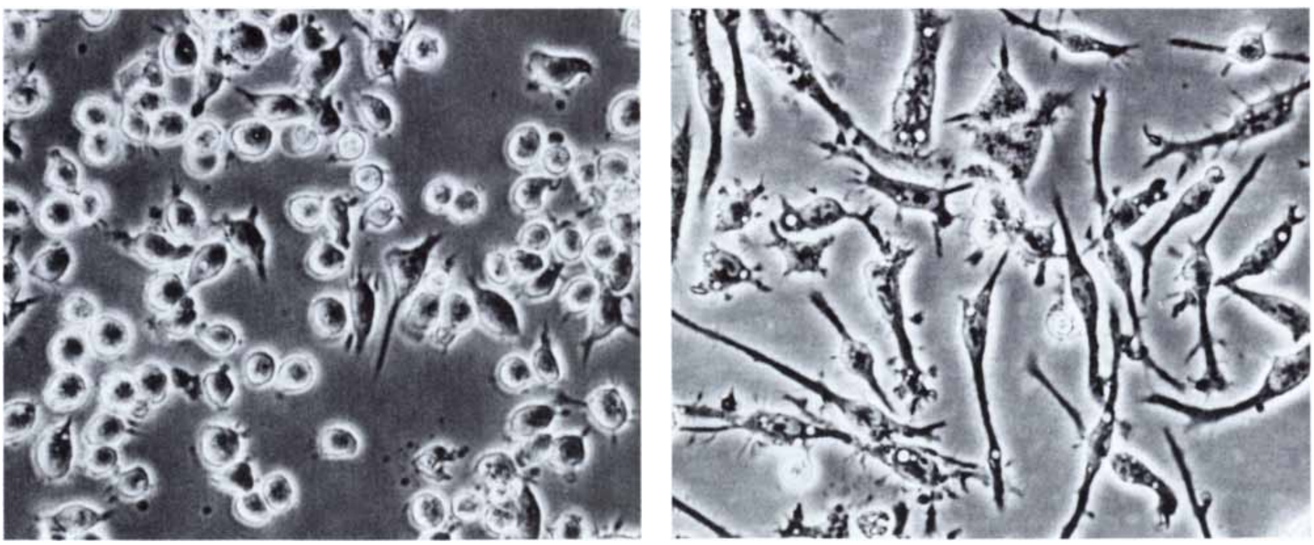



**ENDOTOXIN, the most potent stimulator of TNF production known, resides in the cell wall (*a*) of gram-negative bacteria and protrudes from the outer membrane (*b*). The substance, also known as lipopolysaccharide, or LPS (*c*), is composed primarily of fat (*long squiggles*) and sugar (*hexagons*). Phosphate (*dots*) and another phosphorus-containing compound (*dot with tail*) are also present. The length and number of fatty chains and the specific sugars in the LPS molecule vary with the bacterial species. The endotoxin core includes several sugars (*distinguished by color*); the O-specific chain, a major antigen recognized by the body, consists of a group of sugars (*bracketed*) repeated many (*n*) times. Lipid A is responsible for the antitumor effects (and for many properties) of LPS. The illustration of the LPS molecule is based on the work of Otto Westphal and his colleagues at the Max Planck Institute of Immunobiology in Freiburg, West Germany.**

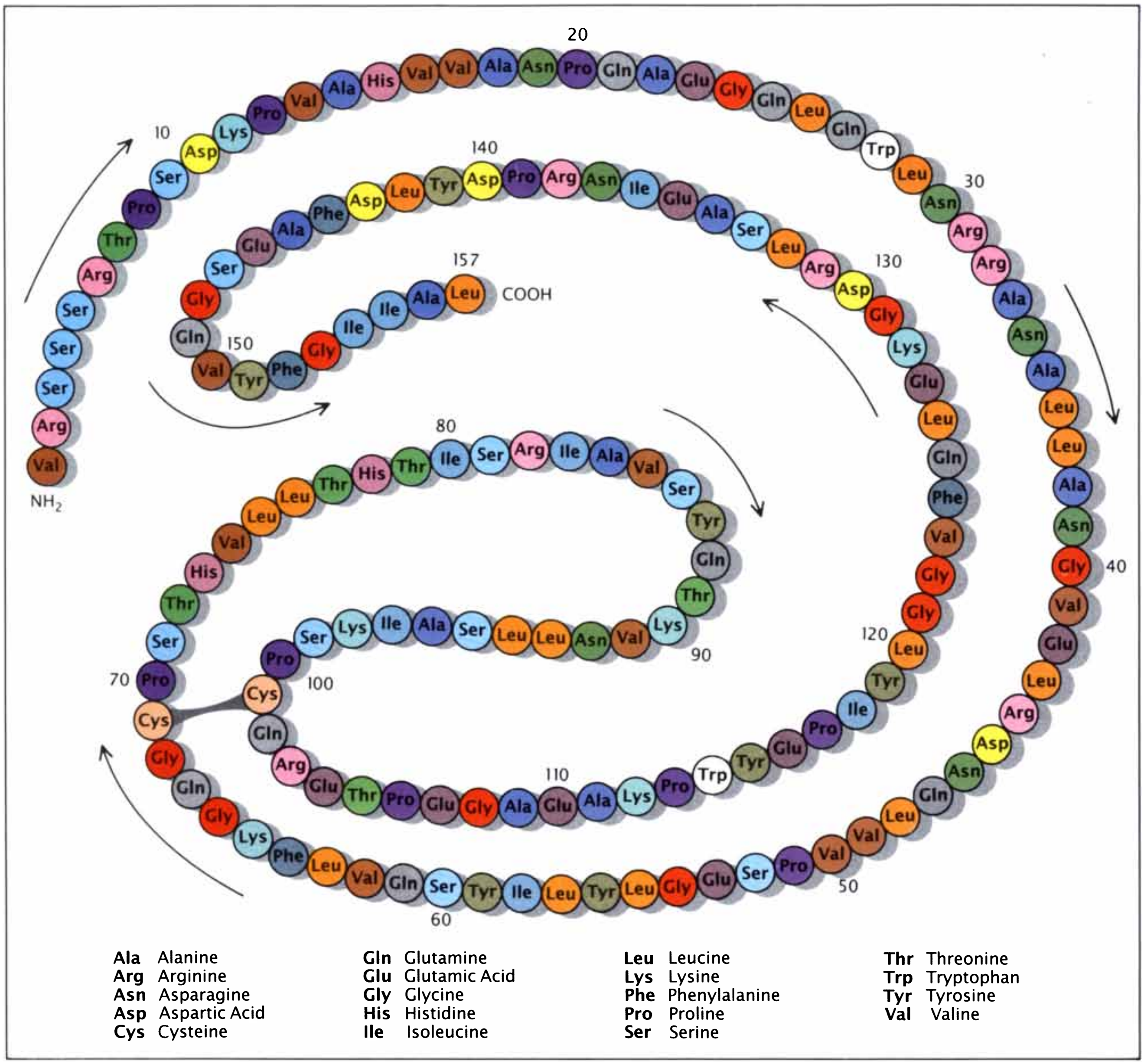
that rabbits, rats and guinea pigs also produce it. The next step was to isolate it from blood serum and define its chemical characteristics so that enough of it could be produced for extensive study. The process of purification was arduous because, although the factor can be detected in blood drawn from animals treated with BCG and LPS, it is usually present in extremely small amounts. Moreover, at each stage of purification we had to satisfy ourselves that the factor was still present. We did so by assaying each isolate for specified activities. For example, the isolate had to cause hemorrhagic necrosis of tumors in animals; it had to kill certain types of cancer cells in the test tube, and the effect had to be produced by progressively smaller quantities at successive stages of purification.

Saul Green, a colleague of ours at Sloan-Kettering, began the purification process and isolated enough of the mouse factor to enable us to demonstrate that the effect on animal tumors and the effect on malignant cells in the test tube were the work of a single substance rather than of two associated substances, one active in the body and the other active in the test tube. We also demonstrated that the substance is a protein. Katsuyuki Haranaka, first as a member of our group and then with Nobuko Satomi of the University of Tokyo, continued the effort and eventually succeeded in obtaining a single polypeptide—pure tumor necrosis factor—from the blood of both mice and rabbits.

Meanwhile Danielle N. Männel and Stephan E. Mergenhagen of the National Institute of Dental Research and N. Matthews of the University of Wales proved that macrophages do indeed produce tumor necrosis factor. In search of a human cell type that could be coaxed to produce high quantities of the factor in culture, Williamson, Carswell and I, along with Berish Y. Rubin of the New York Blood Center, then screened many types of human cells and found that defensive cells other than macrophages also secreted a factor that exhibited the known activities of tumor necrosis factor.



**MACROPHAGES (*left*) spread and enlarge (*right*) when they are exposed to LPS, indicating that they have become activated. The cells, which play a central role in inflammation and immunity and can kill cancer cells, secrete dozens of factors, including tumor necrosis factor, that carry out many of their activities. For example, Hans Schreiber and his colleagues at the University of Chicago and Genentech, Inc., have shown that TNF is an important mediator of cancer-cell killing by activated macrophages.**




**MOLECULE of human tumor necrosis factor is a protein consisting of 157 amino acids. The amino acid sequence was determined in 1984, the year the gene encoding TNF was first cloned by several groups associated with biotechnology companies.**

Our work with the mouse and human factors soon enabled us to uncover two features of tumor necrosis factor that may have important implications for treatment. We found that tumor necrosis factor and interferon—which was also known to have antitumor effects—act synergistically. Exposure of cancerous cells to both factors together results in a far greater cell kill than would be expected if the individual effects of the two substances were simply added together. We also found that human tumor necrosis factor, like the mouse variety, lacks species specificity: it kills cancer cells from mice as well as from human beings. Interferon, in contrast, is species-specific.

The decade-long effort to purify tumor necrosis factor finally culminated in 1984 with the cloning of the gene, the identification of the protein's amino acid sequence and the production of large quantities of the factor by several groups associated with biotechnology companies, including David V. Goeddel and his colleagues at Genentech, Inc., and Walter Fiers and his co-workers at the State University of Ghent and at Biogen SA. Since then there has been an explosion of information about the factor's activities.

It is now clear that tumor necrosis factor elicits a remarkable range of reactions in the body and that it will be some time before a complete picture of its normal functions is constructed. Nevertheless, research has shown that tumor necrosis factor is crucial to inflammation and immunity, is secreted early in these processes and stimulates defensive cells to produce many other polypeptides.

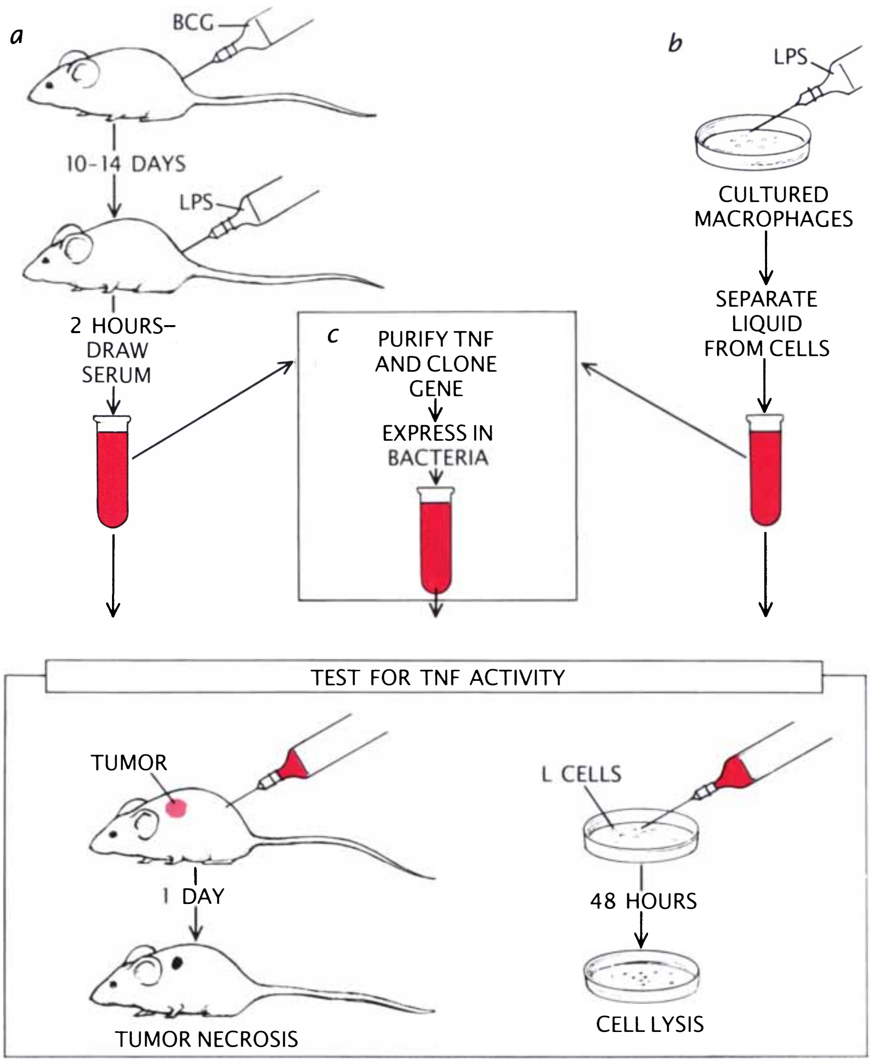
In order to understand the role of tumor necrosis factor in inflammation and immunity, one must have a sense of the events set in motion by an injury, whether it results from mechanical trauma, from chemicals or from infection. What follows is a highly simplified outline and necessarily omits a host of important actors and events. I should point out that, in theory, inflammation can be viewed as the aspect of the defensive function that confines and repairs injury, and immunity as the aspect that specifically neutralizes invading microbes and confers immunity to future infection by the same invader. In reality the two are inseparable: most of the cells and molecules that defend the body are involved in both inflammation and immunity.

In the first phase of the response to injury white blood cells known as polymorphonuclear leukocytes, or granulocytes, leave the flowing blood and adhere to endothelial cells, the cells that line blood vessels. The endothelial cells spread apart somewhat and enable the granulocytes to pass into the injured tissue, where they ingest and destroy any microbes that may have entered the wound. (When large numbers of the white cells accumulate and die at the inflamed site, they form pus.)

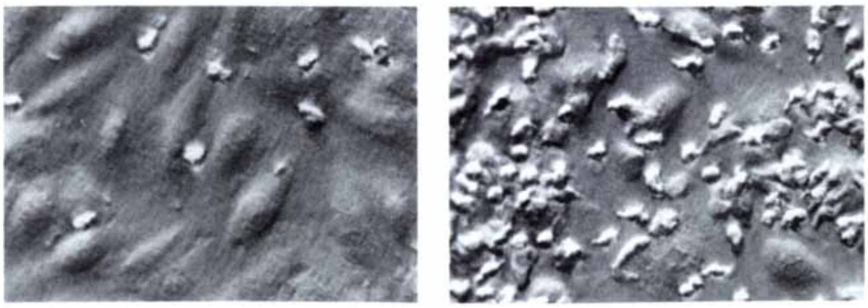
Macrophages soon join the battle in force and replace the granulocytes as the predominant cell type at the injury. The macrophages swallow and destroy bacteria, particularly those coated with antibody, and damaged cells. At the same time other blood cells also become more active: *T* lymphocytes (white cells that mature in the thymus) proliferate and arouse other defensive cells, including *B* lymphocytes, which divide, differentiate and secrete antibodies in quantity. As the infection is controlled, connective-tissue cells called fibroblasts and other cells begin to repair damaged tissue.

Tumor necrosis factor starts to exert its many effects once macrophages, which also release the cytokines known as interleukin-1 and colony-stimulating factors, take the offensive. Interleukin-1 is unlike tumor necrosis factor in structure but has many of the same activities. Colony-stimulating factors cause the bone marrow to produce blood cells, such as polymorphonuclear leukocytes and monocytes (which are precursors of macrophages).

Michael P. Bevilacqua, Jordan S. Pober and Michael A. Gimbrone, Jr., of the Harvard Medical School have found that both tumor necrosis factor and interleukin-1 stimulate endothelial cells to synthesize molecules that increase the adhesion of granulocytes to the surface of blood vessels. Tumor necrosis factor also has a direct effect on the granulocytes, increasing their attachment to the vessel wall and their migration into damaged tissue. In addition, Michael A. Palladino, Jr., of Genentech and Carl




**TNF-CONTAINING MATERIAL can be harvested from the blood of mice that have been injected first with the bacterium bacillus Calmette-Guérin (BCG) and then with LPS (*a*) or from the fluid produced by cultured macrophages mixed with LPS (*b*). Purification of such materials made it possible to clone the TNF gene and insert it into bacteria that produce the factor in quantity (*c*). Showing that TNF is present in a preparation (*bottom*) requires demonstrating that the collected material can cause hemorrhagic necrosis of tumors in animals and can kill L cells (a strain of malignant cells) grown in culture.**



**ONE ACTIVITY of tumor necrosis factor can be demonstrated readily in the test tube. Endothelial cells (*large bumps*), which line blood vessels, do not ordinarily interact with polymorphonuclear leukocytes (*bright bodies*)—a type of white blood cell that digests bacteria—when the cells are mixed (*left*). Exposure of the endothelial cells to TNF or interleukin-1, another macrophage product, causes the cells to synthesize molecules that promote the adherence of leukocytes (*right*). A similar phenomenon presumably occurs in the body: in response to an injury, TNF and other factors promote the attachment of leukocytes to the blood-vessel wall, aiding their migration into injured tissue. Michael A. Gimbrone, Jr., of the Harvard Medical School provided the micrographs.**

F. Nathan of the Cornell University Medical College have demonstrated that the factor is one of the most potent signals for stimulating granulocytes to produce toxic oxygen products that destroy bacteria.

Tumor necrosis factor, along with interleukin-1, also plays a part in the activation of *T* lymphocytes. These cells in turn produce interleukin-2 (a growth factor for *T* and *B* cells), gamma interferon (which further activates macrophages), several other factors that trigger the multiplication of *B* cells, and colony-stimulating factors [*see illustration on opposite page*]. Activated *T* cells produce a substance known as lymphotoxin as well; this cytokine is quite similar to tumor necrosis factor in both structure and function. It was discovered by Gale A. Granger and T. W. Williams of the University of California at Irvine and Nancy H. Ruddle and Byron H. Waksman of the Yale University School of Medicine.

Like the *T* lymphocytes, *B* lymphocytes are affected by tumor necrosis factor and interleukin-1. The factors regulate the production of antibodies by *B* cells; they also cause many cell types, including endothelial cells and fibroblasts, to secrete still more of the colony-stimulating factors.

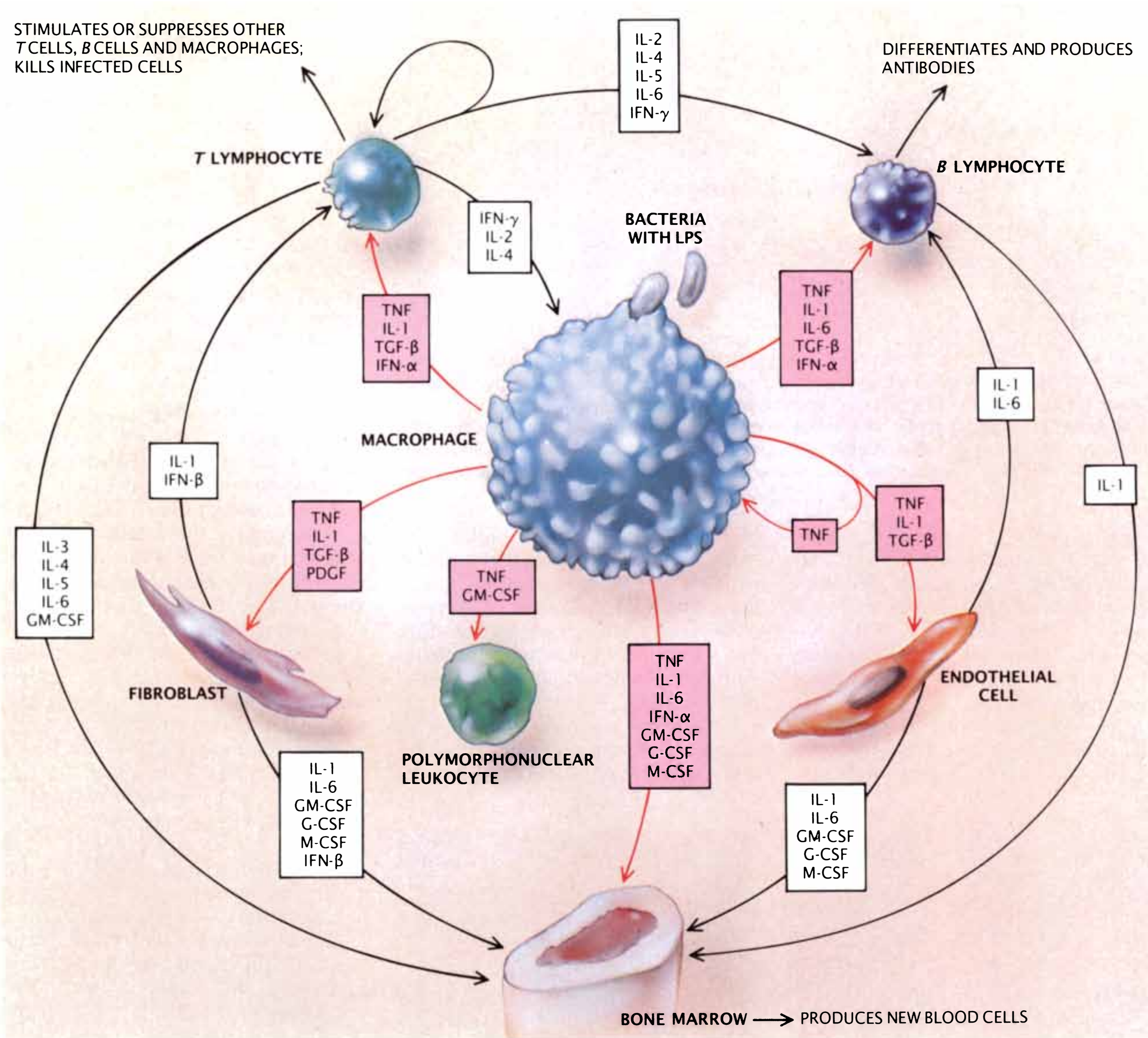
Certain systemic effects of the inflammatory and immune responses can also be traced in part to tumor necrosis factor and other cytokines. For instance, tumor necrosis factor and interleukin-1 act on the temperature centers of the brain to produce fever. Another effect—the increased circulation in the blood of so-called acute-phase proteins—is an indirect result of the release of tumor necrosis factor. The acute-phase proteins, produced by the liver, are thought to increase the efficiency of inflammatory reactions, and they appear to be regulated by a recently described factor called interleukin-6. This factor is synthesized by a variety of cell types in response to tumor necrosis factor, interleukin-1 or interferon.

| FAMILY | MEMBERS | OTHER NAMES |
|---|---|---|
| Interferons (IFN) | IFN-α | Leukocyte Interferon |
| | IFN-β | Fibroblast Interferon |
| | IFN-γ | Immune Interferon |
| Tumor Necrosis Factors (TNF) | TNF | TNF-α, Cachectin |
| | Lymphotoxin | TNF-β |
| Interleukins (IL) | IL-1 α, IL-1 β | Endogenous Pyrogen, Lymphocyte-Activating Factor, Leukocyte Endogenous Mediator, Hemopoietin 1 |
| | IL-2 | *T*-Cell Growth Factor |
| | IL-3 | Multipotential CSF, Mast Cell Growth Factor |
| | IL-4 | *B*-Cell Stimulatory Factor 1 (BSF-1) |
| | IL-5 | *T*-Cell Replacing Factor (TRF), Eosinophil Differentiation Factor, *B*-Cell Growth Factor-II (BCGF-II) |
| | IL-6 | *B*-Cell Stimulatory Factor 2 (BSF-2), Interferon-$\beta_2$ Hepatocyte-Stimulating Factor (HSF) |
| Colony-Stimulating Factors (CSF) | Granulocyte Macrophage-CSF (GM-CSF) | CSF-2 |
| | Granulocyte-CSF (G-CSF) | Pluripoietin |
| | Macrophage-CSF (M-CSF) | CSF-1 |
| | Erythropoietin | |
| Other Growth and Regulatory Factors (GF) | Epidermal Growth Factor (EGF) | |
| | Fibroblast Growth Factor (Acidic- and Basic-FGF) | |
| | Insulin-like Growth Factor-1 (IGF-1) | Somatomedin C |
| | Insulin-like Growth Factor-2 (IGF-2) | Somatomedin A |
| | Nerve Growth Factor (NGF) | |
| | Platelet-Derived Growth Factor (PDGF) | |
| | Transforming Growth Factor-α (TGF-α) | |
| | Transforming Growth Factor-β (TGF-β) | |

**PROTEIN FACTORS involved in inflammation, immunity and the growth and inhibition of cells are often grouped into families. There is considerable overlap in the activities of the factors, however: structurally unrelated molecules, such as tumor necrosis factor and interleukin-1, elicit many of the same effects. Adding to the complexity, individual factors can cause cells to secrete other factors and can potentiate or antagonize one another's effects. Lymphotoxin, which is in the TNF family, is produced by *T* lymphocytes: white blood cells critical to the killing of infected cells. It is grouped with TNF because it is structurally and functionally related to the TNF produced by macrophages.**

Inflammation and immunity, like all other normal reactions of the body, are meant to preserve or restore health. They can nonetheless cause a range of uncomfortable symptoms. The typical effects were identified by the Roman physician Aulus Cornelius Celsus: *rubor* (redness), *tumor* (swelling), *calor* (heat) and *dolor* (pain). Symptoms can vary widely, however, and include such discomforts as the stuffy, runny nose of a cold or an allergy. It generally is not the microbes invading the body that make one feel ill but the body's response to those invaders. If the defensive response is too vigorous or goes on too long—as it does when an infection is chronic—it can actually do permanent harm. Such is the case with rheumatoid arthritis, in which chronic inflammation can produce debilitating effects. A vigorous inflammatory response can also cause shock and death. Recent work implicates tumor necrosis factor as a cause of many negative effects associated with the defensive response.

For instance, studies have suggested that tumor necrosis factor might contribute to cachexia: a condition marked by a profound weight loss stemming from the disappearance of fat deposits and the shrinkage of muscles. This disorder sometimes



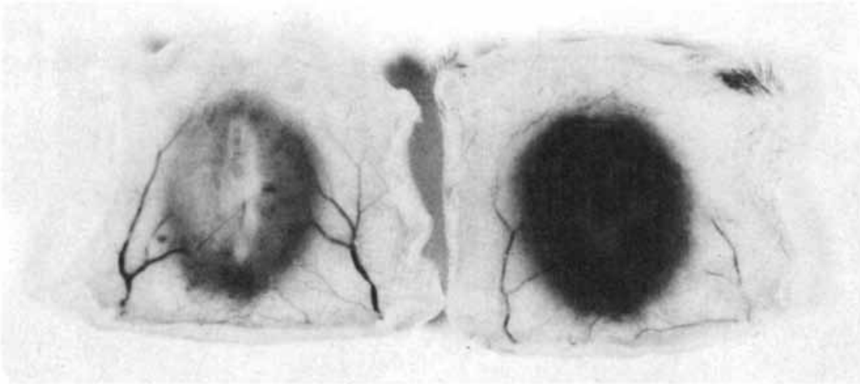

**TUMOR NECROSIS FACTOR (*colored arrows*) secreted by an LPS-stimulated macrophage (*center*) activates an array of immune defenses. Together with interleukin-1 (IL-1) and many other macrophage products, TNF influences such cells as polymorphonuclear leukocytes (PMN's), *T* lymphocytes, antibody-producing *B* lymphocytes, endothelial cells, fibroblasts (connective-tissue cells) and the blood-producing cells of the bone marrow. These cell types all have specific roles in fighting infection and in limiting and repairing injury, and they carry out their tasks in part by secreting factors, including interleukins, interferons (IFN-α, -β and -γ), growth factors (GF) and colony-stimulating factors (CSF). This drawing, as complex as it is, omits much.**

develops when the body fails to overcome an infection; it is also observed in some patients with cancer.

Several years ago Bruce Beutler and Anthony Cerami of Rockefeller University found that a factor produced by activated macrophages inhibits an enzyme, lipoprotein lipase, that is crucial for the normal storage of fat and is suppressed in cachectic individuals. When the factor, which the workers called cachectin, was purified and its amino acid sequence determined, it was found to be identical with tumor necrosis factor. Subsequent work showed that interleukin-1 and interferon also inhibit lipoprotein lipase activity in the test tube. The role of these three cytokines and other factors in causing cachexia in patients with chronic infection or cancer needs further definition.

Investigations into the ability of tumor necrosis factor to mediate the effects of LPS provided other indications of the factor's ability to do ill as well as good. Injection of enough tumor necrosis factor in mice and other animals has been found to cause tissue injury, shock and death, just as LPS does. Yet at lower doses the factor exhibits the protective activities associated with low doses of LPS: it protects mice against bacterial infections, lethal doses of X rays and tumor growth.

Studies of parasitic diseases in animals have also been revealing. On the one hand, mice injected with tumor necrosis factor are more resistant to certain forms of malaria, and the factor has been shown to stimulate macrophages and other blood cells to kill the parasites that cause Chagas' disease and schistosomiasis. On the other hand, Pierre Vassalli and his colleagues at the University of Geneva have recently shown that

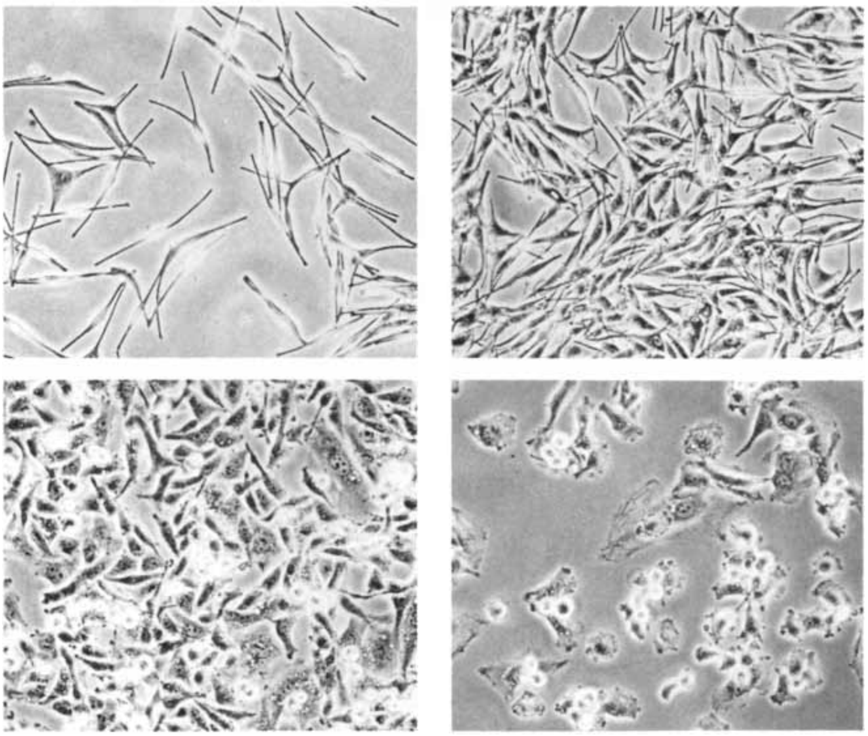


TINY BLOOD VESSELS normally feed a tumor (*left*), but the vessels inside the tumor are destroyed (*right*) and bleed within hours after a mouse is injected with tumor necrosis factor. Because such selective vascular damage deprives a tumor of oxygen and nutrients, most of its cells soon die. Much evidence suggests that vascular destruction is the major way TNF leads to hemorrhagic necrosis of tumors. Edward A. Havell and Robert J. North of the Trudeau Institute, Inc., in Saranac Lake, N.Y., provided the photograph.

the substance has a role in the death of mice with malaria that has invaded the brain: it appears to mediate a lethal inflammation of the brain tissue. The mice do not die if they receive an antibody that neutralizes the brain-damaging effect of tumor necrosis factor.

Another indication of the potential toxicity of tumor necrosis factor comes from human beings. A. Waage of the University of Trondheim in Norway has discovered that patients with severe meningococcal infections who have relatively high levels of the factor in their blood are more likely to die from shock than patients who have no detectable levels of the polypeptide.

With tumor necrosis factor as with many other factors produced by the body, there seems to be a fine line between benefit and harm: an agent that is helpful in the local control of injury and infection may be toxic when it is released in large amounts or in the wrong place. The growing awareness of the harmful potential of tumor necrosis factor has given rise to interest in the design of drugs that will block its action when its detrimental effects outweigh its protective ones.



MELANOCYTES, skin cells that produce the pigment melanin (*top left*), proliferate markedly when exposed to tumor necrosis factor in the test tube (*top right*). In contrast, cells from melanoma, a skin cancer (*bottom left*), stop growing after exposure to TNF (*bottom right*), as Yuko Arita and Magdalena Eisinger of the Memorial Sloan-Kettering Cancer Center have shown. Such findings indicate that in addition to damaging tumor blood vessels, TNF might exert a direct, selective toxicity on cancer cells in the body.

In spite of the rapid accumulation of new information about tumor necrosis factor, my colleagues and I still do not fully understand how it causes the effects for which it was named: hemorrhage and necrosis of tumors. One partial explanation has to do with the factor's directive role in inflammation and immunity. Even if tumor necrosis factor were not directly toxic to tumors, the *T* cells and other cells it stimulates and the other cytokines those cells secrete when they are activated do sometimes join forces to destroy tumors.

Beyond this generalized antitumor role, however, there is clear evidence that tumor necrosis factor has a more immediate effect on cancer. In animal studies it has been shown to damage the blood vessels that nourish tumors. This damage reduces the flow of blood and oxygen to the tumor cells, which then starve and die.

In contrast, it appears that in noncancerous tissue the cytokine actually plays an important role in normal angiogenesis (the formation of new blood vessels during development or growth, or the replacement of injured vessels in existing tissue). Marijke Fràter-Schröder of the University of Zurich and S. Joseph Leibovich of Northwestern University and their co-workers have demonstrated that injection of tumor necrosis factor into an animal induces normal endothelial cells to grow and join together to form new vessels. Tumor necrosis factor is not alone in exhibiting angiogenic activity; four other cytokines do the same. Nature, it seems, has chosen not to construct a different mediator or set of mediators for every process in the body but rather to call on the same molecules repeatedly for quite different purposes.

Why would blood vessels in a tumor respond to tumor necrosis factor so differently from those in normal organs? Investigators need to learn more about the regulation of angiogenesis before the question can be answered. One idea holds that tumor cells themselves, or defensive cells attracted to the tumor, may produce some factor that renders the tumor vessels susceptible to damage induced by tumor necrosis factor.

In addition to damaging a tumor's blood vessels, tumor necrosis factor in the body might also kill cancer cells directly, as it has long been known to do in the test tube. In one survey of cells from more than 60 different human cancers, my colleagues and I demonstrated three distinct responses to the factor: cell death, inhibition of cell growth and no effect. Approximately one-third of the tested cells fell into each category. Cells from breast cancer appeared to be most sensitive to the killing action of tumor necrosis factor, whereas cells from melanoma (a type of skin cancer) responded to the substance by growing at a reduced rate. We also tested normal cells and found no inhibitory effect. In fact, Jan T. Vilček of the New York University School of Medicine has demonstrated that tumor necrosis factor actually stimulates cultured fibroblasts to grow more vigorously.

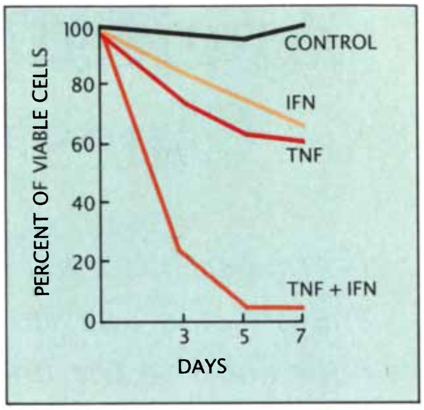
Because the first step in the action of many substances that influence cell activity is binding to specific receptors on the surface of the cell, Rubin, Vilček and others have independently looked to see whether the differential responses of cancer cells to tumor necrosis factor could be correlated with the presence or absence of receptors for the polypeptide. They did find factor-specific receptors but detected no relation between the number of receptors and the cells' responsiveness to tumor necrosis factor. Characteristics other than the number of receptors, then, must determine what the response to bound tumor necrosis factor will be.

One characteristic that might make a cell particularly sensitive to the toxicity of tumor necrosis factor is a reduced ability to repair damage caused by the factor. When George E. Gifford of the University of Florida College of Medicine exposed cells to agents that inhibited their essential functions, such as the synthesis of RNA, he greatly enhanced the sensitivity of the cells to the toxic effects of tumor necrosis factor; even cells that would ordinarily not be affected could be killed. These results imply that cells do have a mechanism for repairing injury caused by tumor necrosis factor and that when the mechanism is absent (as it may be in certain cancer cells) or compromised (as it apparently was by Gifford's agents), cells exposed to tumor necrosis factor die. What actually causes the cell death is not known, but there is evidence that the cytokine activates intracellular enzymes that liberate highly reactive molecules. It may be the action of these molecules that injures and finally kills cancer cells.

A major goal of the extensive current research into the activities of tumor necrosis factor is the development of treatments for cancer. Preliminary clinical trials of the factor are now under way at many medical centers around the world. When the polypeptide has been administered so that it circulates throughout the body, only a few patients have had tumor inhibition at the doses studied to date. In Japan and Germany the factor has been injected directly into cancers. A number of the patients in those trials have had tumor regression, in some instances complete. There are side effects. Like other cytokines under study as cancer treatments, tumor necrosis factor can cause fever, chills, lethargy and a drop in blood pressure. Many patients feel as if they have come down with the flu.

Although clinical trials of tumor necrosis factor and other individual cytokines are aimed at revealing what each substance can do by itself, it seems likely that combinations of cytokines, or of cytokines with other substances, will be needed in order to produce the most successful cancer treatments. Indeed, on the basis of the fact that tumor necrosis factor and interferon act synergistically, Janice L. Gabrilove and Herbert F. Oettgen of Sloan-Kettering are now testing a therapy that couples these two agents.

One of the preliminary findings is that certain doses of the combined agents cause patients to have sharp pain at the tumor site, where there had been no pain before. We do not yet know the cause, but such acute pain is suggestive of damage to blood vessels. We are now exploring the possibility that the treatment is damaging the blood vessels that feed the tumors. A number of other trials of combination therapies are planned as well: tumor necrosis factor will be paired with interleukin-2 (which has promise as an anticancer agent), chemotherapeutic agents, monoclonal antibodies or radiation.




**INTERFERON AND TNF have a synergistic effect on human breast-cancer cells. When the factors are mixed together, they destroy more malignant cells (*orange*) than would be killed if the independent effects of the interferon (*yellow*) and the TNF (*red*) were added together. Such synergy indicates that cancer treatments combining the two substances will probably prove more effective than therapies consisting of either substance alone.**

Studies have now confirmed that William Coley's decision to treat cancer patients with microbes made good sense; when his toxins were successful, they almost certainly induced macrophages in the human body to produce tumor necrosis factor and other factors that in combination exerted anticancer effects. Why then has Coley's approach been ignored by most clinicians for so many years?

For one thing, an understanding of how the toxins worked required an understanding of inflammation and immunity; it has only been in recent years that critical molecules involved in these processes have been identified and isolated. In addition, the response to Coley's toxins varied widely from patient to patient, with many people not benefiting at all.

From our current perspective it seems likely that patients who did not respond may have been unable to produce tumor necrosis factor or the other cytokines that activate inflammation and immunity and are destructive to cancer. Now that tumor necrosis factor and many of the other factors elicited in the course of infection have been identified, it may be possible to employ them as new, more effective versions of Coley's toxins.

We are just beginning. If we are fortunate, the new treatments will consistently arouse the body's natural anticancer forces and produce the tumor regressions that have fueled the imagination of generations of cancer researchers. The forces exist; the task ahead is to find ways to unleash them.

# Ancient Magnetic Reversals: Clues to the Geodynamo

*Is the earth headed for a reversal of its magnetic field? No one can answer this question yet, but rocks magnetized by ancient fields offer clues to the underlying reversal mechanism in the earth's core*

by Kenneth A. Hoffman

For well over a century geophysicists have observed a steady and significant weakening in the strength of the earth's magnetic field. Indeed, if this trend were to continue at the present rate, the field would vanish altogether in a mere 1,500 years. Most investigators are inclined to think that the decay is merely an aspect of the restlessness inherent in the field and that the field will recover its strength. Yet one cannot dismiss out of hand the possibility that the weakening portends a phenomenon that has recurred throughout geologic time: the reversal of the geomagnetic field.

Which of these two scenarios is correct? The answer lies concealed 3,000 kilometers below the earth's surface within the outer core, a slowly churning mass of molten metal sandwiched between the mantle of the earth and the solid inner core. It is now generally accepted that the earth's magnetic field is generated by the motion of free electrons in the convecting outer core. This theory supposes the core behaves like a self-sustaining dynamo, a device that converts mechanical energy into magnetic energy. In the geodynamo the earth's rotation, along with gravitational and thermodynamic effects in and around the core, drives the fluid motions that produce the magnetic field [see "The Source of the Earth's Magnetic Field," by Charles R. Carrigan and David Gubbins; SCIENTIFIC AMERICAN, February, 1979].

Although the basic principles of dynamo action are well established, geophysicists do not yet understand the thermodynamics, fluid mechanics and electrical properties of the earth's interior well enough to construct a universally accepted model of the geodynamo. Yet its workings can be glimpsed indirectly by observing the present-day field. These measurements yield many details of the short-term behavior of the field, such as its shape and "secular variation," or ordinary fluctuation. To study the activity of the dynamo over aeons one must turn to the paleomagnetic record—the ancient magnetism frozen into rocks from the time of their formation.

Indeed, paleomagnetic evidence led to the first proposal that the earth's field has reversed itself, put forward in 1906 by the French physicist Bernard Brunhes. Brunhes was intrigued by the discovery of rocks that were magnetically oriented in the direction opposite to the earth's field. His startling suggestion was furiously debated for more than five decades. It was not until the early 1960's, at about the time J. S. B. Van Zijl and his colleagues published the first detailed study of a paleomagnetically recorded field reversal in lavas from South Africa, that the idea was accepted by the scientific community at large. Today it is a fundamental tenet of geophysics that the earth's magnetic field can exist in either of two polarity states: a "normal" state, in which north-seeking compass needles point to the geographic north, and a "reverse" state, in which they point to the geographic south.

In the 1960's studies of radiometrically dated lavas yielded a consistent log of past polarity changes, including no fewer than nine major reversals in the past 3.6 million years, the most recent of which occurred 730,000 years ago [see "Reversals of the Earth's Magnetic Field," by Allan Cox, G. Brent Dalrymple and Richard R. Doell; SCIENTIFIC AMERICAN, February, 1967]. The time scale of polarity transitions has since been extended back nearly 170 million years.

Paleomagnetic records show that the geomagnetic field does not reverse instantaneously from one polarity state to the other. Rather, the process involves a transition period that typically spans a few thousand years. Hence for perhaps 98 percent of the time the field is stable and its shape is well understood. But for the remaining 2 percent of the time the field is unstable and its shape is not obvious. The foremost task for geophysicists in my field has been to chronicle the behavior of the reversing field—its shifting shape and fluctuating intensities—based on the sometimes faint and complex record of past events, imprinted in stone. The findings provide an invaluable probe into the hidden mechanisms of the geodynamo.

### Clues to the Field's Geometry

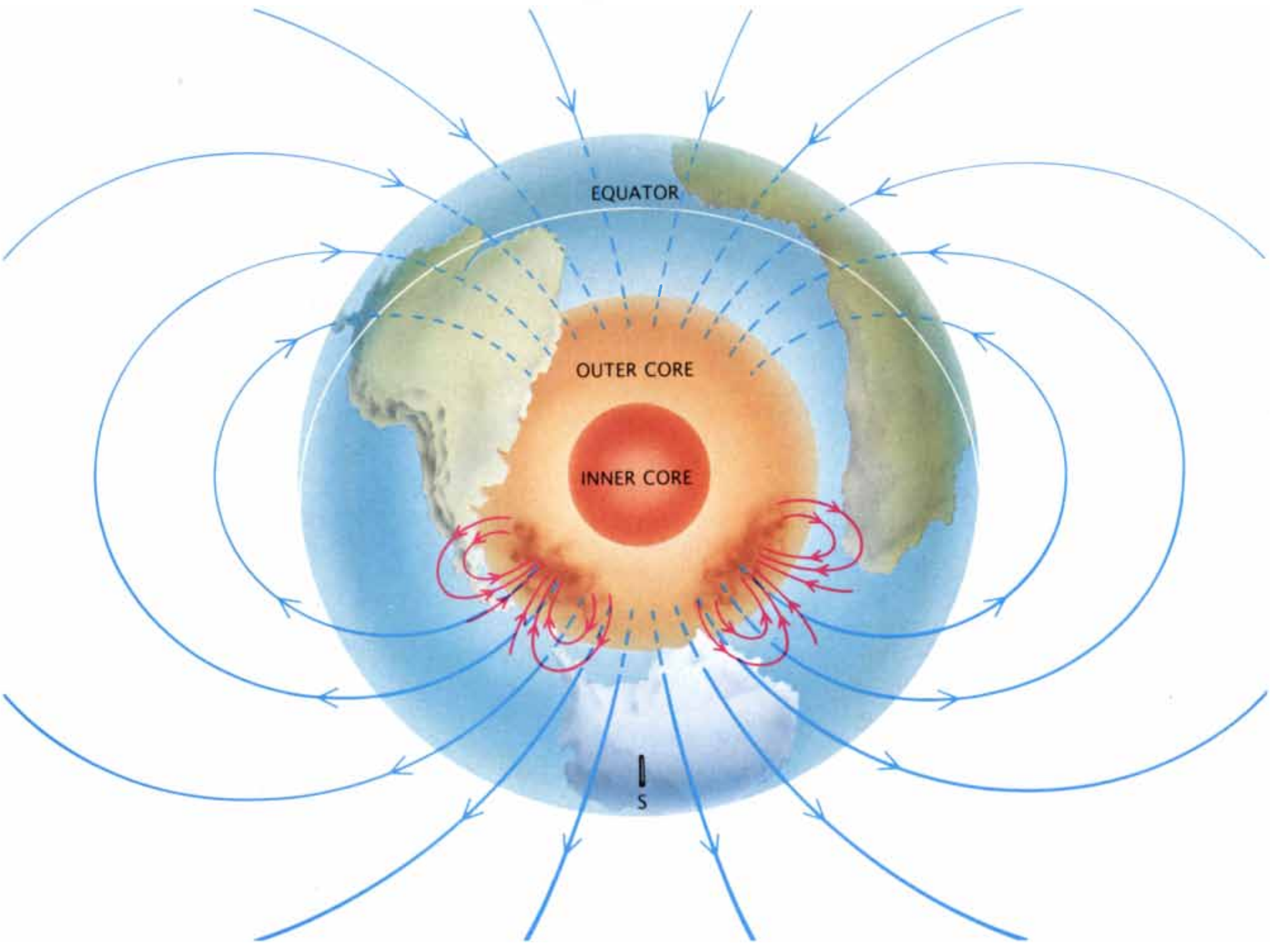
The paleomagnetic record has enabled investigators to deduce the geometry of ancient fields during both stable and unstable periods. It is well known that during times of stable polarity the magnetic field of the earth is dominated by a dipole shape, as though there were a bar magnet in the core, slightly tilted away from the earth's axis of rotation. During normal polarity a free compass needle—one that can swing in three dimensions—would everywhere point northward, dipping into the ground in the Northern Hemisphere and aiming skyward in the Southern Hemisphere. The angle of dip depends on the latitude of the compass. Conversely, at times of reverse polarity the needle would point to the south, tilting upward in the Northern Hemisphere and downward in the Southern Hemisphere.

But what of the field during the process of reversal? This question can be resolved by examining records of the same reversal from sites scattered around the globe. For each recorded magnetic direction it is possible to determine the location of the magnetic pole from the angle of dip and the horizontal orientation [*see top illustration on page 79*]. Hence by examining the change in paleomagnetic field direction at various intervals during a transition, one can track the virtual, or suspected, geomagnetic pole (VGP) as it travels from one polarity to the other. Moreover, if paleomagnetic records from different sites yield the same locations for the path of the pole, one can conclude that the field was indeed dipolar. But if the different records suggest wildly dissimilar VGP paths, one must conclude that the geometry of the reversing field was more complicated than a dipole.

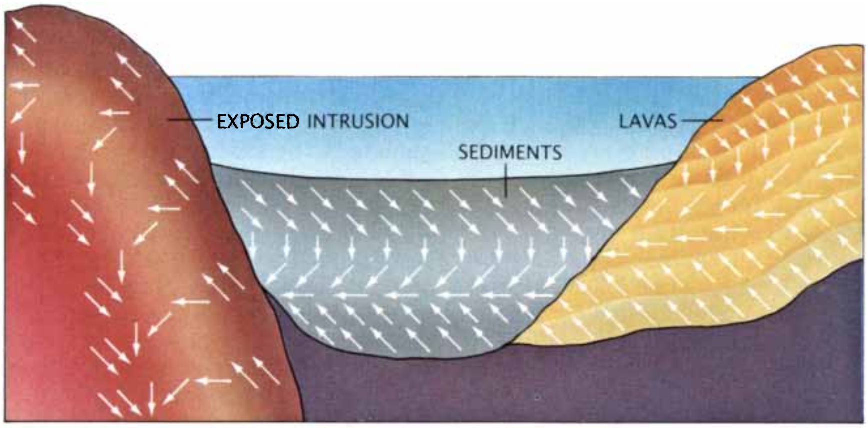
John Hillhouse and the late Allan Cox, both then at Stanford University, were the first to attempt such an analysis. They studied a record they had discovered in the sediments of dry Lake Tecopa in California. The sediments chronicled the most recent reversal, the polarity transition that took place some 730,000 years ago and brought the field from the Matuyama reverse epoch into the Brunhes normal epoch. Noting that the VGP path obtained from their data was quite different from the path associated with a marine-sediment record from Japan, they concluded that the configuration of the transitional field was complex and predominantly nondipolar. Similar paleomagnetic records of the same reversal from other locations also indicated quite disparate VGP paths [*see top illustration on page 80*].

Assuming that these data are reliable, can one analyze them in some other way to help determine the geometry of the transitional field during the Matuyama-Brunhes reversal? Since the reversing field cannot be assumed to be dipolar, the VGP method described above is of limited use. One needs a way to analyze the data without making any prior assumptions about the shape of the field. The method I now employ is to consider only the direction of the paleomagnetic field vectors at each location and to plot the path of the vectors on the surface of a "directional sphere" [*see top illustration on page 79*].

For the Matuyama-Brunhes reversal there are now many records from separate locations in the Northern Hemisphere. One can therefore compare the data by plotting all the directional paths on the same sphere [*see top illustration on page 80*]. When this is done, an interesting feature emerges: during the change in polarity from reverse to normal, most of




**MAGNETIC LINES OF FORCE (*blue*) emanate from the earth's molten outer core in this see-through view. David Gubbins of the University of Cambridge has determined that patches of oppositely oriented field lines (*red*) exist near the tip of Africa and of South America. These "core spots" can account for the observed weakening of the earth's field. The spots are growing in size and strength and are moving southward. Gubbins suggests that this may be the process that eventually reverses the earth's field.**




**ROCKS can record the direction (*arrows*) of the earth's magnetic field at the time of their formation. Igneous rocks contain magnetic grains whose magnetic moments become oriented with the prevailing field as the rock cools. Lava flows (*yellow*) cool rapidly and provide the most accurate "snapshots" of the paleomagnetic field. Because of the irregularity of eruptions, such records may contain significant gaps. Intrusions (*brown*), formed by magma cooling underground over thousands of years, provide a more continuous record, but because they take a longer time to cool, the recorded field orientation may be averaged over some time. Intrusions cool inward, so that the paleomagnetic record is oldest near the surface and youngest in the interior. Sediments (*gray*) contain magnetic grains that become aligned with the field and fixed in place as the sediment consolidates. Sediments take time to "lock in" their magnetic remanence.**

the intermediate field vectors—those that stray more than 30 degrees from either of the axial dipole field directions (or "pseudopoles")—sweep along the "underside" of the plot. That is, if one imagines that all over the Northern Hemisphere north-seeking compass needles are swiveling from geographic south to north as the field reverses polarity, most of the needles would flip downward rather than upward. Moreover, the needles would not stray far from the north-south vertical plane. The implication is that the transitional field does not change much in the east-west direction and hence that it is roughly symmetrical about the earth's rotational axis.

### "Hot Spots" in the Geodynamo?

If all sites in the Northern Hemisphere recorded similar field behavior during the Matuyama-Brunhes reversal, one would strongly suspect that the underlying activity in the earth's core was also roughly symmetrical about the earth's rotational axis. Such axisymmetry can be explained by supposing the reversal process begins within latitudinal—that is, axisymmetric—bands in the core fluid [*see bottom illustration on page 80*]. Such bands would be associated with local field lines oriented in a direction opposite to that of the overall field. The bands might grow, flooding their way through the rest of the core and bringing about a reversal of the entire field.

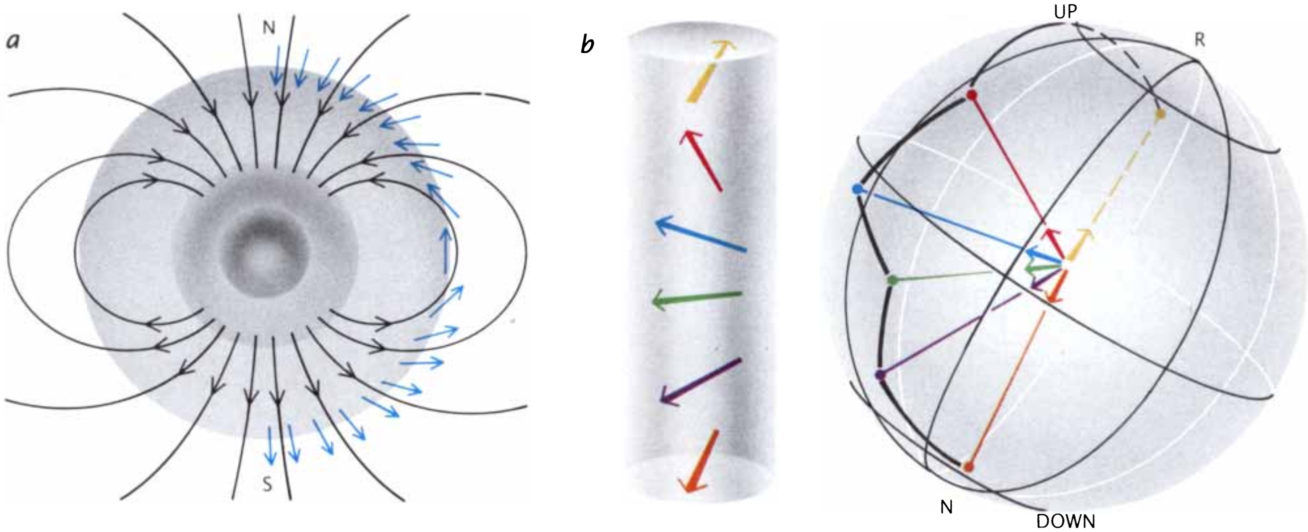
The Matuyama-Brunhes data impose constraints on such a theory: they are consistent with models in which this process begins near the equatorial plane of the core or in its southern hemisphere but are inconsistent with a process that begins in the core's northern hemisphere [*see bottom illustration on page 80*]. Before one can decide whether this model is plausible it will be necessary to obtain reliable paleomagnetic data from additional sites, particularly in the Southern Hemisphere.

In the meantime, support for this model comes from the investigation of the ongoing decline of today's field. David Gubbins of the University of Cambridge surmises that the waning strength of the dipole field results from the growth and intensification of regions at the surface of the outer core from which emerge magnetic field lines, or flux, whose direction is opposite to present-day polarity. These regions are at high southern latitudes of the core, below the tip of Africa and of South America [*see illustration on preceding page*]. They are thought to be associated with particularly hot parts of the lowermost mantle, directly above the outer core. If this trend continues, Gubbins suggests, the sense of the dipole can ultimately reverse.

Gubbins' findings enable one to speculate that the reverse-to-normal Matuyama-Brunhes event and the possible onset of a normal-to-reverse transition today are both consequences of a similar process, initiated in the southern hemisphere of the core. If they are, the field lines of the earth should eventually change direction in a manner similar but antipodal to the behavior observed for the Matuyama-Brunhes reversal. In other words, one would expect compass needles around the world to trace out a path largely confined to the vertical plane. In contrast to the behavior of hypothetical compasses during the Matuyama-Brunhes reversal, however, compass needles in both hemispheres should rotate upward instead of downward.

Yet what reason is there to believe the reversing dynamo can repeat itself in this way? Are there particular "hot spots" within the core that trigger geomagnetic field reversals? Actually there exists very good paleomagnetic evidence that the triggering process in the core can remain essentially unchanged over the time span of several reversals. The most striking evidence comes from a series of reversals recorded in marine sediments from the island of Crete. Of four recorded transitions (two of which were reverse-to-normal and two normal-to-reverse) reported by Jean-Pierre Valet and Carlo Laj of the French National Center for Scientific Research, the three oldest events


**TIME SCALE of magnetic reversals for the past 170 million years was deduced from the magnetic field pattern observed over the ocean crust. The spreading ocean floor pre-**




**DIPOLE FIELD has a clearly defined shape (*a*). Hence for a given paleodirection one can deduce the location of the magnetic pole, or virtual geomagnetic pole (VGP), simply from the field vector's horizontal direction and angle of dip. One can analyze a reversal by plotting the path of the VGP as the field changes polarity. This method is not wholly suited for more complex fields. A "directional sphere" (*b*) makes no assumption about the shape of the field. The sphere is imagined to enclose a particular site. One can picture an observer standing at the site, plotting on the sphere's surface the directions indicated by a free compass needle. In practice one can take field directions recorded in strata of rock, such as the vertical sediment core depicted here, and trace out the path (*black line*) of the field vector during a reversal. "Up" and "down" are reference points corresponding to vertically upward and downward. The "pseudopoles" (*N, R*) represent directions associated with pure axial dipole fields that have normal and reverse polarity. The "equator" defines all directions that are 90 degrees from each pseudopole, or midway between them.**

exhibit directional field behavior that indicates a similar underlying core mechanism. The two reverse-to-normal events trace out paths that are antipodal to the normal-to-reverse event. (In contrast to the Matuyama-Brunhes reversal, these events were characterized by strong east-west movements.)

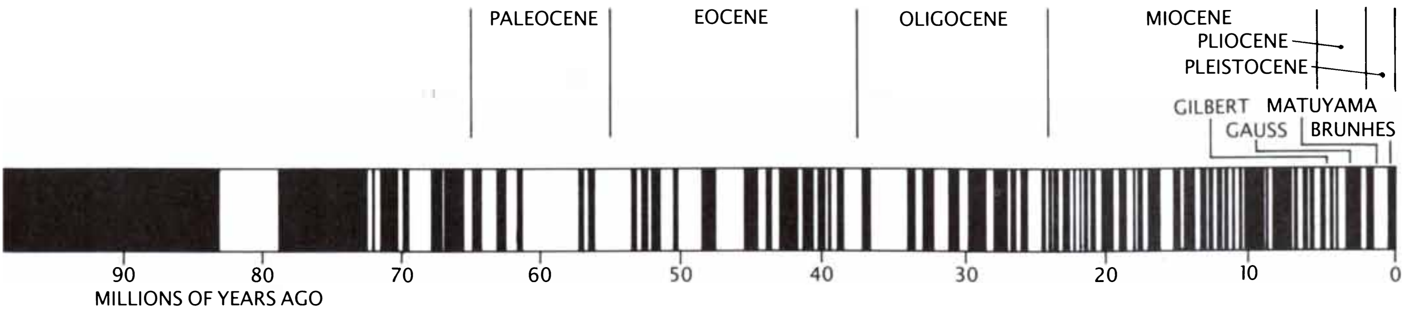
The most reasonable interpretation of this finding is that the geometry of the reversal process in the core did not change throughout the time interval—well over a million years—spanned by the three events. Demonstrating that other sets of contemporaneous paleomagnetic records show a similar pattern would increase confidence in the interpretation, however.

It must also be pointed out that just as there is repeatability in the reversal process, there is also variability. For example, the fourth and youngest of the transition records from Crete displays directional changes that are distinct from those in the older records, suggesting that the core process changes its spatial characteristics from time to time. Other back-to-back reversals reported by Scott W. Bogue and Rob S. Coe of the University of California at Santa Cruz and by Bradford M. Clement and Dennis V. Kent of the Lamont-Doherty Geological Observatory show back-to-back reversals in which the field directions follow similar rather than antipodal paths. These findings suggest that after a transition the dynamo sometimes "rewinds" like a film running backward.

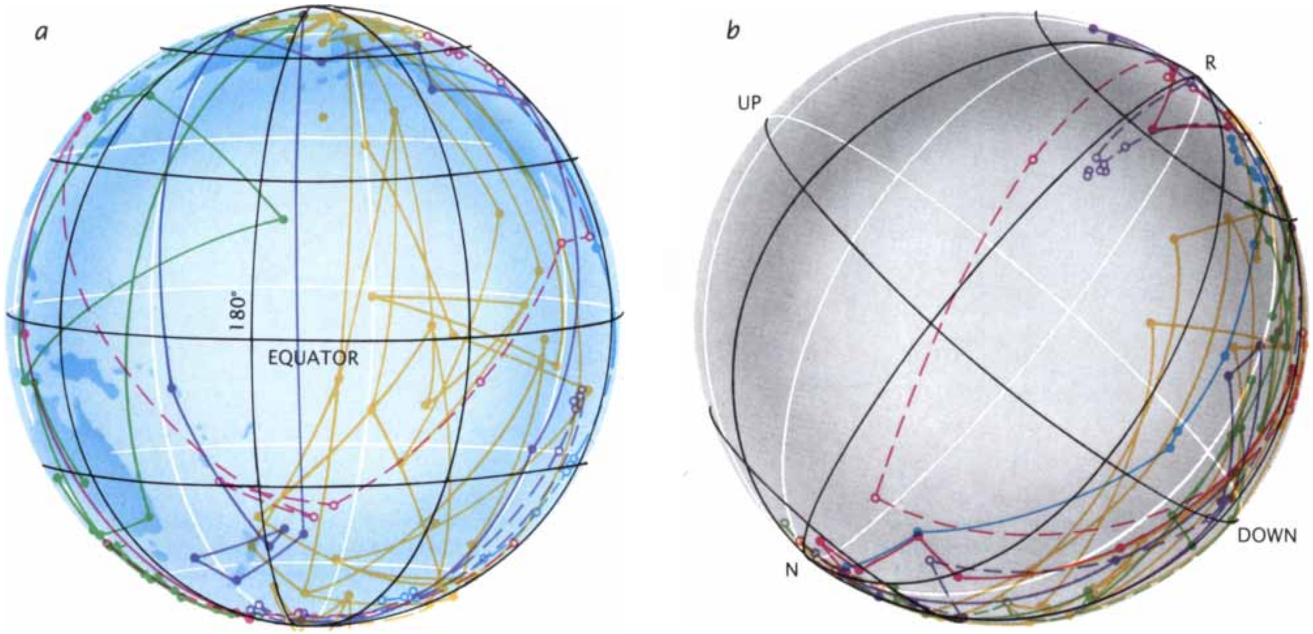
### Excursions and Rebounds

The "rewinding" reversals make one wonder whether the dynamo sometimes fails to complete the reversal process and ends up backtracking. A similar question is raised by the observation of geomagnetic excursions, or large leaps in field direction, a phenomenon that has long been noted in the paleomagnetic literature. The chief question about these excursions is whether they are produced in times of abnormal secular variation and have no relevance to the phenomenon of polarity reversal or, alternatively, are manifestations of unsuccessful, or aborted, reversal attempts.

Evidence that there may indeed be a link between certain geomagnetic excursions and polarity reversals comes from a 34-million-year-old sequence of lavas from the Liverpool Volcano in eastern Australia. I have found that these basalt flows record a series of three apparently rapid departures from reverse polarity to essentially the same intermediate field




**serves a log of ancient fields in basalts formed by magma welling up along deep-sea ridges. The pattern of reversals varies greatly. During the Cretaceous period there was a 35-million-year hiatus. The time scale is based on the work of Allan Cox.**

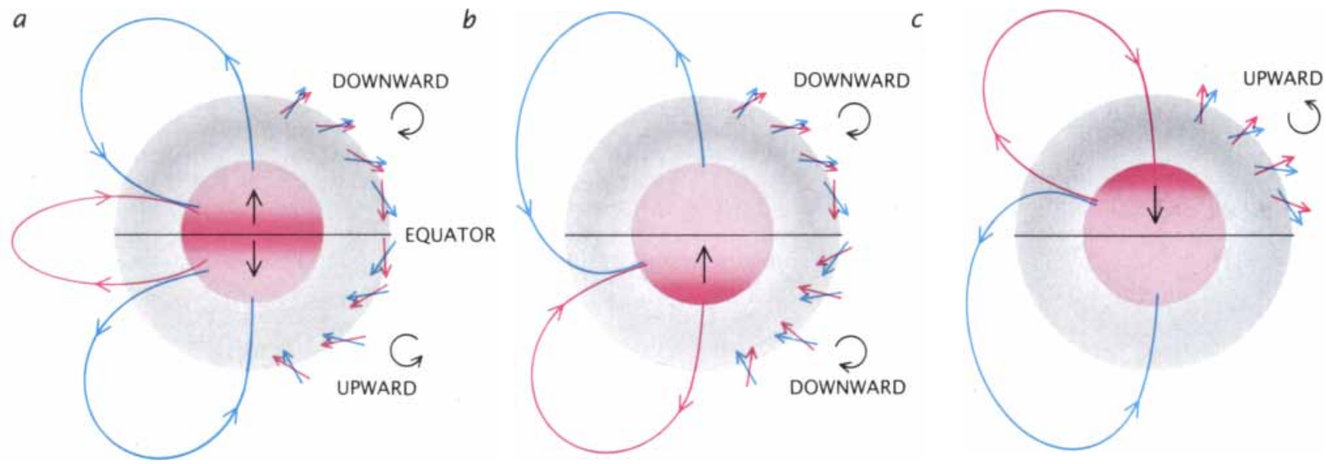



**PALEOMAGNETIC RECORDS** of the Matuyama-Brunhes reversal are analyzed by the VGP method (*a*) and by the directional method (*b*). The paths are based on data from five northern-latitude sites: Japan (*green*), Maui (*purple*), California (*yellow*), the North Atlantic (*blue*) and West Germany (*red*). If the reversing field were a dipole, all the VGP paths should coincide. But in fact the paths are widely scattered, proving that the field was nondipolar. Directional analysis reveals that during the transition the field vectors at all sites tended to tilt through a downward direction. This suggests east-west movements were small during reversal and hence the process was roughly symmetrical about the earth's axis. Maui data (unpublished) courtesy of Rob S. Coe.

orientation [*see illustration on page 82*]. The first two of these advances appear to be excursions because the lavas that immediately follow record the same (reverse) polarity direction that prevailed before the excursion. The lavas extruded directly after the third departure, on the other hand, possess normal polarity, indicating a successful reversal.

The Liverpool data are consistent with the contention, first made by John Shaw of University College Cardiff in Wales, that the reversal process may involve an intermediate dynamo state. This state seems to act as a kind of springboard from which reversal attempts, both successful and unsuccessful, are made. Such relatively stable intermediate field directions, or "hang-up" points, are in fact quite common and suggest that certain field geometries may offer energy states that are preferred by the geodynamo during reversal events.

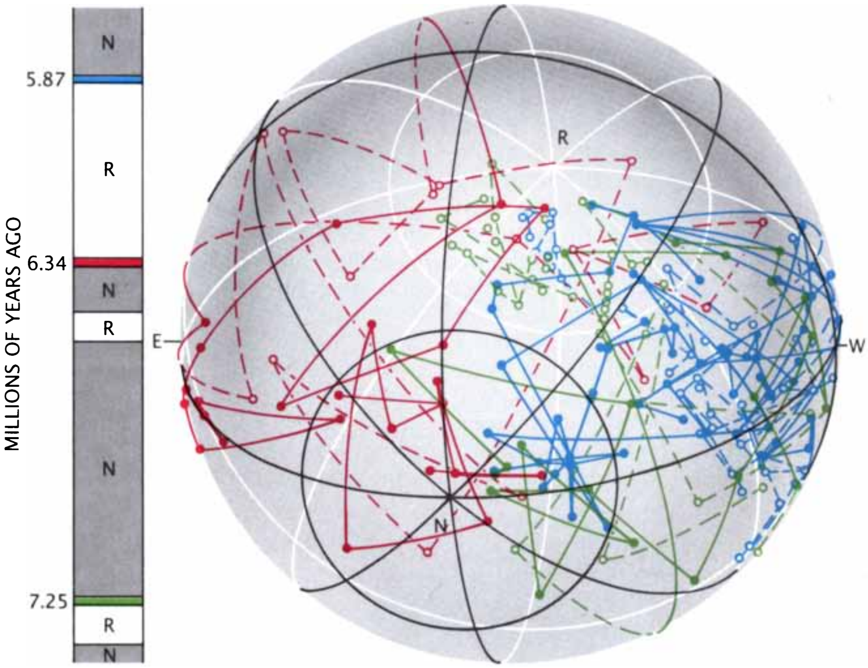
The first investigator to describe apparently stable intermediate directions was the late Norman D. Watkins of the University of Rhode Island, who in the late 1960's made a study




**AXISYMMETRY** during the Matuyama-Brunhes reversal can be explained by a process that begins with a latitudinal band (*dark red*) of core fluid that produces a field opposite to the dominant field. The field shown here begins in the reverse dipole state (*blue arrows*); the arrows point, like north-seeking compasses, toward the geographic South Pole. An equatorial band spreading northward and southward (*a*) would cause field vectors (*red arrows*) in the Northern Hemisphere to rotate downward (*black arrow*). A band starting near the geographic South Pole and expanding northward (*b*) would also cause field vectors to tilt downward. But a band near the geographic North Pole spreading southward (*c*) would result in an upward rotation, contrary to the observed pattern. Data from the Southern Hemisphere would reveal whether the reversal process began near the equatorial plane, in which case field vectors would rotate upward, or in the south, in which case field vectors would rotate downward.

of the 15-million-year-old reverse-to-normal transition recorded in the lavas from Steens Mountain in southeastern Oregon. The extent of this lava record is without parallel: it contains numerous, successive flows that preserve the field direction at some 55 distinct times throughout the reversal. Because of the rarity of such complete, high-resolution paleomagnetic records of a reversing field, the Steens Mountain lavas have remained the focus of an intensive investigation.

The work has proved well worth the effort. The study, carried out by Coe, Michel Prévot of Montpellier, France, and Edward A. Mankinen and C. Sherman Grommé of the U.S. Geological Survey, provides the most detailed account of both the field's directional behavior and its changes in intensity during a polarity transition. Their result suggests that the core process is rather complicated. A free compass needle at the Steens Mountain site at first would have traced out a generally downward path. But when the field reached a particular intermediate orientation, it apparently made a rapid directional jump and attained normal polarity. The new polarity, however, did not last for long, because what is observed next is yet another directional jump in which the field vector rebounded to the very same intermediate orientation. After the rebound the field direction swept out a loop to the east while again approaching normal polarity. This time the polarity change was successful.




**THREE SEQUENTIAL REVERSALS recorded in sediments from Crete all seem to have been initiated in the same manner. The oldest reversal (*green*) and the youngest reversal (*blue*), both of which took the field from reverse polarity to normal polarity, followed paths on the west side of the sphere. In contrast, the middle reversal (*red*) from normal to reverse polarity followed an eastward path. This antipodal behavior suggests that spatially the same process was played out in the core during all three events.**

### Stop-and-Go Dynamo

The Steens Mountain lavas give the distinct impression that changes in polarity involve a kind of stop-and-go behavior of the dynamo. Furthermore, the investigators argue that because the observed jumps return the field vector to the same intermediate position, the changes result not from gaps in the record but from rapid shifts, perhaps taking only a few years, in the direction of the field. Support for this idea comes from the discovery of significant differences in paleomagnetic field direction between the faster-cooling edges and slower-cooling center within a single lava flow. If these results are reliable, they suggest that certain mechanisms for rapidly changing the magnetic flux must play a dominant role in the core dynamo during reversal.

The best way to test our confidence in this interpretation is to compare records such as the one from Steens Mountain with contemporaneous records, preferably from a nearby site, obtained from material that provides more continuous data than intermittent lava flows do. Unfortunately transition records of any kind are hard to find. A reversal from a similar period, however, has left a detailed record at the Tatoosh Intrusion from Mount Rainier, not far from Steens Mountain. Although the reverse-to-normal polarity change recorded in this intrusion is most likely not the same event recorded at Steens Mountain, the two sets of data share certain features in common.

Michael D. Fuller and his fellow workers at the University of California at Santa Barbara reported that early directional changes appearing in the Tatoosh Intrusion record are dominated by what appears to be a quite rapid movement to an intermediate position, where the field more or less stabilized. The field then underwent an unsuccessful attempt to reverse: normal polarity was approached and then lost as the field direction returned to about the same intermediate position in a pattern resembling the rebound seen in the Steens Mountain record. Following that aborted attempt the polarity change was successfully completed.

Thus two different types of record—the lava record from Steens Mountain and the Tatoosh Intrusion record—both support the contention that during reversal attempts the field can move rapidly as well as remain stationary and can undergo directional rebounds and unsuccessful attempts to complete the reversal process. Remarkably similar behavior in another reversal from the same geologic epoch has just been reported by Laj and his colleagues, who obtained their data on the island of Zakinthos in Greece from yet a third type of paleomagnetic record, marine sediments.

So far I have focused exclusively on the directional aspects of the reversing field, and yet a complete comprehension of the underlying dynamo activity also requires one to know about accompanying changes in field strength, or paleointensity. The latter is considerably more difficult to determine than direction. Nevertheless, many sites have now yielded intensity data, which indicate that intermediate orientations tend to be associated with weak field strengths, sometimes as low as 10 percent of the stable-polarity field strength.

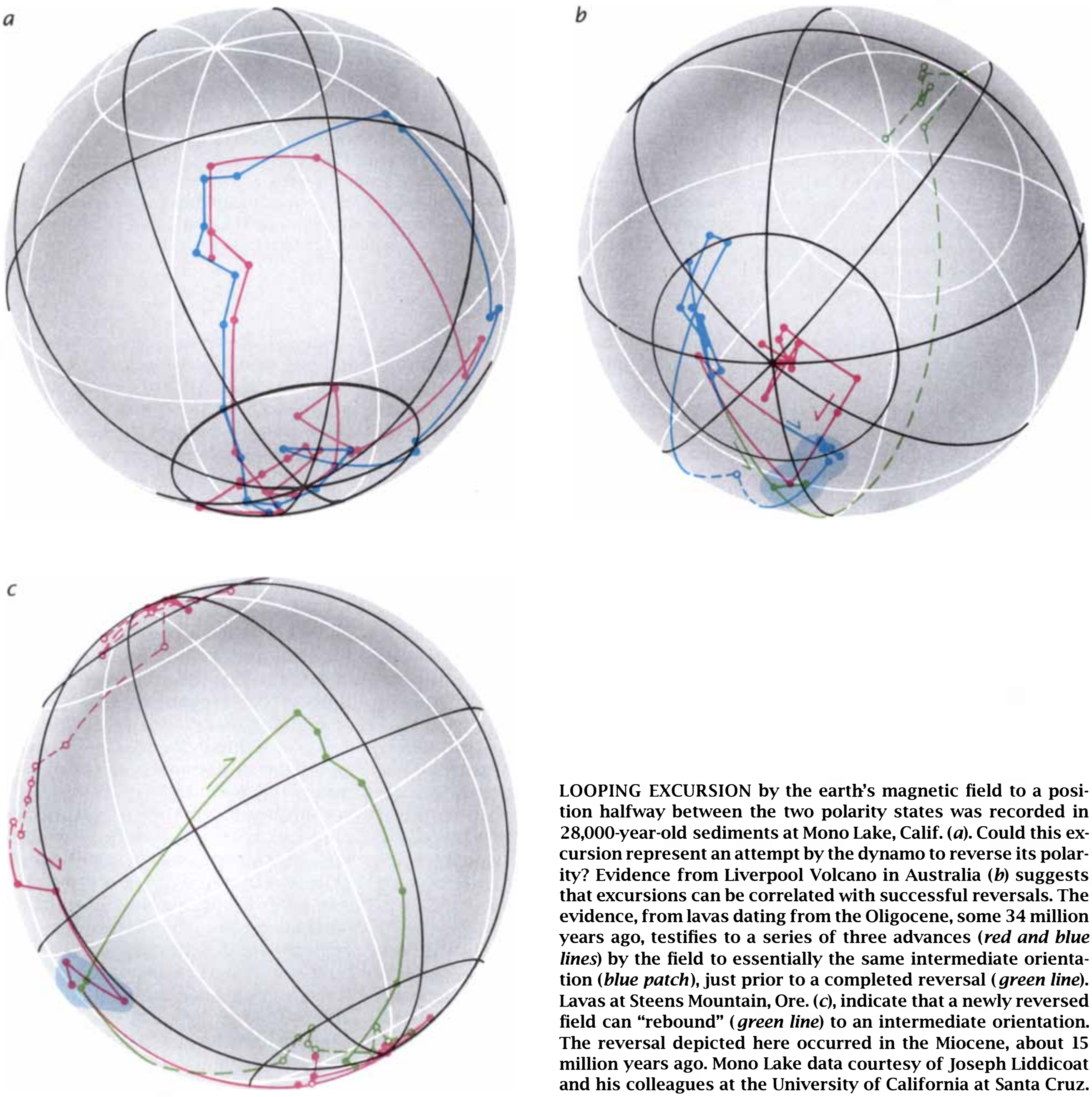
The most detailed information on

intensity changes comes once again from the reversal event recorded at Steens Mountain. The field recovered to pretransition intensities when it approached full reverse polarity during its first, unsuccessful attempt, prior to the rebound. The intensity plunged again as the field retreated to the intermediate position. During the final stage of the reversal the intensity seems to have varied widely, sometimes becoming stronger than the ordinary field, as the dynamo regained equilibrium in the new polarity state. My own preliminary results from studies of a reversal recorded in lavas on the Hawaiian island of Molokai indicate that paleointensities fluctuated to more than three times the pretransition value as the dipole field recovered.

### Two Core Processes

The paleomagnetic record is beginning to provide a picture of a reversing geodynamo that is capable of both slow and rapid changes in direction and intensity. First principles of physics allow two basic dynamo mechanisms, a slow mechanism and a potentially fast one, by which the pattern of magnetic field lines emerging from the core can change. One mechanism rests on "flux diffusion," which makes it possible for magnetic lines of force to move through the core fluid from higher to lower flux concentrations. Diffusion is a passive process that provides a way for the field to decay but not regenerate.

The dipole field today, however, is decaying about 10 times faster than can be accounted for by diffusion alone. Clearly there must be a more active process that builds up and alters the field dynamically. This mechanism is provided by "frozen-in flux," in which the magnetic flux lines are carried along with the flowing core fluid. The more closely the core fluid resembles a perfect conductor of electricity, the more rigidly the field lines move along with it.



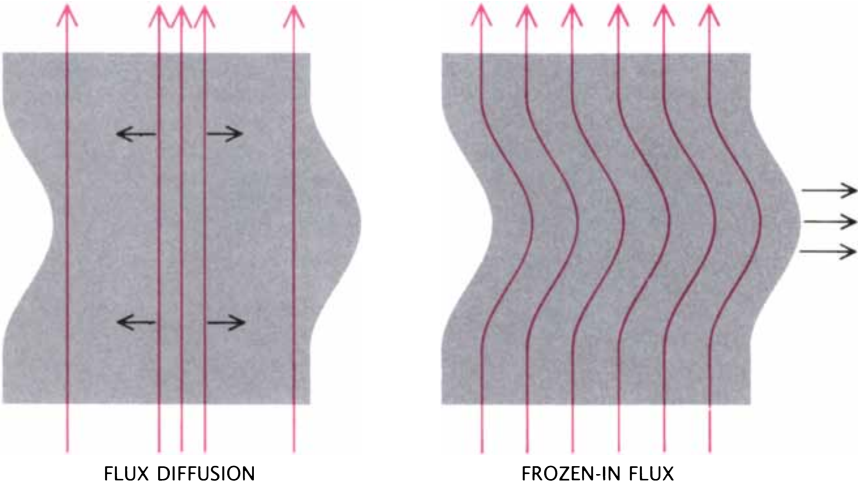

**LOOPING EXCURSION by the earth's magnetic field to a position halfway between the two polarity states was recorded in 28,000-year-old sediments at Mono Lake, Calif. (*a*). Could this excursion represent an attempt by the dynamo to reverse its polarity? Evidence from Liverpool Volcano in Australia (*b*) suggests that excursions can be correlated with successful reversals. The evidence, from lavas dating from the Oligocene, some 34 million years ago, testifies to a series of three advances (*red and blue lines*) by the field to essentially the same intermediate orientation (*blue patch*), just prior to a completed reversal (*green line*). Lavas at Steens Mountain, Ore. (*c*), indicate that a newly reversed field can "rebound" (*green line*) to an intermediate orientation. The reversal depicted here occurred in the Miocene, about 15 million years ago. Mono Lake data courtesy of Joseph Liddicoat and his colleagues at the University of California at Santa Cruz.**

Based on observations of the present-day field, frozen-in flux appears to dominate short-term changes—changes that occur over a period of a few decades.

A successful theory of the geodynamo, and in particular the mechanism of polarity reversal, will rest in part on determining the relative importance of these two processes. Again the paleomagnetic data are beginning to yield a reasonably consistent picture. The reversing geodynamo appears to take a very active role at times and a rather passive role at other times. It is thought that during the rapid directional jumps unusually high accelerations in fluid flow drive the frozen-in flux into continuously changing patterns. In times when the field is relatively quiet, such as the hang-up periods, flux diffusion must play a larger role.

In some way that is not yet fully understood, gravity and the earth's rotation, acting on density differences within the core fluid, provide the driving forces behind the geodynamo and control this dichotomy of dynamo processes. The reversal phenomenon may be triggered when something disturbs the convection pattern of the core fluid, and with it the magnetic flux. Phillip L. McFadden of the Bureau of Mineral Resources, Geology and Geophysics in Australia and Ronald T. Merrill of the University of Washington suggest that the triggering process is intimately related to the way the outer core vents its heat into the mantle. For example, heat transfer could create hotter (rising) or cooler (descending) blobs of material from the inner and outer boundaries of the fluid core, thereby perturbing the main convection pattern.

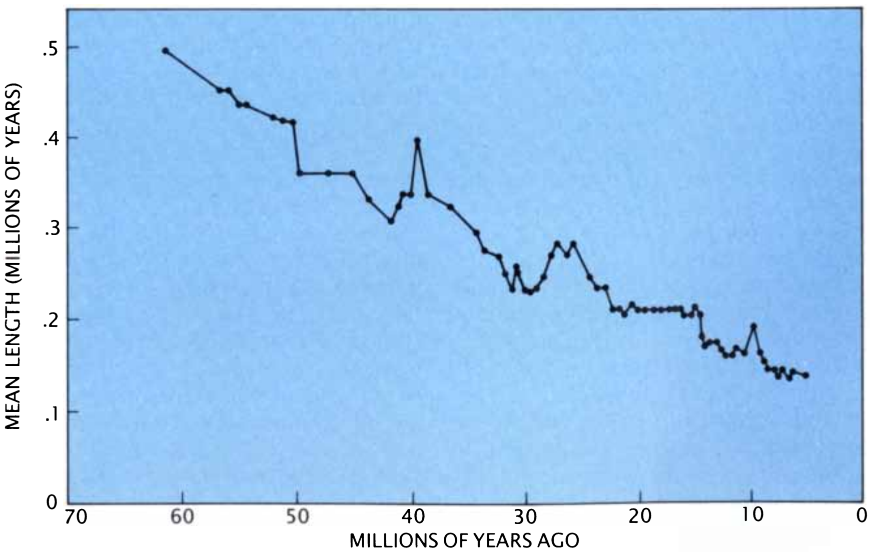
Among the more unorthodox and controversial theories is the "asteroid-impact hypothesis" proposed by Richard A. Muller and Donald E. Morris of the Lawrence Berkeley Laboratory. Their scenario begins with the impact of an asteroid or other extraterrestrial object large enough to send a great cloud of dust into the atmosphere. A kind of "nuclear winter" is seen as resulting, during which the polar ice caps would grow, abruptly altering the distribution of water on the earth's surface and with it the moment of inertia of the earth. The rotational acceleration of the mantle would then be increased to conserve angular momentum, causing friction and turbulence near the core-mantle boundary and initiating a reversal of the geomagnetic field.




**FIELD LINES AND CORE FLUID can interact in two ways. Flux diffusion occurs when field lines (*red*) from regions of high field density spread to regions of low field density. This process accounts for field decays. Frozen-in flux occurs because the field lines are dragged along with the highly conductive core fluid and so can build up local fields.**

How well do these two hypotheses account for such observations as the long-term increase in the frequency of reversal? In support of their model, Morris and Muller argue that the gradual cooling of the average ocean temperature would enable progressively smaller asteroid collisions (which occur more frequently) to induce ice-cap growth and reversals. But theories that depend on extraterrestrial intervention seem less convincing than theories such as the one offered by McFadden and Merrill, accounting for observed features of the geodynamo based solely on the thermodynamic state of the core and its effect on the deep mantle.

Geophysics today stands at a historic juncture. For the past three decades workers have combed the globe for the ghostly traces of ancient field reversals, correlated evidence from scattered sites, tried to sift reality from a less than ideal record and pieced together bit by bit a picture of the reversing field. Their results now challenge theorists to develop geodynamo models that can account for the observed behavior. As more paleomagnetic evidence is brought to light, it will play a critical role in determining which of these models are plausible and which ones are not.




**DWINDLING length of the average polarity interval, when the earth's field is in the normal or the reversed state, reveals that polarity reversals have been taking place at an increasing rate. A complete geodynamo model will have to account for this trend.**

# The Platypus

*This ancient mammal is surprisingly specialized. It has mechano- and electroreceptors on its beak for detecting prey and is better than many placental mammals at regulating its body temperature*

by Mervyn Griffiths

The platypus was brought to the attention of the Western world in 1799, when George Shaw of London described it in *The Naturalists Miscellany,* where he wrote: "Of all the Mammalia yet known it seems the most extraordinary in its conformation; exhibiting the perfect resemblance of the beak of a Duck engrafted on the head of a quadruped.... At first view it naturally excites the idea of some deceptive preparation by artificial means."

Shaw guessed that the platypus, whose feet were webbed, "must be a resident in watery situations; that it has the habits of digging or burrowing in the banks of rivers or underground and that its food consists of aquatic plants and animals." He gave it the name *Platypus anatinus* (flatfooted ducklike animal), which was later changed to *Ornithorhynchus anatinus* (ducklike animal with a bird's snout).

His description was based on meager evidence: a single skin and a sketch that are thought to have been sent to Newcastle-upon-Tyne from Australia by Captain John Hunter of the Royal Navy, who became acquainted with the platypus while serving a term as governor of the British penal colony in New South Wales. Nevertheless, Shaw's early account was surprisingly accurate.

Even now a casual observer would remark that the platypus has the appearance of several animals rolled into one. Covered in beautifully soft fur, this graceful amphibious mammal has a tail not unlike that of a beaver and a bill not unlike that of a duck. Its feet are webbed and, although it has mammary glands with which it suckles its young, it does not give birth but lays eggs that are incubated and hatch outside the body. Adult males weigh only 1.7 kilograms (3.8 pounds) and measure 50 centimeters (20 inches) from the tip of the beak to the tip of the tail. Females are even smaller: 43 centimeters (17 inches) in length and .9 kilogram (two pounds) in weight.

The platypus is one of three species belonging to two families—the Ornithorhynchidae and the Tachyglossidae (the latter includes two genera of spiny anteaters called echidnas)—that together are known as the Monotremata, or egg-laying mammals. The name Monotremata (from the Greek, *mono* meaning one and *trema* meaning hole) refers to the fact that these animals have a single external opening, or cloaca, for both excretion and reproduction.

Although the monotremes are highly specialized, they represent a branch of mammals that is quite ancient. They are more closely related to marsupial and placental mammals than to any group of reptiles, yet they have retained a surprising number of ancestral reptilian traits over the course of evolution and possess an interesting mosaic of mammalian and reptilian characteristics. When their similarity to reptiles was first noted early in the 19th century, there was much confusion within the scientific community; some even suggested that the platypus might be a missing link between mammals and reptiles.

In 1802 the English anatomist Sir Everard Home published a description of the platypus in which he stated that male platypuses, like reptiles, have internal testes and that females have a cloaca. The latter finding was particularly interesting because it was a trait known to occur in ovoviviparous lizards: those that retain their eggs in the genital tract until the young are just ready to hatch.

The evolutionary link between reptiles and monotremes was reinforced in 1826, when the German anatomist Johann F. Meckel reported that the platypus has several bones in its pectoral girdle (the coracoids, epicoracoids, clavicles and median interclavicle) that are known only from fossil therapsids: extinct mammal-like reptiles thought to be the stem group from which all mammals evolved. At the same time he reported finding mammary glands in the female, an anatomical trait that is unquestionably mammalian.

The way in which a platypus reproduces remained a mystery, however, until some 82 years after Home's description, when an energetic young Scottish investigator, W. H. Caldwell of the University of Cambridge, set out for Queensland. Intrigued by Home's description of the reproductive tract, he wanted to settle the question of ovoviviparity in the monotremes. Soon after arriving in southern Queensland he found the answer he had been seeking: the platypus and the echidna do not bear young as other mammals do but lay eggs, like reptiles. Moreover, he found that monotreme eggs, like reptile eggs, are meroblastic: only the nuclei and cytoplasm at the upper pole, or anterior end, of the egg divide; the yolk does not.

Since that time additional data linking the platypus to its reptilian ancestors have come to light. Even at a cellular level the platypus is more similar in some ways to reptiles than it is to marsupial and placental mammals. Peter D. Temple-Smith of the Australian National University and Frank N. Carrick and R. Leon Hughes of Queensland University found that platypus sperm is long and slender, with a filiform, or threadlike, head, much like the sperm of reptiles; the arrangement of the subcellular elements called microtubules, however, is typical of mammalian sperm.

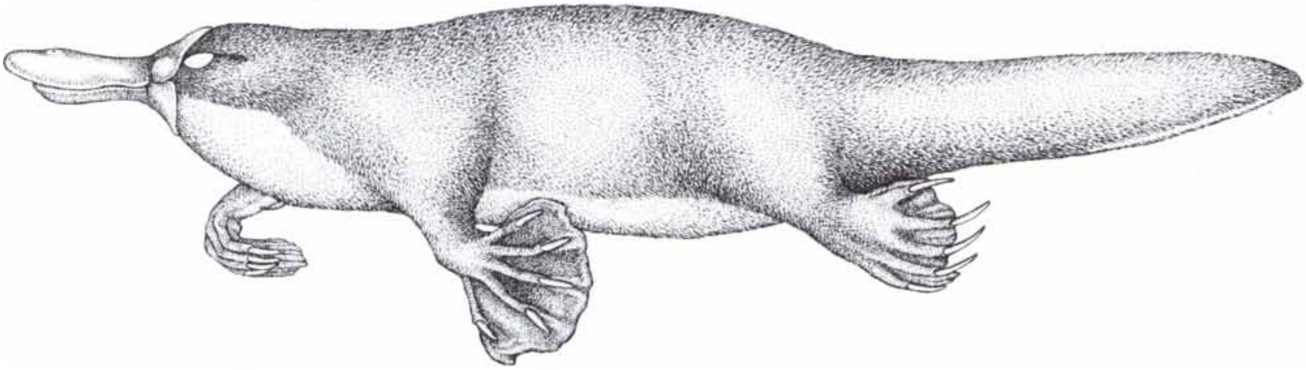
Monotreme chromosomes also reflect a mixture of reptilian and mammalian traits. Although these animals are unique in having two categories of chromosomes, large and small, the large ones (macrochromosomes) are typical of those found in mammals, whereas the small ones (microchromosomes) are similar to those found in many species of reptiles and do not occur in other mammals.

At the turn of the century the platypus was almost rendered extinct by hunters seeking its thick, soft fur. Today it is heavily protected throughout its range with severe penalties imposed on anyone caught with one, and as a consequence the species has experienced a remarkable comeback in numbers. Surveys published by Temple-Smith, Tom Grant of the University of New South Wales, Gordon Stone of the Queensland National Parks and Wildlife Service and me in 1973, 1983, 1985 and 1988 show that the platypus is now present, and even abundant, in waters east of the Great Dividing Range—from northern Queensland to the southern part of Tasmania.

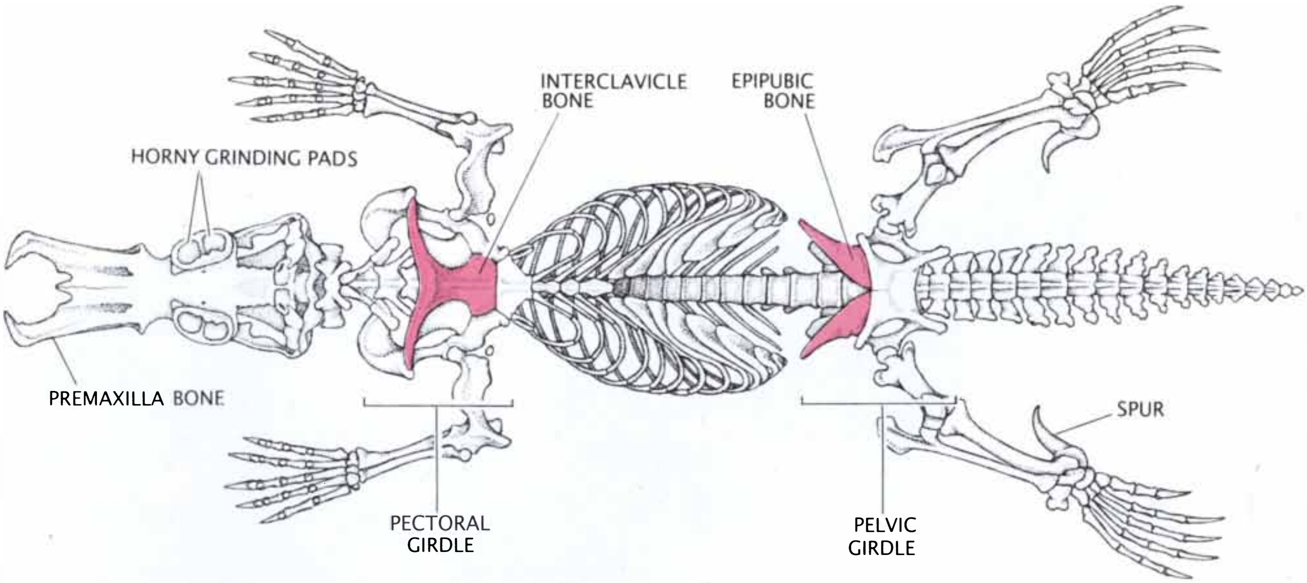
The platypus has been the focus of much attention in recent years, with many investigators studying its biochemistry and physiology in the laboratory, yet few have studied the animal in its natural habitat, in large part because it is so elusive. Endemic to freshwater streams and lagoons in eastern Australia, it is most active at night, when it dives for bottom-dwelling invertebrates such as insect larvae, shrimps and bivalve mollusks. During the day the platypus retires to a burrow it has constructed in the bank of the waterway, where little is known of its behavior. In spite of these drawbacks to its study, the image of the platypus as an evolutionary leftover, midway between reptiles and mammals and poorly adapted to modern life, is disappearing; instead it appears to be a highly specialized and successful mammal.

Life-history data collected by Grant over the past 15 years have provided a wealth of information on the longevity of the platypus and its movements. By suspending gill nets in the

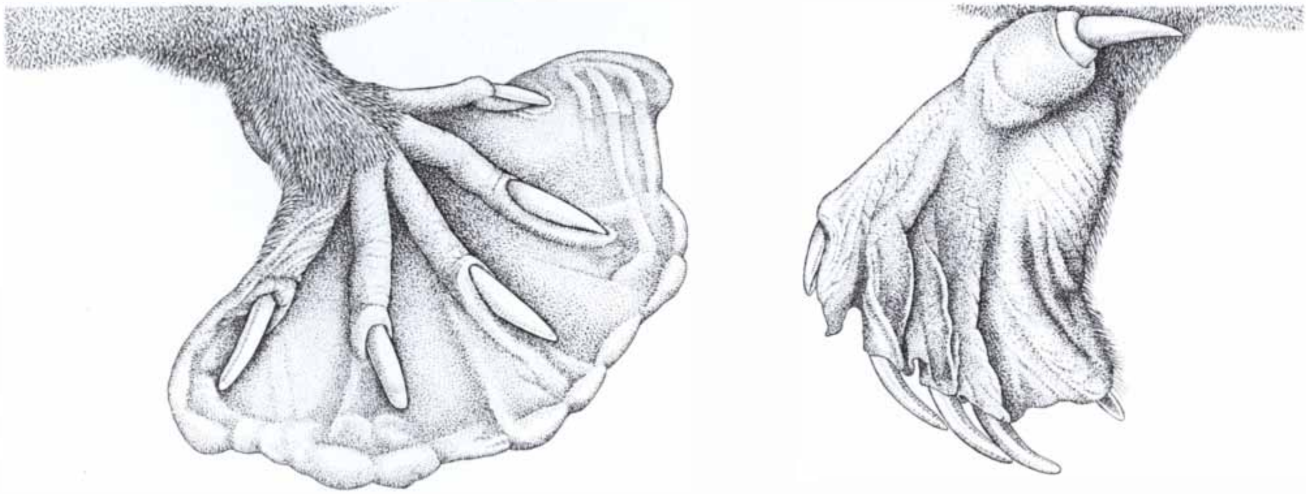
**PLATYPUS is amphibious in its habits, spending several hours a day foraging for invertebrate prey on the bottom of freshwater rivers and lakes. During a dive, often in murky water, the animal closes its eyes, ears and nostrils. To find its prey and navigate around obstacles, it relies on thousands of receptors in its duck-like snout that respond both to touch and to electric currents.**



STREAMLINED BODY of the platypus is efficient both in the water and on land. When underwater, as is shown here, the animal maneuvers about, propelling itself with paddlelike forefeet. On land it moves forward by gripping the substrate with its claws.




SKELETON of a modern platypus (in ventral view) is partly mammalian and partly reptilian. Like all mammals, the platypus has seven cervical vertebrae, but other features (*color*), including the cervical ribs, the interclavicle of the pectoral girdle and the epipubic bones of the pelvic girdle, resemble those of a reptile. The premaxilla bone has been extended to form a bill, and horny grinding plates on the jaw replace teeth. Juveniles have degenerate molars that are replaced by the horny pads at weaning time.



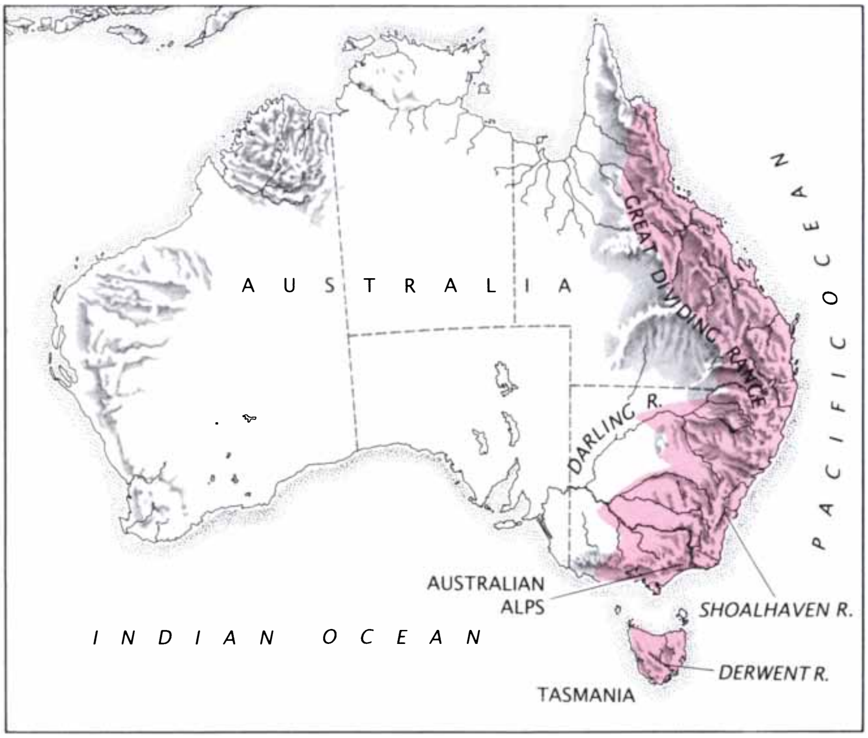
FORE AND HIND FEET of the platypus are short and close to the body. The fore foot (*left*) is broadly webbed and is the animal's main means of propulsion in the water. On land the webbing is retracted, exposing strong nails with which the animal grips the terrain and excavates its burrow. The hind foot (*right*) has claws that also provide traction; in addition males have a hollow spur near the ankle that connects to a venom gland and delivers a deadly toxin during aggressive encounters with other animals.

water, Grant was able to trap the animals, fit them with a numbered metal band on the hind leg and then release them from the spot where they were captured. He has marked and recaptured 468 platypuses to date along 11 kilometers of the Shoalhaven River in New South Wales. Because of Grant's work it is now known that a platypus lives for up to 12 years, successfully breeds into old age (Grant has captured females at least 11 years old that are still lactating) and may spend most of its life in the same body of water.

Temple-Smith has done much to elucidate the processes of reproductive physiology in the platypus. The animals are seasonal breeders; in summer and early fall the reproductive organs in both males and females are small and inactive. During May, June and July the glands associated with reproduction begin to enlarge and secrete hormones. At the same time there is a distinct change in the temperament of the males: they become aggressive, engaging in frequent combat with other males.

Each male has a sharp, inwardly directed hollow spur on the hind leg, up to 1.5 centimeters long and filled with a powerful toxin. These spurs are used in combat against other males; Temple-Smith believes they may play an important role either in establishing territories or in competition with other males for access to a female. The toxin is produced in a special gland called the crural gland at the top of the leg. In August, when crural-gland development and spermatogenesis have reached their peak, the males exhibit maximum levels of aggression.

The toxin secreted by the crural gland is unquestionably powerful: an animal the size of a dog dies after being spurred, apparently from respiratory and cardiac arrest. Temple-Smith tested the sublethal effects of the toxin by injecting .05 milliliter of it into his forearm; his account confirms previous reports of the agonizing pain experienced by human beings who have been accidentally spurred by a male platypus.




RANGE of the platypus is limited to freshwater systems in eastern Australia. There the animals can be found—sometimes in relative abundance—both east and west of the Great Dividing Range, where they may inhabit streams with temperatures varying from near-freezing in the Australian alps to subtropical at the northern end of their range.

Courtship among platypuses is an elaborate ritual, involving various acrobatic activities in the water. Ronald Strahan and David E. Thomas of the Taronga Park Zoo have often seen a male grab the tail of a receptive female. The two then swim slowly in circles for several days, after which copulation takes place. Following copulation from one to three (usually two) yolk-laden eggs between four and 4.5 millimeters in diameter are shed into the left fallopian tube. (The right ovary in the platypus is nonfunctional.) It is in the fallopian tube that fertilization takes place and the egg acquires its first shell layer.

Sex determination takes place at the time of fertilization. The male is heterogametic, that is, he produces two kinds of sperm, as humans do, one with a *Y* chromosome and the other with an *X* chromosome. Carolyn E. Murtagh and Geoff Sharman of Macquarie University found that platypus and echidna males differ from all other mammals in that sex is determined by the presence of a multivalent *XY/XX* complex: at meiosis in the male the *X* and *Y* chromosomes are associated with small autosomal chromosomes, four unpaired and two paired.

After the first shell layer has been deposited, the egg passes to the uterus, where a second layer is added. There the egg increases in size, absorbing secretions from the uterine glands that nourish the developing embryo. When the egg has grown to 12 millimeters in diameter, a third and last layer of shell is added. The oval egg is now encased in a protective outer shell; when it is laid, it measures approximately 14 millimeters at its middle and 17 millimeters from end to end. Although the exact gestation period is not known, it is thought to span about 20 days.

A female prepares for the arrival of her eggs by constructing a special nesting burrow that is deeper and more elaborate than the normal burrow. The Australian naturalist David Fleay has seen a pregnant female carrying grass and leaves, which are tucked between the tail and the abdomen, to a terminal chamber in the nesting burrow, where the vegetation provides moisture and cushioning for the young.

Egg laying in the platypus has never been seen, but it is surmised that the female platypus sits on her rump so that the eggs, which are covered with a sticky secretion, pass directly from the cloaca to the abdomen. The female then curls her tail around them, incubating them with the warmth of her body. Even when the ambient temperature is as low as 20 degrees Celsius, the temperature of the incubatorium the female creates is maintained at 31.5 degrees C.

Although the incubation period is not known, it is probably similar to that of the echidna, which lasts for 10 or 11 days. The hatching process is also thought to be the same for both the platypus and the echidna. Their hatchlings are the same size and have a sharp egg tooth on the upper jaw (as do many reptiles), which enables them to tear open their rubbery shells. When an echidna hatches, it is only about 1.5 centimeters long and weighs about 380 milligrams. At this stage it looks much like a newborn marsupial: its forelimbs are well developed and its hind limbs are rudimentary.

The tiny suckling platypuses are held securely against the mother's abdomen by her tail, which she curls around them. Here they imbibe milk from her two nipples, which are covered with fur but are otherwise similar to the nipples of other mammals. The milk is produced in a pair of mammary glands of extraordinarily large size: in a female that measures 43 centimeters from snout to tail, the lactating glands may be almost one-third her length, or 13.5 centimeters long.

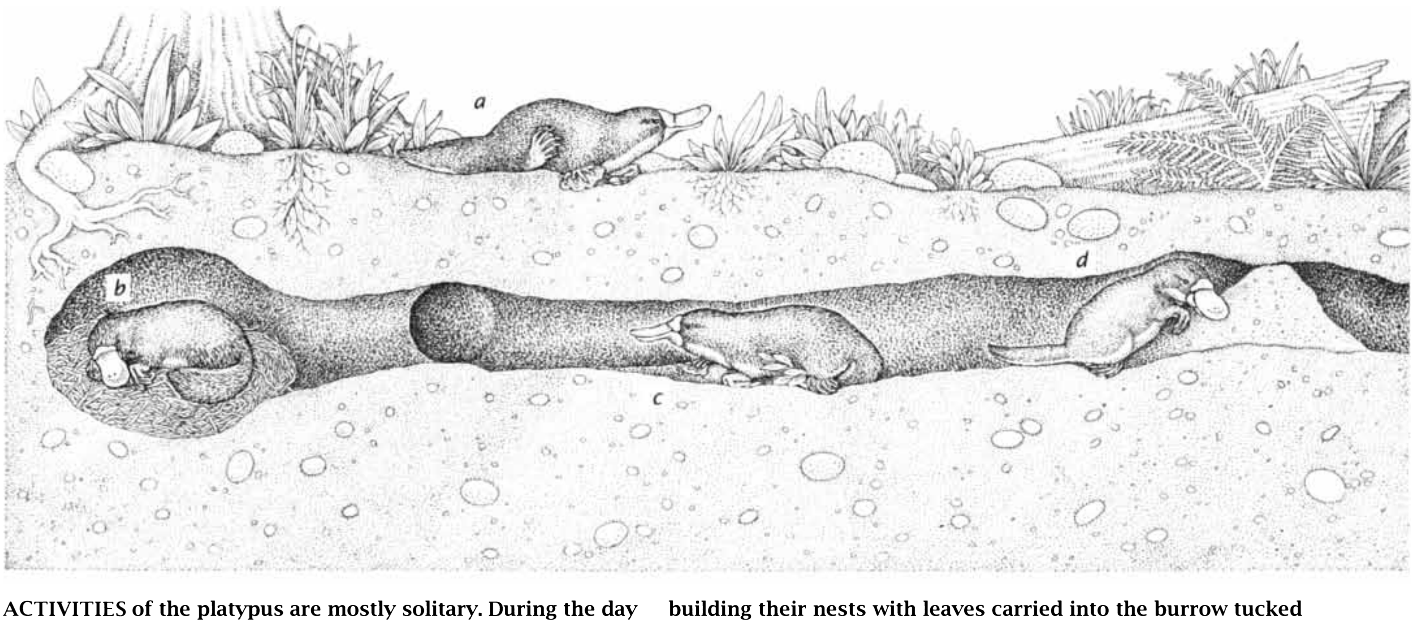
As early as 1831 Lieut. the Hon. Lauderdale Maule of the 39th Regiment of the British army demonstrated, with the help of soldiers stationed in New South Wales, that the platypus has mammary glands that produce true milk. Nevertheless, in 1959 *Physiological Review* published the statement that monotremes produce a fatty exudate that is licked up by the young from hairs covering pores in the skin—an erroneous finding often cited in popular accounts. Any question about the nature of milk production in the platypus was settled by research I carried out with my colleagues Gutta Schoefl and Carmel Teahan of the Australian National University, Michael Messer of the University of Sydney, Robert Gibson of Flinders University and Tom Grant. Our work demonstrates that the mammary glands of the platypus are as well developed and elaborate as they are in most mammals.

Each gland consists of a collection of club-shaped lobules, bound together by connective tissue, that converge to form a fan-shaped mass. As is the case for all mammary glands, the lobules are filled with innumerable alveoli, hollow sacs covered by a network of filamentous myoepithelial cells. A second layer of epithelial cells lines the inside of the alveoli, and it is these cells that secrete the various milk components. As in other milks, the components are casein (the primary protein), whey proteins, carbohydrates, fat globules and minerals.

In response to suckling at the teat, the pituitary gland secretes the hormone oxytocin, which then enters the bloodstream. When oxytocin reaches the mammary glands, it causes the myoepithelial network to contract, thereby raising the intra-alveolar pressure and forcing milk to the exterior. This phenomenon, known in the dairy industry as milk let-down, occurs in all mammals and can be artificially triggered with oxytocin injections. About five minutes after a platypus is injected there is a copious flow of milk, which can then be analyzed biochemically. Because only females that have laid eggs produce milk, studies of milk let-down provide a gauge of how many females in a given year are breeding. In this way investigators have determined that female platypuses do not breed every year.

Monotreme milk, like the milk of marsupials, differs from that of placental mammals in being particularly rich in iron. It contains from two to three milligrams per 100 grams; in contrast, the milk of humans and cows contains only .05 milligram of iron per 100 grams. Such high levels are apparently necessitated by the fact that the liver of newly hatched monotremes and marsupials is much too small to store an adequate supply of iron. During the months when their sole sustenance is milk, an infant must therefore get the iron it needs from its mother.

After being weaned (limited data suggest that this occurs around the age of three and a half months), a young platypus begins diving for food much as adults do, snapping it up from the bottom and storing it in one of its two cheek pouches. When both pouches are full, the platypus swims to the surface, where it trans-




**ACTIVITIES of the platypus are mostly solitary. During the day platypuses live in burrows along the water's edge. Sometimes they bask in the sun, grooming themselves with their hind feet (*a*). Females guard their eggs and young in nesting chambers (*b*), building their nests with leaves carried into the burrow tucked between the tail and the abdomen (*c*). For protection from predators, platypuses often plug the end of their burrows with a mound of dirt (*d*). At night they are active and take to the water,**

fers the pouch contents to its mouth, grinding the prey between horny pads in its upper and lower jaws.

By removing the contents of the cheek pouch with a long spoon, Robert Faragher and a team from the New South Wales State Fisheries department found that the platypus varies its diet according to the season. In mid-December it feeds primarily on insect nymphs and larvae that live at the bottom of streams and lakes. Of these, caddisfly larvae are by far the commonest. As the year progresses, the percentage of caddisflies in the diet drops and the platypus supplements its food supply with small shrimplike crustaceans.

The ability of the platypus to locate its prey in murky water—while avoiding all kinds of obstacles—is truly remarkable. During the minute and a half it remains underwater the platypus is essentially cut off from the outside world: its eyes and ears are situated in a muscular groove, the lips of which close tightly, and its nostrils, at the dorsal surface of the end of the snout, also close. How under these circumstances does the animal find its prey?

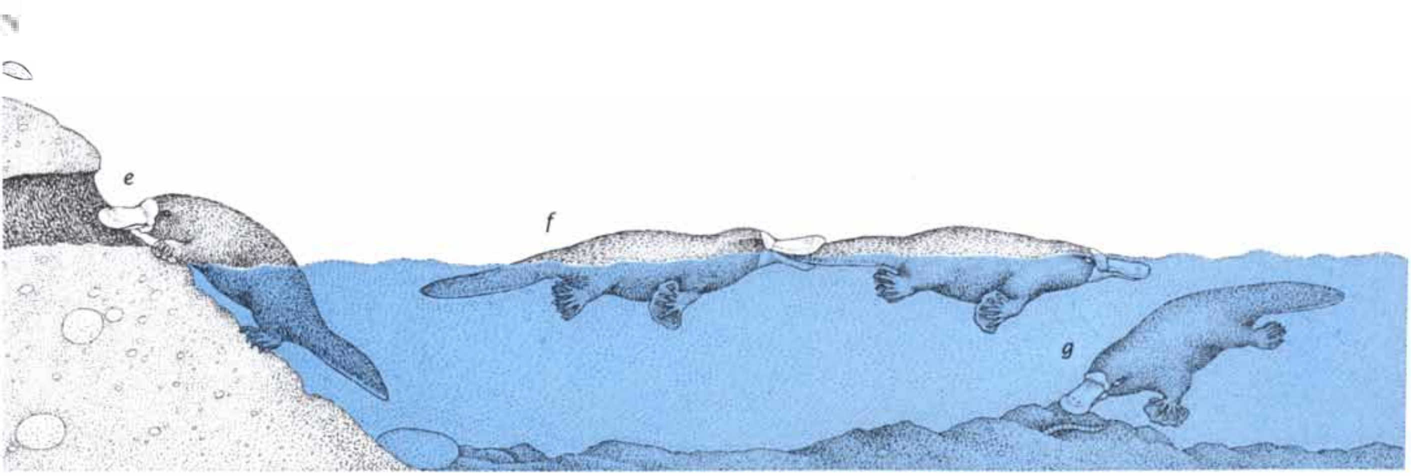
Recent investigations of the feeding behavior of the platypus have led to some astonishing findings. As long ago as 1884 it was known that the exquisitely soft surface of the bill is perforated with openings that contain sensitive nerve endings. Neurophysiological studies by Ros C. Bohringer of the University of New South Wales, Hennig Scheich of the Technical University in Darmstadt, West Germany, and Uwe Proske of Monash University revealed not long ago that within these pores there are two kinds of sensory receptors: mechanoreceptors, which are tiny pushrods that respond to tactile pressure, and electroreceptors, which respond to weak electrical fields.

Bohringer and her colleague Mark Rowe discovered that tactile stimulation of the pushrods sends nerve impulses to the brain, where they evoke an electric potential over an area of the neocortex much larger than the one stimulated by input from the limbs, eyes and ears. They concluded that the bill must be the primary sensory organ for the platypus. That finding was supported by studies showing the bill is extraordinarily sensitive to tactile stimulation: stimulus with a fine glass stylus sent a signal by way of the fifth cranial nerve to the neocortex and from there to the motor cortex. Presumably nerve impulses from the motor cortex then induced a snapping movement of the bill. But it was still not clear how the animal locates its prey at a distance.

Scheich and his colleagues carried out experiments showing that when a platypus feeds, it swims along, steadily wagging its bill from side to side at the rate of two or three sweeps per second, patrolling in this manner until prey is encountered. It thereupon switches to searching behavior, characterized by erratic movements of the bill over a small area at the bottom, which is followed by homing in on the object and seizing it with a rapid snap of the jaws.

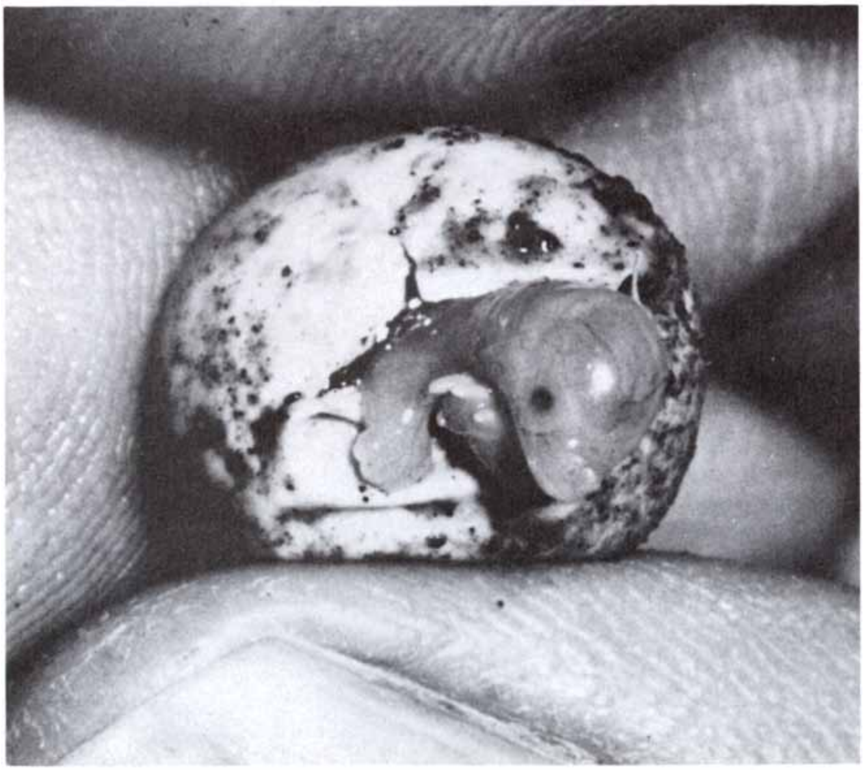
These experiments show that the animal clearly orients to objects in the water, but how it accomplishes that task was not known. How was it able to distinguish a suitable prey item from other objects on the bottom? A series of electrophysiological experiments provided the answer.

Scheich and his colleagues hypothesized that a sensory system based on electroreception as is found in some fishes and sharks might exist in the platypus. Experiments showed this to be true. The investigators found they could trigger the switch from patrolling to searching behavior in the platypus by creating a dipole electric field in the water with the aid of a small 1.5-volt battery. The platypus, sensitive to the weak electric current that was created, rapidly oriented toward the battery at a distance of 10 centimeters and sometimes as much as 30 centimeters. Once the battery was detected, the platypus would inevitably attack it as if it were food. Investigators then discovered that the tail flicks of freshwater shrimp also produce weak electric fields and elicit an identical response. The assumption is that all




**where they feed and mate. After foraging a platypus may return to its burrow to rest (*e*). Courtship is an elaborate ritual and frequently involves tandem swimming: a male grabs a female by her tail and the two circle about for several days before mating (*f*); feeding takes place mostly on the bottom, where prey is detected by receptors on the platypus's bill (*g*). This illustration is a composite of behaviors; it is not meant to suggest that the animal is communal or that the activities occur at the same time.**

**ECHIDNA *Tachyglossus aculeatus,* also known as a spiny anteater, is one of three kinds of living monotremes, or egg-laying mammals, endemic to Australia and New Guinea. Unlike the platypus, the echidna is terrestrial, feeding mostly on termites and ants, which it extracts from their nests with its long, sticky tongue. Interspersed with its fur are barbless spines, which give it a superficial resemblance to the hedgehog.**



**MONOTREMES, in this case an echidna, emerge from their eggs after about 10 days of incubation. The hatchlings have a sharp tooth at the tip of the upper jaw with which they tear open the rubbery eggshell. A similar egg tooth also occurs in some reptiles.**

the invertebrates on which the platypus feeds must do so too.

The means by which the platypus detects weak electric currents remained a mystery, however. When Scheich and his team exposed the platypus bill to an electric stimulus, it became apparent from evoked-potential experiments that the region of the brain being stimulated was an area of the neocortex adjacent to the area stimulated by the mechanoreceptors. Moreover, it was found that maximum sensitivity was recorded at some but not all of the pores in the skin of the bill.

Proske and his team discovered that receptors inside those pores were insensitive to probing with a glass stylus but responded to low-level electrical stimuli. Conversely, the mechanoreceptors, which were extremely sensitive to touch, responded electrically only when an electric current three orders of magnitude larger was applied.

Detailed examination of the pores in cross section revealed that special nerve endings sensitive to electric currents are indeed present and that they occur in association with mucus-secreting glands. The lubrication provided by these glands apparently ensures that the electroreceptors will not desiccate when the platypus is holed up in its burrow.

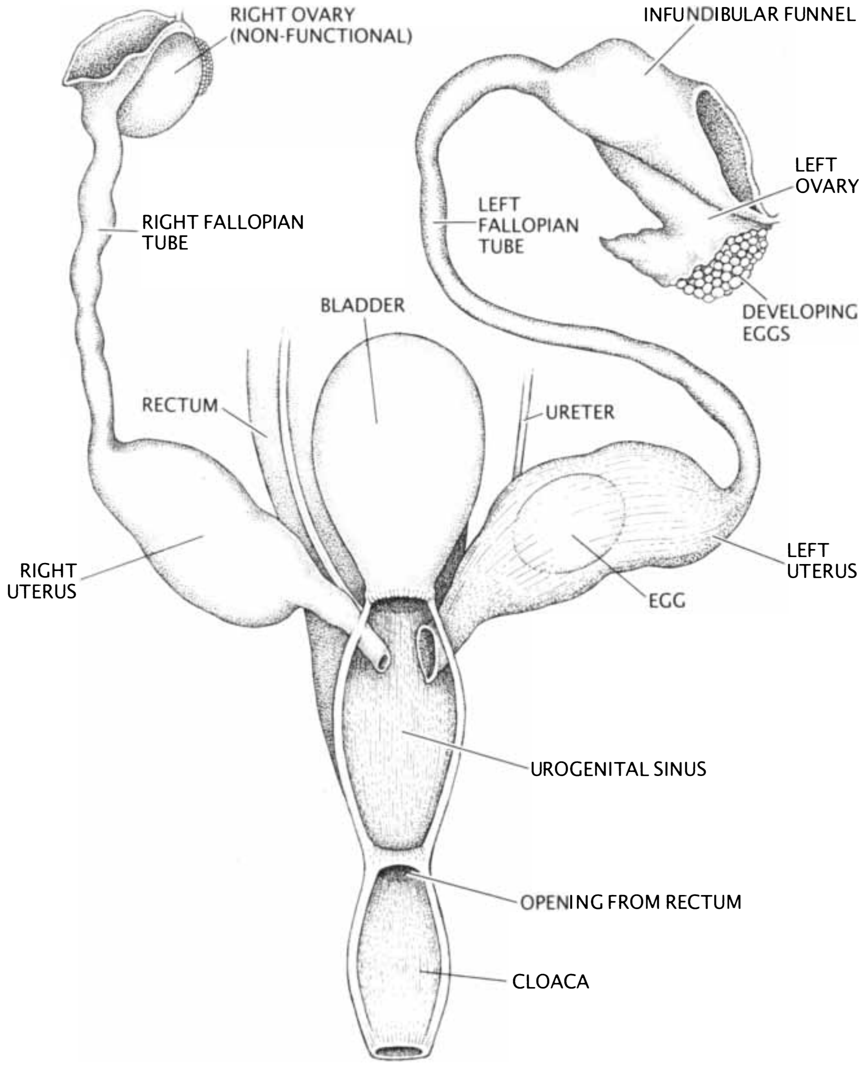
Another remarkable specialization of the platypus is its ability to thermoregulate, which it is able to do more effectively than some placental mammals. Grant and his colleagues have found that the platypus, in spite of its similarities to cold-blooded reptiles, is an excellent thermoregulator under both laboratory and natural conditions. When the platypus forages for food in winter, it must contend with water that may drop to zero degrees C., yet, like any mammal, it maintains its normal body temperature (which in the platypus is 32 degrees) by adjusting its metabolic rate. When the temperature of the air drops from 25 to five degrees, the platypus increases its standard metabolic rate (SMR) 1.8 times; in water that is five degrees it increases its SMR by a factor of 3.2.

The finding that the platypus can maintain an internal body temperature of 32 degrees for several hours even in zero-degree water was a surprising one. That it can do so is related both to its ability to raise its SMR and to the insulative value of its fur, which is higher than that of most mammals. Polar bear and beaver fur, for example, lose from 90 to 96 percent of their insulative value in water; in contrast, platypus fur loses only from 60 to 70 percent of its insulative value. Amphibious placental mammals such as the muskrat *Ondatra zibethicus* and the Australian water rat, *Hydromys chrysogaster,* are poor thermoregulators by comparison. In water of five degrees the water rat's internal temperature drops in a matter of one or two hours from 37 to 28 degrees.

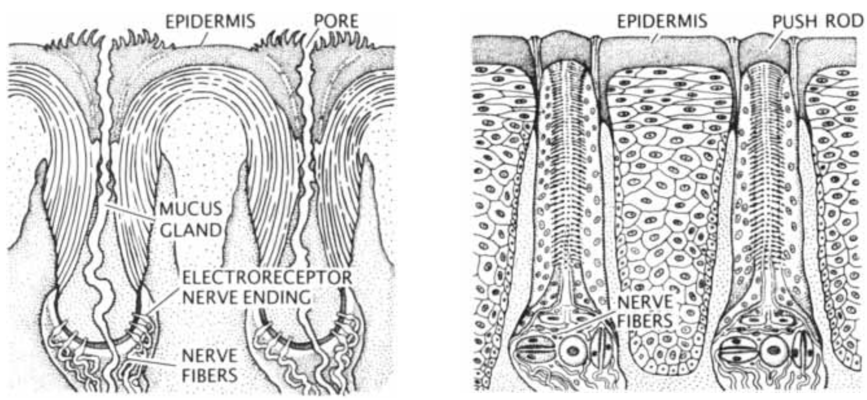
In spite of all that is now known about the platypus, one mystery that remains unsolved is its antiquity. The animal's weird mixture of reptilian and mammalian traits suggests a long evolutionary history, perhaps as long as 150 million years. The fossil record on the animal is notoriously poor, however. That situation can be attributed at least in part to the lack of teeth in the adult platypus, since most mammals are represented in the fossil record on the basis of teeth and jawbones.

In recent years, though, some intriguing fossil remains have been found. First Michael Archer and his colleagues at the University of New South Wales discovered a beautifully preserved skull of an adult platypus in a Miocene deposit some 15 million years old in northwest Queensland. The skull, *Obdurodon dicksoni,* is particularly interesting because it has teeth similar in some respects to those of juvenile platypuses. Such a finding suggests that loss of teeth in the group must have been a relatively recent event. In 1985 a piece of lower jaw containing three molar teeth similar to those of *Obdurodon* was found by Archer, Timothy F. Flannery, Alex Ritchie and Ralph E. Molner in a deposit thought to be early Cretaceous in age, roughly 100 million years old. The fossil, named *Steropodon galmani,* is significant because it sets a minimum date for the age of the monotremes and is the first proof that mammals existed in Australia long before the Tertiary.

Many questions about the platypus and its habits remain unanswered. Although research on the animal continues (studies are currently under way on the biochemistry of lactation and the endocrinology of reproduction), it is now clear—in spite of its antiquity and its mosaic of reptilian and mammalian characteristics—that the platypus is a highly specialized animal, well adapted to its amphibious existence in the waterways of eastern Australia.




**REPRODUCTIVE SYSTEM of monotremes is unique among mammals: its anatomy is more reptilian than it is mammalian. Eggs are shed by the left ovary (the right ovary is nonfunctional) into the fallopian tube, where they are fertilized. They then pass to the uterus, where they acquire a rubbery shell, and they are eventually laid through the cloaca. As in reptiles, both the rectum and the bladder open into the cloaca; all other placentals perform excretory and reproductive functions through separate openings.**




**SENSORY RECEPTORS in the bill of the platypus enable it to navigate and to locate prey underwater. Electroreceptors (*left*) respond to weak electric currents generated by the movement of small bottom-dwelling invertebrates; when electroreceptors are stimulated, the animal swims toward the signal. Tactile stimulation by the prey causes the mechanoreceptors (*right*) to fire; the pushrod is deflected and a nerve impulse is sent to the brain, resulting in a snapping movement of the bill and capture of the prey.**

# Aerogels

*As the name implies, these unusual solids are remarkably light and porous, made more of air than of gel. Their structure endows them with intriguing properties that science has only begun to exploit*

by Jochen Fricke

A jellyfish that has washed up on the shore dries in the sun and leaves very little collapsed tissue behind. Like the jellyfish, most gelatinous substances will shrink to 10 percent of their original volume when they dry in the air. In spite of the shrinkage such air-dried gels are still about 50 percent empty space. The gels are riddled with pores that make them half as dense as a comparable nonporous solid.

In the early 1930's S. S. Kistler of Stanford University found a way to dry gels without shrinking them. By extracting the fluid from a wet gel under pressure and at a high temperature he was able to produce extremely light materials with porosities as high as 98 percent. Kistler called the solids aerogels. The first ones he produced were translucent slabs made from silica, the major component of ordinary glass. Later Kistler succeeded in fabricating aerogels of alumina, tungsten, iron and tin oxides, nickel tartrate, cellulose, gelatin, egg white and rubber. He saw no reason the list could not go on indefinitely.

Thirty years later a research group in France demonstrated that aerogels could serve as components of particle detectors in high-energy physics experiments. The French team also devised a new way to make the gels that took hours rather than weeks. Since then aerogels have attracted a small but rapidly growing scientific community whose interest is drawn by the unique properties of these curious materials as well as by their potential applications.

As Kistler had noted, the thermal conductivity of aerogel tiles is about 100 times smaller than that of fully dense silica glass. The combination of this excellent insulating ability with high transparency suggests that aerogels might act as superinsulating spacers in window systems or as shields to trap solar energy in passive collectors. Furthermore, recent work in my laboratory at the University of Würzburg has revealed that sound moves through aerogels even more slowly than it does through air. This property could tremendously improve the efficiency of devices that emit ultrasonic waves in order to gauge distances, such as the autofocus systems in cameras.

Most of the acoustic and mechanical properties aerogels exhibit can be related to the intriguing structure of the delicate gel body. We are just beginning to explore how this structure can be altered in the course of synthesis to optimize the performance of aerogels in practical systems. Before discussing the application of aerogels, I shall therefore describe the art of making them.

Kistler's aerogels started out as a solution of sodium silicate called water glass that, when mixed with hydrochloric acid, gelled in about 24 hours. Kistler rinsed the acid from the gel with water to yield what is known as an aquagel, and then submerged the gel body in alcohol to make an "alcogel."

Kistler knew that as soon as liquid begins to evaporate from a gel, surface tension creates concave menisci in the gel's pores. A meniscus is the crescent-shaped contour on the surface of a liquid contained in a column such as a test tube. As evaporation continues, the menisci retreat into the gel body, compressive forces build up around its perimeter and it contracts. Eventually surface tension causes the collapse of the gel body, just as it does in the case of the stranded jellyfish [*see top illustration on page 94*].

In order to prevent surface tension from building up, Kistler dried the gel in a pressure vessel called an autoclave. When the temperature and pressure in the vessel are increased above a critical threshold, the liquid is transformed into a "supercritical" fluid in which every molecule can move about freely and surface tensions cease. Without surface tension menisci do not form. The fluid is slowly released from the autoclave, leaving the gel body behind. Finally, the temperature in the autoclave is lowered to the ambient temperature and the aerogel is removed.

The critical parameters governing this drying procedure depend on the liquid employed. A gel wetted with methanol, for instance, requires temperatures of about 240 degrees Celsius (water boils at 100 degrees C.) and pressures of about 81 bars (comparable to the pressure exerted by 250 feet of water). The gel body does not experience any compressive forces, and so it retains its original size and shape. The silica gels Kistler produced had densities of between 30 and 300 kilograms per cubic meter, corresponding to porosities of between 86 and 98 percent.

Although Kistler succeeded in drying gels without shrinking them, his method of preparation is time-consuming and tedious. In particular, washing the aquagel and then changing the solvent to alcohol is a laborious business. Yet the gel has to be bathed in alcohol because water would dissolve the gel structure at high temperatures. When in the early 1960's Stanislas J. Teichner and one of his students at the University of Lyons tried to duplicate Kistler's results, it took them weeks to prepare just two aerogel samples.

The student, whose thesis rested on the aerogel research, was thoroughly dismayed at the prospect of spending years procuring the material he needed. He ended up having a nervous breakdown. Soon after he recovered, a new method for gel preparation was introduced that took only one day. The technique became the standard one.

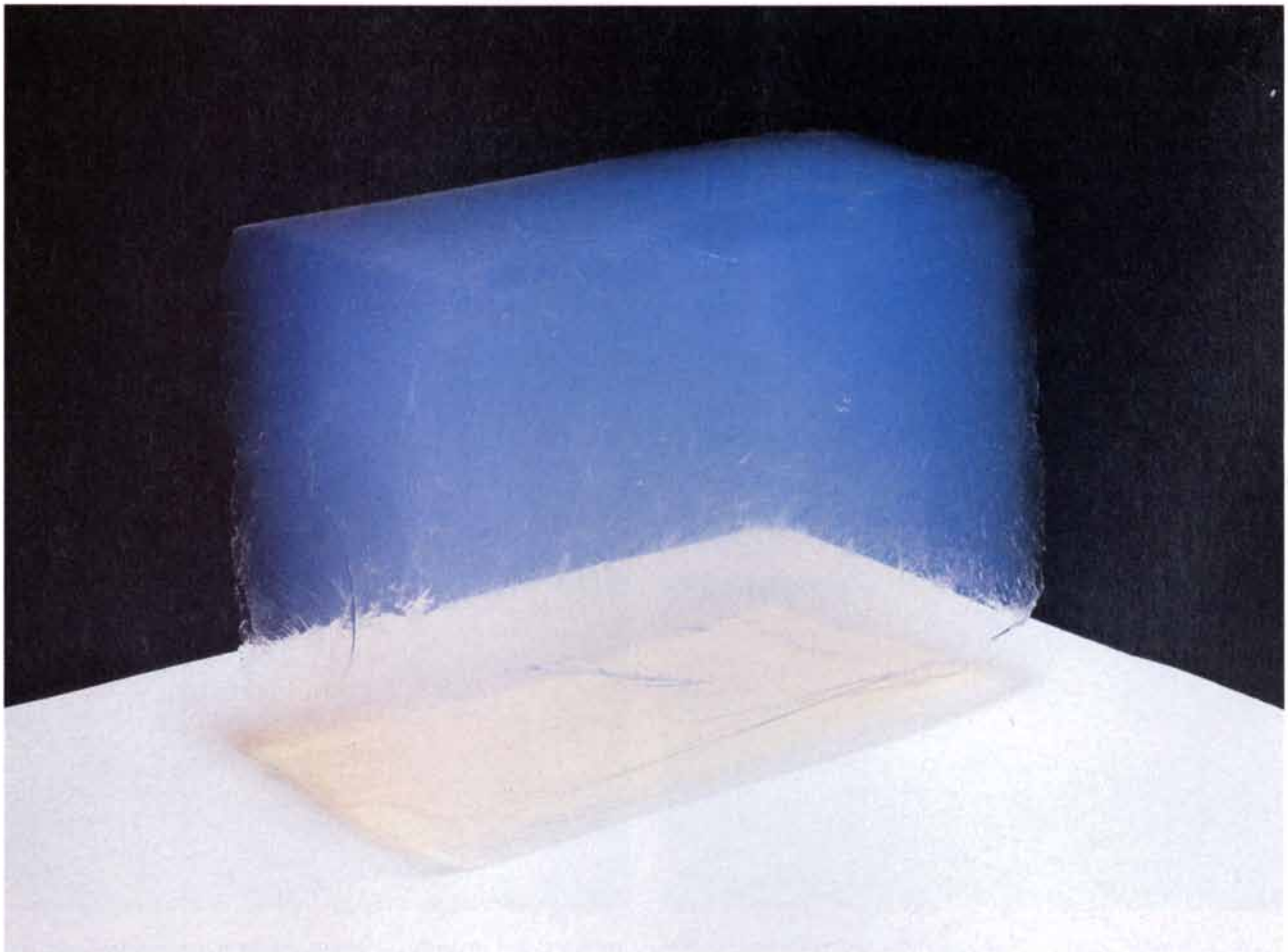
To make a silica aerogel according to Teichner's method, a colorless liquid chemical called tetramethoxysilane (TMOS) is mixed with water; when an acid or base catalyst is added, silicic acid and methanol are produced. Molecules of the silicic acid start to aggregate and to form microscopic clusters. These clusters in turn stick together to form larger entities, finally building up the coherent gel body [*see illustration on page 95*]. The structure and density of the resulting aerogels can be influenced by the *p*H value, the temperature and the concentration of the ingredients employed in the synthesis, which is called a sol-gel process.

Other methods of aerogel synthesis have been explored. A research group at the University of California at Berkeley makes pieces of aerogel from tetraethoxysilane (TEOS) instead of the poisonous TMOS. The Berkeley group, which is led by Arlon J. Hunt, also substitutes carbon dioxide for the alcohol that is traditionally used as the fluid phase. Carbon dioxide is nonexplosive and needs to be heated only above 31 degrees C. during drying. That methodology is much slower than the French technique, however, because the exchange between alcohol and carbon dioxide takes a long time.

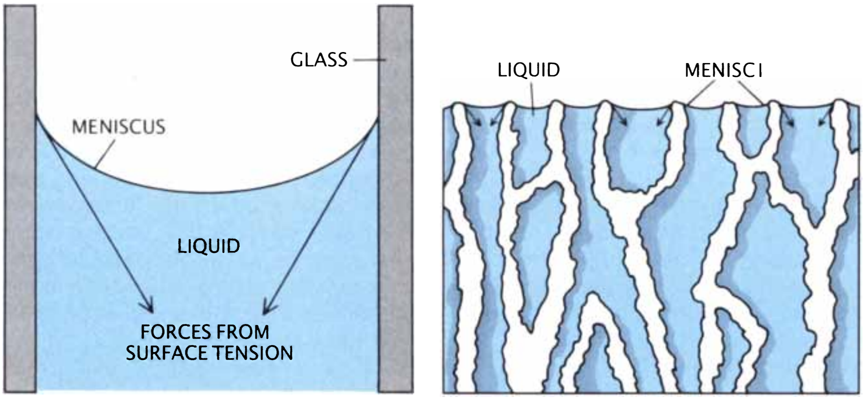
At B.A.S.F. in Ludwigshafen, West Germany, pellets rather than tiles of aerogels are fabricated by spraying an acid and water-glass solution from a mixing jet into a flask, where the pellets are washed and then supercritically dried. The resulting grains have diameters of a few millimeters and densities of some 200 kilograms per cubic meter. The granular aerogels are less expensive to manufacture than monolithic tiles.

The French group's incentive for reviving research on aerogels came from a government initiative to find porous materials with which to store oxygen and rocket propellants. The subsequent development of aerogels, however, was most strongly promoted by their utility in detectors of Cerenkov radiation. Cerenkov radiation is emitted by charged particles such as pions, muons and protons moving in a medium at a velocity greater than the speed of light in that medium. Such particles give off light in a cone-shaped front surrounding their flight path. This electrical shock wave is comparable to the acoustic shock front around an airplane flying at supersonic speed.

The velocity of the particle can be deduced from the angle of the cone; that is, the angle between the wavefront and the particle's flight path. The speed of light in a given medium is determined by the medium's index of refraction, and the index of refraction for aerogels happens to be in the range that can be covered neither



**AEROGEL FRAGMENT is translucent. On a white background it has a pale yellow tint; against a dark background it looks milky blue. The low-density structure that gives rise to these optical effects also slows sound waves and enhances thermal insulation.**




**SURFACE TENSION shrinks gels when they dry in the air. The same tension creates a meniscus on the surface of a liquid contained in a tube (*left*). Compressive forces exert a powerful inward pull (*arrows*). Menisci also form in the pores of a gel bathed in liquid (*right*). As the liquid evaporates, the menisci retreat and draw the gel body inward.**

by compressed gases nor by liquids. Aerogels have already been incorporated in several Cerenkov detectors worldwide. Indeed, this is currently the widest practical application of the material.

Two groups adopted Teichner's method to produce aerogel tiles for such detectors. Günter Poelz and his co-workers at the Deutsches Elektronen-Synchrotron (DESY) in Hamburg supplied a detector there with about 1,700 liters of silica aerogel. At the University of Lund in Sweden, Sten Henning and Guy v. Dardel generated about 1,000 liters of aerogel for an experiment at the European laboratory for particle physics (CERN). To make large 60-by-60-centimeter tiles, the Swedish group set up a facility with a 3,000-liter autoclave. In 1984 the main gasket of the autoclave failed and more than 1,000 liters of methanol leaked out into the building. The explosion that ensued destroyed the entire facility. Undaunted, the Swedish team built a new facility with an equally voluminous autoclave and expects to resume production soon.

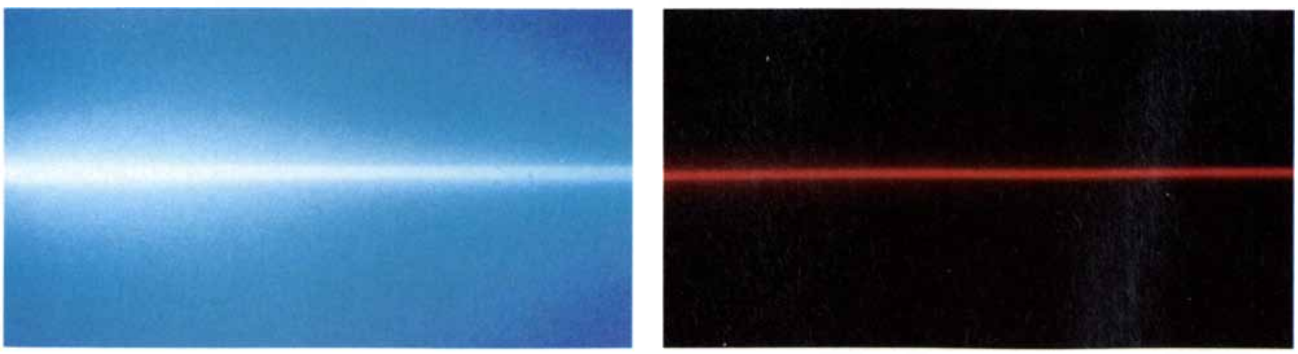
Today the main goal of aerogel research is to produce large amounts of tiles or pellets cheaply for use as spacers in windows and in passive solar collectors. The utility of aerogels in this context rests on their excellent thermal insulation, which is obviously a result of their high porosity. The source of this porosity can best be visualized if I address the unusual structure of the gels.

The first clues to the size of the structural entities in aerogels came from evidence of Rayleigh scattering, the phenomenon responsible for the scattering of sunlight in the atmosphere. Rayleigh scattering occurs in a medium when it contains inhomogeneities that are smaller than the wavelengths of visible light. The inhomogeneities deflect rays of light; the efficiency of scattering decreases dramatically as the wavelength of the light increases. Consequently red light, which has wavelengths near 700 nanometers, is scattered only about one-tenth as strongly as blue light, which has wavelengths of approximately 400 nanometers (a nanometer is a billionth of a meter).

That is why the setting sun, seen through the atmosphere, looks yellow or red, and why at the sun's zenith the sky looks blue. Similarly, aerogel tiles show a yellowish hue when they are held against the light and a bluish tint when they are held against a dark background. Hence it appears that the structural entities in aerogels are smaller than the wavelengths of visible light; in fact, from precise measurements of these optical effects one can deduce that the largest entities are roughly 50 nanometers in size.

The microscopic technique that is most reliable for structural analysis of aerogels involves bouncing X rays or neutrons off the gel body at very small angles. This technique makes it possible to probe aerogel structure from atomic levels to the 50-nanometer range. It is also possible to glean information on structure by observing the influence of gas pressure on the material's thermal properties. Data from these various methods are not always easy to interpret. The composite model that investigators have been able to contruct for base-catalyzed aerogels is represented by spheres, clusters and chains [*see illustration on opposite page*].

In silica aerogels made from TMOS under basic conditions, the densest structures are nonporous spheres with diameters of less than one nanometer. These "primary particles" are as dense as ordinary silica glass: about 2,000 kilograms per cubic meter. The primary particles cluster together and form porous "secondary



**RAYLEIGH SCATTERING in aerogel tiles gives clues to the material's structure. It occurs when light, generated in this case by a laser beam, passes through a medium containing inhomogeneities that are smaller than the wavelengths of visible light. The particles that make up aerogels are of a size that causes the short wavelengths of blue light (*left*) to scatter much more than the longer wavelengths of red light (*right*). These photographs were made by Wolfram Görlitz of the University of Würzburg.**

particles" with diameters of about two nanometers. These clusters have half the density of silica glass; their surface areas average a remarkable 600 to 800 square meters per gram. Given these figures, one should not be surprised at the large amounts of moisture aerogels can absorb.

The delicate aerogel skeleton is built up between two and 50 nanometers. As far as anyone can tell, it consists of chains that branch and contact one another to create the exceptionally porous macrostructure. The degree of branching and of interchain contact depends on the aerogel's density and probably accounts for most differences in optical, thermal and acoustic properties.

I have already touched on the fascinating optical properties of aerogels. The basics of their thermal properties were established by Kistler's early work and—more than 20 years ago—by the investigations of M. G. Kaganer of the All-Union Scientific Research Institute of Oxygen, Cryogenic and Compressor Machinery Construction in Moscow. In the past five years my thermal physics research group at Würzburg has provided more detailed analyses.

The thermal losses typical of an insulating system are measured in terms of a quantity called the thermal loss coefficient. This coefficient gives the power in watts that penetrates an area of the material measuring one meter by one meter when the temperature difference across the material is equal to one degree C. In my group Dietrich Büttner, Peter Scheuerpflug and Ulrich Heinemann found that an evacuated layer of low-density aerogel only 15 millimeters thick has a thermal loss coefficient of about .5 watt per square meter and degree C. for temperatures between 10 and 20 degrees C. In comparison, a single pane of ordinary glass would have a thermal loss coefficient of approximately six watts per square meter and degree C.

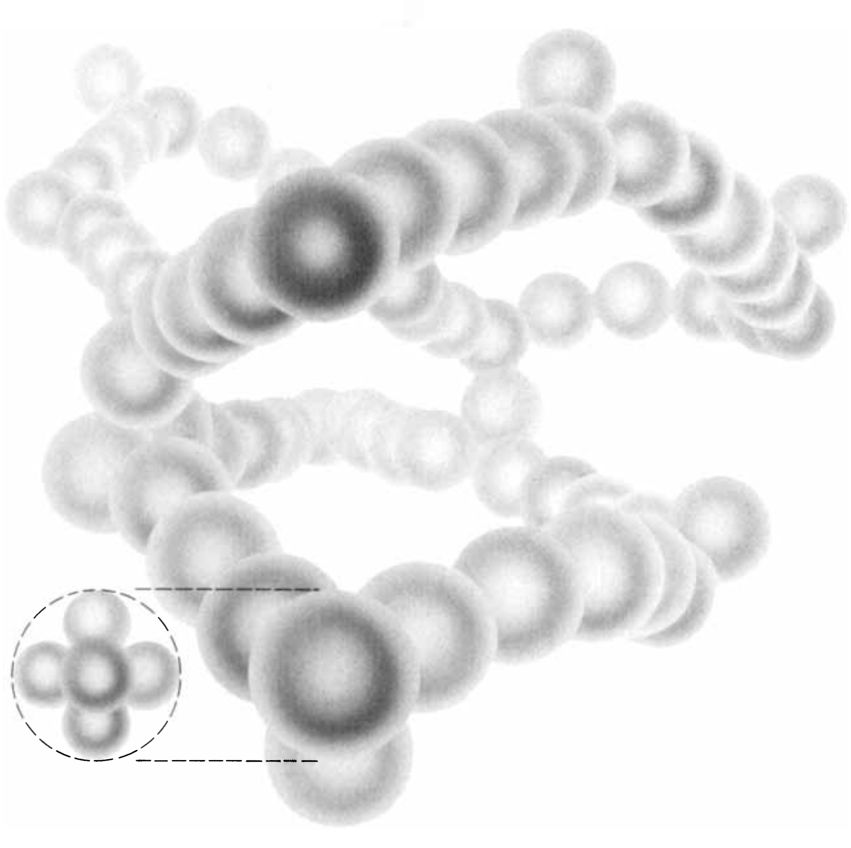
Even the best available window systems, made of double-pane, silver-coated glass and filled with argon gas, experience thermal losses three times as high as systems incorporating evacuated aerogel spacers would. Hence aerogel layers in windows could greatly reduce heating bills in cold countries, particularly in regions where there is ample sunshine in the winter. Windows with aerogel spacers would look slightly frosted and so would not replace ordinary windows universally. The fact that in a couple of years several glass manufacturers expect to produce aerogel window systems commercially attests to the savings these systems could yield, even with limited application.

In a similar way aerogels could improve the passive use of solar energy. A blackened house wall, thermally protected by a layer of aerogel, could trap the heat provided by solar radiation. Because the aerogel coat insulates so well, most of the heat would pass through the wall and into the house [*see illustration on next page*]. The insulating wall would appear to be gray because of reflected and scattered light and would be effective with eastern, southern or western exposure.

In order to optimize the performance of aerogels in such insulating and solar-collection systems, the factors influencing thermal losses must be taken into account. The thermal loss associated with a material depends on the efficiency of thermal transport within the material. In evacuated, disperse mediums such as aerogels, heat can be transported in two ways. First there is thermal conduction along the contacting chains in the delicate aerogel structure. This contribution lessens with increasing porosity (or decreasing density), but it shows only a weak temperature dependence. The other means of transfer, however, is strongly dependent on temperature: it is transfer by thermal radiation.

Thermal radiation in the form of photons emanates from all warm bodies and is absorbed or scattered by surfaces obstructing the photons' paths. With rising temperatures the spectrum of thermal radiation gets shifted toward shorter wavelengths: very hot bodies such as the sun emit photons with wavelengths in the visible spectrum, whereas walls, windows and the human body radiate photons with infrared wavelengths measuring roughly from five to 20 micrometers (a micrometer is a millionth of a meter).

Silica aerogels can absorb infrared radiation from walls and windows



**PRESUMPTIVE AEROGEL STRUCTURE is built up from molecular clusters. The smallest, "primary particles," are spheres less than one nanometer (one billionth of a meter) across; they have the same density as ordinary, nonporous glass. These spheres cluster together into so-called secondary particles, whose density is about half that of glass. The secondary particles are strung out in chains to create the delicate and extremely porous aerogel skeleton, whose density is typically 5 percent that of ordinary glass.**

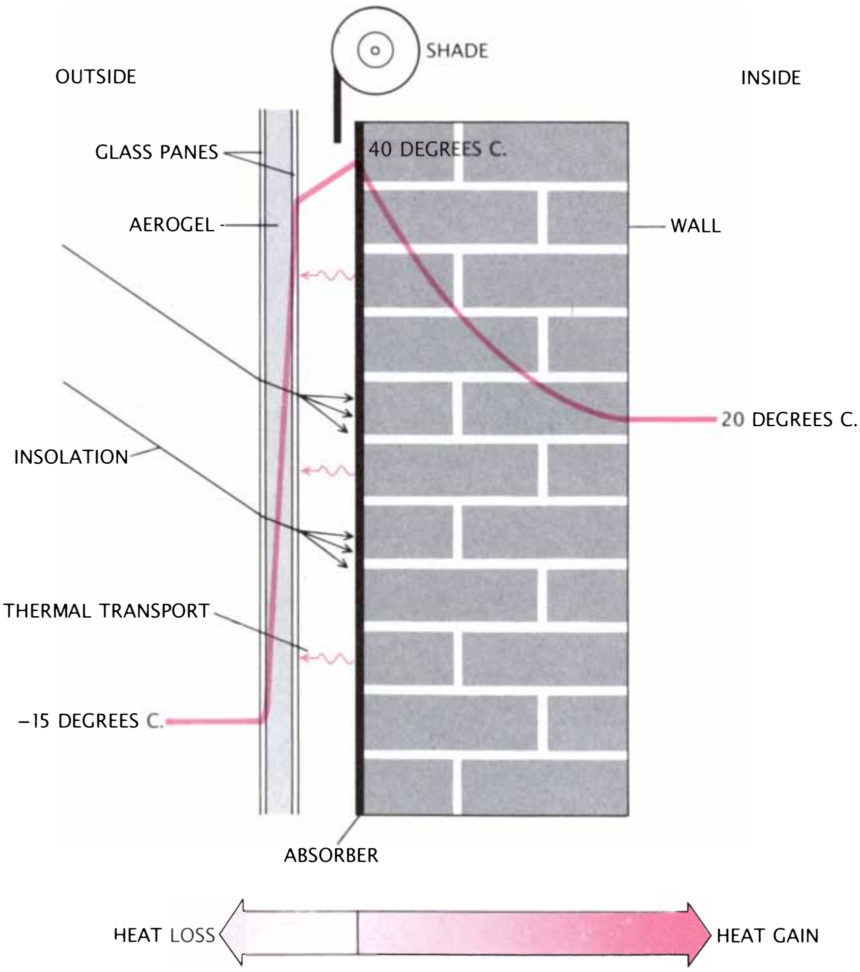
quite well at wavelengths above five micrometers and are therefore able to block thermal radiation effectively at ambient temperatures. Between three and five micrometers, however, their absorption is low. Consequently the insulating properties of aerogels deteriorate as temperatures rise and increasingly more heat can penetrate—and escape—the aerogel layer. This phenomenon was investigated in detail by Roland Caps in my research group.

In window spacers the problem of radiation leakage can be reduced by lining one of the inner sides of the two glass panes with a "low emissivity" layer of silver. Such layers, which are already employed in double-pane, argon-gas-filled window systems, emit about 10 times less thermal radiation than normal glass surfaces.

The variation of the thermal loss coefficient with gas pressure might likewise pose a problem to would-be manufacturers of aerogel window systems. With aerogel tiles, thermal losses rise significantly only above 100 millibars; a window system evacuated to one millibar at production could maintain a vacuum below that threshold for 30 years if the leak rate were less than 100 micrograms of air per day per square meter of glass area. Such a leak rate might be attained with durable, high-quality organic rim seals such as those used in argon-gas-filled systems.

The thermal loss coefficient of aerogel pellets shows a much greater sensitivity to air pressure. The coefficient escalates considerably at pressures as low as a tenth of a millibar. Windows with pellet spacers would therefore probably have to be sealed with metal or glass rims, which are only in the development stage. Pellet spacers filled with argon gas are simpler to make but do not insulate as well as evacuated pellet layers.

Finally, before aerogel layers can be implemented in insulating systems the problem of "creeping" will have to be overcome. Creeping is the tendency of aerogels to change their shape when they are subjected to a constant external load. Joachim Gross in my group discovered and is exploring this phenomenon, which probably is related to the water adsorbed on the expansive internal surface of the material. It might be possible to minimize the effects of creeping by baking off adsorbed water after the aerogel is made.




**TRANSPARENT INSULATION is provided by an aerogel layer sandwiched between two panes of glass. Such insulation on the wall of a house could increase dramatically the feasibility of passive solar collection. The wall is covered with a black absorptive material that converts the sun's rays into heat. Because aerogel is a much better insulator than the wall, most of the heat generated penetrates the wall and enters the house instead of escaping to the outside. In summer a shade could be drawn over the absorber.**

My group at Würzburg has also done extensive studies on the acoustic properties of aerogels. Our first measurements of sound velocity in aerogels were carried out about four years ago. We sent pulses of ultrasonic waves through aerogel tiles using a piezoelectric transducer, a device that vibrates when an alternating electric field is applied to it. When we first examined our results, we hesitated to believe what we saw: sound travels in aerogels at speeds of from 100 to 300 meters per second! Although such low velocities are typical for materials that are easily compressed, such as rubber, they are surprising in a material made of the same components as ordinary glass, which can transmit sound waves at speeds of 5,000 meters per second.

Even more remarkable is Gross's discovery that the velocity of sound decreases in aerogels when they are subjected to mechanical stress. Most solids get stiffer when they are compressed, and so sound travels at greater speeds. Aerogels instead become softer, due perhaps to changes in the skeleton's conformation.

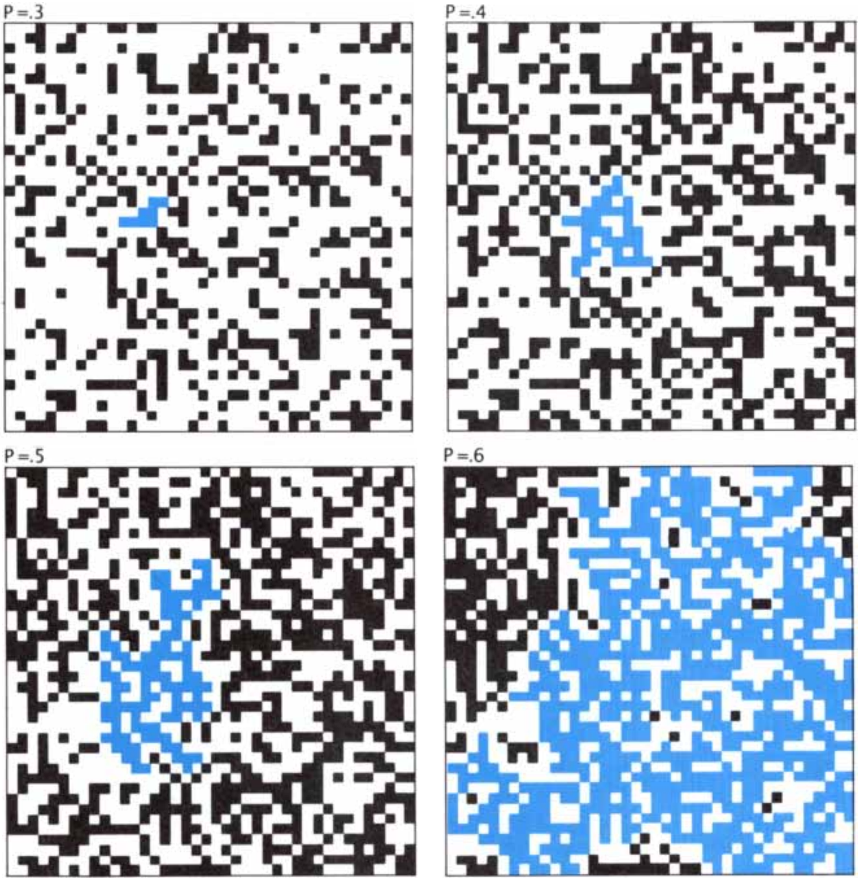
The combination of low sound velocity and low density means that in aerogels the product of the two quantities, called the acoustic impedance, will also be low. In fact, the impedance of aerogels is intermediate between the impedance of ceramic piezoelectric transducers and that of air. As a consequence aerogels could help to boost the efficiency of such transducers, which are common components of cameras and of robotic systems that gauge distance by emitting ultrasonic waves.

A thin aerogel layer spread on the surface of the transducer could act as a buffer between the drastically different impedance values of air and transducer, thereby reducing the reflection of sound waves. We estimate that aerogel layers could increase by a factor of 100 the ultrasonic energy emitted by piezoelectric devices. So far, however, aerogels have not been implemented in these devices, and so their actual utility has not been ascertained.

I do not want to end this article without mentioning two other interesting ramifications of aerogel structure. The first is a phenomenon in aerogels called percolation. The expression, from the Latin word meaning "to flow through," can refer to water dripping through ground coffee or to an electric current flowing through a highly porous conducting body. In an aerogel system it refers to thermal conduction by the gel body. In order for percolation to occur the gel body must be coherent; during the sol-gel process there is a certain critical point at which the density of molecular clusters is sufficient to form a coherent body.

This "percolation threshold" can be estimated with the help of computer simulations that build simple chessboard patterns in two or three dimensions. Sites meant to represent the gel clusters are randomly painted black, with a parameter called occupation probability that is correlated with gel density. The higher the occupation probability is, the more likely it is that any given site will be "occupied," or painted black. In this scheme a low occupation probability corresponds to a small density of clusters in the solution; with increasing probability, more clusters appear until the percolation threshold is surpassed and the clusters make up a coherent field. The percolation threshold for aerogels is extremely low, as can be surmised from their low density.

The chessboard simulations have also helped to reveal that at densities above the threshold the mechanical properties of aerogels "scale" with the density excess; that is, the properties vary with the difference between the actual density and the critical density taken to a power larger than 1. For instance, my colleague Ove Nilsson demonstrated that thermal conduction in an aerogel sample changes when the sample is compressed; in a certain range of compression, conduction is proportional to the excess density taken to the power of 1.6. In my laboratory at Würzburg we have found that the elastic modulus of aerogels, which influences their acoustic properties, scales even more dramatically, with the excess density taken to the power of 3.7.




**COMPUTER-GENERATED CHECKERBOARD PATTERNS simulate the sol-gel process. Light boxes represent the solution, dark boxes the gel. Interconnected sections of the checkerboard are shown in color. The probability that a square will be painted black corresponds to the time elapsed since gel formation began. At a critical probability *P* the squares in the checkerboard become contiguous; likewise in the sol-gel process there is a critical threshold at which the clusters that make up the gel body become coherent. These patterns were created by Manfred Gronauer in the author's laboratory.**

Scaling laws are important not only from a theoretical point of view but also in terms of optimizing practical systems such as the thermal insulation of aerogel pellet layers in windows. Recently, for example, Ernst Hümmer in my laboratory concluded from our scaling studies that the solid thermal conduction in such systems varies only weakly with the pellet density.

Another intriguing aspect of aerogel structure pertains to its fractality. It seems that the aerogel skeleton is fractal: a self-similar system whose geometric properties repeat at innumerable scale lengths. Dale W. Schaefer and his colleagues at the Sandia National Laboratories were the first to document the fractal nature of aerogels; recently systematic investigations of aerogels made under controlled conditions showed that their fractal behavior changes with the *p*H value and with the concentration of the starting solution. These last investigations were carried out by René Vacher, Thierry Woignier and Jacques Pelous of the University of Montpellier as well as by Eric Courtens of the IBM Zurich Research Laboratory.

The recent studies of aerogels as fractals were contingent on precise control of the sol-gel process. Further understanding of the process in the .1-to-100-nanometer range will also be necessary to tailor the aerogel microstructure to particular applications. Yet aerogels are not just valuable materials in and of themselves: they can also serve as important precursors in the production of optical-quality glass. Proper discussion of the gel-to-glass transition and its potential benefits makes a story in its own right, and as knowledge of these unusual solids grows there will probably be many more stories to tell.

# The Indian Neck Ossuary

*A 1,000-year-old group burial on Cape Cod is the first clue on a trail of evidence implying that the southern New England coast had a settled population centuries before the Pilgrims*

by Francis P. McManamon and James W. Bradley

In September, 1979, National Park Service archaeologists surveying prehistoric remains in the Cape Cod National Seashore were called to a site on Indian Neck, a peninsula in Wellfleet Harbor on outer Cape Cod. A backhoe operator digging a hole for a cottage's septic tank had unearthed human bones. The archaeologists suspected a prehistoric burial; so that the rest of the deposit could be salvaged, the cottage owners allowed us and our colleagues from the Park Service to excavate the remaining material. The excavation began the following week, and after two days of careful fieldwork an unexpected but intriguing burial feature had been exposed.

The backhoe had cut through a collection of bones that proved to contain the remains of at least 56 people—men, women and children—interred together about 1,000 years ago. It is the best-documented example of an ossuary (the technical term for such a deposit of bones) in New England, and one of the best-documented outside northwestern New York State, adjacent parts of Ontario and the area around the Potomac River and Chesapeake Bay. It is also perhaps the strongest evidence to date bearing on what is still an open question about the prehistory of coastal New England: how settled were the native inhabitants in the centuries before Europeans arrived? We believe the ossuary, combined with other archaeological data and the testimony of the first Europeans to explore the region, tips the balance of evidence away from scenarios in which the coast saw no more than occasional visits or seasonal camps and toward a picture of permanent coastal habitation.

That the deposit was prehistoric became evident even before we had dug down to the bone layer through the undisturbed ground next to the backhoe cut. Under a few inches of topsoil we found a midden: an ancient rubbish dump containing shells, animal bone, broken stone tools and the like. The midden's upper layer also yielded a scrap of sheet brass that had served as a cutting tool; at least some of the deposit, then, had been laid down after Europeans had introduced such material. The kind of stone artifacts, however, suggested that much of the midden dated from the final period of prehistory in the Northeast, the Late Woodland, which began about A.D. 900 and ended with the arrival of Europeans. Because the midden lay above the bones (to which it is probably not related) the ossuary had to date from before European contact.

The ossuary itself, at a depth of a little more than a foot, was a startling sight: a dense, compact mound of bones about four feet across. The pile may originally have been twice that long in the north-south direction: a good part of it—perhaps half—had been cut away by the backhoe. The face of the backhoe cut showed that the bone layer was about a foot thick at its thickest point.

At first glance one saw only a jumbled mass, but on close inspection some organization was evident. Craniums tended to be concentrated along the east and west margins of the deposit, with other bones in between. Some of the long bones were arranged in parallel clusters, as if they had originally been bundled. Several clusters later proved to consist of what might be the bones of a single individual: two humeri, two femurs and two tibiae, for example. The fact that most of the bones were no longer connected at the joints suggested that flesh had been scraped off or left to decompose before the bones were placed in the ossuary.

After we had exposed and cleaned the bones of this layer we began removing them for further cleaning and study in the laboratory. In the process we found that many of the bones at the bottom of the layer were slightly charred on their underside. The cause soon became apparent. Underlying the first bone layer was a second layer, containing densely packed bone that had been partially reduced to ash. This mass of cremated remains, perhaps half as big as the first layer in each dimension, showed no signs of internal organization.

Even though these bones had been thoroughly incinerated, the reddish orange sand underlying the cremation layer showed relatively little discoloration. Moreover, the layer contained no charcoal. It therefore seems unlikely that the remains were cremated in the pit itself. Instead the cremation probably took place elsewhere (albeit nearby), after which the remains were placed in the pit while still smoldering. Then the unburned bones, some of them in bundles, were piled on top of the cremated material, which was still hot enough to singe the added bones.

Just when did this mass burial take place, and why? A prehistoric deposit is often dated by the styles of the stone artifacts and pottery fragments found within it. In the ossuary such relics were rare, but one edge of the deposit did yield a large, triangular point made of felsite, a light-colored igneous rock. The point, which may have ended up in the ossuary by chance or may have been placed there intentionally, was of the style known as Levanna, which is characteristic of the Late Woodland period.

Carbon-14 analysis of the bones

confirmed this preliminary dating. Samples of bone collagen (the protein component of bone) and apatite (the mineral component) from the unburned layer yielded ages of approximately 900 years, with an uncertainty of some 200 years. Bone charcoal from the cremation layer was dated to about 800 years ago, with an uncertainty of 200 years. We concluded that the ossuary was built early in the Late Woodland, probably in the 10th or early 11th century A.D.

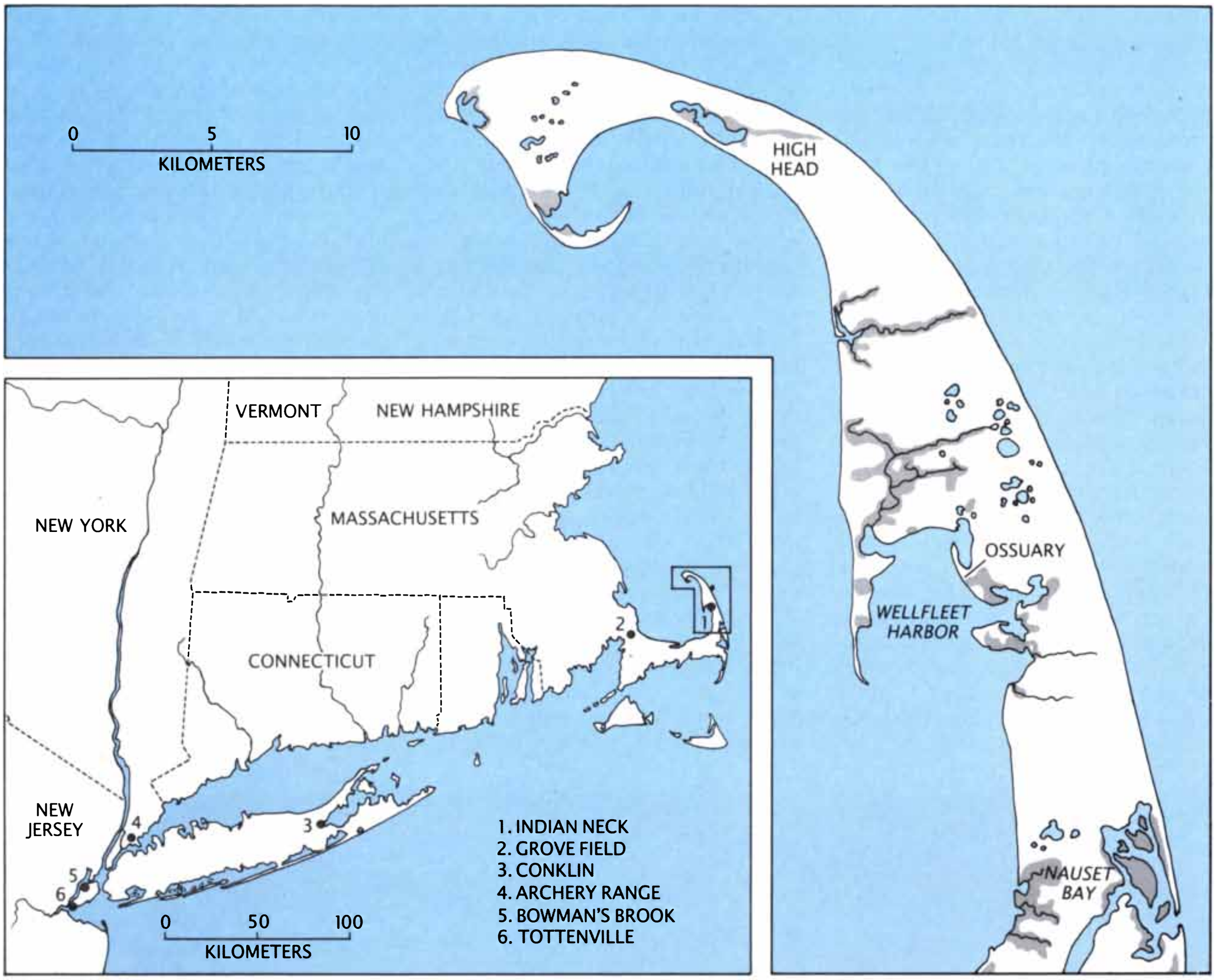
The bones also revealed something about the circumstances of the burial. In a detailed analysis of the bones, Ann L. Magennis of the University of Massachusetts at Amherst counted the remains of 47 people in the unburned layer and at least nine more in the cremated layer. The deposit contained individuals of both sexes (sex was determined for a total of 14 individuals, from the shape of the pelvis) and virtually all ages. The demographics suggested that burial in the ossuary was not limited to people of a specific age or station; rather, the deposit seemed to hold the skeletal remains of everyone in the community who died in a certain period.

Was the ossuary built to hold the dead from a massacre, a famine or an epidemic, as popular speculation held when the site was discovered? Magennis' study of the bones revealed no unusual incidence of trauma and only a modest frequency of the structural changes in bone that result from infectious disease or malnutrition. Indeed, the health of the ossuary population was remarkable compared with some other prehistoric groups, extending even to a low incidence of dental caries. (That finding, together with a high frequency of chipped teeth, prompted Magennis to hypothesize that the ossuary builders' diet included relatively little ground maize but many wild foods.) The ossuary, then, is probably not a monument to a prehistoric catastrophe; more likely it is evidence of a regular mortuary practice.

Removing flesh from the bones before interring them evidently was one part of the practice. Magennis' examination of the unburned bones suggests that the task was left to the process of decomposition, either in a primary burial site or aboveground. She found only two instances of the cut marks that would suggest flesh had been physically cut away from the bones: sets of gouges near the eye sockets of two adult males. The bones in the cremated layer also seem to have been bare when they were incinerated. Rather than serving as a means of cleaning bones before they were interred, Magennis proposes, cremation may have been reserved for a segment of the community that was awarded special burial treatment.

What does the ossuary say about the way of life of the people who built it? Were they nomads, driven from place to place by the shifting availability of food resources, or were they a sedentary people for whom a variety of resources lay within reach of their settlements? The question goes to the heart of an intensifying debate about the nature of the settlement system along the New England coast in the centuries before the arrival of Europeans. In other areas, such as the territory

**TOP OF THE OSSUARY is exposed. The deposit, which dated from the 10th or early 11th century, contained the bones of at least 56 individuals; it is thought to reflect the ceremonial burial practices of a sedentary population. Archaeological excavation began after human bones were unearthed by a backhoe digging the trench at the right in the course of a construction project.**




**OUTER CAPE COD has yielded prehistoric remains at a number of sites: the ossuary on the peninsula of Indian Neck and many concentrations of remains around Nauset Bay and Wellfleet Harbor and on High Head. The artifacts seem to have been deposited by a prehistoric population that lived year-round on the outer Cape, exploiting resources from the ocean, salt marsh, freshwater wetlands and inland woods. The inset map shows locations of known and suspected ossuaries in the coastal Northeast.**

of the Iroquoian people in western New York and southern Ontario, the remains of grain-storage pits and densely clustered dwellings, sometimes ringed by a palisade, give clear evidence of permanent settlements. The archaeological record of the coast, in contrast, is dominated by middens, isolated habitations and a few storage pits; unequivocal signs of permanent habitation are rare.

In recent years, as archaeologists in unparalleled numbers have poked into coastal remains from Maine to Manhattan, three broad models of the coastal settlement pattern have emerged. The first model, put forward by Lynn Ceci of Queens College of the City University of New York, holds that the coast was occupied only sporadically, by people living inland who made occasional summer forays to fish or gather shellfish. In Ceci's picture, sedentary villages would only have sprung up along the coast during the Contact period, as a result of trade with Europeans and the stimulus to manufacture wampum (shell beads), which could be exchanged with inland tribes for the furs prized by Europeans.

A second model has the coast exploited regularly and intensively every summer. Dean R. Snow of the State University of New York at Albany has interpreted coastal sites dating from the Late Archaic period (from about 2500 to 500 B.C.) on Martha's Vineyard, Long Island and the Hudson River estuary as being summer camps. The sites, he argues, are evidence of a pattern of seasonal movement in which the inhabitants regularly shifted their dwellings to exploit a variety of food resources: shellfish at coastal summer camps, for example, and deer in winter forest camps in the interior. Bert Salwen of New York University has suggested that the same pattern prevailed through the entire Woodland period (500 B.C. to A.D. 1500) and into the period of European contact.

A third model envisions permanent residence at some sites along the coast beginning in about 2500 B.C. The inhabitants would also have established temporary camps along the coast for hunting and specialized activities and conceivably seasonal camps inland, and their permanent base may not have been a dense and elaborate compound. Nevertheless, it was occupied continuously. David Sanger of the University of Maine at Orono, who has found evidence of year-round activities at sites on the Maine coast, is a proponent of this view, together with Arthur E. Spiess of the Maine Historic Preservation Commission and Bruce J. Bourque of the Maine State Museum. Kent G. Lightfoot of the University of California at Berkeley and his co-workers recently reported evidence of a similar settlement pattern on eastern Long Island.

We think the Indian Neck ossuary can help us to choose among the models. Even though the ossuary is the best-established instance of such a deposit in New England, it is unlikely to be an anomaly. Contemporaneous accounts suggest, for example, that another ossuary was unearthed in 1911 at Grove Field, near the town of Bourne at the base of Cape Cod. Ossuaries are known or suspected at four sites in coastal New York, an area that may have been culturally similar to coastal New England in late prehistory. If ossuaries were a regular feature of the region's prehistoric culture, it seems reasonable to judge the significance of the Indian Neck deposit by examining the features of other native societies that interred their dead in ossuaries.

The evidence suggests that such burial practices are the hallmark of a sedentary way of life such as the one described in the third model of coastal settlement. The clearest picture of ossuaries and their cultural context comes from the region of the Hurons, in Ontario. Archaeological findings and the detailed accounts left by the first Europeans to penetrate the area—French missionaries of the Recollect and Jesuit orders—indicate that each ossuary was built near a sedentary village, usually within 100 yards of it. The ossuary received remains not only from that village but also from surrounding villages, some of them 10 or 12 miles away.

Modern interpretations of the missionaries' accounts indicate that the ossuaries served as the centerpiece of rituals designed to reassert social ties among villages that had budded from a single, ancestral village as its population grew. Every eight, 10 or 12 years the outlying villagers would collect the remains of their dead relatives and return to the ancestral village. There, amid reunions and feasts, the celebrants would contribute the bones to the ossuary.

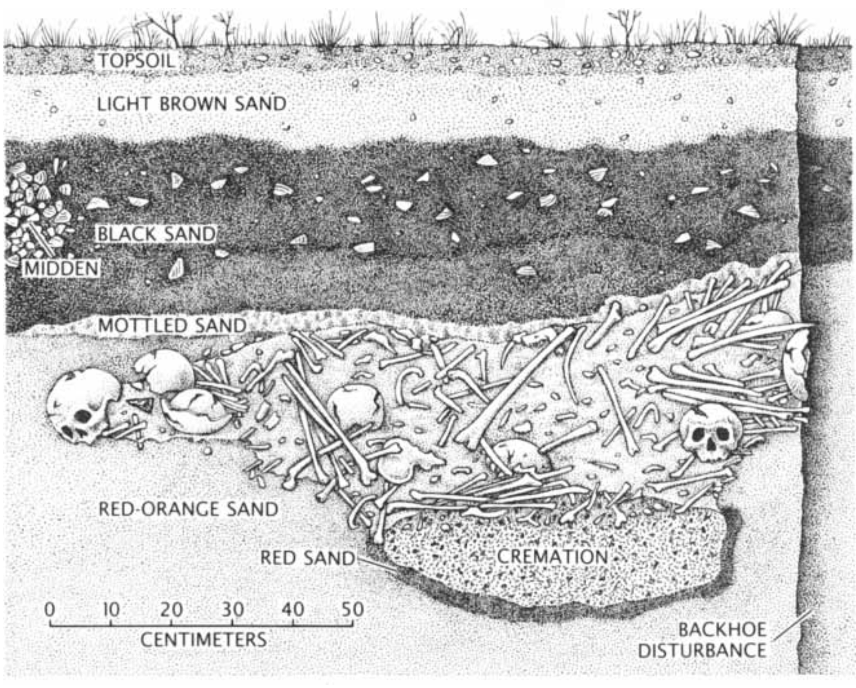
Near the mouth of the Potomac River, along Chesapeake Bay and on the North Carolina coastal plain ossuaries have also been unearthed. These areas lack the detailed descriptions of native life the missionaries left for the Hurons. Archaeological data, however, again record a strong association between ossuaries and settled village life.

To confirm that this same link between ossuary burial and a sedentary way of life existed on Cape Cod, we turned to the broader archaeological record of the Cape. That record is particularly rich for the Late Woodland period, whose beginning roughly coincides with the date of the ossuary. A survey of 779 sites on Cape Cod, Martha's Vineyard and Nantucket done by the Massachusetts Historical Commission found that on the basis of artifact styles more sites could be definitively assigned to the Late Woodland than to any other period. The prehistoric population of Cape Cod may well have been at its largest then.

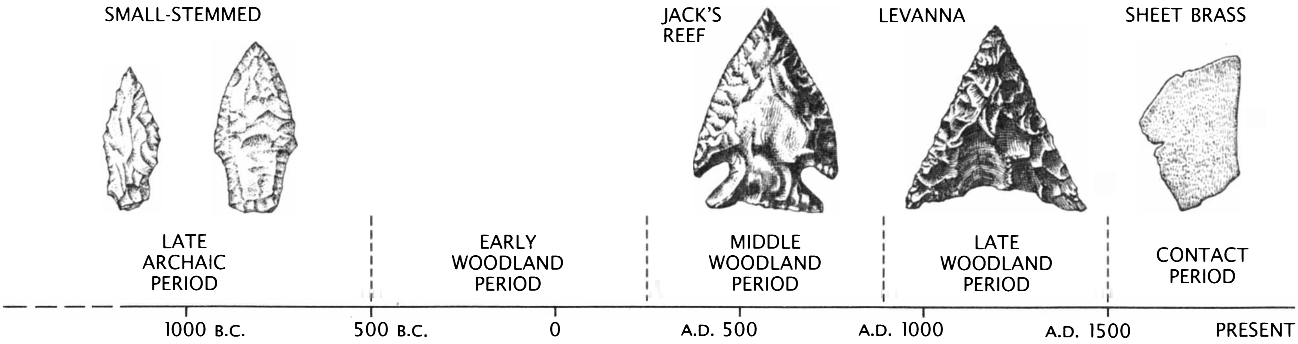
For the outer Cape, clues to that population's way of life come from the Park Service archaeological survey (under whose auspices the ossuary was excavated). The survey, which continued from 1979 through 1986, sought to determine what kinds of human activities took place at prehistoric sites in the National Seashore and how they were organized geographically and seasonally. In answering the last question an analysis of shells recovered from prehistoric middens proved invaluable.

The prehistoric inhabitants of Cape Cod gathered large numbers of the hard-shelled clams called quahogs (*Mercenaria mercenaria*). The age of a harvested clam is often evident in its discarded shell. During each year of its life the mollusk lays down a band of new shell that is etched with hundreds of microscopic growth lines recording the daily tidal cycle. Counting the fine growth lines in a shell's last annual increment can sometimes give a rough indication of the season or even the month of the creature's death. The shells in a prehistoric deposit can thus reveal the timing of shellfish harvesting and hence offer clues to the season of occupation.

Mary E. Hancock and Alison Dwyer, who worked on the Park Service survey during the early 1980's, analyzed shells from 16 concentrations of prehistoric remains in the National Seashore. The concentrations also included various combinations of other relics—stone points and flakes, fire-cracked rock and organic debris




**LEVELS OF THE EXCAVATION include topsoil, several layers of sand, one of them containing a midden—a refuse heap, this one consisting largely of shells—and the ossuary. A mass of unburned bones makes up the ossuary's upper layer, with skulls concentrated at the edges and long bones at the center; the deeper layer consists of cremated bones. At the right is the cut made by the backhoe, which may have destroyed as much as half of the unburned layer. The cross-sectional illustration is based on one by Thomas F. Mahlstedt, now the Metropolitan District Commission Archaeologist in Boston.**




**PREHISTORIC TIME LINE for eastern North America is divided into periods associated with distinct artifacts. The artifacts, which serve for dating archaeological deposits, vary from region to region; the stone points shown here are typical of coastal New England. (It is still unclear what artifacts characterize the Early Woodland period of southern New England.) The ossuary yielded a Levanna point, which dates the burial to the Late Woodland period; the overlying midden contained a scrap of sheet brass, indicating that at least part of that deposit was laid down after the first contact with Europeans, who introduced such material.**

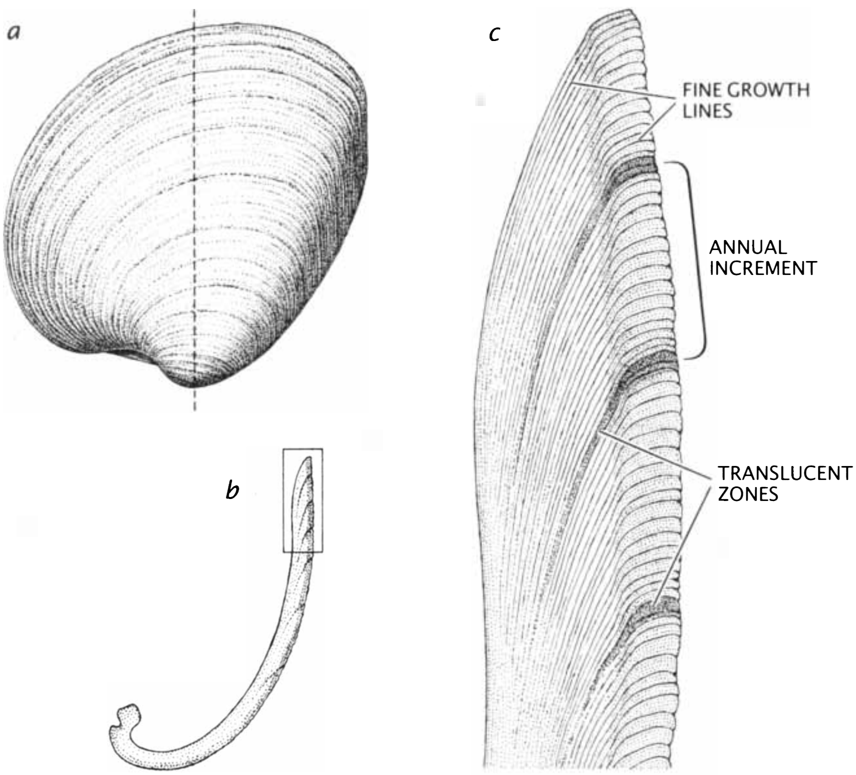
such as charcoal and animal bone. From artifact styles or by carbon-14 dating, each concentration had been assigned to a general time period, from the Late Archaic through the Late Woodland. The analysis showed that in every period of prehistory, shell collecting was most intense during the winter and early spring.

Hypotheses that natives journeyed to the New England coast only occasionally or on a seasonal basis have generally assumed summertime occupation. The seasonality findings suggest at the very least that beginning as early as the Late Archaic the inhabitants were present along the coast in winter. We think it is most plausible to suppose they lived there permanently.

Shellfish collecting did not go on year-round, however, as one might expect if the coast was inhabited continuously; indeed, its timing seems to have become progressively more limited. In two of the concentrations, from the Late Archaic and the Middle Woodland, the growth lines suggest shellfish harvesting continued beyond the winter and early spring into the summer. In all the Late Woodland sites, in contrast, shellfish seem to have been gathered in the winter and early spring exclusively.

The significance of this scheduling shift is obscure, but it may have accompanied the emergence of a practice that is a hallmark of sedentary populations: horticulture. Accounts of life among the Hurons indicate that the tending of corn hills was a time-consuming chore; furthermore, an early observer of native life in New England, William Wood, wrote in 1634 that tending corn and collecting shellfish were both women's tasks. Conceivably shellfish gathering became confined to winter in the Late Woodland because increased attention to crops during the rest of the year left the women with little time for shellfish gathering. Other wild foods would have been scarcest, moreover, in winter.

Artifact concentrations studied in the Park Service survey yielded a second kind of evidence that the outer Cape was populated year-round at least by the Late Woodland and perhaps earlier. It is a reasonable assumption, and one that has been confirmed in studies of other cultures, that the longer a group occupies a single location, the more abundant and diverse its trash and garbage will be. It is also reasonable to expect that whereas a migratory group may leave stone chips, animal bones and other waste within its campsite, a sedentary group will concentrate its refuse in a site at or beyond the edge of the actual habitation area. The dense middens of a sedentary group are known as sec-




**GROWTH LINES in the shell of a quahog (a hard-shelled clam) record the time of year it was harvested. During the winter a living clam grows slowly, laying down a translucent zone of shell material. When the clam begins to grow faster, in February or March, the added material is opaque and the daily growth cycles—set by the tides—are recorded in fine growth lines. The month of a clam's death can often be determined by sectioning the shell (*a, b*) and counting the microscopic lines in the last annual increment (*c*).**

ondary deposits, to distinguish them from the sparser and less diverse primary deposits found at actual habitation sites.

By examining the density of each concentration we could identify it as either a primary or a secondary deposit. Some of the densest secondary deposits consisted almost entirely of shells, however, and so we also distinguished concentrations according to the variety of artifacts they held. We designated deposits that contained high densities of at least two kinds of remains (animal bones and fire-cracked rock, for example), as well as smaller quantities of other artifact types, as general middens; we took them to be strong evidence of a nearby permanent settlement.

We identified 17 general middens in the National Seashore. Most of them date from the Late Woodland, although a few of them represent earlier parts of the Woodland period and the Late Archaic. By the Late Woodland, then, and probably earlier, a cultural system that included permanent settlements had developed on the outer Cape. To judge from the locations of the artifact concentrations and other collections that we studied, the prehistoric settlement was particularly dense near Nauset Bay, around Wellfleet Harbor and at the very tip of the Cape near a bluff called High Head—areas within a few miles of Indian Neck.

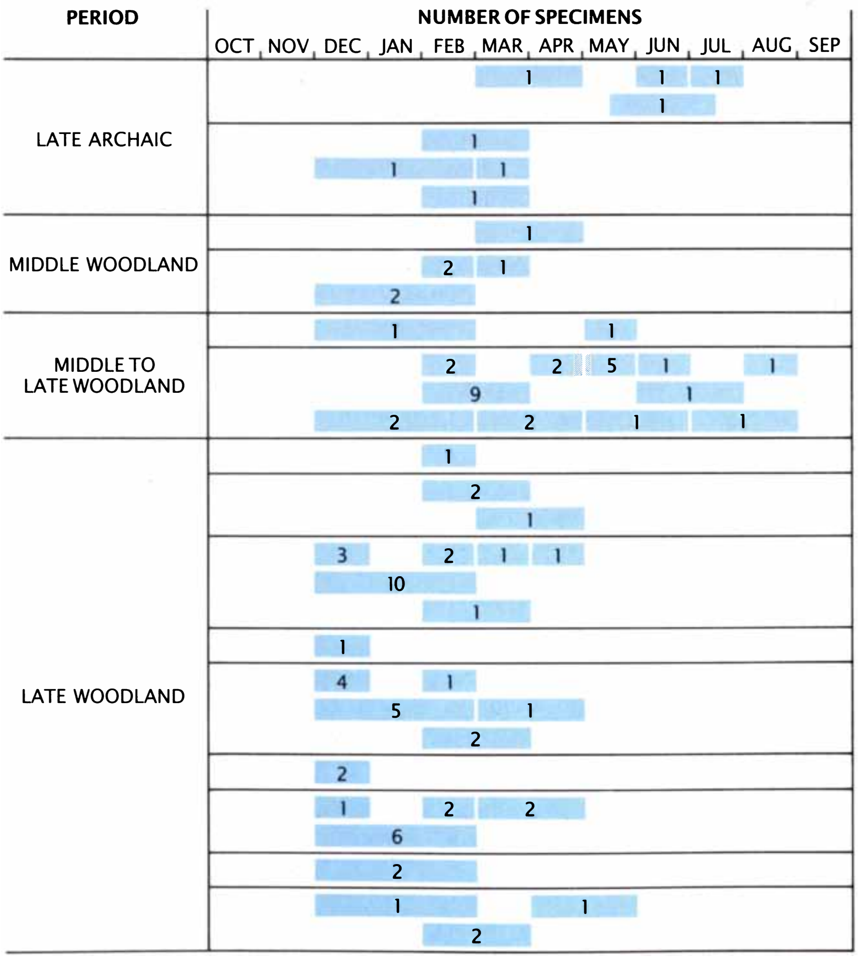
Those sections of the Cape were among the first to be explored by Europeans, at the beginning of the 17th century. The reports of those early visitors seem to confirm the evidence of the archaeological record and the ossuary itself. It is true that the ossuary and many of the artifact concentrations in the Park Service survey predate European contact by centuries. But the Massachusetts Historical Commission survey showed that the archaeological record of the Cape is remarkably consistent, both in variety of sites and in their distribution, throughout the Late Woodland and into the Contact period. What the first Europeans saw is likely to have been representative of the settlement pattern of centuries past.

The earliest detailed reports stem from a French effort to explore and chart the New England coast beginning in the summer of 1605. Samuel de Champlain, later known as the Father of New France, served as the expedition's cartographer. In late July the explorers reached the southern extreme of their voyage, which had begun in the St. Croix River at the present U.S.-Canada border. Their haven was Nauset Bay, which Champlain described as "a bay with wigwams bordering it all around."

The French returned to Nauset Bay in the fall of 1606 to find the shores still populous; as the leader of the party rowed ashore, according to Champlain, "there came to meet him some 150 Indians, singing and dancing in accordance with their custom." Champlain's account suggests that the natives not only remained on the coast longer than a single season but also wintered there. At Stage Harbor, about 20 miles south of Nauset, Champlain noted that "all the inhabitants of this place are much given to agriculture, and lay up a store of Indian corn for the winter." The grain was stored, he reported, in pits dug in sandy hillsides.

The next Europeans to visit the outer Cape, the Pilgrims, saw similar evidence of permanent habitation. The Pilgrims eventually settled in Plymouth, but for more than a month after their landfall in the winter of 1620 the *Mayflower* anchored in Provincetown Harbor, at the tip of Cape Cod. From there the settlers staged three short explorations of the surroundings in order to learn whether the land and the supply of fresh water would support a settlement and whether the disposition of the natives would allow one. Edward Winslow and William Bradford, two of the Pilgrim leaders, recorded what they saw.

In their first exploration, which took them south about 10 miles to




**TIMING of the prehistoric shellfish harvest on outer Cape Cod yields clues to the inhabitants' way of life. The chart shows data derived from growth lines in clam shells found in various artifact concentrations; each number indicates how many shells could be assigned to the month or months spanned by the bar. Throughout prehistory shellfish harvesting was most intense in winter, indicating that the outer Cape was populated in winter (and probably year-round). Some shells from older artifact concentrations may also have been harvested in summer, but shellfish seem to have been gathered exclusively in winter during the Late Woodland period. By then, the authors speculate, an increasing emphasis on horticulture left little time for clamming in the summer.**

**MAP OF NAUSET BAY is from the works of the French explorer Samuel de Champlain, who visited outer Cape Cod in 1605 and 1606. In addition to recording incidents during the French exploration, including skirmishes with the native inhabitants, the map depicts a landscape that gives evidence of a sedentary population: it is patterned with fields and dotted with wigwams.**

the Pamet River, the Pilgrims passed through a landscape that gave every sign of intensive human activity, including horticulture. They reported clearly marked, well-worn paths and land that seemed to have been recently cleared and cultivated, including one tract estimated to cover 50 acres. They noted a series of smaller cornfields that had been cultivated that same year, and near one of them they found the remains of a wigwam. The explorers also found a buried cache of corn, probably some of that year's harvest. Their encounter with the cultural landscape of the New World was completed when Bradford (the future governor of Plymouth Colony) accidentally became snared in a deer trap.

The Pilgrims' later forays turned up graves, food caches containing corn, beans and acorns, fallow fields and unoccupied wigwams, perhaps abandoned as the explorers drew near. Particularly tantalizing, in the light of the ossuary, is their report of "a great burying place": a palisade partially enclosing a cluster of graves, some of which were roofed or surrounded by smaller palisades. It was only at the end of their third expedition that the Pilgrims met the inhabitants face to face. Unfortunately the encounter took the form of an attack on the explorers' camp and had the effect of encouraging the settlers to look for a more hospitable site, which they found at Plymouth.

Like Champlain before them, the Pilgrims provided a glimpse of a fully settled culture in their descriptions of extensive land use, food storage and signs of winter occupation. What those visitors did not report are dense, enclosed villages such as the ones the French missionaries discovered among the Hurons and Iroquois or the village-and-mound complexes that have been excavated in the middle western and southeastern U.S. The fact that such villages are absent both from the accounts and from the archaeological record suggests that native settlements on the outer Cape, even though they were permanent, were small and dispersed—perhaps no more than the clusters of wigwams the European explorers did note.

In the light of the present-day geography of Cape Cod one can easily picture how year-round residents could have sustained themselves there without migrating seasonally in search of food. Occupying sites such as the land bordering Nauset Marsh and Wellfleet Harbor, they had easy access to a variety of microenvironments, ranging from tidal flats and salt marsh to freshwater wetland and wooded upland. Deer, acorns, shellfish, waterfowl and other wild foods, each available in a particular microenvironment at a specific season, contributed to their economy. Horticulture added to but probably did not dominate their diet.

As the quincentenary of Columbus' arrival approaches, the intensive encounter between Europeans and Native Americans that began with his landfall will more than ever be the subject of study and fascination. Understanding the course and results of the encounter requires a clear image of preexisting native culture and its adaptations. We think the Indian Neck ossuary and the archaeological and historical analyses it has stimulated have sharpened that picture for southern New England.

# The Mystery of the Cosmological Constant

*According to theory, the constant, which measures the energy of the vacuum, should be much greater than it is. An understanding of the disagreement could revolutionize fundamental physics*

by Larry Abbott

What determines the structure of space and time in the universe? According to Einstein's general theory of relativity, the geometric properties of space are related to the density of energy (and momentum) in the universe. To understand the structure of spacetime, therefore, we must identify potentially relevant sources of energy and evaluate their contributions to the total energy (and momentum) density. The most obvious energy sources that come to mind are ordinary matter and radiation. A much less obvious source of energy that can have an enormous impact on the structure of the universe is empty space itself: the vacuum.

The notion that the vacuum can be a source of energy may seem counterintuitive. But present theories of elementary particles and forces not only allow for a nonzero vacuum energy density but also strongly suggest that it should have a large value. Is the vacuum energy density really as large as these theories appear to suggest it is?

The answer is most emphatically no. The geometric structure of the universe is extremely sensitive to the value of the vacuum energy density. So important is this value that a constant proportional to the vacuum energy density has been defined. It is called the cosmological constant. If the vacuum energy density, or equivalently the cosmological constant, were as large as theories of elementary particles suggest, the universe in which we live would be dramatically different, with properties we would find both bizarre and unsettling. What has gone wrong with our theories? We do not know the answer to this question at present. Indeed, a comparison of our theoretical and experimental understanding of the cosmological constant leads to one of the most intriguing and frustrating mysteries in particle physics and relativity today.

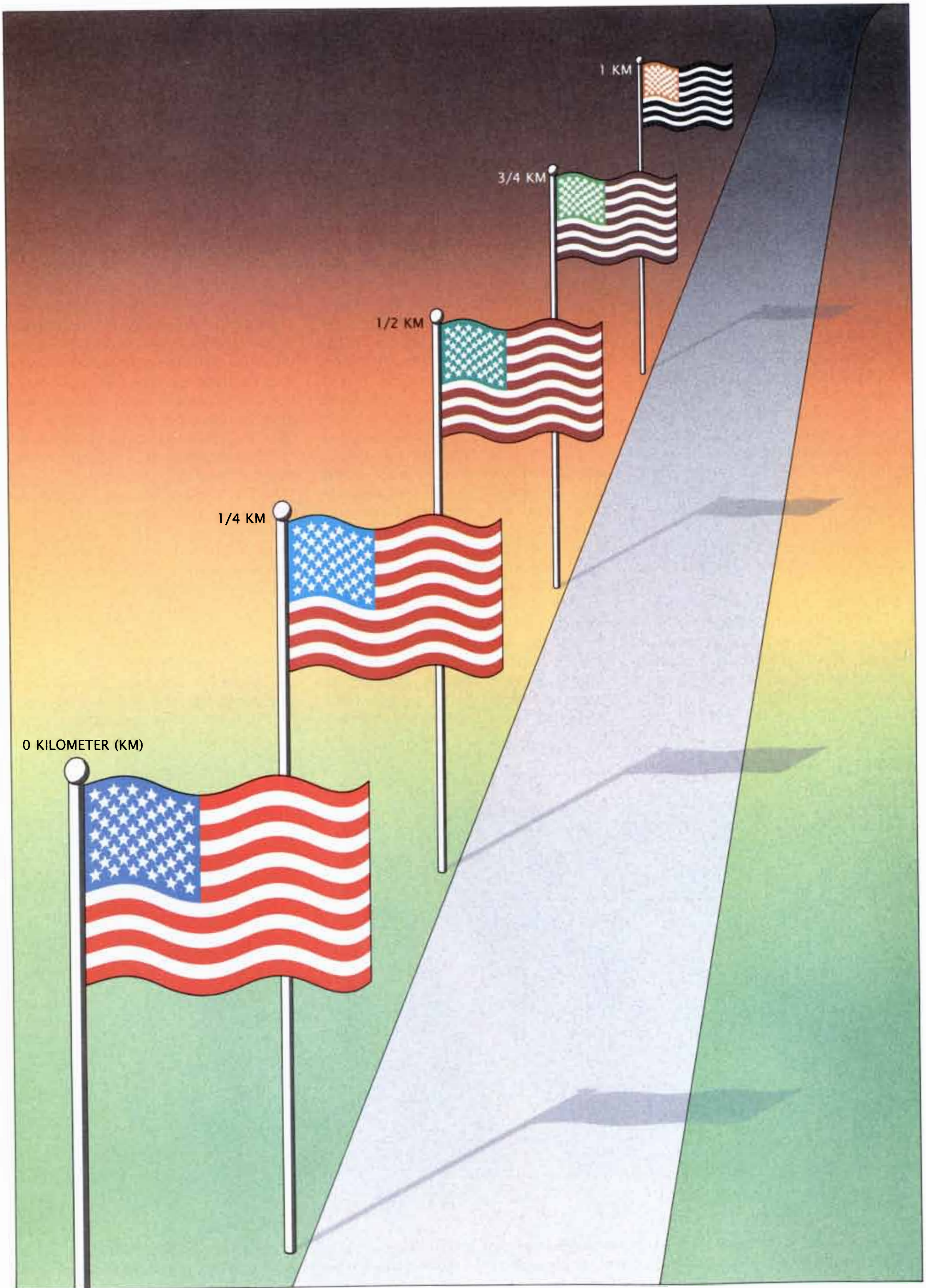
Most people are unaccustomed to the idea that the vacuum might have a nonzero energy density: How can a unit volume of empty space contain energy? The answer in part lies in the fact that, according to quantum mechanics, physical quantities tend to fluctuate unavoidably. Even in the apparent quiet of the vacuum state pairs of particles are constantly appearing and disappearing. Such fluctuations contribute energy to the vacuum.

The notion of a vacuum energy is also unfamiliar because that energy cannot be detected by normal techniques. Energies are usually determined by measuring the change in the energy of a system when it is modified in some way, or by measuring a difference in energy between two systems. For example, we might measure the energy released when two chemicals react. Because of this, energy as we normally define it is a relative quantity. The energy of any state of a system only has meaning in relation to some other state.

By convention, energies are often measured in relation to the vacuum. When it is defined in this way, the vacuum automatically has zero energy in relation to itself. The traditional approach will not work if we want to discuss the energy of the vacuum in an absolute and significant way. We must use a different technique to measure its value.

The only way to establish an absolute measure of energy is by using gravity. In general relativity, energy is the source of gravitational fields in the same way that electric charge is the source of electric fields in the Maxwell theory of electromagnetism. An energy density of any kind, including that produced by fluctuations in the vacuum, generates a gravitational field that reveals itself as a change in the geometry of spacetime. The gravitational field of the earth, for instance, is produced by its rest energy, which equals the mass of the earth multiplied by the square of the speed of light (as given by the famous formula $E = mc^2$). The gravitational field produces a small distortion in the spacetime geometry near the earth, resulting in the attractive force that pulls us all toward the ground. In general relativity the energy density of the vacuum has an absolute meaning, and it can be determined by measuring the gravita-

**UNIVERSE with a large cosmological constant would be vastly different from the existing one. Here an artist has painted a scene as it might appear if the constant were as large as theoretical estimates suggest it could be. The illustration is based on a positive value for the constant on the order of $1/(1 \text{ kilometer})^2$. With such a value the structure of space would be so distorted that the radiation from distant objects would be red-shifted, or shifted toward longer wavelengths. The farther an object is from an observer, the greater the red shift would be. A spectral blue object about a kilometer away would look red; objects more than a kilometer or so away would have such large red shifts that they would be invisible. Distant objects would appear spatially distorted.**



1 KM
3/4 KM
1/2 KM
1/4 KM
0 KILOMETER (KM)

**COSMOLOGICAL CONSTANT = $8\pi G/c^4$ × VACUUM ENERGY DENSITY**

Here $G$ is Newton's gravitational constant and $c$ is the speed of light. Defined in such a way, the cosmological constant has units of 1 over distance squared.

tional field produced not by matter but by the vacuum itself.

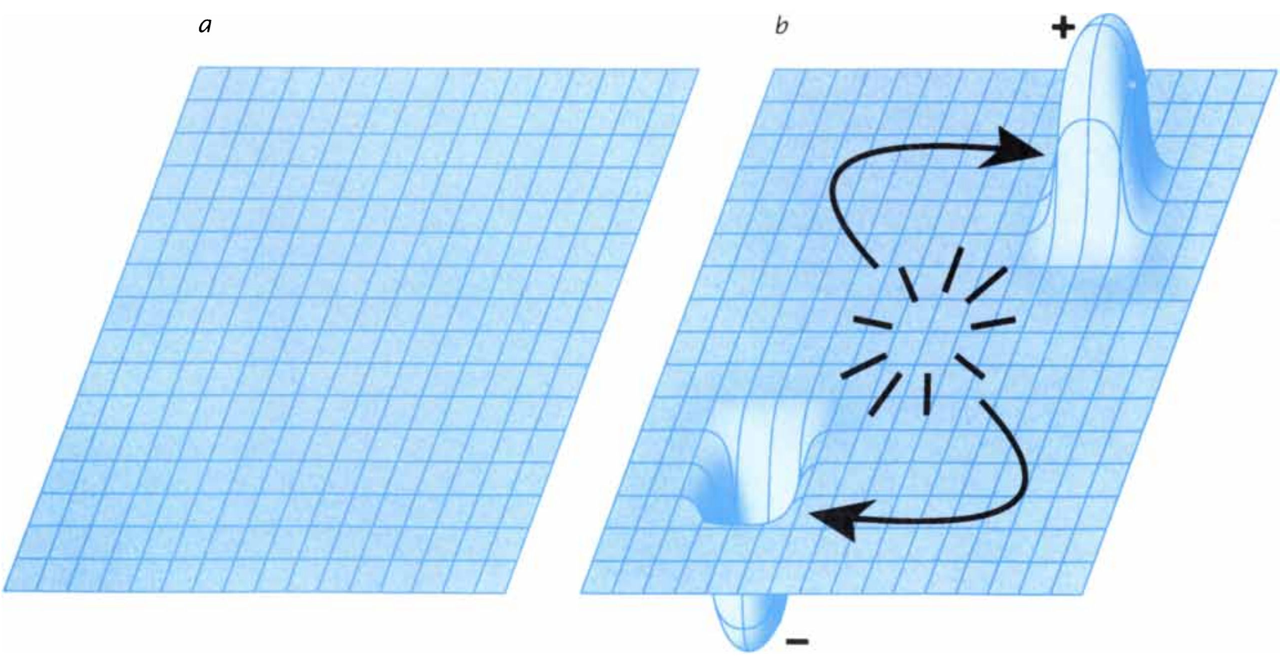
Of course, determining the energy density of the vacuum is tantamount to determining the cosmological constant, since one is proportional to the other. It turns out that the cosmological constant can be assigned units of 1 over distance squared. In other words, the square root of the reciprocal of the cosmological constant is a distance. This distance has a direct physical meaning. It is the length scale over which the gravitational effects of a nonzero vacuum energy density would have an obvious and highly visible effect on the geometry of space and time. By studying the geometric properties of the universe over length scales on the order of that distance, the value of the cosmological constant can be measured.

Physicists have been struggling with the issue of the cosmological constant for more than 70 years. The constant was first introduced by Einstein in 1917 in an attempt to eliminate two "problems" in his original formulation of the general theory of relativity. First, he thought that without a cosmological constant the general theory could not account for a homogeneous and isotropic universe: one that looks much the same everywhere. (It is remarkable that Einstein even cared about such matters in 1917, since at the time there was no evidence that the universe was homogeneous and isotropic, which indeed it is.) Unfortunately Einstein's reasoning was incorrect. In 1922 Alexander A. Friedmann showed that the general theory does allow for a homogeneous and isotropic universe, although not a static one: the universe must be expanding (or contracting). Subsequent astronomical observations have convincingly demonstrated that models based on Friedmann's work accurately describe the large-scale structure of the universe.

Einstein was also dissatisfied with his original formulation because the theory did not provide an explanation of inertia. He believed that by adding a cosmological constant he might produce a theory capable of relating the inertial properties of matter directly to the distribution of energy and momentum in the universe, in a manner first suggested by the Austrian physicist and philosopher Ernst Mach. The hope was dashed soon after Einstein's paper appeared by an argument advanced by the Dutch physicist Willem de Sitter, who discovered the spacetime we shall discuss.

After such an ignominious start it is not surprising that in 1923 Einstein wrote, perhaps somewhat bitterly, "away with the cosmological term." As we shall see, it has not been so easy to eliminate the cosmological constant—it has survived to frustrate many theoretical physicists since Einstein. George Gamow has written that Einstein felt "the introduction of the cosmological term was the biggest blunder he ever made in his life," but once introduced by Einstein "the cosmological constant...rears its ugly head again and again."

At the present time we would appear to be in an excellent position to address the issue of the cosmological constant, because we possess one of the most successful physical theories ever developed, namely the standard model. The standard model is the rather unimaginative name given to a collection of theories that successfully describes all the known elementary particles and their interactions. The remarkable ability of the standard model to interpret and predict




QUANTUM FLUCTUATIONS are among the phenomena that contribute to the energy density of the vacuum (*a*). According to quantum mechanics, the values of physical quantities tend to fluctuate unavoidably. As a consequence pairs of so-called virtuthe results of an enormous range of particle-physics experiments leaves it unchallenged as a model for particle physics (at least up to the highest energies accessible to current particle accelerators).

The standard model is a quantum field theory. This means that for every distinct type of fundamental particle in nature there exists a corresponding field in the model used to describe the properties and interactions of that particle. Thus in the standard model there is an electron field, a field for the photon (the electromagnetic field) and a field for each of the known particles.

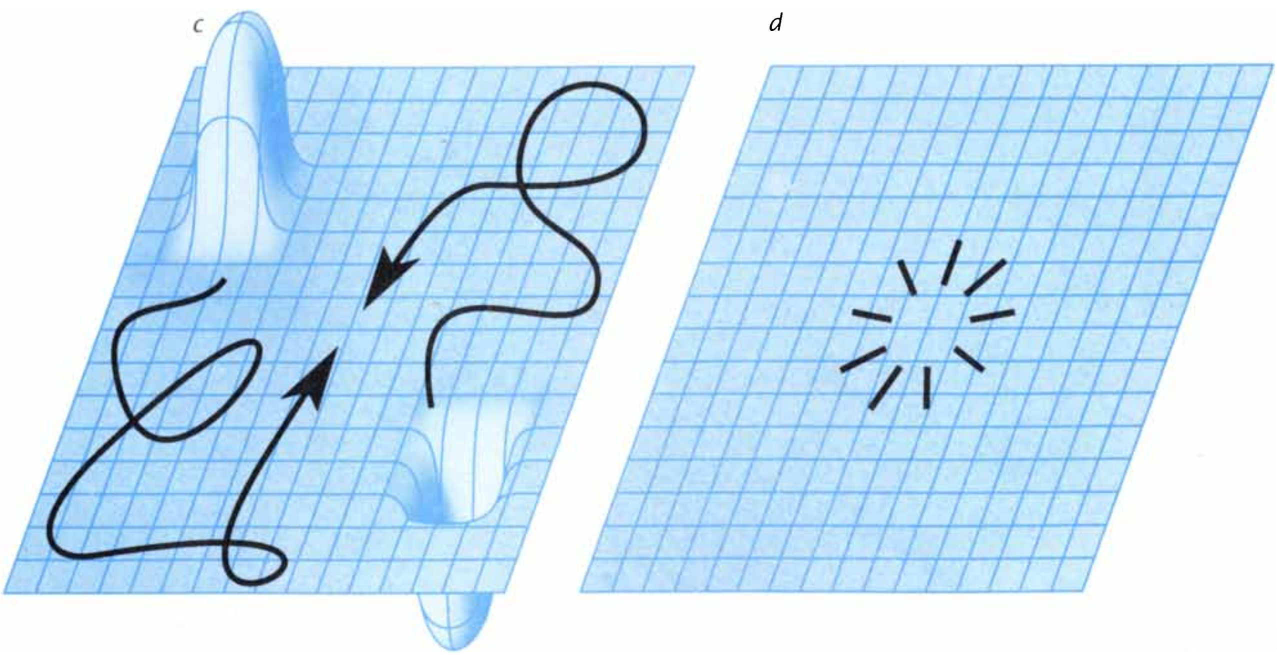
The standard model depends on a fairly large number of free parameters: numbers that must be determined by experiment and fed into the theory before definite predictions can be made. Examples of free parameters include the values of the masses of the particles and numbers characterizing the strengths of their interactions. Once the numbers have been determined the model can be used to predict the results of further experiments, and it can be tested on the basis of its predictions. In the past such tests have been spectacularly successful.

The free parameters of the standard model will play a central role in our discussion. Although the standard model is highly successful, the fact that it depends on such a large number of free parameters seriously limits its predictive power. The model, for example, predicts that an additional particle called the top quark remains to be discovered, but is unable to provide a value for its mass, because this is another free parameter of the theory. A key challenge in particle physics today is to develop a more powerful theory based on a smaller number of free parameters that nonetheless incorporates all the successes of the standard model. Such a theory would be able to determine the values of some of the parameters that cannot be predicted by the standard model. In their search for such a theory, physicists are constantly looking for relations among the parameters of the standard model that might reveal a deeper structure. As we shall see, the cosmological constant will provide us with such a relation, but in this case we shall get more than we bargained for.

In the standard model, as in any quantum field theory, the vacuum is defined as the state of lowest energy, or more properly as the state of least energy density. This does not imply that the energy density of the vacuum is zero, however. The energy density can in fact be positive, negative or zero depending on the values of various parameters in the theory. Regardless of its value, there are many complex processes that contribute to the total vacuum energy density.

In essence the total energy density of the vacuum is the sum of three types of terms. First there is the bare cosmological constant: the value the cosmological constant would have if none of the known particles existed and if the only force in the universe were gravity. The bare cosmological constant is a free parameter that can be determined only by experimentally measuring the true value of the cosmological constant.

The second type of contribution to the total energy density of the vacuum arises in part from quantum fluctuations. The fields in the standard model, such as the electron field, experience fluctuations even in the vacuum. Such fluctuations manifest themselves as pairs of so-called virtual particles, which appear spontaneously, briefly interact and then disappear. (Each pair of virtual particles consists of a particle and its corresponding antiparticle, such as the electron and the positron, which have identical masses but opposite electric charges.) Although virtual particles cannot be detected by a casual glance at empty space, they have measurable impacts on physics, and in particular they contribute to the vacuum energy density. The contribution made by vacuum fluctuations in the standard model depends in a complicated way on the




al particles can appear spontaneously in the vacuum (*b*), interact briefly (*c*) and then disappear (*d*). Here fluctuations are depicted in an abstract and highly symbolic manner. Each pair of virtual particles consists of a particle and corresponding antiparticle.

masses and interaction strengths of all the known particles.

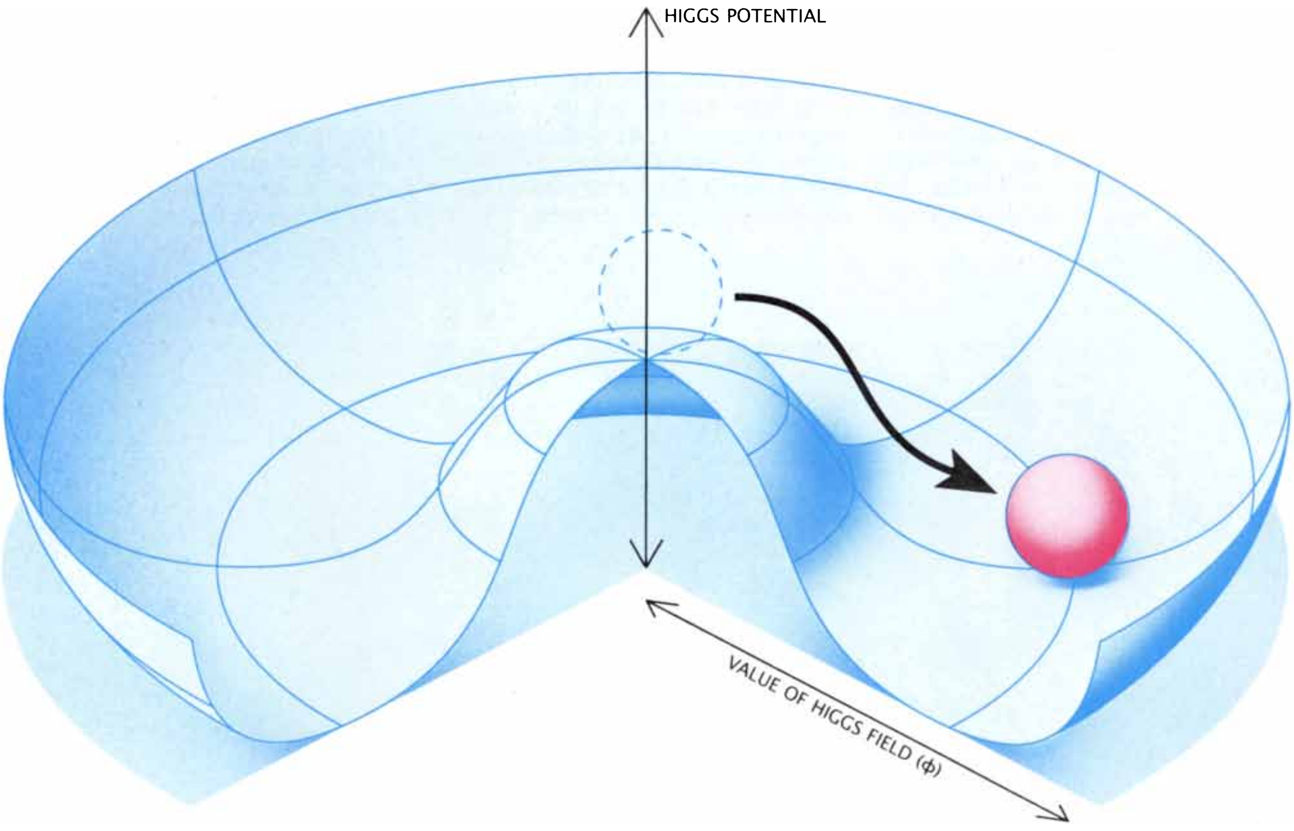
The second type of term also depends on at least one additional field known as the Higgs field, which represents a massive particle, the Higgs boson, that has not yet been detected. The Higgs field should have a particularly dramatic effect on the energy density of the vacuum state [see "The Higgs Boson," by Martinus J. G. Veltman; SCIENTIFIC AMERICAN, November, 1986].

The last type of term that must be included is essentially a fudge factor representing the contributions to the vacuum energy density from additional particles and interactions that may exist but we do not yet know about. The value of this term is of course unknown.

The cosmological constant is determined by adding together the three terms we have discussed. Our ability to predict its value using the standard model is frustrated by the existence of the bare cosmological constant—a free parameter that can be determined only by carrying out the very measurement we are attempting to predict—and by the sensitivity of the vacuum energy to unknown physics. All is not lost, however, at least not yet. Although all the terms that go into making up the cosmological constant depend in a complicated way on all the parameters of the standard model, the values of many of the terms can be fairly accurately estimated. The constituents of protons and neutrons, the "up" and "down" quarks, contribute an amount of about 1/(1 kilometer)$^2$ to the cosmological constant, for instance, and the Higgs field contributes an even larger amount, roughly 1/(10 centimeters)$^2$.

Each of the terms that contributes to the cosmological constant depends on the parameters of the standard model in a distinct and independent way. If we assume that the parameters of the standard model are really free and independent (an assumption we are continually checking in our search for deeper structure), it seems unlikely that these apparently unrelated terms would cancel one another. As a consequence it seems reasonable to assume that the total cosmological constant will be at least as large as or larger than the individual terms we can compute. Such an argument is too crude to predict whether the cosmological constant should be positive or negative, but we would conservatively estimate that its magnitude should be at least 1/(1 kilometer)$^2$, that it could well be something on the order of 1/(10 centimeters)$^2$ and perhaps that it is even larger. In other words, we expect the gravitational effects of a nonzero vacuum energy density to appear as distortions in spacetime geometry over distances of one kilometer or less.

It does not require any sophisticated experimentation to show that the theoretical estimate we have just



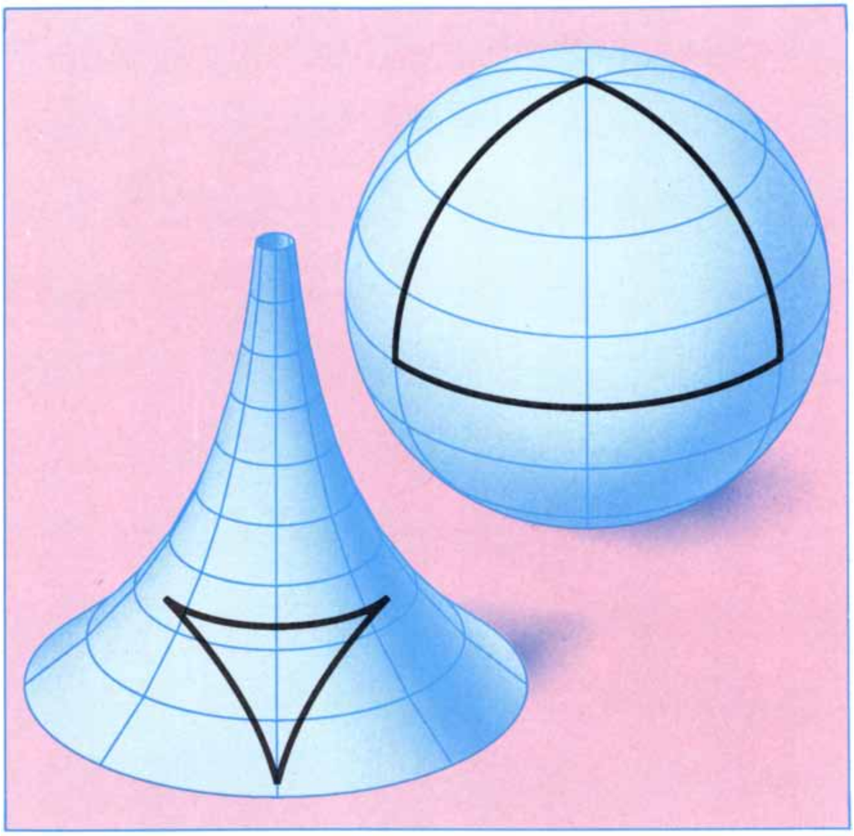

**HIGGS FIELD, if it exists, would make a particularly large contribution to the energy density of the vacuum. The Higgs field is the conjectured field corresponding to the particle called the Higgs boson, which is thought to give rise to particle masses. Here the Higgs potential—the part of the vacuum energy density that depends on the value of the Higgs field—is plotted against the value of the field, $\phi$. Although the Higgs potential is completely symmetric about the vertical axis, the vacuum must break the symmetry by choosing a certain position in the trough (*ball*). Such a selection is known as spontaneous symmetry breaking, and it plays a key role in the standard model: the theory that describes elementary particles and their interactions.**

given is wildly wrong. We all know that ordinary Euclidean geometry provides a perfectly adequate description of space over distances much greater than one kilometer. While walking around the block none of us has ever noticed large distortions in the spacetime structure of our neighborhood. If the magnitude of the cosmological constant were as large as our standard model estimate, ordinary Euclidean geometry would not be valid over distance scales of one kilometer or even less. If the cosmological constant were negative with a magnitude of $1/(1 \text{ kilometer})^2$, then the sum of the angles of a triangle with sides on the order of one kilometer would be significantly less than 180 degrees, and the volume of a sphere of radius one kilometer would be significantly greater than $4\pi/3$ cubic kilometers.

A positive cosmological constant of order $1/(1 \text{ kilometer})^2$ would have even more bizarre consequences. If the cosmological constant were that large, we would not be able to see objects more than a few kilometers away from us owing to the tremendous distortions in spacetime structure. In addition, if we walked farther than a few kilometers away from home to see what the rest of the world looked like, the gravitational distortion of spacetime would be so great that we could never return home no matter how hard we tried.

What if the cosmological constant is nonzero but quite small? In this case we would have to look over large distances to see its effects on spacetime structure. Of course, we cannot draw triangles the size of the universe and measure their angles, but we can observe the positions and motions of distant galaxies. By carefully charting the distribution and velocities of distant galaxies, astronomers can deduce the geometric structure of the spacetime in which they exist and move.

It has long been recognized that the dominant source of gravitational distortion in the spacetime geometry of the universe at large scales appears to be the energy density of matter and not that of the vacuum. Although the energy density of matter and that of the vacuum both affect the geometric structure of the universe, they do so in different and distinguishable ways. Numerous observations have shown that the galaxies in the universe are moving away from one another, a fact that is one of the cornerstones of the expanding universe in the "big bang" cosmology currently accepted. The ordinary gravitational attraction among galaxies tends to slow this expansion. As the galaxies get farther away from one another their gravitational attraction weakens, and so the rate at which the expansion slows decreases with time. Thus the effect of ordinary matter on the expansion of the universe is to decelerate the expansion at an ever decreasing rate.

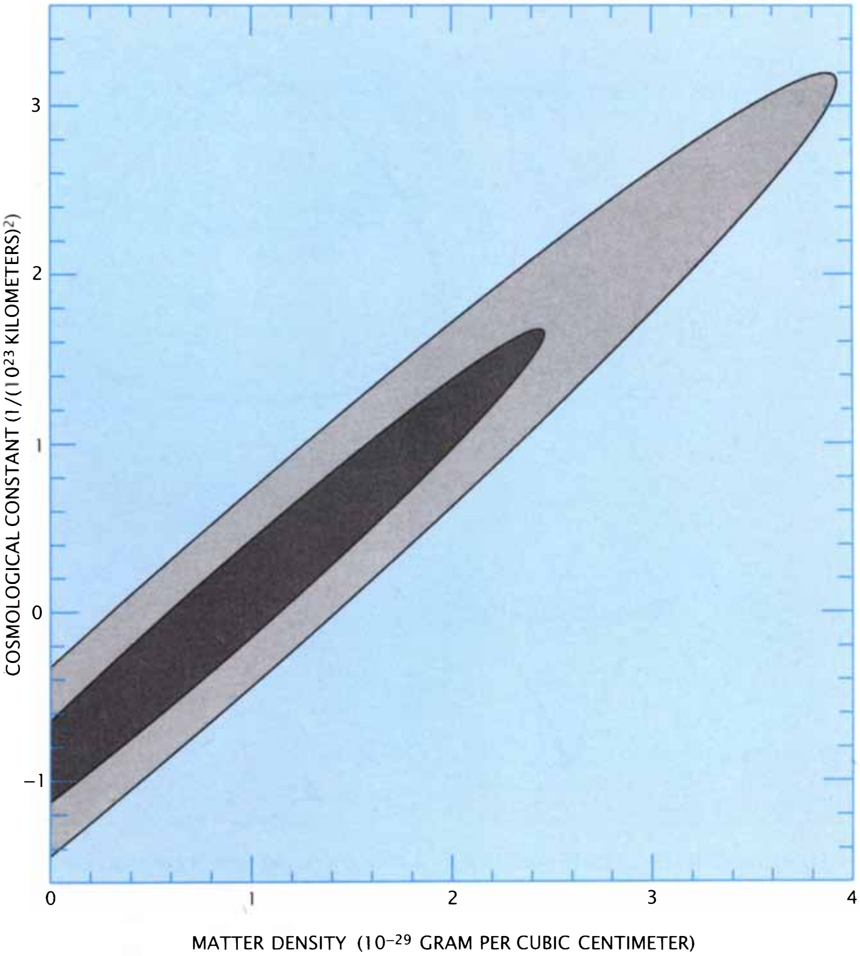


GEOMETRIC DISTORTIONS produced by a nonzero cosmological constant can affect both space and time. Here the effect on spatial geometry is shown, assuming that the distortions are independent of time. A negative cosmological constant would produce a space with negative constant curvature (*left*); a positive constant would produce positive constant curvature (*right*). (The positive case corresponds to the illustration on page 107.) In a space with negative curvature the sum of the angles of a triangle would be less than 180 degrees; with positive curvature the sum would be greater than 180.

What effects would a nonzero cosmological constant have on the expansion rate of the universe? A negative cosmological constant would tend to slow the expansion of the galaxies, but at a rate that is constant, not decreasing with time. A positive cosmological constant, on the other hand, would tend to make the galaxies accelerate away from one another and increase the expansion rate of the universe. Comprehensive studies of the expansion rates of distant galaxies show no evidence for either a positive or a negative cosmological constant.

A good example of how astronomers can measure the geometry of the universe and look for a nonzero cosmological constant is provided by the recently published work of Edwin D. Loh and Earl J. Spillar of Princeton University. Their survey counts the numbers of galaxies in regions of a specific size at various locations in space. If we assume that on the average the number of galaxies per unit volume is the same everywhere, then by counting galaxies in a region we are estimating the volume of that region. By measuring volumes of regions far from us we are determining the relation between distance and volume over very large scales and at earlier times, since the light from distant galaxies takes a long time to reach us—billions of years in the case of this survey.

Although such surveys contain many subtle sources of potential error, the results differ so startlingly




**COSMOLOGICAL CONSTANT has been probed by counting the number of galaxies in regions of the universe and thereby determining the geometry of those regions. The graph plots allowed values of the cosmological constant versus the matter density of the universe. (The black area corresponds to values that are allowed with a confidence of 67 percent; the gray area is a region of 95 percent confidence.) The units are approximate, but the graph shows that the magnitude of the cosmological constant must be less than about $1/(10^{23}$ kilometers$)^2$, some 46 orders of magnitude smaller than the value predicted on the basis of the standard model. The graph is from an analysis by Edwin D. Loh of Princeton University, based on work with Earl J. Spillar, also of Princeton.**

from our theoretical estimate that errors of a factor of two or even 10 are fairly insignificant. All galactic surveys agree that there is no evidence for any spacetime distortions due to a nonvanishing cosmological constant out to the farthest distances accessible to astronomers, about 10 billion light-years, or $10^{23}$ kilometers. This implies that the magnitude of the cosmological constant must be smaller than $1/(10^{23}$ kilometers$)^2$. Our theoretical estimate suggesting a magnitude larger than $1/(1$ kilometer$)^2$ is incorrect by, at the least, an astonishing factor of $10^{46}$. Few theoretical estimates in the history of physics made on the basis of what seemed to be reasonable assumptions have ever been so inaccurate.

The stupendous failure we have experienced in trying to predict the value of the cosmological constant is far more than a mere embarrassment. Recall that the basic assumption we used to obtain our estimate was that there are no unexpected cancellations among the various terms in the sum determining the total energy density of the vacuum. This expectation was based on the assumed independence of the free parameters of the standard model. Clearly this assumption is spectacularly wrong. There must in fact be a miraculous conspiracy occurring among both the known and the unknown parameters governing particle physics, so that the many terms making up the cosmological constant add up to a quantity more than 46 orders of magnitude smaller than the individual terms in the sum. In other words, the small value of the cosmological constant is telling us that a remarkably precise and totally unexpected relationship exists among all the parameters of the standard model, the bare cosmological constant and unknown physics.

A relationship among the free parameters of the standard model is just what we seek in our quest to discover deeper and more predictive theories. How could such a complex relationship among what we thought were free and unconstrained parameters arise, and what does it mean?

In answering this question it is well to keep in mind two examples from an earlier period in the history of physics. In the mid-19th century the speed of light had been measured and theories existed describing electric and magnetic phenomena, but it had not yet been shown that light propagation is an electromagnetic effect. Several physicists noticed, however, a curious relation between the speed of light and two parameters that enter into the equations for electric and magnetic phenomena. In modern notation what they noticed was that the electromagnetic permittivity constant $\epsilon_0$ and the magnetic permeability constant $\mu_0$ could be combined in the form $\sqrt{(1/\epsilon_0\mu_0)}$, yielding a quantity that is numerically equal to the measured velocity of light (at least within the rather large experimental errors of that time).

The workers appreciated the fact that this was either a miraculous numerical coincidence or evidence of a fundamental and as yet undiscovered relation between electromagnetic phenomena and light. James Clerk Maxwell was also aware of this numerical curiosity, and it served as an important inspiration for him in showing, through the set of equations now bearing his name, that the propagation of light is indeed profoundly related to electric and magnetic phenomena.

Does the remarkable relation among the parameters of the standard model implied by the small value of the cosmological constant suggest that a wonderful unifying theory awaits our discovery? Before jumping to such a conclusion, I should like to relate another example from the history of electromagnetic theory.

After Maxwell had incorporated light propagation into electromagnetic theory it was generally assumed that light waves traveled through a medium known as the ether. Using an interferometer, Albert A. Michelson and Edward W. Morley attempted to measure the velocity of the earth as it traveled through the ether. They found that the relative velocity was zero: the velocity of the earth and the velocity of the ether were identical. This is another relationship involving what was then thought to be a fundamental parameter of nature, namely the velocity of the ether. Did the discovery point the way to a unified theory relating a fundamental property of electromagnetism to the motion of the earth?

Although the idea that the ether drifted with the earth was suggested, the zero result of the Michelson-Morley experiment is actually explained by Einstein's special theory of relativity, which showed that the conception of the ether being used in that era was inconsistent with the symmetries of space and time. No theory providing a fundamental relation between the velocity of the ether and something as idiosyncratic as the velocity of the earth has survived. That is hardly surprising. The velocity of the earth is affected by many things—the shape and size of its orbit around the sun, the mass of the sun and the motion of the sun in the galaxy, for instance—that seem completely unrelated to issues in the theory of electromagnetism. There is no fundamental relation between the velocity of the ether and the velocity of the earth because the ether itself as the 19th-century theorists imagined it does not even exist.

In both examples a surprising relation between parameters of nature foreshadowed dramatic and revolutionary new discoveries. We have every reason to believe the mysterious relation implied by the vanishingly small value of the cosmological constant indicates that discoveries as important as these remain to be made. The two examples we have considered are quite different. The first relation, which involves two parameters of electromagnetism and one from light propagation, is what physicists today would call a "natural" relation: one that involves a small number of well-known parameters. The existence of a natural relation may indicate that a unifying theory exists, and, more important, it suggests that such a theory can be discovered.

The second example, in which the velocity of the ether was related to the velocity of the earth, is what today would be called an "unnatural" relation: one that involves many parameters, some of which are unknown or even unknowable. It seems unlikely, for instance, that we will ever know and understand all the many factors that determine what the velocity of the earth is in relation to the distant galaxies. Any unified theory developed to account for an unnatural relation would have to explain the values of many known and unknown parameters all at once. It seems quite unlikely that such a theory could be discovered even if it did exist.

Our example indicates that an unnatural relation suggests a deep misunderstanding about the essence of what is being measured and related, rather than the existence of an underlying unified theory. As a consequence an unnatural relation may point to an even more dramatic revolution in our thinking than a natural one would.

If we discount the possibility that the vanishingly small value of the cosmological constant is accidental, we must accept that it has profound implications for physics. Before we launch into constructing new unified models, however, we must face the dilemma that the relation implied by the vanishing of the cosmological constant is unnatural. The miraculous cancellations required to produce an acceptably small cosmological constant depend on all the parameters relevant to particle physics, known and unknown. To predict a zero (or small) value for the cosmological constant, a unified theory would face the imposing task of accounting for every parameter affecting particle physics. Even worse, achieving a sufficiently small cosmological constant requires that extremely precise (one part in $10^{46}$ or more) cancellations take place; the parameters would have to be predicted by the theory with extraordinary accuracy before any improvement in the situation regarding the cosmological constant would even be noticeable. Constructing such a theory, even if it does exist, seems to be an awesome if not impossible task.

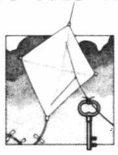
Although certain theories of the "ether drift" variety have been proposed, most efforts concerning the cosmological constant now focus on finding the underlying misunderstanding, the missing piece of the standard model or the misconception about the vacuum, which once understood will either eliminate the problem or at least turn it into a natural one. As long as the problem of the cosmological constant remains unnatural, the only hope we have for finding a solution is to stumble on an all-encompassing theory capable of accounting for all particle-physics parameters with nearly perfect accuracy. If we can change the relation required to produce an acceptably small vacuum energy density into a natural one, then, even though we have not yet accounted for its value, we at least reduce the issue of the cosmological constant to a more manageable problem involving a reasonable number of known parameters that only have to be predicted with a moderate degree of accuracy. There is little to report to date about this effort. In spite of a lot of hard work and creative ideas we still do not know why the cosmological constant is so small.

Even though nature does not, in the words of Aristotle, "abhor a vacuum," perhaps it does abhor a vacuum that is not empty. By introducing the ether in the early days of electromagnetic theory, Maxwell and others cluttered the vacuum with a hypothetical fluid that had complex properties. Michelson and Morley showed that this view of the vacuum was inconsistent with experimental reality, and Einstein showed that it was inconsistent with the symmetries of the universe.

Quantum field theories also fill the emptiness of the vacuum, this time with quantum fluctuations and fields rather than ether. These modern forms of clutter are consistent with the special theory of relativity, but they seem to cause problems when they are viewed in the framework of the general theory. With the mystery of the cosmological constant, perhaps we are again paying the price for dumping too much into the vacuum. The standard model, which has a large number of fluctuating quantum fields including a Higgs field, is a particularly egregious polluter of the vacuum. There is no doubt that the resulting theory is a beautiful and highly successful structure, but it may be based on a conception of the vacuum or of spacetime that is flawed. It is our challenge to repair that faulty foundation without destroying the towering edifice we have built on it.

# THE AMATEUR SCIENTIST

*Icicles ensheathe a number of puzzles: just how does the water freeze?*



by Jearl Walker

Those icicles that were drooping from eaves and telephone lines not long ago may have appeared to be simple constructions, but in fact their shapes and the way they develop have long perplexed people who investigate them. Why is an icicle usually cone-shaped, with a tip no wider than a few millimeters? Why does a narrow, liquid-filled tube of ice extend several centimeters up the center of the icicle from the tip? (You can probe the tube with a toothpick.) What accounts for the white line that traces the central axis of the column? Why do horizontal ribs spaced about a centimeter apart develop along the sides of the column? Why is the ice solid in some places but spongy in others? What bends and twists some icicles?

The physics involved in any freezing of water is always richly complex. In a simple model the freezing interface between liquid and ice consists of dendritic fingers of ice that stretch into the liquid. Along the fingers molecules in the liquid gradually join the crystal structure of the ice by giving up part of their energy and becoming immobile. The lost energy, called heat, is conducted to some region colder than the freezing interface—often to nearby cold air.

Icicles grow by such a process. One way they begin to grow has been studied by Norikazu Maeno and Tsuneya Takahashi of Hokkaido University. Picture a slowly melting blanket of snow on a rooftop. When water first trickles over the edge of the roof, it forms a pendent drop in the cold air. As the sides of the drop begin to freeze, forming a thin shell of ice, the heat released by the freezing is transferred to the air and to the roof edge.

As more meltwater runs down the ice shell, part of it freezes on the way, widening the fledgling icicle. The rest joins the drop dangling at the bottom of the structure [*see upper illustration on page 116*]. The gradual freezing of the sides of the drop lengthens the icicle. If the drop ever becomes too large—somewhat wider than five millimeters—it falls, but additional meltwater soon adds a new drop. As long as the supply of meltwater continues, the icicle grows wider and longer. The tip, whose width is set by the diameter of the pendent drop, remains narrow.

H. Hatakeyama of the Tokyo District Meteorological Observatory and S. Nemoto of the Meteorological Research Institute in Tokyo reported another way an icicle might be initiated [*see lower illustration on page 116*]. The top part of the initial drop may freeze uniformly, creating a horizontal freezing interface that moves downward. If the water supply is feeble and the roof edge is cold, all the water may freeze rather than just a thin shell. The structure may then lengthen in steps as successive meltwater drops develop at the bottom and freeze solid. If, however, there is enough water to maintain a pendent drop, at some stage in the growth of the icicle the sides of the drop will freeze and form an ice shell, as in Maeno and Takahashi's scheme.

Whenever an ice shell forms, the liquid inside the shell from then on freezes only slowly. According to Lasse Makkonen of the Technical Research Center in Espoo, Finland, the heat released by the internal freezing is conducted through the ice to the top of the icicle (called the "root") and then to the edge of the roof. The conduction is so gradual that the internal freezing interface may move down the central axis of the icicle very slowly; if the interface is well separated from the root, as it is in the case of a mature icicle, it may even be stationary.

From the interface to the tip of the icicle, liquid is trapped in a narrow ice tube. In spite of its weight the liquid is stable, in part because of the surface tension between it and the tube's walls. In addition the tube is so narrow that chance disturbances along the bottom of the water column or in the pendent drop are usually insufficient to allow air to seep up into the tube to drain the liquid. In typical winter temperatures the interior freezing interface can reach the tip, leaving a completely frozen icicle, only if the meltwater supply is cut off and growth at the tip ceases.

The external surface of the ice is sheathed by a thin layer of liquid [*see illustration at top right on page 117*]. Freezing at the external ice-liquid interface is rapid, because the heat released by the freezing is quickly conducted through the liquid and lost to the air. (Maeno and Takahashi find that the liquid sheath on active icicles is no thicker than .1 millimeter.) The temperature at the freezing interface is the freezing point of water, which is zero degrees Celsius for pure water but may be lower if the water is impure. The temperature in the rest of the water layer is slightly lower than the freezing point, a condition known as supercooling. The coldest water borders the air, which may of course be considerably colder than the water.

Charles A. Knight of the National Center for Atmospheric Research in Boulder, Colo., points out that icicles can grow in air warmer than the freezing point provided the air has little water vapor. The scarcity of water vapor promotes the evaporation of water from the external surface of the water sheath. As water molecules escape from the liquid, they carry away energy, supercooling the water surface. The chilled water acts as a heat dump, into which water freezing along the ice surface sends its heat.

Why is the heat released along the internal freezing interface not conducted horizontally through the thin layer of ice separating it from the air? The reason is that both the internal and the external freezing interface are at the freezing point of water. With no temperature difference, conduction is eliminated. If the internal interface is to advance, the released heat must be conducted to the root of the icicle.

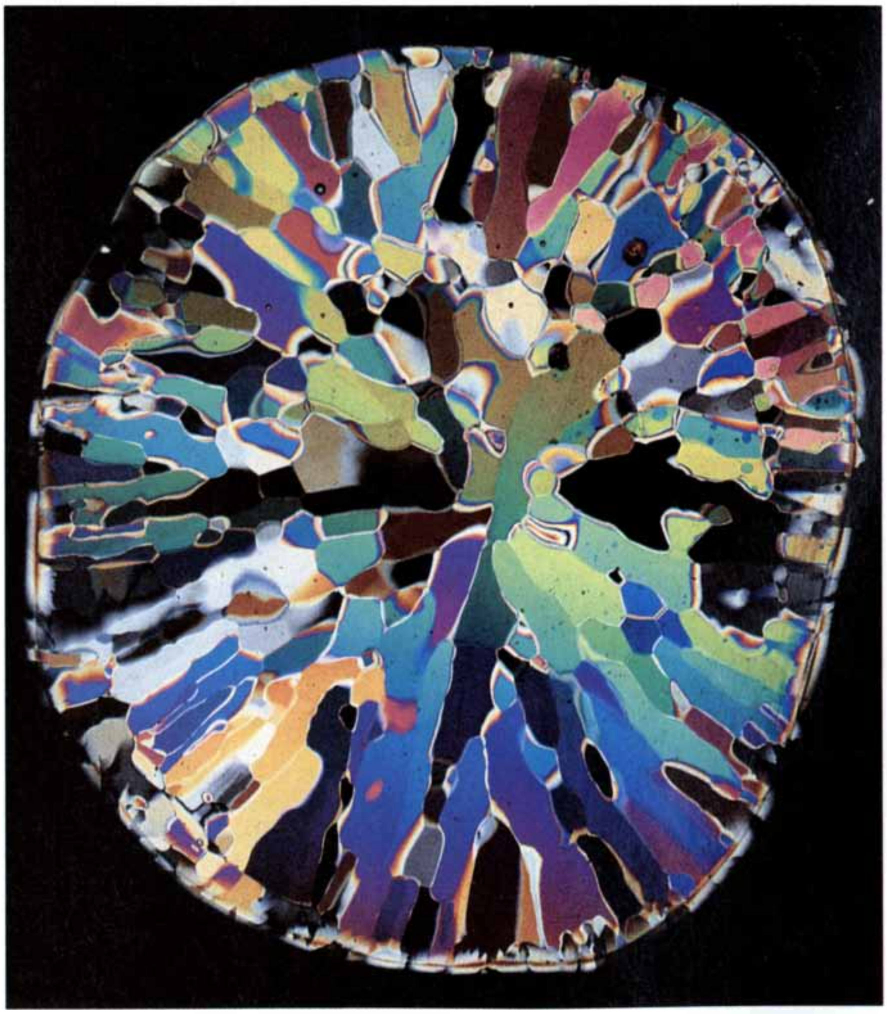
When the water in or on an icicle freezes, air is driven out of solution, forming bubbles imprisoned in the ice. The most striking bubbles are the tiny ones that develop along the central axis as the internal freezing interface descends. When the icicle is illuminated by white light, some of the light scatters from the bubbles, giving the appearance of a white line along the central axis.

According to Maeno and Takahashi, an icicle lengthens between eight and 32 times faster than it widens. I think one reason for the dissimilar growth rates is that more water collects at the tip than at any place along the sides. Another reason may have to do with the way ice crystals grow. The basic geometry of an ice crystal is a thin hexagonal plate. The central axis perpendicular to the plane of the plate (the basal plane) is called the *c* axis. As molecules join the plate, it grows more rapidly in the basal plane than along the *c* axis. If crystals in an icicle are oriented with the *c* axis pointed radially outward, perpendicular to the freezing interface and thus approximately perpendicular to the central axis of the icicle, then the icicle should be expected to lengthen faster than it widens.

Early research about the dominant orientation of icicle crystals was often contradictory. Some investigators even argued that there is no dominant orientation—that an icicle is a hodgepodge of small, randomly oriented crystals. In an attempt to settle the matter, Robert A. Laudise and Robert L. Barns of the AT&T Bell Laboratories in Murray Hill, N.J., set out with a group of young volunteers to examine icicles collected from nearby houses. Icicles with diameters larger than about an inch were sawed from their perch. The specimens were then sawed again to provide both horizontal and vertical cross sections. Working in a cold room, the investigators reduced each slice to a thickness of about an eighth of an inch by rubbing it on a thick plate of aluminum that had been warmed with hot water.

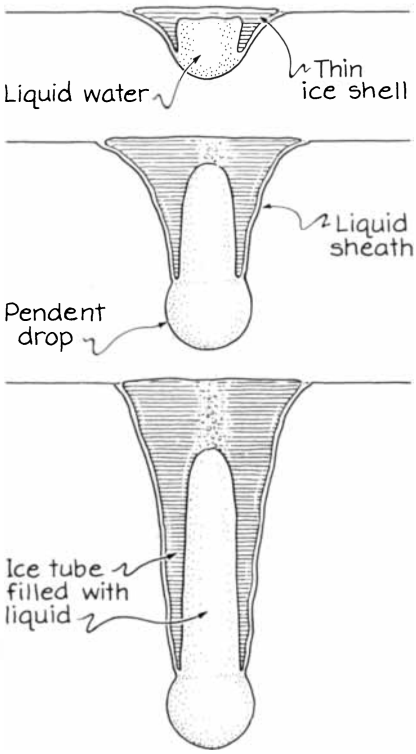
Inspection involved two "crossed" polarizing filters mounted on a platform that had a glass top. (The arrangement was described in this department in July, 1986.) In this procedure one filter is laid on the glass and a specimen slice is placed on it; a second filter is laid over the slice and is rotated so that its direction of polarization is perpendicular to that of the first filter. Lamplight scattered from white cardboard under the platform becomes polarized when it passes through the lower filter.

In the absence of the icicle specimen, the light could not pass through the second filter. When, however, the light passes through an ice crystal, its polarization is rotated and so—de-

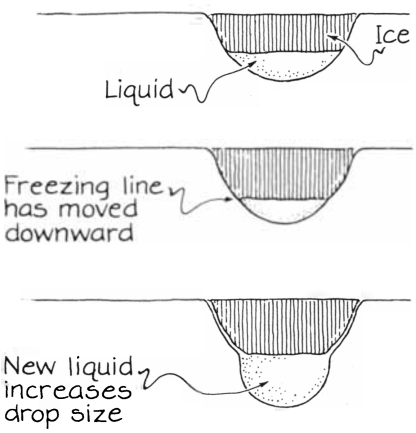


***Ribbed icicles* (left) *and a section photographed through crossed polarizing filters***

pending on the extent of the rotation—all or some of the light does pass through the second filter. When you peer down through the arrangement, the individual ice crystals differ in brightness depending on how much each one rotates the polarization of the light passing through it. If the slice is thin enough, the regions are even colored differently. If the slice is rotated about the vertical, the brightness of each crystal changes unless it happens to have its *c* axis approximately along your line of sight, in which case it remains dark throughout the rotation. By thus inspecting an icicle slice with polarized light, you can detect both the size of the crystals and the orientation of their *c* axis.




*The thin-shell mode of growth*




*The uniform-freezing mode of growth*

When Laudise and Barns examined specimens from some 60 icicles, they found that the width of the crystals varied from less than .8 millimeter to more than 20 centimeters. Some icicles consisted of a multitude of small crystals randomly oriented; others had large single crystals whose *c* axis seemed never to be aligned with the icicle's central axis.

Knight later added the idea that the orientation of the *c* axis of the large crystals may result from recrystallization as the icicle grows. As heat is supplied from the freezing, initially small and randomly oriented crystals may change their boundaries and merge, becoming oriented as Laudise and Barns observed. Variations in the air temperature or the supply of sunlight and meltwater may also play a role. Knight cautions, however, that crystal orientation may actually have little influence on the shape of an icicle. Instead the dissimilar growth rates may be due to the way heat is removed from an icicle, somewhat as modeled by Makkonen.

The relation between the growth rate of an icicle and the rate at which meltwater is supplied is surprising. According to Maeno and Takahashi, the rate of widening is largely independent of the supply rate; the rate of lengthening, on the other hand, is actually slowed by an increase in the supply rate. The first finding suggests that the liquid sheath maintains a constant thickness, thereby regulating the rate of heat conduction through it and the rate of freezing on the ice surface. The second finding is more difficult to understand. Is the lengthening slowed by heat that is released when a larger supply of water arrives at the drop and then is supercooled? Does the more frequent detachment of drops reduce the chance that drops at the tip will freeze, or does it somehow disrupt the ice growth?

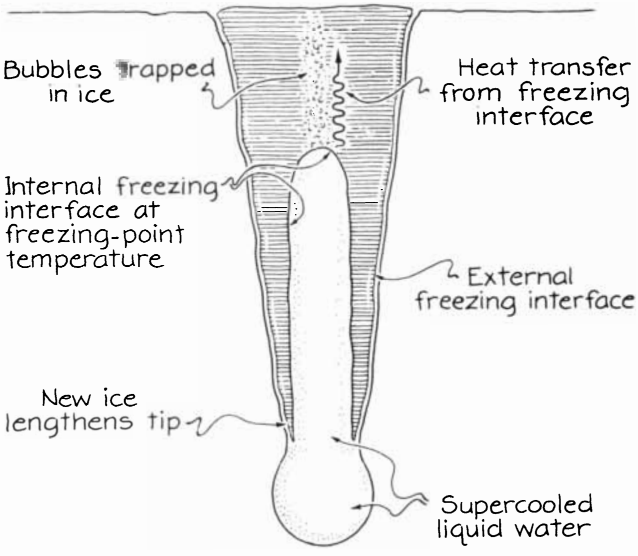
Ira W. Geer of the State University College at Brockport in New York found that an icicle tip is sometimes so irregular that air spurts up into the ice tube, dumping the liquid. The tube is soon refilled by liquid that has drained to the tip along the exterior of the icicle and then is drawn up the tube by surface tension. Usually a large air bubble remains trapped in the upper reaches of the tube, later becoming buried in ice as the interior freezing interface descends.

Geer investigated the growth of an icicle by photographing it every five minutes for about 90 minutes. From the negatives he constructed a composite drawing to learn whether the horizontal ribs along the sides of the icicle migrate vertically in the course of growth. He found that the ribs are approximately stationary; they are regions that grow outward faster than the intermediate hollows do. Maeno and Takahashi offer two explanations for the fast growth of the ribs. The liquid layer is thinner on a rib than it is in the adjacent hollows, allowing the heat from the freezing on the ice surface to be conducted quickly to the external surface of the water. The rib is also more exposed, enhancing heat transfer to the air and perhaps the evaporation of water.

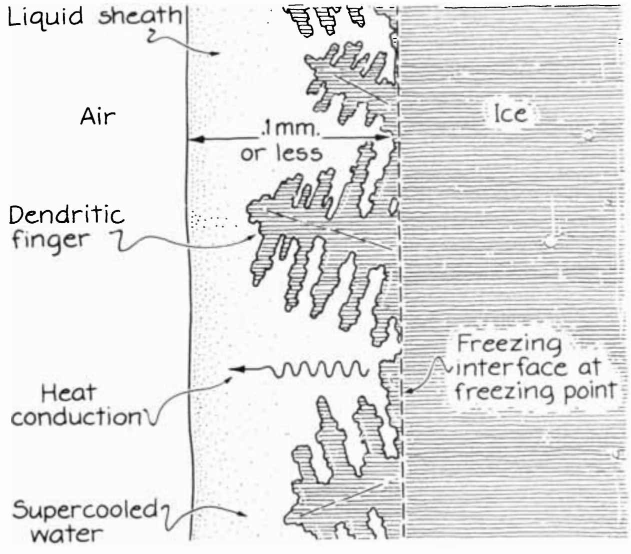
Evidence for the rapid growth of the ribs can be seen in horizontal sections cut from an icicle. A section from a rib region displays abundant air bubbles collected in bands that resemble the growth rings in a cross section cut from a tree. Each band in the icicle is created when the rib grows outward, trapping air bubbles in ice before they can escape. Presumably the bands are linked to variations in air temperature and water supply.

Knight found that the ribs are usually solid ice, whereas the hollows are often spongy: they consist of liquid separating sheets of ice that are extensions of the basal planes of the crystals. He reports that he could insert a knife blade a centimeter or more into the spongy sections, forcing out some of the liquid. Knight suggests that the spongy sections develop during periods when heat loss by the icicle is rapid and the water supply is abundant, but he thinks more study is needed.

Mature icicles also develop vertical ridges, which are often complex and branched. They may be due to wind that increases the removal of heat, and therefore the rate of freezing, on one side of an icicle. Another factor is the possible asymmetry of the water supply. On a wide icicle water may flow along a narrow trail instead of wetting the icicle uniformly. Geer noticed that water draining along a dormant icicle tends to follow one track as long as the track remains wet. The ridge it produces may grow outward five millimeters or more before the water veers off onto a new track. A drop of water may follow a wet




*The ice tube at the tip*




*Heat transfer through the water sheath*

track because it encounters less surface tension there than it encounters on a dry, icy region.

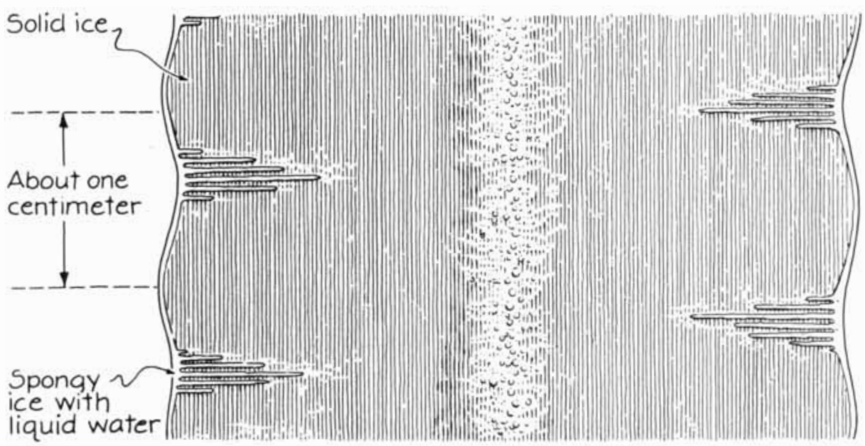
Icicles are sometimes decorated with tiny spikes extending from the surface. Maeno and Takahashi measured spikes about a millimeter wide and as much as 20 millimeters long. The spikes are created when the water sheath freezes during a meltwater drought. When liquid water trapped under a skin of ice begins to freeze, its expansion can rupture the skin, ejecting the remaining liquid outward from the surface; the ejected liquid freezes to form a spike. Are the spikes more likely to form in the initially spongy ice where liquid is slow to freeze?

I noticed a different kind of spike on an icicle tip. The collapse of the liquid bridge between a detaching drop and the liquid remaining on the tip frequently left a thread of water that immediately froze. The ice thread, less than a millimeter thick and only two or three millimeters long, was twisted—evidence of the rapid freezing. A touch of my bare finger would melt the thread, and so would the formation of the next pendent drop.

Icicles are often crooked, asymmetric or misaligned from the vertical. Maeno and Takahashi offer several reasons for such growth. Complex shaping can develop from an asymmetric supply of water, particularly on wide icicles. A steady wind can also bias the growth by pushing the pendent drops leeward and by increasing the heat loss on the windward side of the icicle. If the icicle grows on a branch, the gradual addition of weight bends the branch, curving the icicle as the fresh growth at the tip continues to follow the vertical. Similarly, if the icicle grows on the edge of a roof, partial melting of the ice and snow on the roof may allow the layer to creep and curve over the edge, changing the direction of the icicle's growth. To this list of explanations one should add the possibility that wind-driven snow may accumulate on the icicle and distort its shape.

Several helpful suggestions for the study of icicles are available [see "Bibliography," page 126]. For example, Geer wonders how icicle growth might be influenced if an impurity such as salt, soap, alcohol or food coloring were added to the water supply. What would the bubbles be like in an icicle grown with unsalted carbonated water? When a natural icicle becomes dormant, how does sublimation gradually smooth its surface?

More projects can be added to the list. How might a nail embedded in the ice tube at the tip alter the icicle's later growth? What is the temperature distribution inside a mature icicle? What factors determine the length of the ice tube and the size and spacing of the horizontal ribs? Does infrared radiation from nearby warm objects or air turbulence on the leeward side of an icicle bias its shape? How do icicles grow from seawater or freshwater spray? If you are prepared for painstaking work, you might try to correlate the growth rate of an icicle with the ambient air temperature and the wind speed. I shall be interested to hear about your observations of these often ignored sculptures of ice.




*Horizontal ribs along the sides of the icicle*

# COMPUTER RECREATIONS

## *The invisible professor holds a chalk-talk session on the display monitor*

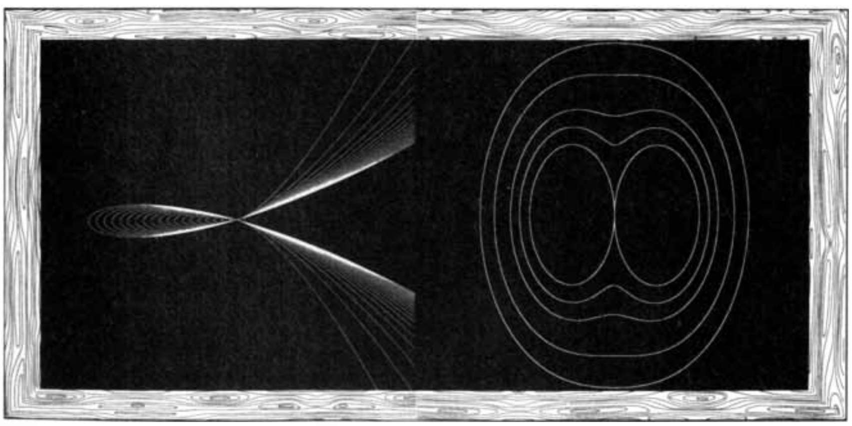
by A. K. Dewdney

If the darkened screen of a computer display is like a chalkboard, who is drawing the curves that arc gracefully across its matte surface? The hand must belong to the invisible professor, an electronic incarnation of all the academicians who have ever lovingly sketched the curves of analytic geometry. These curves have romantic and mysterious names such as the witch of Agnesi, the kampyle of Eudoxus and the nephroid of Freeth. They include parabolas, cycloids, spirals and serpentines. Nearly forgotten when mathematics rushed to become general and geometry became non-Euclidean, the curves remain a testament to the forms that follow function in physics. The curves mark beautiful discoveries, some famous, some less than famous, by mathematicians of past centuries.

The invisible professor is a collection of the simplest computer programs imaginable. In drawing the curves it captures their beauty with a piquancy their discoverers could scarcely have imagined; a single, bright point is the new chalk. It traces an equation called Tschirnhausen's cubic, for example, by gliding onto the screen from the upper right, swooping through the origin, then looping up and back through the origin. The point burns its way to the lower right-hand corner of the display, where it is extinguished [*see left half of illustration below*].

If asked, the invisible professor will even draw a family of curves. The equation of Tschirnhausen's cubic has a special constant symbol. When different numbers are substituted into the symbol, different versions of the curve are drawn: the screen is filled with variations on a theme. The only thing missing is the professor's frail old voice: "It will be seen at once, from a mere glance at the equation, that this curve is a generalized, semicubical parabola."

The invisible professor can also render transcendental landscapes as undulating sheets of parallel curves and hint at complex oceans by making curvilinear distortions of grids. Such programs, which have been written by lovers of geometry, are distributed for a price.



***Twenty versions of Tschirnhausen's cubic* (left) *and five of the hippopede* (right)**

In order to lay down a minimal background on how to draw curves, it is my turn to be an invisible professor. The humble circle will serve as an example of a curve. Three principal formulations of it will lead to curve-drawing programs.

The first and most familiar formulation was encountered by the majority of us in high school:

$$x^2 + y^2 = c^2$$

The equation distinguishes points that lie on the circle from those that do not. A point that has Cartesian coordinates $x$ and $y$ lies on the circle if the square of $c$ (the radius of the circle) results when $x$ and $y$ are squared and added. Whenever the program is run, the user supplies a value for $c$.

The formula can be employed by a circle-drawing program if the curve is divided into an upper section and a lower one. The program runs through a series of $x$ values, plotting the corresponding $y$ values. To be useful, $y$ is best unsquared. Consequently the equation must be solved in terms of $y$. The two branches differ only in sign:

$$y = +\sqrt{c^2 - x^2}$$
$$y = -\sqrt{c^2 - x^2}$$

This formulation is already rather awkward, however. Perhaps polar coordinates would serve better.

The Cartesian coordinates $x$ and $y$ measure distance vertically and horizontally from a central point called the origin. Polar coordinates, $r$ and $\theta$, establish a point quite differently. The point lies $r$ units from the origin, and the line joining it to the origin makes an angle $\theta$ with the horizontal. The Cartesian equation of a circle is simply $r = c$. No matter what the angle $\theta$ may be, the distance $r$ must be $c$. Unfortunately few readers will be using languages that employ polar coordinates for plotting.

The third principal presentation involves so-called parametric equations. For many curve-drawing programs these are ideal:

$$x = c \cdot \cos(t)$$
$$y = c \cdot \sin(t)$$

Here a parameter $t$ runs from zero to 360 degrees. At each degree, say, the $x$ and $y$ values are plotted. Readers not familiar with the functions sine and cosine may at least discover how to use them by consulting their language manual.

At the algorithmic level a bareboned version of the program I call CIRCLE might appear as follows:

```
input c
for t← 1 to 360
  x← c•cos(t)
  y← c•sin(t)
  point (x,y)
```

The exact form employed will not differ greatly from the formulation shown. Similarly, there will be a plot command somewhat like the one immediately above.

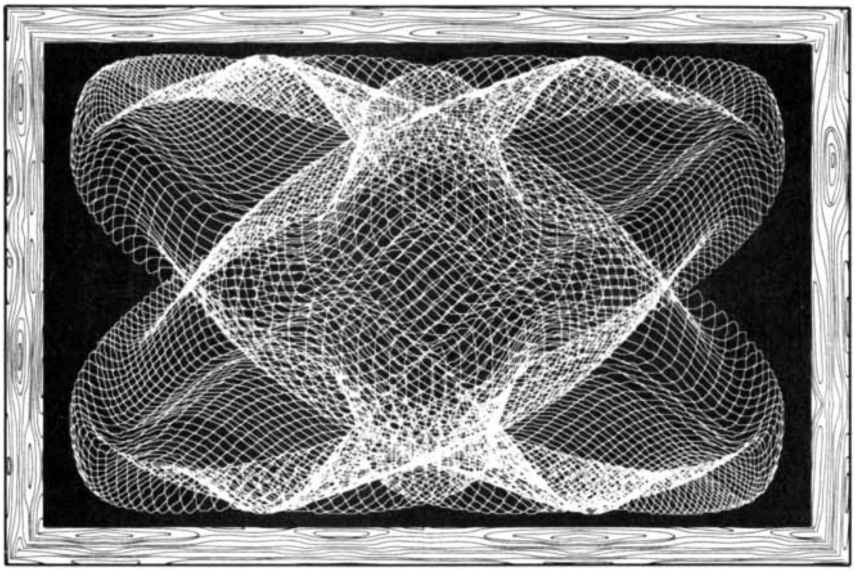
The circle-drawing program will not, of course, be quite as simple. For one thing, most personal computers have the origin in one corner of the screen. Offsets must be added to place the circle in the center. A screen that has 200 points horizontally and 150 points vertically, for example, will need the following modification in the program:

```
x← c•cos(t) + 100
y← c•sin(t) + 75
```

The invisible professor can now be incarnated through a host of little programs that all share the same basic structure just outlined. Only the equations are different. Tschirnhausen's cubic, for example, has parametric equations that do not involve sine or cosine:

$$x = 3a(t^2 - 3)$$
$$y = at(t^2 - 3)$$

The miniprogram called TSCHIRNHAUSEN therefore employs a parameter $t$ that runs from a minimum value to a maximum value, both determined by screen size. The 150-by-200 screen mentioned above, for instance, will accommodate the curve from $t = -4.4$ to $t = +4.4$ when $a$ is equal to 1. No digital computer can vary the parameter $t$ continuously through such a range of values; instead it must increase $t$ by increments small enough to give the impression of continuity. The chalk must not squeak on the board. Because this particular board has 200 points horizontally, a reasonable step size for $t$ might be obtained by dividing the range of values for $t$, namely 8.8, into 200 equal steps. The calculation yields .044. On the other hand, the curve is not traced out evenly. Some parts are drawn faster than others, resulting in separation of points. Fortunately computation of this kind is so cheap (and fast) that we can drop the step size to .01 in a perfectly cavalier fashion:



*Stanley S. Miller's MADNESS illustrates serendipity in constants*

```
input a
for t← -4.4 to +4.4 step .01
  x← 3a(t² - 3) + 100
  y← at(t² - 3) + 75
  point (x,y)
```

Another simple way to improve the smoothness of the curves involves the use of line segments. TSCHIRNHAUSEN and any of the other curve-drawing programs described here can be modified so that each time a new point is obtained a line-drawing command joins it to the point previously plotted. The coordinates of the old point are then replaced by those of the new. To the eye even a coarsely polygonal curve may look graceful.

In either version of the program one can add a new outer loop that varies the constant $a$ from .1 to 2 in steps of .1. In this case the invisible professor will draw a family of 20 beautifully nested cubic curves that would make Tschirnhausen proud.

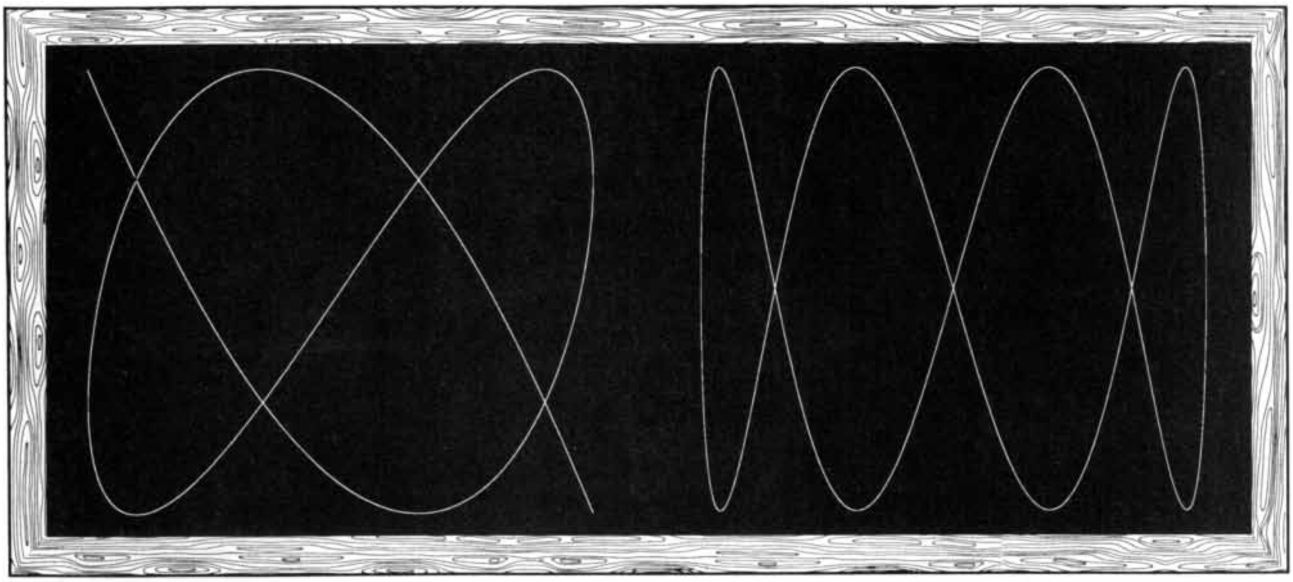
The hippopede, first investigated by the Greek philosopher Proclus in about A.D. 475, has two forms. It can appear as an 8 lying on its side or it can assume a dumbbell shape, depending on the values of its two constants, $a$ and $b$ [*see right half of illustration on opposite page*]. The parametric equations are only slightly complicated:

$$x = 2\cos(t)\sqrt{ab - b^2\sin^2 t}$$
$$y = 2\sin(t)\sqrt{ab - b^2\sin^2 t}$$

The parameter $t$ runs from −180 to +180 degrees. A small economy of computation is available in that the same square-root function appears in both equations. As a result the program HIPPOPEDE needs to compute the function only once, store it as a variable called *temp* and then multiply $2\cos(t)$ and $2\sin(t)$ by *temp*. Here again readers setting up their own invisible professor must consider the step size. Although it depends to some degree on $a$ and $b$, the screen size and other factors, a step size of one degree will work well in most cases. Again, different values of the constants $a$ and $b$ can be tried, anywhere from 1 to 10 with stops either at integer or at fractional values. The wise old Athenian owl will be found staring out from the screen if for each of four curves $a$ is set to 20, 25, 30, 40 and 50 and $b$ is held at 20.

Spirals, among the most beautiful forms found in nature, are second nature, so to speak, for the invisible professor. It is no surprise that they grow best in the polar medium. A point on the spiral of Archimedes has the polar coordinates $(r,\theta)$. The equation is, as mathematicians say, trivial: $r = a\theta$. The angle $\theta$ begins at 0 and can run as high as one's screen (and the constant $a$) will allow. Step size can be 1 or less. If $a$ is chosen to be .01, the spiral will be tightly wound; at .1 it will be much less so.

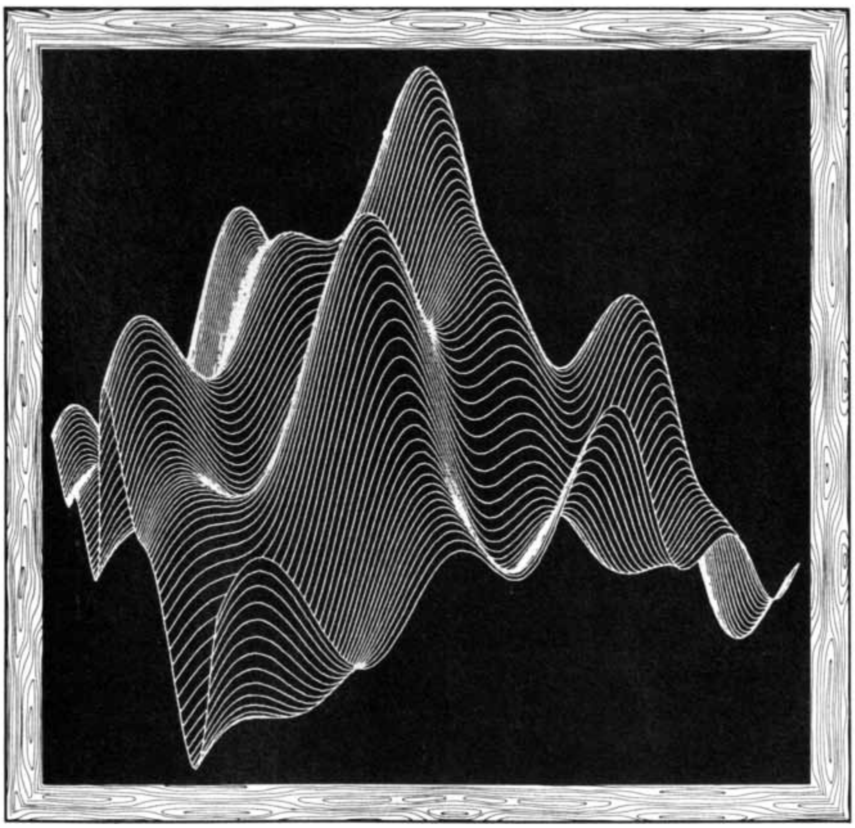
The invisible professor can also easily handle the famed logarithmic spiral (also called the equiangular spiral for the sake of people who are frightened by the idea of logarithms). The spiral inhabits seashells and the heads of sunflowers. The invisible professor uses the polar equation $r = e^{a\theta}$, where the transcendental



*Two Bowditch curves, one of which retraces itself*

number $e$ can be taken as 2.7183. The step size used with the Archimedean spiral can be used here, but readers must be warned about the awesome speed with which the exponential function grows: even when $a$ is equal to .01, the curve does not stay tightly wrapped for long.

There are hundreds of curves in a handy little paperback called *A Catalog of Special Plane Curves,* by J. Dennis Lawrence. The invisible professor will thrive on these forms. For more advanced programmers it is an interesting project to assemble all the curve-drawing programs one can write into a single package the user can access through a menu.



*A surface constructed from simple functions*

In some cases it is not necessary to incorporate several distinct programs into a single package. The magic of generalization leads to a single program that generates all the curves involved. It becomes simply a matter of selecting values for various constants that appear in the master formula. The circle, for example, is just one member of a class of forms called the Bowditch curves, or Lissajous figures, that arise from a single parametric equation:

$$x = a \cdot \sin(ct + d)$$
$$y = b \cdot \sin(t)$$

These curves are familiar to owners of cathode-ray oscilloscopes. Electronic signals that vary sinusoidally will produce such curves on the tube.

The constants $a$ and $b$ define a rectangle $a$ units long and $b$ units high. Within this rectangle a Bowditch curve will weave its wild dance. The constants $c$ and $d$ are the ones that make a difference. When $d$ is equal to 0, for example, values for $c$ of 3/5 and 1/4 produce the two curves shown above. If $c$ happens to be rational, a Bowditch curve will come back to its starting point, tracing itself for as long as the invisible professor does not run out of breath. There is no point in worrying about whether the numbers in one's computer are rational; a digital computer accepts only such numbers. (The foregoing remarks apply when the units of $t$ are radians. To convert the equations into degrees, one must multiply the input of the function *sin* by $\pi/180$.)

If the parameter $t$ starts at 0 degrees, a Bowditch curve begins somewhere on the $x$ axis. In general it weaves back and forth, in and out, before rejoining itself. But how large will $t$ get before that happens? Here is a small puzzle for rational readers. For puzzle dodgers, let $t$ get no bigger than 3,600 in steps of 1 and stick to rational numbers whose components are small. For the curious, humorous or simply insane, open-ended adventures are available; exhaust the old professor by setting no upper limit to $t$. One can then decide, by watching the drawing of the curve in process, how long it takes to rejoin itself. The loop one uses must be mechanically constructed in this case. At the bottom of the loop place a simple "go to" statement that redirects execution to the top of the loop. To save the elderly academician from an eternity at the chalkboard, it would be a kindness to insert an "if" statement in the loop, so that if a key is pressed, the program goes to an instruction outside the loop.

Among the factors prompting the invisible professor to come out of hiding was a letter from Stanley S. Miller of Concord, Mass. Miller, a management consultant by profession, has fallen in love with the trails left by various functions. It all began with a commercially available program called CURVES, about which I shall have more to say below.

"One late, rather winy evening," he writes, "I hit the wrong button and produced something entirely different from any cycloids I had generated before...I thought, at first, that some of the Brie I had been eating had got into my computer, but I ran it through again with different parameters [that is, constants], and produced a close relative." The encounter with serendipity gave Miller a glimpse of the profound power and variety possible with a curve-generating program that is reasonably general. I include a trophy of his hunt on page 118. It looks like a three-dimensional net thrown out to ensnare impressionable minds. If the curve needs a name, let it be called Miller's MADNESS. Its parametric equations have an innocent simplicity about them:

$$x = \sin(.99t) - .7\cos(3.01t)$$
$$y = \cos(1.01t) + .1\sin(15.03t)$$

I have no idea how large $t$ must become to complete the plot. Readers who want to view Miller's MADNESS on their own computer are advised to use the unending loop described above; the units of $t$ are radians.

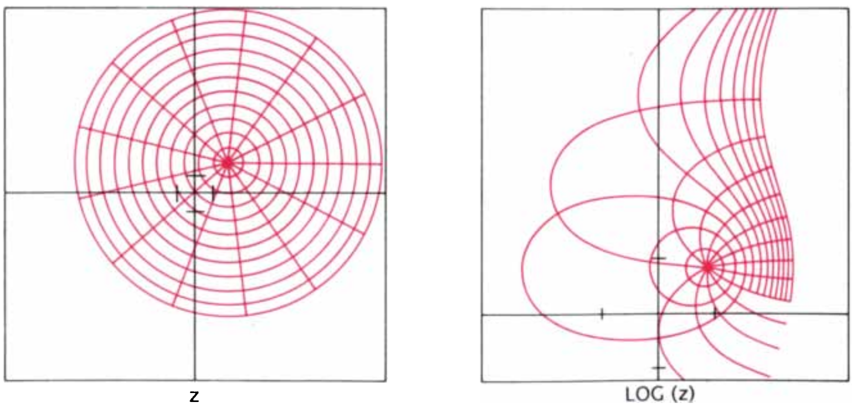
When programming MADNESS, the reader may as well make it general by putting constants in front of the four trigonometric functions. This will put a virtually infinite space at their disposal for exploration.

CURVES, the program that inspired Miller, is a highly general program produced by Bridge Software, P.O. Box 118, New Town Branch, Boston, Mass. 02258. Filled with useful options and sophisticated functions, it represents something like the ultimate invisible professor. But there is more. Bridge Software also sells a program called SURF, an output sample of which is shown at the bottom of the opposite page. Users of this program specify a function of two variables and then watch as the program draws a landscape in three dimensions, complete with mountains, valleys and hidden lines. Neither program is too pricy.

Bridge Software is not the only small company producing mathematical software. For lovers of complex functions, Lascaux Graphics, 3220 Steuben Avenue, Bronx, N.Y. 10467, has created a program called $f(z)$. Many will remember the Mandelbrot set featured in this department as recently as last November. Perhaps a few readers will remember that all the action there took place in the complex plane. The functions produced by Lascaux Graphics map the complex plane onto itself. We cannot visualize a four-dimensional space but we can almost feel a complex function when its effect on the complex plane is clearly seen; a regular polar grid [*see illustration below*] embedded in the plane is transformed by a complex function into something of curvilinear grace, like currents in a mathematical sea. Somewhat more expensive than the CURVES program, $f(z)$ is a bargain for those awaiting initiation into the higher mysteries of the chalkboard.

Nanotechnology, the topic of this department in January, drew responses from skeptics and enthusiasts alike. An amusing version of skepticism arrived from the Humanist Association of Oklahoma, an organization represented by Clinton L. Wiles of Oklahoma City. Wiles puts little store in the promise of eternal life that some enthusiasts see in nanotechnology. In particular, the nanosubmarine I took some delight in describing may never be built, much less deliver eternal life, according to Wiles. He writes: "Being skeptical enough not to have bought any supernatural eternities, we'll suggest that anyone investing in nanotechnology expectations be prepared to wait awhile for a return on the investment.... Even if submarines in the bloodstream is a godless idea, we'll file it as an incredible claim requiring incredible proof, right next to the supernatural eternal-life claims." Wiles would be glad to receive any and all opinions on the subject at P.O. Box 94043, Oklahoma City, Okla. 73143.

Eric Drexler is among the principal defenders of nanocomponents. Responsibly enough, he carefully avoids any claim beyond the ultimate prolongation of human life at the hands of his "engines of healing." Science-fiction author Paul Preuss of San Francisco admires Drexler and has studied some of the possibilities nanotechnology seems to offer. Preuss's latest book, *Human Error*, describes a hypothetical, self-replicating nanomachine that is accidentally ingested by a brilliant but unlovable investigator at a biotechnology company on the West Coast. The investigator changes in surprising ways and even learns about love.




***A complex function* (left) *and its logarithm* (right)**

# BOOKS

## *A celebration of design: Voyager, nautilus, the Veterans Memorial, the periodic table*

by Philip Morrison

**THE COMPLETE GUIDE TO RUTAN AIRCRAFT,** by Don and Julia Downie. Third edition. Tab Books Inc. (paperbound, $14.95). **VOYAGER,** by Jeana Yeager and Dick Rutan, with Phil Patton. Alfred A. Knopf, Inc. ($19.95).

During World War II the Navy trained many young pilots out on the glaciated flatlands near Oshkosh, Wis. The vast airfield they left behind is now host every August to 15,000 aircraft and three-fourths of a million people, who gather for the world's largest aviation meet. It is not for the supersonic fighters or the jet transports the truest devotees have come; this is the Fly-In Convention of the Experimental Aircraft Association. The somewhat eclectic EAA has 100,000 members throughout the world. It remains at heart the expression of those remarkable amateurs, most of them linked in some other way with the aviation industry, who fly aircraft they have built themselves. Some 700 of those creations are certified yearly by the FAA to fly at large, a third or more of all new U.S. light aircraft.

Most of the homebuilts are slowly and carefully constructed from designs, some even from kits of parts, for prototypes that were first seen and admired at Oshkosh. The engineers and designers who conceive and produce the planes are a small group of professional innovators. These two books sum up the remarkable work of one such man, Burt Rutan (an original even among such originals), and his gifted coterie. The guide is a knowing and personal introduction to the long, demanding work in the shop and to flight itself. It is a homebuilder's look at the Rutan way. (To build a flyable aircraft at home requires a few thousand hours of work and $10,000 to $20,000 for engines and equipment; it is easy to understand that many more are begun than are in fact completed.)

The second book, a candid account of adventure and despair, meant for general readers, sets out years of single-minded effort and one week aloft in danger and aching sacrifice by the pilot couple, Jeana Yeager and Dick Rutan. At the end of 1986 they filled up the tanks of Burt Rutan's most famous design, the one-of-a-kind *Voyager,* and in nine nonstop days flew it out and back around the bulge of the world.

*Voyager* is much the largest of some 10 Rutan Aircraft Factory designs; all of them share family traits. They are built of composites, white, sleek and slender. This is not the old way of canvas stretched taut over spruce, nor the sturdy but intricate structures of Rosie the Riveter, a bucket of rivets, bent metal sheet, spars of forged aluminum. Layers of woven fibers, glass, graphite and aramid polymers are bound into a single immobile yet forgiving structure by the right mixture of epoxy resin polymer and its hardeners. Cores of foamed urethane shaped by hand tools give form to the light, strong skins of wings and tanks and fuselage. The scales control every part; tyros make heavy sections, and then learn to remake them. The homebuilder and the FAA inspector alike can see through the translucent forms to check the layers of lay-up and matrix.

Apt Rutan aerodynamics go with the new materials. Most of these craft are pushers; instead of a tail they have small canard wings in front of the main airfoil. The arrangement confers simple control, efficient low-drag flight and virtual immunity to stalling. These birds have no tail to drag. (A few pages by Burt Rutan give a clear, qualitative account of how all this happens.)

On top of his Dodge, Burt Rutan carries well-instrumented models. His rather low-speed wind tunnel is as long as the faster highways of southern California after midnight. Inspired designer though he is, at 44 he is no amateur but a trained aero engineer and a fine pilot. Burt's brother, Dick, a few years older, is a virtuoso test pilot, strong-minded veteran of 100 fighter missions against the little MiG's over Vietnam, and for years a record setter in one after another of Burt's prototypes. Jeana Yeager, Dick's partner, copilot and coauthor—their book is written in short passages first for one voice and then for the other—is equally skilled and sharply defined in character. Big *Voyager* was built out of just 11 sketches by Burt, and plenty of printout from his PC; Jeana created all the formal drawings from that material.

This group raised and spent a million dollars the hard way, without compromise of personal principles. They would accept neither foreign corporate sponsors, cigarette manufacturers nor any who set excessive constraints on their design, materials or route. Volunteer professionals of many specialties and a large number of individual donations complemented the well-known avionics firm, the fuel and lube supplier, an audio firm and a few others who took a chance with expensive equipment or with money.

The round-the-world plane realized the famous old flight-range formula of Louis Breguet. Long range means large lift-to-drag ratio, big fuel load for low structural weight, and high efficiency of prop and engine. "If the *Voyager* was easy to fly, Burt didn't do his job right." Stiff, pitching steadily, controls ineffectual, comfortless and crowded, *Voyager* gave up everything to haul fuel with low drag and minimal weight. But it did the pilots' joint will: after a heavy, dragging takeoff (during which wingtip fittings rubbed off on the runway), *Voyager* circumnavigated against unfavorable weather, returning to Edwards Air Force Base with 2 percent of the fuel it had carried off the two-mile runway.

The chief design deficiency seems to have been the complicated fuel plumbing, a little of it misrouted. Two engines fed by 17 tanks—gas was stored in every volume not otherwise filled—meant that the flight crew lost track of fuel use in spite of their elaborate routines to log and interpret the minimal gauges. These fliers were never free; the radio link to the Mojave was a bond tighter than they liked, but indispensable. The drama of technical brilliance and personal adventure, sociology and character, is honestly played against the looming backdrop of a powerful paramilitary industry, itself the gainer from the work.

Dick presents one unexpected insight. "All of Burt's airplane designs have been...an extension of his fascination with building model airplanes.... The legacy of modeling is physically visible.... The methods live on...in the stage of design after the plans are drawn, the period of 'tweaking,' the 'cut and fit' method that only composites make possible. Burt always wanted...the same control over the design as model airplane building afforded.... Sometimes he was in love with the design...to the exclusion of market considerations." We hear snatches of an old tune; this is the tale of an artist and his trusting, daring friends.

**THE NATURAL HISTORY OF NAUTILUS,** by Peter Douglas Ward. Allen & Unwin, Inc. ($34.95).

The chambered nautilus is found here and there over a wide stretch of ocean, from the Andaman Islands westward to Fiji and beyond, and from Japan south to New Caledonia. Once their kind were the masters of the ancient seas: 10,000 species and more are found in the fossil record. Now there are extant only five, quite similar species of this, the last of the shelled cephalopods. Creatures of the middle depths, they live between 150 and 400 meters down, usually along the undersea slopes of some steep coral forereef, in a band a couple of kilometers wide out from the coral shore.

This book is the work of Peter Douglas Ward, a paleontologist who has taken a field trip in time; he has gone to the modern ocean to follow living examples of the ammonite fossils he seeks to understand. His own studies and those of others conducted over some two decades have brought a certain understanding of this remarkable survivor, whose way of life is now unique. Most of what is described is recent work, documented here with graphs and photographs. There are imploded shells and X rays of air-filled chambers. There are closeups and diagrams. One illustration shows the eyes, tongue and beaky jaw of this strange pedigreed head, from which sprout about 90 tentacles, sensory, grasping and reproductive.

These beautiful animals are rare in warm surface waters; night scuba dives on the reef fringe will disclose them. For several years nautiluses have been kept for observation in aquariums, such as the one at Noumea in New Caledonia. Even more recently Ward and others have tracked nautilus individuals in the wild for as long as a week. The surface trackers "began to appreciate the problems of anti-submarine warfare." Graphs map the vertical position of a few animals followed around the steep walls of the reef of Palau, running deep by day, shallower at night, possibly deep again at the full of the moon. Telemetry enabled the investigators to track the animals. A small ultrasonic transmitter affixed to the shell, drawing power from a tiny lithium battery, encoded the pressure readings of a strain gauge and pulsed them up for two weeks from the submarine depths to alert hydrophones above.

The eight-inch-diameter nautilus forages vigorously over hundreds of kilometers, covering a few kilometers a day. It can rise and dive swiftly, nearly vertically at a rate of a couple of meters per minute. It appears to dwell in the cool waters below the reef system, taking live hermit crabs from the deep slopes for food, and scavenging as well. Its enemies are not well known; there are big reef fish with jaws powerful enough to take the nautilus. The beast propels itself by a flexible siphon that can point in almost any direction. Swimming with head trailing is fastest. In this mode the nautilus's graceful rounded shell experiences minimal drag; it is also shaped and aligned to minimize pinwheeling, induced by the strong pulses of the off-center jet. Most fossil forms must have been far poorer and less stable swimmers.

This living submarine has a hard protective shell, an elegant pressure hull that fails by implosion at a pressure depth of some 70 atmospheres, below the working range of many a steel submarine. This hull is a complex interlayered composite of horny protein bolstered with calcareous microneedles. The lovely chambers are fully or partly air-filled; the X-ray pictures prove it. A complicated tube

*Beginning its world flight,* Voyager *leaves California behind* links all the chambers. The nautilus, like a submarine, operates at near-neutral buoyancy, all but weightless, a few grams of net weight out of a total displacement of a pound or two. It must keep moving to avoid sinking to unsafe depths. Its pumps are osmotic ones, and the residual fluid in its partly filled ballast tanks is somewhat modified seawater, low in magnesium and sulfate.

With powerful pumps and strong tanks, a submarine takes care to keep both from rising toward its enemies and from sinking to a pressure depth it cannot sustain. So does the nautilus. In addition it has to do what no manmade submarine has ever done: the nautilus must grow. It grows from a hatchling swimmer an inch across, equipped with seven tiny chambers, to a big adult three orders of magnitude greater in volume.

Each of the couple of dozen successively larger chambers is thought to be first filled by the animal's body, at ambient seawater pressure, and then emptied as the animal grows and moves out. During this process gas slowly replaces fluid as the enclosing partition of the shell becomes strong enough to support the pressure difference between chambers. It is still uncertain whether or not the gas is ever at sufficiently high pressure to help resist the stress of the deep. Growth is slow, for the chemical pumps work slowly, particularly at depth. The nautilus cannot use its gas chambers to adjust density in its everyday motion; it empties its tanks to match buoyancy to the bigger, thicker shell only in accord with the deliberate rates of growth. The statelier mansions take longer to grow and longer to pump dry. The so-called growth rings once thought to mark daily or monthly increments of shell growth in fossil specimens now seem to have little to do with the passage of time.

A nautilus lives from hatching to maturity for some 10 or 12 years and then, unlike other cephalopods, survives for years the start of its reproductive period. The unornamented but beautiful nautilus produces rather few, large eggs. The ornate ammonites of the ancient past had thinner shells, much ribbed and adorned, perhaps to strengthen them against fishy predators. They laid far more eggs and grew much faster.

The last word has not yet been said, but Ward speculates that "the impression is one of the ammonites being forced into ever deeper water, where their...advantages of high growth rates become blunted by increasing...pressure, and hence the necessity to take ever longer to empty chambers." The chambered nautilus had found the way, a victory for patience and clean design. Those fancy ammonites did not survive the Cretaceous.

**Drawing the Future: A Decade of Architecture in Perspective Drawings,** by Paul Stevenson Oles. Van Nostrand Reinhold Company ($37.95).

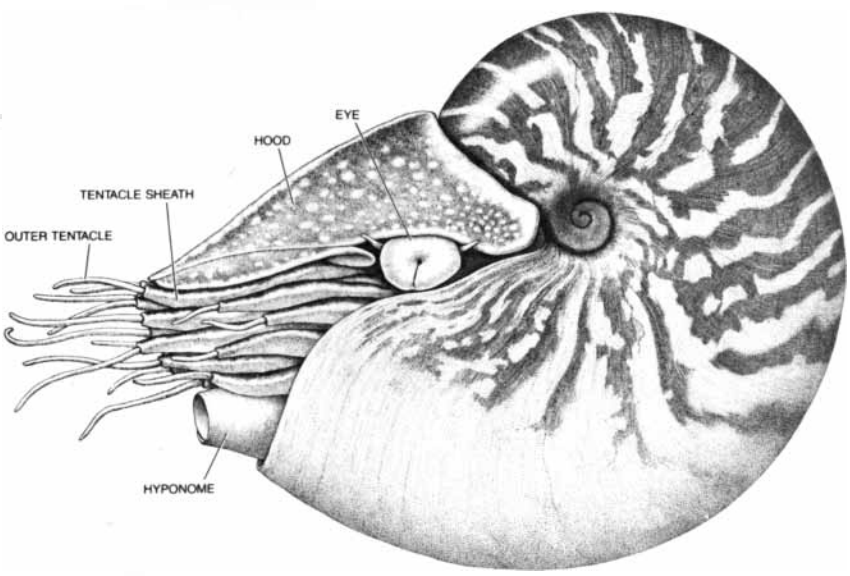
The future here is not the airy future that year by year recedes from us, like that of dreamers both before Gatsby and after him, but one that usually comes within grasp. Indeed, many of the 100 drawings here are fair renderings of buildings designed by American architects from 1976 to 1986 and now quite tangible in brick, glass and steel. Others remain on paper, a variety of projects still to come and some projects and contest entries that did not find realization.

All of them were made with meticulous honesty of form and purpose and rendered with a dazzling technique. This presentation shows the delighted eye—particularly of inexpert client or public—what it cannot glean from a big sheaf of technical working drawings. Such plans and elevations are dense with arrows and annotations, all pressed flat into two dimensions. The drawings instead unify the parts of a complex whole, and they convincingly unfold the third dimension largely by an utter mastery of the clues of light and shadow. One journal calls Oles "the profession's favorite architectural renderer." The collection of perspectives is a pleasure to eye and mind; "visual poetry can also be practical," in the phrase of the foreword.

Possibly the showiest of these black-and-white plates offers a night view of Kennedy Airport seen appropriately during an approach by air, at a low altitude of 1,000 to 2,000 feet. We look toward the center of the big ring road that the main airline terminals flank. But this is the future, JFK 2000. In the midst of the ring is shown a proposed central customs and information hall, able to accommodate 45 million passengers a year and distribute them through a web of enclosed pedestrian transport ways to their chosen airlines.

The form is accurate; it was controlled from the architects' plans and elevations, augmented by a photograph of the scaled model. The context was not left to whim; helicopter photographs were made at twilight to fix the existing structures and to provide the artist with the patterns of light shed. He then proceeded to reckon and supply the glow of the proposed building and its ribbons of outliers. Luminous present and future were matched, and "the amount of moisture in the atmosphere produced a specific degree of flare" for each point source of light that gives the scene a haunting reality, even to the subtle reflection of the lights of distant suburbs in the waters of the bay beyond the runways. The magical drawing was eight days' work, white pencil on black board.

About a dozen examples are given.




*Tentacled head and coiled shell of the nautilus*

Some are rendered—with attention to verisimilitude—as they might look by night and day or in different seasons. The proposed buildings are large and small; they are located in New England, New York, Washington, several cities in Texas and other U.S. cities. Some are abroad: London, Paris, Canberra and Buenos Aires. Always there is a true framework: a computer perspective made from the plans, or a model. There are street scenes, interiors and textures and reflections in wide variety. The artist is above all aware of the extent to which small clues add highlights and ground, and his technique has made him able to foresee quite accurately how the entire scene will look, documented by calculation of sun angle, light reach and repeated reflection. Nor are these drawings mechanical; this is personal pencilwork, interpretation of a space-time volume by a gifted hand and mind.

The new courtyard of the Louvre with its glazed pyramid is shown in three drawings, under gray Paris skies on a wet night. The drawings are as classical as the old palace, and as responsible as the carefully designed glazed framework.

Two other drawings show the Vietnam Veterans Memorial by Maya Ying Lin, a 21-year-old Yale undergraduate whose proposal won in competition among 1,400 entries. By now her simple, somber and yet healing work has deservedly become a focus of public devotion. Her competition drawings, however, eloquent to the wise and expert jury, were "very abstract and legible only to professionals." The public, in particular the sponsoring veterans, were "bewildered" by the spare design. Oles—himself a bested competitor—quickly made a pair of views, one in snow, one in summer green, spending two days on each (the originals were in color; all the reproductions in this book are in black and white). These beautiful renderings "provided a catalyst" for defeating proposed changes in the design and helped to win over many who had not yet appreciated what the purity of Lin's modernism might bring.

Here is a remarkable way to view a defined future, a prediction of architecture yet to come in something of the way a successful theory predicts the outcome of a well-planned experiment.

**ELSEVIER'S PERIODIC TABLE OF THE ELEMENTS,** compiled by P. Lof. Elsevier Science Publishing Co., Inc., P.O. Box 1663, Grand Central Station, New York, N.Y. 10163 ($17.50).

This presentation of the elements occupies more than 12 square feet of heavy paper. It bears some two dozen black-bordered rectangles, filled with colorful patterns in a variety of sizes and forms. At first the poster evokes the modern Dutch master Piet Mondrian. A closer look suggests kinship with still older painters, those Dutch and Flemish masters who crowded canvases in tireless detail with the painted pleasures of an unstinted harvest feast: too much! Yet just as those overflowing old images charm us by their mastery and redeem their good-humored superfluities by an honest delight in plenty, so does this dense new graphic compilation by the atomists of Amsterdam. Dr. Lof, scrupulous in citing specific references to some 90 sources of data and aided by a set of international advisers, has gathered here a bumper harvest of the atoms.

Even the advertising brochure has style: "Do you need to know the melting point of beryllium, the crystal structure of germanium, the mass of the electron, the price of yttrium, the heat capacity of iron, the half-life of cobalt-60, the charge of the strange quark, or the nuclear spin of tin-119?" They are all here and of course there is much more.

The large central panel is a color-coded periodic table of the elements from hydrogen out to undiscovered element 118, accompanied by the usual atomic weights and electron configurations and, less usual, the discoverer and date of discovery. The smaller frames that surround the center carry out the theme across a flood of variations, always in the adopted tabular pattern and color code. Most frames also offer a striking graph of some property related to atomic number. Nuclear and particle data are included. There is a clever mapping of all stable isotopes, and data are provided on radioactive ones that have lifetimes of five days or more. The information also encompasses nuclear gamma-ray and magnetic-resonance energies. Even mere supersymmetrical hopes for particles, with too cute names such as squark and slepton, are duly listed here.

Element abundances are particularly well treated. There are data for the solar atmosphere, the solar system as a whole, the earth's crust and mantle and the Orgueil chondrite. In one small periodic table is entered the human intake of all the atoms, placed in four classes by the weight taken each day. Another unusual table and graph compile calculated electron energy levels for the outermost orbitals, atom by atom.

The cheapest pure element in commerce is abundant and useful iron, at a dime per kilogram in 1984 dollars. (The elements are here priced by weight for a purity of better than 99 percent. These price tags announce no bargains; they are probably often quotations for small quantities, extrapolated without reduction to kilogram lots.) Gold and industrial diamonds are marked at about $11,000 per kilo; the rarest earth, lutetium, useful as a reaction catalyst in the petrochemical industry, is the dearest stable element, at $75,000 per kilo. Tritium and long-lived plutonium 239 are not listed. The most expensive cited are two short-lived radioactive isotopes, radon and berkelium 249, both quoted at about $1 million million a kilogram.

The peaks of molar volume against atomic number stand out very well among many such periodic graphs. There is a little unexpected history—Mendeleev's handwritten draft of his first scheme—and some unexpected practicality: an element classification according to principal hazardous property. Five pure elements (Be, Cr, Ni, As and Cd) are "established carcinogens"; others poison, burn, explode, radiate or corrode.

Nine graphed correlations, some lacking entries for many elements, are hard to resist as exercises for the reader. Why is antimony (and three of its kin) so anomalously high on the fine plot of heat of fusion versus melting temperature? Why is lithium 100 times commoner in a meteorite than in the photosphere of the sun? Other properties fill the curves without exceptions.

The chart is not without flaw. It is rather big to consult with ease, even when posted. The printed colors are far from painterly. The official temporary Latinate names and symbols for the seven claimed elements beyond the last accepted actinide (lawrencium 103), and for eight more extrapolated elements not yet even claimed, make a conspicuously ugly jargon; that cannot be helped. The entire chart is a madly rich challenge to students from high school on and full of information useful to professionals. This dazzlingly intricate poster, chiefly well-rounded truth and not mere human opinion, deserves to spread its amusing surfeit of learning across many a wall.

# BIBLIOGRAPHY

*Readers interested in further exploration of the subjects covered by the articles in this issue may find the following lists of publications helpful.*

## CONTROLLING INDOOR AIR POLLUTION

INDOOR AIR POLLUTION: A PUBLIC HEALTH PERSPECTIVE. John D. Spengler and Ken Sexton in *Science*, Vol. 221, No. 4603, pages 9–17; July 1, 1983.

INDOOR AIR QUALITY AND HUMAN HEALTH. Isaac Turiel. Stanford University Press, 1985.

RADON AND ITS DECAY PRODUCTS IN INDOOR AIR. Edited by William W. Nazaroff and Anthony V. Nero, Jr. John Wiley & Sons, Inc., 1988.

## SQUEEZED LIGHT

SQUEEZED LIGHT. Bernard Yurke and Richart E. Slusher in *Optics News*, Vol. 13, No. 6, pages 6–9; June, 1987.

PRECISION MEASUREMENT BEYOND THE SHOT-NOISE LIMIT. Min Xiao, Ling-An Wu and H. J. Kimble in *Physical Review Letters*, Vol. 59, No. 3, pages 278–281; July 20, 1987.

SQUEEZED STATES OF THE ELECTROMAGNETIC FIELD. *Journal of the Optical Society of America B*, Vol. 4, No. 10; October, 1987.

SQUEEZED-LIGHT-ENHANCED POLARIZATION INTERFEROMETER. P. Grangier, R. E. Slusher, B. Yurke and A. LaPorta in *Physical Review Letters*, Vol. 59, No. 19, pages 2153–2156; November 9, 1987.

## TUMOR NECROSIS FACTOR

CELLULAR FUNCTIONS IN IMMUNITY AND INFLAMMATION. Edited by Joost J. Oppenheim, David L. Rosenstreich and Michael Potter. Elsevier North-Holland, Inc., 1981.

TUMOR NECROSIS FACTOR (TNF). Lloyd J. Old in *Science*, Vol. 230, No. 4726, pages 630–632; November 8, 1985.

TUMOR NECROSIS FACTOR AND RELATED CYTOTOXINS. Edited by Gregory Bock and Joan Marsh. John Wiley & Sons, Inc., 1987.

SECRETORY PRODUCTS OF MACROPHAGES. Carl F. Nathan in *Journal of Clinical Investigation*, Vol. 79, No. 2, pages 319–326; February, 1987.

TUMOR NECROSIS FACTOR/CACHECTIN AND RELATED CYTOKINES. Edited by Benjamin Bonavida, George E. Gifford, Holger Kirchner and Lloyd J. Old. S. Karger AG, Basel, in press.

## ANCIENT MAGNETIC REVERSALS: CLUES TO THE GEODYNAMO

THE EARTH'S MAGNETIC FIELD: ITS HISTORY, ORIGIN AND PLANETARY PERSPECTIVE. Edited by Ronald T. Merrill and Michael W. McElhinny. Academic Press, Inc., 1983.

REVERSALS OF THE EARTH'S MAGNETIC FIELD. J. A. Jacobs. Adam Hilger, Ltd., 1984.

HOW THE GEOMAGNETIC FIELD VECTOR REVERSES POLARITY. Michel Prévot, Edward A. Mankinen, C. Sherman Grommé and Robert S. Coe in *Nature*, Vol. 316, No. 6025, pages 230–234; July 18, 1985.

MECHANISM FOR GEOMAGNETIC POLARITY REVERSALS. David Gubbins in *Nature*, Vol. 326, No. 6109, pages 167–169; March 12, 1987.

SEQUENTIAL GEOMAGNETIC REVERSALS RECORDED IN UPPER TORTONIAN MARINE CLAYS IN WESTERN CRETE (GREECE). Jean-Pierre Valet, Carlo Laj and Cornelis G. Langereis in *Journal of Geophysical Research*, Vol. 93, No. B2, pages 1131–1151; February 10, 1988.

## THE PLATYPUS

THE BIOLOGY OF THE MONOTREMES. Mervyn Griffiths. Academic Press, Inc., 1978.

THE PLATYPUS. Tom Grant. New South Wales University Press, 1984.

FIRST MESOZOIC MAMMAL FROM AUSTRALIA—AN EARLY CRETACEOUS MONOTREME. Michael Archer, Timothy F. Flannery, Alex Ritchie and R. E. Molnar in *Nature*, Vol. 318, No. 6044, pages 363–366; November 28, 1985.

THE ORIGIN OF EGG-LAYING MAMMALS. Zofia Kielan-Jaworowska, A. W. Crompton and Farish A. Jenkins, Jr., in *Nature*, Vol. 326, No. 6116, pages 871–873; April 30, 1987.

## AEROGELS

INTRODUCTION TO PERCOLATION THEORY. Dietrich Stauffer. Taylor & Francis, 1985.

AEROGELS. Edited by J. Fricke. Springer-Verlag, 1986.

ON GROWTH AND FORM: FRACTAL AND NON-FRACTAL PATTERNS IN PHYSICS. Edited by H. Eugene Stanley and Nicole Ostrowsky. Martinus

Nijhoff Publishers, 1986.

### THE INDIAN NECK OSSUARY

RECONSTRUCTION OF DEMOGRAPHIC PROFILES FROM OSSUARY SKELETAL SAMPLES: A CASE STUDY FROM THE TIDEWATER POTOMAC. Douglas H. Ubelaker. Smithsonian Institution Press, 1974.

THE ARCHAEOLOGY OF NEW ENGLAND. Dean R. Snow. Academic Press, Inc., 1980.

METHOD AND THEORY IN COASTAL NEW YORK ARCHAEOLOGY: PARADIGMS OF SETTLEMENT PATTERN. Lynn Ceci in *North American Archaeologist,* Vol. 3, No. 1, pages 5-36; 1982.

CHANGING VIEWS OF ABORIGINAL SEASONALITY AND SETTLEMENT IN THE GULF OF MAINE. David Sanger in *Canadian Journal of Anthropology/ Revue Canadienne d'Anthropologie,* Vol. 2, No. 2, pages 195-203; Spring, 1982.

HISTORIC AND ARCHAEOLOGICAL RESOURCES OF CAPE COD AND THE ISLANDS. Massachusetts Historical Commission, Boston, 1987.

### THE MYSTERY OF THE COSMOLOGICAL CONSTANT

EXPANDING UNIVERSES. Erwin Schrodinger. Cambridge University Press, 1956.

THE LARGE-SCALE STRUCTURE OF SPACE-TIME. S. W. Hawking and G. F. R. Ellis. Cambridge University Press, 1973.

THE ACCIDENTAL UNIVERSE. P. C. W. Davies. Cambridge University Press, 1982.

'SUBTLE IS THE LORD...': THE SCIENCE AND THE LIFE OF ALBERT EINSTEIN. Abraham Pais. Oxford University Press, 1982.

### THE AMATEUR SCIENTIST

ARE ICICLES SINGLE CRYSTALS? R. A. Laudise and R. L. Barns in *Journal of Crystal Growth,* Vol. 46, No. 3, pages 379-386; March, 1979.

ICICLES AS CRYSTALLIZATION PHENOMENA. Charles A. Knight in *Journal of Crystal Growth,* Vol. 49, No. 1, pages 193-198; May, 1980.

THE NOT-SO-ORDINARY ICICLE. Ira W. Geer in *Weatherwise,* Vol. 34, No. 6, pages 257-259; December, 1981.

### COMPUTER RECREATIONS

A CATALOG OF SPECIAL PLANE CURVES. J. Dennis Lawrence. Dover Publications, Inc., 1972.

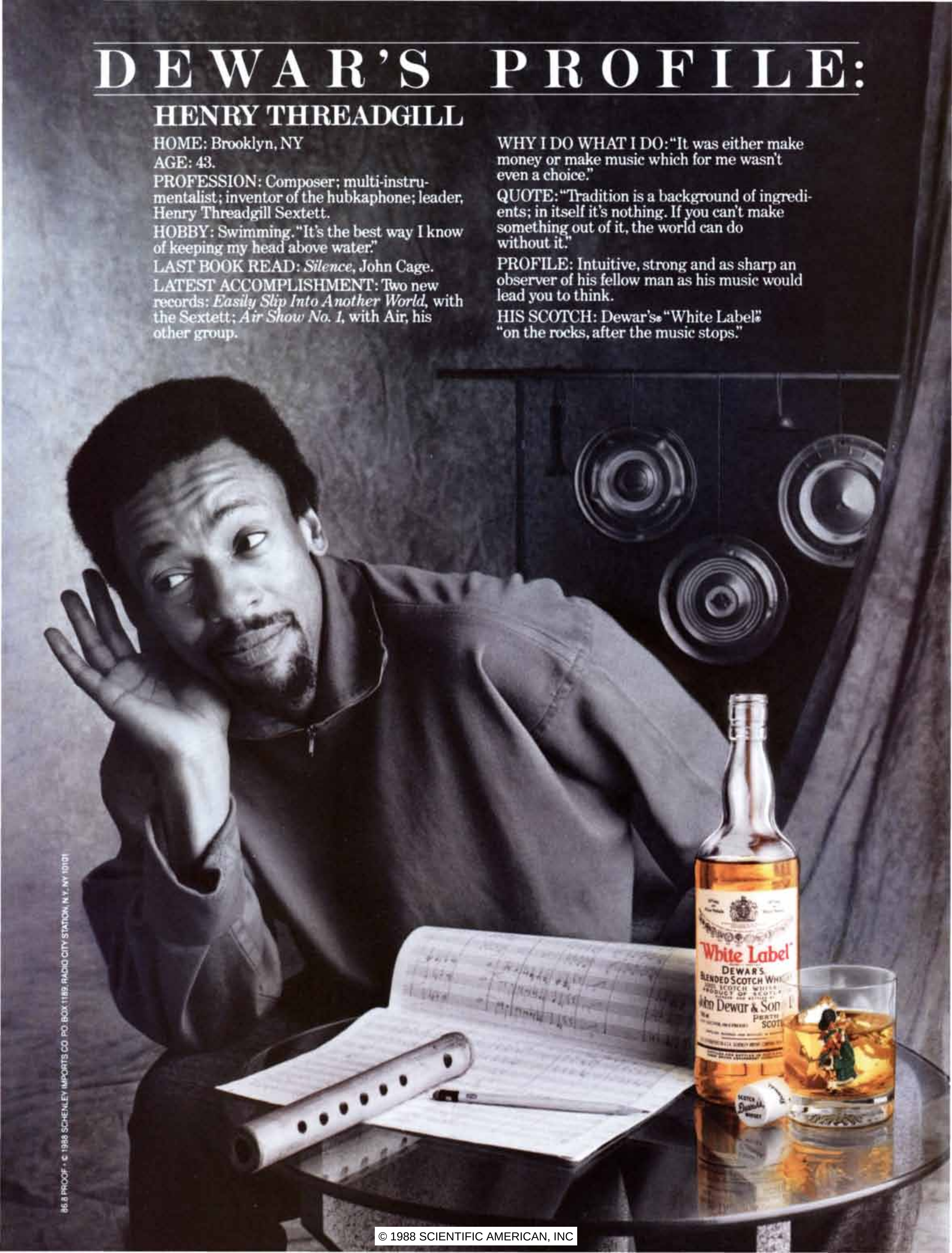
# Beretta GT.
# It bites when cornered.

Downshift the fluid 5-speed stick. Lean into the turn. And *dig in* with the available Z51 Performance Handling Package. Fat Goodyear Eagle GT + 4s, coil springs, shocks, struts and front and rear stabilizers combine to create a gripping experience. Taut, yet smooth and sophisticated. Beretta GT.

That's the nature of this beast. It's a feeling of confidence on the curves that translates into a sticky .83g on the skidpad with a professional driver. But controlled conditions can't tell the whole story. The feel of a meaty wheel in your hands. The rumble of a throaty exhaust in your ears. The envy of every driver in your rearview mirror. Beretta. It's not just the newest shape in performance coupes.

It's a whole new species.


Let's get it together... buckle up.
GM 6/60
QUALITY COMMITMENT PLAN
Crisp, tight handling from available Z51 Performance Handling Package.
Powerful Multi-Port Fuel Injected V6.
Smooth-shifting, Getrag-licensed 5-speed.
0-60 MPH in 9.1 seconds. (Professional driver on test track.)
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:
HENRY THREADGILL
HOME: Brooklyn, NY
AGE: 43.
PROFESSION: Composer; multi-instrumentalist; inventor of the hubkaphone; leader, Henry Threadgill Sextett.
HOBBY: Swimming. "It's the best way I know of keeping my head above water."
LAST BOOK READ: *Silence*, John Cage.
LATEST ACCOMPLISHMENT: Two new records: *Easily Slip Into Another World*, with the Sextett; *Air Show No. 1*, with Air, his other group.
WHY I DO WHAT I DO: "It was either make money or make music which for me wasn't even a choice."
QUOTE: "Tradition is a background of ingredients; in itself it's nothing. If you can't make something out of it, the world can do without it."
PROFILE: Intuitive, strong and as sharp an observer of his fellow man as his music would lead you to think.
HIS SCOTCH: Dewar's® "White Label"® "on the rocks, after the music stops."
White Label
DEWAR'S
86.8 PROOF • © 1988 SCHENLEY IMPORTS CO. PO BOX 1189, RADIO CITY STATION, N.Y., NY 10101