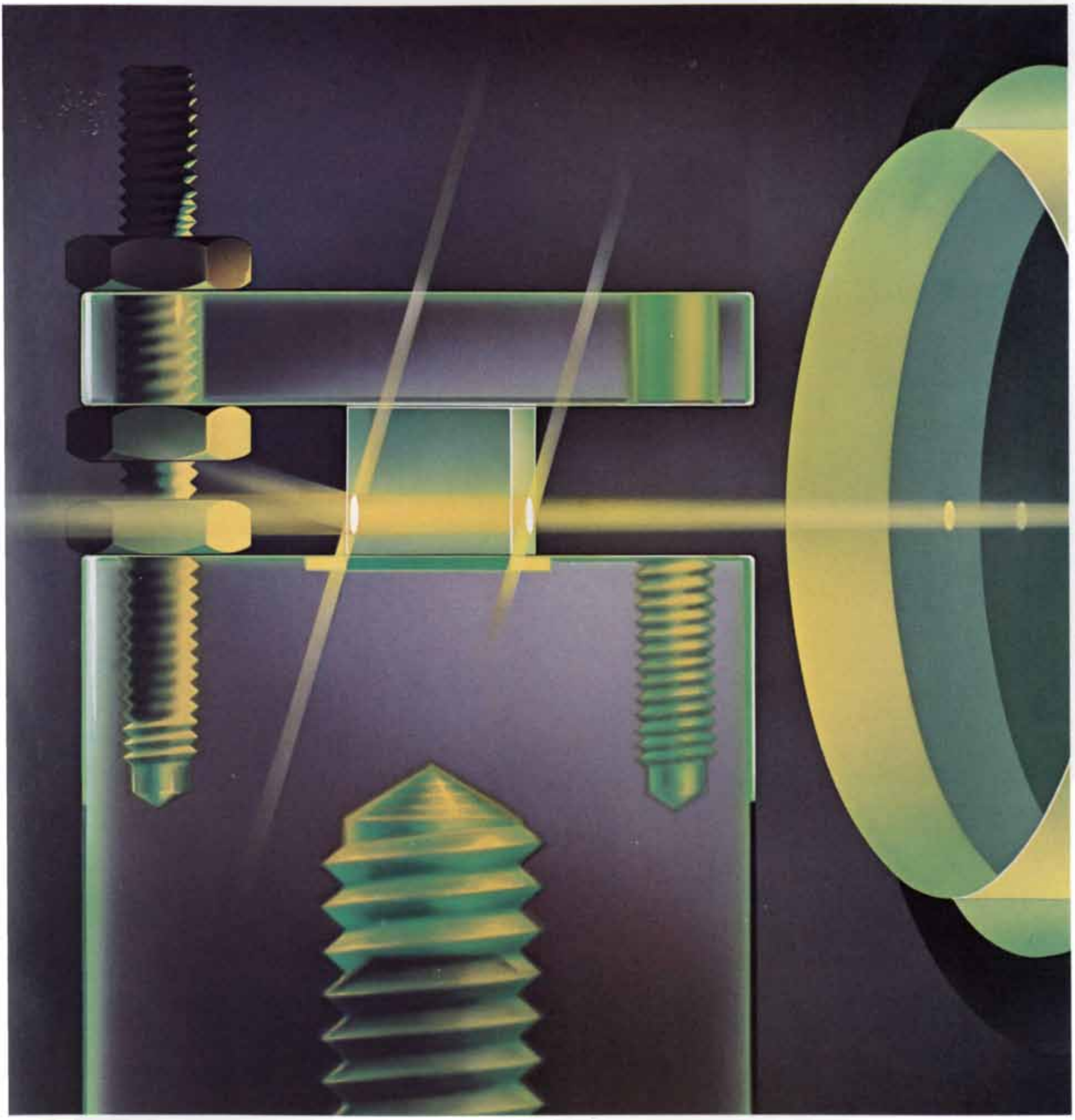
SCIENTIFIC AMERICAN



TIME-REVERSED LIGHT

$2.50

January 1986

FORD'S NEW DESIGN PHILOSOPHY:
TO ANSWER YOUR NEEDS
BY LISTENING TO
WHAT YOU WANT.

FORD TAURUS.

THE GOAL WAS TO BE BEST IN CLASS.

Even as the first designer or engineer began work, over 400 specific objectives were being established for Taurus.

The list was based on how you drive an automobile...what you want from it...what you need from it. The list included: How comfortable the seat belts are to wear; the effort required to open and close the doors; the convenience of checking the oil.

The idea was this: If we design each part, assembly and feature to make the car better to use, then we will have designed a better vehicle overall.

THE RESULT IS TAURUS.

A front-wheel-drive sedan that is responsive and rewarding to the driver. That provides room. That offers thoughtful features to driver and passenger alike. A unique sedan where design and engineering come together in one functional unit.

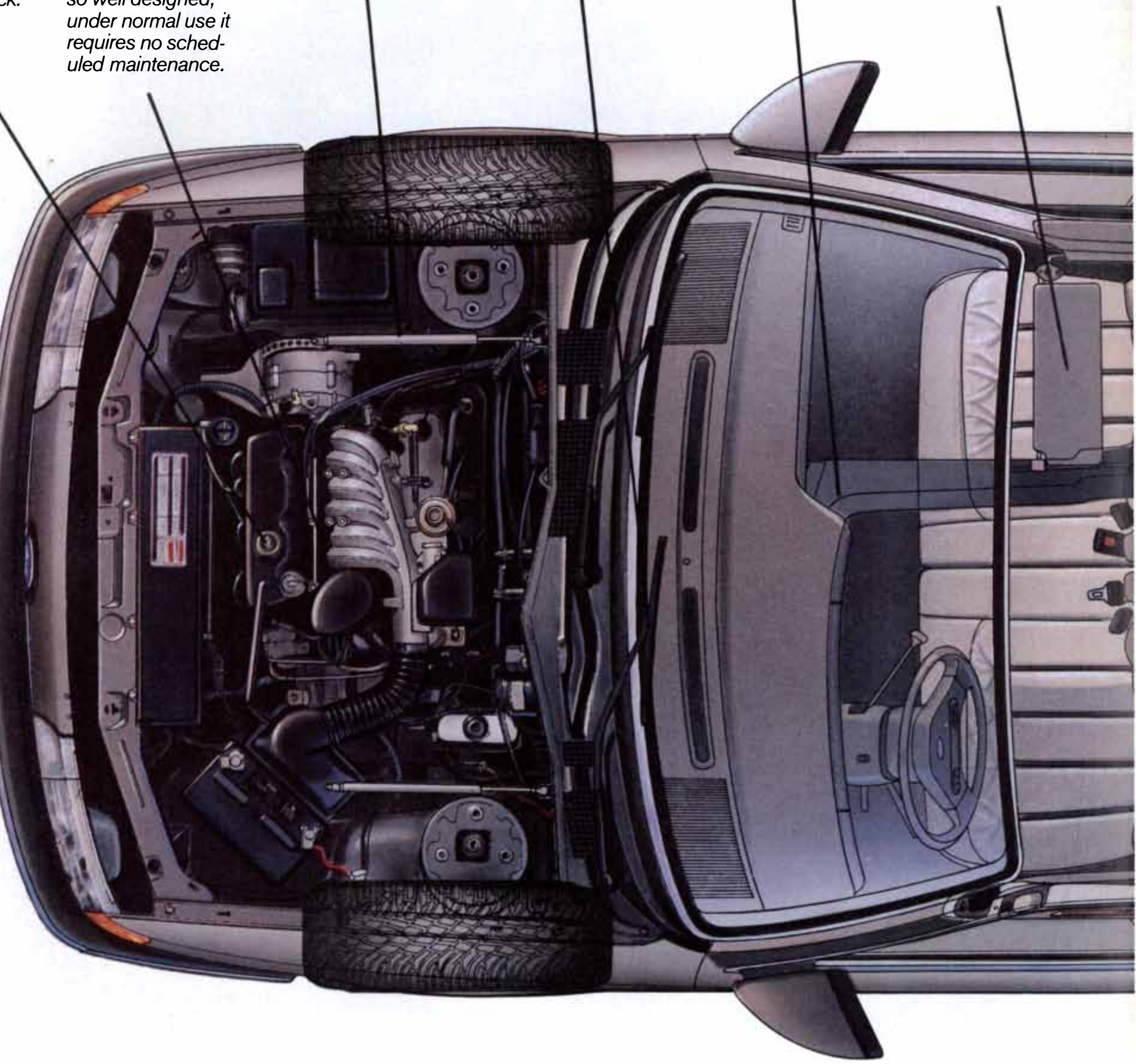
Taurus. A car so well thought out even the shape is part of its dedication to function. So you can judge it not only by how good it looks, but how well it works.

CONSIDER THE SHAPE.

Even the shape of Taurus is designed to help it work better. The shape manages the flow of air to help press the tires to the pavement for positive road holding. The flush-fitting side windows not only contribute to the overall integrity of the design, but reduce wind noise as well.

Taurus is now available at your Ford Dealer. You're invited to see it for yourself. And when you do, we think you'll have a feeling that the car was designed to be exactly what you're looking for. Because it is.

Ford



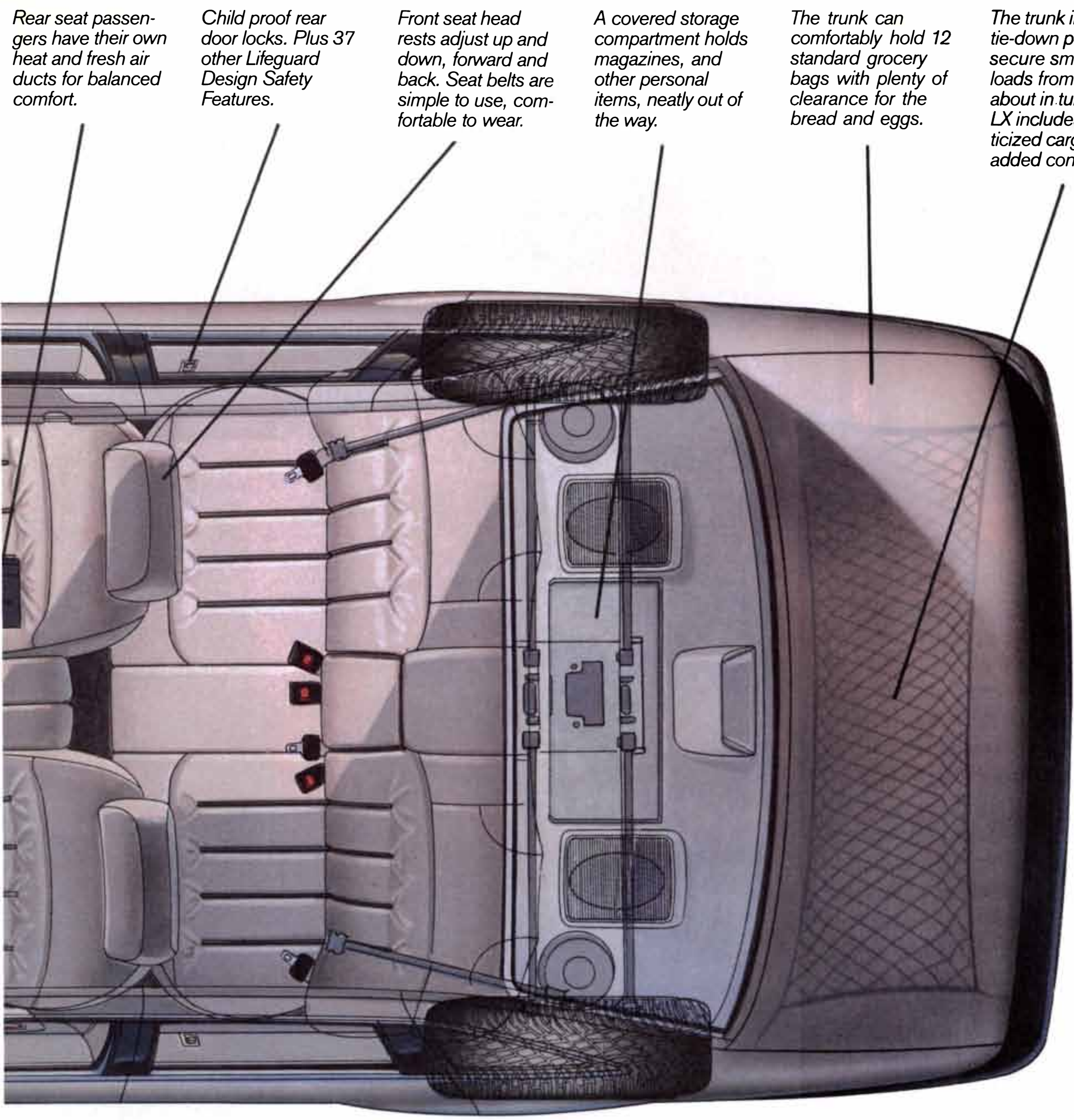
FORD'S NEW CONCEPT IN ENGINEERING BEGAN WITH YOUR NEEDS FIRST.

FORD TAURUS.

ENGINEERING... FOR PEOPLE.

Taurus was not simply engineered to work well, but also to be easy to use.

This means making sure Taurus is a car that is easy to live with day to day. That is satisfying to own, maintain and operate. And a car that accommodates the needs of the driver, the needs of the passengers,



Buckle up—Together we can save lives.

that provides ample room for luggage.

THE LIFETIME SERVICE GUARANTEE.

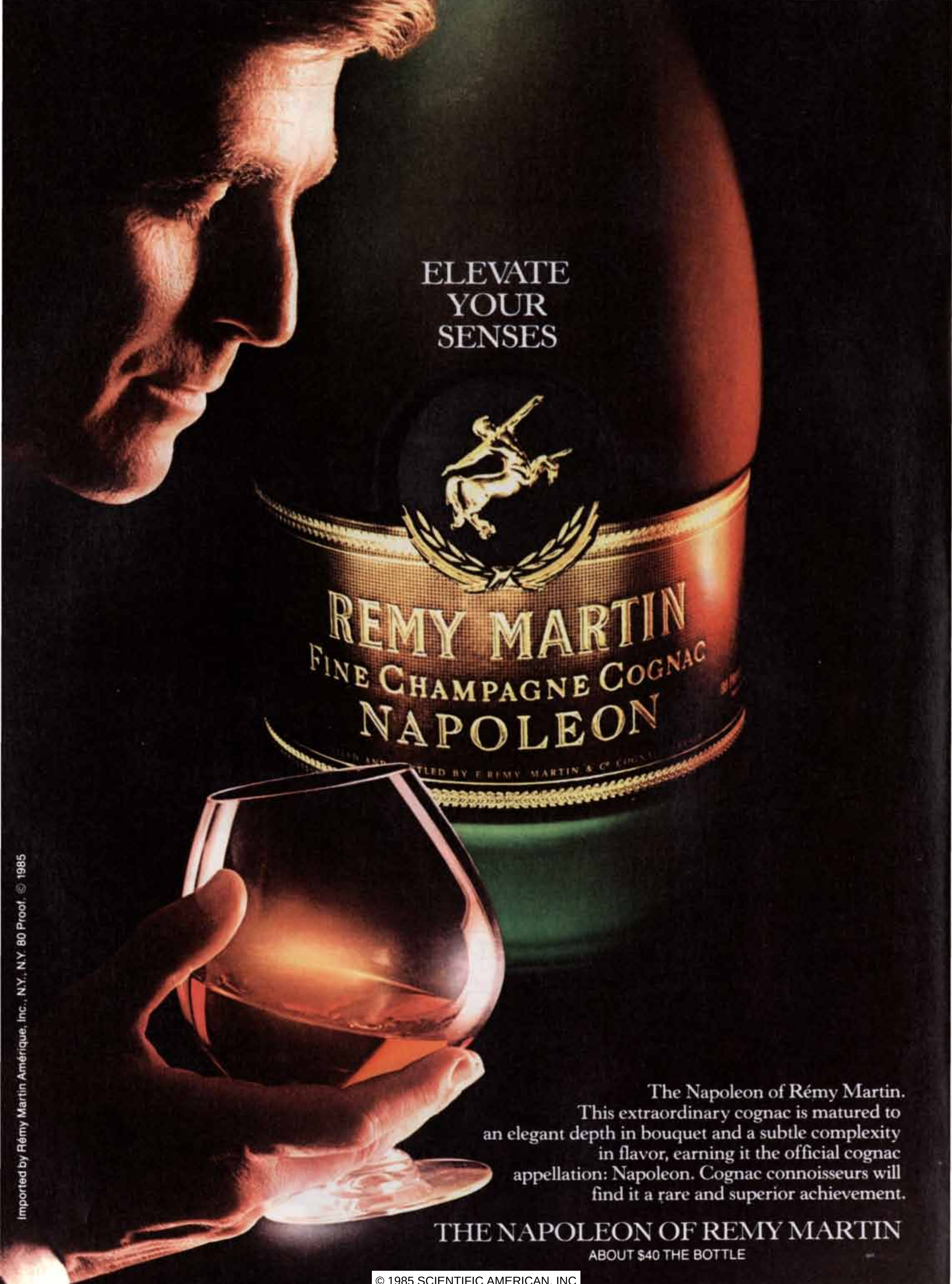
Participating Ford Dealers stand behind their work, in writing, with a free Lifetime Service Guarantee for as long as you own your Ford car. Ask to see this guarantee when you visit your participating Ford Dealer.

BEST-BUILT AMERICAN CARS.

At Ford Quality is Job 1. A 1985 survey established that Ford makes the best-built American cars. This is based on an average of problems reported by owners in a six month period on 1981-1984 models designed and built in the U.S.

Have you driven a Ford... lately?

Ford


ELEVATE
YOUR
SENSES
REMY MARTIN
FINE CHAMPAGNE COGNAC
NAPOLEON
The Napoleon of Rémy Martin.
This extraordinary cognac is matured to
an elegant depth in bouquet and a subtle complexity
in flavor, earning it the official cognac
appellation: Napoleon. Cognac connoisseurs will
find it a rare and superior achievement.
THE NAPOLEON OF REMY MARTIN
ABOUT $40 THE BOTTLE
Imported by Rémy Martin Amérique, Inc., N.Y., N.Y. 80 Proof. © 1985

Established 1845 SCIENTIFIC AMERICAN January 1986 Volume 254 Number 1

ARTICLES

32 **SPACE SCIENCE, SPACE TECHNOLOGY AND THE SPACE STATION, by James A. Van Allen** The author argues that a manned space station will hurt the space-research program.

40 **GROWTH, DIFFERENTIATION AND THE REVERSAL OF MALIGNANCY, by Leo Sachs** Proteins have been isolated that induce differentiation, halting some cancer cells' growth.

48 **THE STRUCTURE OF COMET TAILS, by John C. Brandt and Malcolm B. Niedner, Jr.** Observations of Halley should illuminate the interrelations of the plasma tail and the solar wind.

74 **APPLICATIONS OF OPTICAL PHASE CONJUGATION, by David M. Pepper** "Time-reversed" light waves can be manipulated to probe the structure of atoms or track a satellite.

84 **MINERAL DEPOSITS FROM SEA-FLOOR HOT SPRINGS, by Peter A. Rona** Ores are formed when seawater heated by the earth's mantle percolates through broken crustal rock.

94 **THE CHEMICAL DEFENSES OF HIGHER PLANTS, by Gerald A. Rosenthal** Some compounds poison or repel herbivores; others mimic insect hormones, interfering with growth.

100 **RADIOCARBON DATING BY ACCELERATOR MASS SPECTROMETRY, by Robert E. M. Hedges and John A. J. Gowlett** A new technique makes dating far smaller samples possible.

108 **KIN RECOGNITION IN TADPOLES, by Andrew R. Blaustein and Richard K. O'Hara** The ability of Cascades frog tadpoles to recognize siblings they have never known may be genetic.

DEPARTMENTS

8 LETTERS
13 50 AND 100 YEARS AGO
14 THE AUTHORS
16 COMPUTER RECREATIONS
27 BOOKS
60 SCIENCE AND THE CITIZEN
120 THE AMATEUR SCIENTIST
126 BIBLIOGRAPHY

SCIENTIFIC AMERICAN (ISSN 0036-8733), PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. COPYRIGHT © 1985 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, 415 MADISON AVENUE, NEW YORK, N.Y. 10017.

SCIENTIFIC AMERICAN

CORRESPONDENCE

Offprints of more than 1,000 selected articles from earlier issues of this magazine, listed in an annual catalogue, are available at $1.25 each. Correspondence, orders and requests for the catalogue should be addressed to W. H. Freeman and Company, 4419 West 1980 South, Salt Lake City, UT 84104. Offprints adopted for classroom use may be ordered direct or through a college bookstore. Sets of 10 or more Offprints are collated by the publisher and are delivered as sets to bookstores.

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, MA 01970. Specify CCC Reference Number ISSN 0036-8733/84. $1.25 + 0.00.

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

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

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

Name ____________________

New Address

Street ____________________

City ____________________

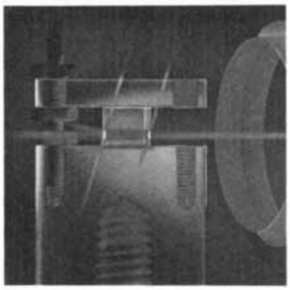
State and ZIP ____________________

Old Address

Street ____________________

City ____________________

State and ZIP ____________________



THE COVER

The painting on the cover depicts a phase-conjugate mirror, a device that can correct optical distortions. Such a mirror returns any beam of light to the point of origin in a "time-reversed" sense, so that the rays of light retrace their original trajectories. This phase-conjugate mirror is a crystal of barium titanate, measuring five millimeters on each side. A beam of light from a laser at the right enters the crystal. Some of the light emerges from the crystal at the left. Most of it, however, interacts with the atoms in the crystal, producing a phase-conjugate, or "time-reversed," beam that travels backward to the right. Phase-conjugate beams can be exploited to compensate for atmospheric turbulence, perform optical computing functions, probe atomic structure and track moving satellites (see "Applications of Optical Phase Conjugation," by David M. Pepper, page 74). The painting is based on an experimental setup designed by Ravinder K. Jain, Ross A. McFarlane, Knut E. Stenersen and Cliff W. Olson of the Hughes Research Laboratories.

THE ILLUSTRATIONS

Cover painting by Hank Iken

Page	Source
17	A. K. Dewdney
22–24	Ilil Arbel
25	Ilil Arbel (*top*), Hale Observatories (*bottom*)
28	American Philosophical Society
33	Jet Propulsion Laboratory, courtesy of the IBM Gallery of Science and Art, New York (*top and bottom left*); Louis A. Frank, University of Iowa (*bottom right*)
34–35	Edward Bell
36	Gabor Kiss
39	McDonnell Douglas Corporation
41	Leo Sachs, Weizmann Institute of Science
42–43	Ilil Arbel
44	Leo Sachs, Weizmann Institute of Science
45–47	Ilil Arbel
48	Joint Observatory for Cometary Research
50	Uwe Fink, University of Arizona
51–54	George V. Kelvin, Science Graphics
55	Yerkes Observatory
56	Edward Bell (*top*), American Museum of Natural History (*bottom*)
75	Jack Feinberg, University of Southern California
76–77	Ian Worpole
78	Jack Feinberg and Robert W. Hellwarth, University of Southern California
79	Gilmore J. Dunning, Richard C. Lind, William C. Keeler and Cliff W. Olson, Hughes Research Laboratories (*top*); Ian Worpole (*bottom*)
80–82	Ian Worpole
85	Kyung-Ryul Kim, Scripps Institution of Oceanography
86–87	Andrew Tomko
88–91	Hank Iken
92	Andrew Tomko
95	Isao Kubo, University of California at Berkeley
96–99	Patricia J. Wynne
101	Research Laboratory for Archaeology and the History of Art, University of Oxford
102–106	Andrew Christie
109	Andrew R. Blaustein and Richard K. O'Hara, Oregon State University
112–116	Tom Prentiss
120–125	Michael Goodman

SCIENCE/SCOPE®

Enemy submarines have nowhere to lurk now that the U.S. Navy has deployed a totally new passive sonar system. The Surveillance Towed Array Sensor System (SURTASS), now operational, is an array of miniaturized hydrophone listening devices towed behind a dedicated T-AGOS ship. It acquires and transmits acoustic information to shipboard processors, while shore stations analyze the data to detect and classify targets. A SURTASS preproduction development program is under way at Hughes Aircraft Company to replace the present large array with one having a smaller diameter. This new version will simplify storage and handling, as well as allow for a faster towing speed.

Advanced satellites will provide communications to the world's shipping and offshore industries later this decade. The International Maritime Satellite Organization (INMARSAT), a group of 43 countries, plans to launch the first of the spacecraft in 1988. The new series will accommodate the increasing demand for services, which is growing as fast as 60% a year. Each spacecraft will be able to carry at least 125 simultaneous transmissions. More than 3,300 vessels are equipped to use the INMARSAT satellite system. Users include operators of oil tankers, liquid natural gas carriers, off-shore drilling rigs, seismic survey ships, fishing boats, passenger liners, and tug boats. British Aerospace will build three satellites, with INMARSAT having an option to purchase six more. Hughes, which in 1976 built the world's first maritime communications satellite, will provide the communications electronics for the second-generation spacecraft.

An Amraam missile bored through radar clutter to intercept a drone aircraft target in the second guided launch of the full-scale development program. The test firing was the third consecutive launch of the advanced medium-range air-to-air missile, under development by Hughes for the U.S. Air Force and Navy. An F-15 launched the missile in a "look-down, shoot-down" tail-aspect attack while flying at Mach 0.9 approximately 16,000 feet above the desert floor at White Sands Missile Range. The QF-100 target flew at Mach 0.7 only 1,000 feet above the ground. The Amraam flew the first part of its flight under control of its on-board inertial reference unit, using target coordinates provided in prelaunch by the F-15's Hughes APG-63 radar. The missile then switched to its own active radar for guidance and tracked the drone through the heavy ground clutter to intercept.

Controller-thrust aircraft engines will be tested at NASA's Lewis Research Center with three heat monitors. Three Probeye® thermal video systems will monitor, measure, and analyze heat from the engines while they are operated within the environmental test chamber at the research center. Each Probeye system consists of an infrared imager, image processor, and TV display, all of which interface with a computer. Hughes designed and produces the Probeye system.

Hughes Missile Systems Group, in Canoga Park, California, an attractive suburb of Los Angeles, offers engineers and scientists a unique environment that features challenging projects and professional recognition. Our primary work is in advanced missile and missile systems design. Openings are in analog/digital circuit design, software engineering; high-voltage power supply design; microwave process, transmitter and radome design; electro-optical design; IR imaging sensors; focal plane arrays; quality assurance engineering; operations analysis; mechanical engineering; and production design and systems engineering. Send resume to Hughes Missile Systems, Engineering Employment, Dept. S2, 8433 Fallbrook Ave., Canoga Park, CA 91304. Equal opportunity employer. U.S. citizenship required.

For more information write to: P.O. Box 45068, Dept. 76-6, Los Angeles, CA 90045-0068

HUGHES
AIRCRAFT COMPANY

LETTERS

To the Editors:

This letter pertains to the article "Yellow Rain," by Thomas D. Seeley, Joan W. Nowicke, Matthew Meselson, Jeanne Guillemin and Pongthep Akratanakul [SCIENTIFIC AMERICAN, September, 1985], in which Seeley et al. claimed that evidence for the use of trichothecenes as biological-chemical warfare agents in Southeast Asia could be attributed to the feces of indigenous honeybees. Perhaps it is best to define "yellow rain" before we proceed. The authors, according to my interpretation, defined it as a yellow substance that is found in various substrates and is indistinguishable from bee feces. The best example given was the yellow spots on leaves collected by undisclosed sources in Southeast Asia. If we accept that definition, then indeed the authors have confirmed what is already found in the entomological literature, namely that certain honeybees defecate in flight and, moreover, that bee feces contain spent pollen.

The leaves the authors had collected or had been given indeed contained apian defecations, and it is clear that none of those leaves when analyzed by any laboratory showed any trace of trichothecenes (nor should they have). The work of H. Cohen and Gordon A. Neish (1984) demonstrated that pollen does not support any significant growth of *Fusarium* fungus or the production by the fungus of trichothecene toxins. A scientific treatment would have cited these data (a serious omission). On the other hand, if we accept "yellow rain" as being the substance, variable though it may be, that has been associated with aerial attacks in Hmong villages, that yielded trichothecenes and that caused signs and symptoms including headache, diarrhea, hemorrhaging, skin rash and in some cases death, then that is my working definition and the etiology of the substance is a more serious matter.

I must emphasize here that we cannot ignore the testimony given by Hmong refugees at both the Senate and the House hearings on biological warfare as being folklore and hearsay. I have listened to intelligent witnesses describe the plight of their people. No deaths have been attributed to apian cleansing flights by eyewitnesses, and it is clear that the Southeast Asian people are intelligent enough in their ecological orientation to understand the obvious habits of honeybees that share their environment.

The leaves collected in attack areas in 1981 in Southeast Asia and analyzed in July, 1981, in my laboratory at the University of Minnesota did not contain yellow spots as portrayed in the article by Seeley et al.; however, they did contain *T*-2 toxin, deoxynivalenol (DON) and nivalenol. All the results were confirmed by mass spectrometry and submitted to journals for review by peer scientists before publication in reputable journals. The data are clear and unequivocal. In a similar manner, analyses were made of powderlike yellow scrapings from two independent attack areas verified by witnesses. One set of scrapings was analyzed in my laboratory and the other by Joseph D. Rosen at Rutgers University. Both analyses showed excessive amounts of trichothecenes, ranging from 27 to 143 parts per million of *T*-2 toxin, diacetoxyscirpenol (DAS) and zearalenone. The last substance is not a toxin but is more appropriately known as a phytoestrogen. Nevertheless, it is produced by some of the same *Fusarium* species that produce the trichothecene toxins. Both reports were published in separate journals after having been reviewed by peer analytical chemists, who judged the information acceptable and appropriate for publication.

Blood and urine samples were also collected from alleged victims of attacks. (Witnesses substantiated that the attacks were due to low-flying aircraft or to artillery and not to low-flying squadrons of honeybees.) In at least 200 blood and urine samples collected in 1981 and 1982, analysis showed about 30 percent were positive for *T*-2 toxin or *HT*-2 toxin; of those collected in 1983, 8 percent were positive, and of those collected in 1984–85 (the time of de-escalation of attacks) fewer than 1 percent were positive. In no case was toxin found in those samples designated as control samples. Collaborative studies of our analytical methodology have shown any error that may occur in our analyses would be on the conservative side; that is, in a biological mixture we might miss actual concentrations of less than 10 parts per billion of *T*-2 or *HT*-2 toxin. Hence we may be reporting less than is actually present. More important, no false positives were found.

The data of Roy Greenhalgh, J. David Miller, Neish and H. Bruno Schiefer (1984) as well as data of my own have shown that isolates of *Fusarium* species collected in Southeast Asia do not produce significant quantities of trichothecenes when they are grown on a rice substrate under ideal laboratory conditions of adequate moisture and carbon supply and an absence of competition from other fungi. Moreover, they do not colonize pollen and produce toxins. On the other hand, we found copious amounts of *T*-2, DAS and zearalenone in scrapings of yellow powder, which according to my experience can be accounted for only by the intervention of man. The analytical data are backed up by mass spectra and are unequivocal.

From the data presented in the article I can only conclude that honeybees in Southeast Asia, like their relatives in other parts of the world, defecate in flight, a fact already well reported in the scientific literature and not one that is in the forefront of research. Yellow spots on leaves the authors analyzed were indeed bee feces and had nothing to do with chemical attacks on the Hmong in their villages. Seeley et al. made no attempt to verify that the samples of leaves came from areas where aerial attacks were taking place.

One of the authors questioned 16 groups of people in the Ban Vinai camp about their familiarity with leaves spotted with feces of *Apis dorsata*. Thirteen of the 16 groups did not show any familiarity with the leaves, two groups said the spots were *kemi* (poison) and one group identified the spots as bee feces. From these data the authors inferred that some of the Hmong identify bee feces as the alleged agent of chemical warfare. The aforementioned statistical sampling hardly warranted such a conclusion.

The only task left to Seeley et al. in order to maintain credibility with the scientific community is to attack the methods of the analysis we performed—an area in which they have no professional competence. The fact is that we as well as other laboratories have found copious amounts of trichothecenes in samples from attack areas. (The leaf samples did not have theatrical yellow spots on them.) There is no precedent for the natural occurrence of the large amounts of trichothecenes the samples contained. Moreover, we as well as Rosen's laboratory independently and with different methods found large quantities of trichothecenes in yellow powder samples. In addition our laboratory found trichothecene residue in the blood, urine and body tissues (heart, lung and kidney) of victims, and this cannot be explained away with logic. Let us see cold scientific data presented by Seeley et al. on the analyses of samples from victims and controls, data that can pass the scrutiny of reviews by reputable journals. Until that time the prose written by Seeley, Nowicke, Meselson, Guillemin and Akratanakul remains in the realm of political polemics.

CHESTER J. MIROCHA

(*Seeley et al. reply on page 10.*)

To the Editors:

Chester Mirocha clearly does not understand the results and implications of our study of yellow rain. We have sought to determine the nature of the alleged chemical-warfare agent consistently described by the U.S. Government as yellow and rainlike. The report of Secretary of State Alexander M. Haig in March, 1982, stated that "the majority of the attacks were carried out by aircraft spraying a yellowish substance which 'fell like rain.'" Later that year, the report of Secretary George P. Shultz affirmed that "descriptions of the 1982 attacks have not changed significantly from descriptions of earlier attacks. Usually the Hmong state that aircraft or helicopters spray a yellow rain-like material on their villages and crops." And most recently, in 1985, an article coauthored by a Central Intelligence Agency official with primary responsibility for yellow-rain studies reiterated that "aerial attacks, usually by spray, dispersed yellow to yellow-brown liquid or semi-solid particles that fell and sometimes sounded like rain when striking thatched rooftops" (C. J. Stahl, C. C. Green and J. B. Farnum, 1985).

The Government's description of the alleged agent is well supported by the evidence. Nearly all interviews with the Hmong regarding the residue of the presumed agent describe it as yellow, and nearly all samples of it are also yellow. Contrary to Mirocha's impression, our working definition is that yellow rain is a yellowish material alleged to have been deposited in chemical attacks in Southeast Asia. Moreover, the samples we have studied include the yellow scrapings analyzed by Mirocha and Rosen and defined by them as yellow rain.

To the best of our knowledge all samples of the yellow material examined under the microscope, whether spots or scrapings, are without exception found to be composed mainly of pollen. They include samples from more than 30 alleged attacks, collected from 1979 on. They specifically include the yellow scrapings reported by Mirocha and Rosen to contain trichothecene mycotoxins. A leaf sample and a sample of pond water are also reported by Mirocha to contain the toxins. But leaves and water cannot be the chemical-warfare agent itself, since no one asserts that the alleged attack aircraft release such materials. Thus, if one accepts the well-supported and unchanging assessment of the U.S. Government that the alleged chemical-warfare agent leaves yellow residues, one must explain why its principal ingredient is pollen. Our explanation, based on evidence published in *Nature* and in SCIENTIFIC AMERICAN, is that the material is bee feces. We believe this conclusion is generally accepted by knowledgeable Government experts in the U.S. and in allied countries.

This poses a critical problem for the trichothecene-warfare hypothesis. If the yellow material is bee feces and therefore is not released from aircraft, then the purported trichothecenes would have to be dispersed in some other physical form. What could it be? Certainly not a gas: trichothecenes have negligible vapor pressure. Not an aerosol either: aerosols do not fall and the alleged witnesses do not describe the very dense ground-level clouds that would be required. What witnesses (including witnesses to several of the alleged attacks from which Mirocha's samples come) are in fact reported to say is that the agent is sprayed by aircraft and leaves a yellow residue. But no known samples of such deposits, including those reported by Mirocha and Rosen to contain trichothecenes, have been shown to be anything other than bee feces.

Therefore in order to rescue the trichothecene-warfare hypothesis one must reject the U.S. Government's description of the alleged agent as yellow and rainlike. It is then necessary to postulate a different material, one that actually is released from aircraft, contains toxins and frequently falls on villagers in lethal quantities but that, strangely enough, has gone unnoticed. No one has come up with a plausible hypothesis for what such an elusive material could be. Unless an explanation can be devised, and supported with solid evidence, the trichothecene-warfare hypothesis collapses.

What explanation is then left for the reports of trichothecenes? Occasional or sporadic natural occurrence is not ruled out. Mirocha does not mention the finding of Greenhalgh and his colleagues (1985) that one out of four *Fusarium* isolates from a leaf collected in Thailand produced 1,000 parts per million (dry weight) of trichothecenes when grown on glucose-yeast medium. But the U.S. Army's Chemical Research and Development Center has found no trace of the toxins in any of the more than 80 samples from alleged attacks that it has analyzed. These consistently negative findings, and similar results from other laboratories, provide strong evidence against common natural occurrence of the toxins in such materials and also against their use as chemical-warfare agents.

Mirocha characterizes his evidence for trichothecenes as "unequivocal." Based on our own laboratory experience with mass-spectrometric analysis of trace components in natural materials, we would reserve this adjective for cases in which three principles of forensic chemistry are met. First, results for critical samples must be unambiguously confirmed by independent analysis. Second, adequate control samples must be analyzed concurrently, and the results must be fully disclosed. Third, the integrity of the samples must be assured. None of these conditions was met in the present case.

Mirocha asserts without evidence that the Hmong would generally recognize bee feces for what they are. If he were right, the Hmong would have practiced massive deception since handing in the first of many bee feces samples to U.S. officials in 1979. Actually, it appears that Asian villagers often cannot identify bee feces. None of the Hmong we interviewed did so, nor did most of the Thai villagers we also interviewed. It is not surprising that villagers misinterpret the fecal showers of honeybees flying too high to be seen or heard. In 1976 yellow showers in Jiangsu, China, mystified rural communities and were erroneously suspected of being toxic until university scientists identified them as bee feces (Zhang Zhong-ying et al., 1977, and personal communication). Two other cases on record could be similar misperceptions of bee feces. Last year yellow showers were attributed by Thai villagers to imagined Burmese atom bomb tests. And, in 1964 at the United Nations Security Council, the government of Prince Norodom Sihanouk charged that U.S. and Vietnamese aircraft were spraying Cambodian villages with lethal yellow powders.

Maintaining a critical approach to scientific data can be difficult when an issue becomes politically charged. In 1982 Mirocha began an article reviewing his yellow-rain research with this assertion: "During the last 6 to 7 years, chemical and/or biological warfare has been waged in Laos, Kampuchea and Afghanistan resulting in the death of 75–100,000 human beings." Considering that there is not a single unambiguous case on record of a medically authenticated victim, living or dead, nor is there either a single chemical canister or a chemical munition, what hard data justify his assertion?

In short, the evidence for mycotoxin warfare fails the test of critical examination.

THOMAS D. SEELEY

JOAN W. NOWICKE

MATTHEW MESELSON

JEANNE GUILLEMIN

Don't Be Bound To A Single Solution.

Too many vendors have a vested interest in tying you into proprietary technology. Which, down the road, can leave you with a rather knotty problem.

The solution is flexibility. From a vendor whose only vested interest is solving your problem.

Enter Corvus. Exit chaos.

The battleground is your office. The dilemma: how to successfully manage information. So people will interact more efficiently. Dissimilar machines and systems will work together. And capital equipment will be shared by departments and work groups.

The single solutions you're offered are local-area networks or multi-user computers. And while they are fighting each other, you have to survive the war.

Unless you talk to Corvus.

Because our expertise crosses both networking and multi-user systems, we'll apply whichever solution best suits your purposes. Instead of which best suits ours.

You'll get a system-wide solution. One that's cost-effective. And has the flexibility to give you the benefits of multiple technologies.

If you have to solve a tightening office productivity problem, call us toll-free at **(800) 4-CORVUS**. We'll show you how our multiple technology approach is the most cost-effective way to free the ties that bind.

CORVUS

Corporate Headquarters: 2100 Corvus Drive, San Jose, CA 95124. (408) 559-7000. Telex: 278976.

European Offices: Corvus (U.K.), Ltd., #7 Fairmile, Henley-on-Thames, Oxfordshire, England RG9 2JR. Phone: 0491 571100. Telex: 847282. Geneva Corvus Systems, S.A., 47A, Rue de Lausanne, CH-1201 Geneva, Switzerland. Phone: 327289. Telex: 27699.


Grand Enchantment.
Grand Marnier
Liqueur
Triple Orange
Triple Orange
PARIS — FRANCE
750 ml (25.4 fl.oz) · 80° Proof
Gallardo
For gift delivery anywhere call 800-CHEER-UP (except where prohibited by law).
Product of France. Made with fine cognac brandy 80 proof. © 1985 Carillon Importers, Ltd., N.Y.

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

JANUARY, 1936: "Radium is the best known of a group of chemical elements which have the property of spontaneously emitting various kinds of radiation in the process of an atomic disintegration. The radiations emitted, especially the gamma rays, have proved useful in the treatment of cancer. These naturally occurring radioactive substances are of great rarity and consequently of great expense. Therefore the recent discovery of means of making other elements artificially radio-active marks an advance in physics which may easily prove to be a great boon to medicine."

"The first airport directory of the United States, dated 1935, has appeared. It comprises over 200 pages, listing airports by states, giving main dimensions and facilities and including photographs from above of all the main landing points. Perhaps some day maps and airport directories will be distributed from aerial service stations in the way automobile maps are given away by the oil companies."

"There has been developed at the Metropolitan Museum of Art in New York a small, readily portable, ultraviolet-ray-producing apparatus. This equipment was developed by James J. Rorimer, Curator of the Department of Medieval Art. It has proved very satisfactory in the routine examination of works of art. It will also be valuable for the stamp collector, jeweler, geologist and any one using the long-wave ultra-violet waves."

"Highway construction has come in for mature consideration during the past year. Three, four and five lane highways built only a few years ago have come to be termed 'Death Highways,' and rightly so. The realization has been reached that, in congested areas where traffic on super-highways is heavy, something more than a white line is needed to keep traffic separated as it moves in opposite directions. Parkways are the coming thing."

"From an acreage of 50,000 in 1907 to more than 5,000,000 in 1935 is the record made by soybean cultivation in this country. Thus production of this valuable farm product, brought here from the Orient, has increased a hundredfold in less than 30 years."

SCIENTIFIC AMERICAN

JANUARY, 1886: "The report of the United States Geological Survey shows that the mining industries of the United States are assuming giant proportions. Not less than $800,000,000 is invested in mining enterprises as productive capital, over 400,000 people are furnished employment and the mineral product of the United States for the year 1884 had a value of $413,104,620. In terms of value the chief products were bituminous coal, pig iron and anthracite coal."

"An experiment of the greatest importance to the commercial world is now being made on the east coast of England by the Telegraph Construction and Maintenance Company. For the last eight months the company has had several of its best operatives located in the neighborhood of the Naze. These gentlemen are hourly in communication by telephone with a lightship which is anchored ten miles out. An ordinary telephone cable has been laid from Walton-on-the-Naze to the lightship, and telephone and telegraph apparatus have been affixed to both ends. It was considered improbable that the human voice would be conducted ten miles, especially in rough weather; but this has been now proved to be thoroughly practicable. Telephonic communication with lightships would be a boon to mariners and merchants. It is stated by the gentlemen engaged at Walton that the telephone will act over twice ten miles; and there is no reason why some day it should not act over much greater distances."

"An interesting paper has been communicated to one of the California scientific societies on the fossil wood which is found in different localities throughout the State. This silicified wood is stated to be a variety of quartz; the wood fiber is gradually replaced by quartz, leaving the form of the wood intact, so much so that sections cut and placed under a microscope show the characteristic grain of the wood, by which the genera may often be determined, and sometimes the species."

"Starvation, semi-starvation, 'banting,' alkalies, purgatives, Turkish baths, exercise and the thousand and one ways of reducing corpulency to respectable dimensions still leave a large section of our stout population in despair. M. Germain See comes to the rescue. 'Oh, ye massive fat ones desiring to be made lean, eat not much meat, but drink enormously of tea.'"

"Street locomotion by steam has just made a great stride in the domain of practice. Hitherto we have been accustomed to seeing heavy locomotives, weighing several thousand pounds, hauling carriages at a speed much less than that of the horse and resembling road-rollers for crushing stones more than anything else. Now Messrs. Dion, Bouton & Trepardoux have succeeded in manufacturing steam vehicles of all sorts and of all dimensions, from the tricycle up to the largest omnibuses and merchandise vans. Our engraving represents one of their steam phaetons—a vehicle of remarkable elegance, lightness and strength. The speed of the carriage is 18 miles per hour. Its weight with six passengers and a stoker is 3,960 pounds."

A steam carriage for street traffic

THE AUTHORS

JAMES A. VAN ALLEN ("Space Science, Space Technology and the Space Station") is the discoverer of the radiation belts surrounding the earth that are generally referred to by his name. He got his Ph.D. in physics from the University of Iowa in 1939. After stints at several research institutions and active duty as a gunnery officer with the Pacific Fleet during World War II, he returned to Iowa, where he is professor of physics and chairman of the department of physics and astronomy. Van Allen's career has run parallel to the development of the U.S. space science program. He discovered the Van Allen belts in 1958 during the mission of *Explorer 1,* the first successful U.S. earth satellite. He was principal investigator for the first studies of the radiation belts of Jupiter carried out by means of a probe traveling through those belts and was one of the discoverers of the radiation belts of Saturn. He was the chairman of the working group that developed the Voyager missions as well as of the group that developed the *Galileo* mission. He is currently serving as principal investigator for the *Pioneer 10* and *Pioneer 11* projects and as interdisciplinary scientist on the *Galileo* mission.

LEO SACHS ("Growth, Differentiation and the Reversal of Malignancy") is Otto Meyerhof Professor of Biology and head of the department of genetics at the Weizmann Institute of Science in Israel. He earned his Ph.D. in plant genetics from the University of Cambridge in 1951. In 1952 he moved to the Weizmann Institute, where he established the department of genetics and virology. In addition to the subject of the current article Sachs's research interests include prenatal diagnosis and the mechanisms of carcinogenesis.

JOHN C. BRANDT and MALCOLM B. NIEDNER, JR. ("The Structure of Comet Tails"), are members of the laboratory for astronomy and space physics of the NASA-Goddard Space Flight Center. Since 1975 they have been collaborating in work on the large-scale structure of comets and the associated plasma physics. Brandt's Ph.D. in astronomy was awarded by the University of Chicago in 1960. He is comet scientist for the *International Cometary Explorer* spacecraft mission and a member of the Astro-Halley science team; the Astro-Halley team will monitor Astro-1, a package of ultraviolet and visible-light telescopes that is scheduled to be carried aboard the space shuttle in March, 1986. Niedner's Ph.D. in astronomy was granted by Indiana University in 1979. He is chairman of the Astro-Halley science team.

DAVID M. PEPPER ("Applications of Optical Phase Conjugation") is staff physicist in the optical-physics department at the Hughes Research Laboratory. The son of survivors of the Holocaust, he received his undergraduate education at the University of California at Los Angeles and his Ph.D. in applied physics from the California Institute of Technology in 1980. In addition to the subject of the current article his interests include optical information processing, quantum electronics and laser chemistry. Pepper is an adjunct professor of mathematics and physics at Pepperdine University.

PETER A. RONA ("Mineral Deposits from Sea-Floor Hot Springs") is senior research geophysicist at the National Oceanic and Atmospheric Administration and adjunct professor of marine geology and geophysics at the University of Miami. He got an A.B. at Brown University in 1956 and an M.S. in geology at Yale University the next year. Thereafter he went to work for the Standard Oil Company of New Jersey (now part of the Exxon Corporation) as an exploration geologist. After about two years there he decided that the future of the earth sciences lay under the oceans and moved to Columbia University's Hudson Laboratories to get experience in oceanography. He supplemented that experience by earning his Ph.D. in marine geology from Yale in 1967. For the past 25 years he has been directing oceanographic expeditions to the Atlantic, Pacific and Indian oceans.

GERALD A. ROSENTHAL ("The Chemical Defenses of Higher Plants") is professor of biological sciences and toxicology at the University of Kentucky. He went to the State University of New York at Syracuse as an undergraduate, continuing his studies at Duke University, which granted his Ph.D. in plant physiology and biochemistry in 1966. After three years at Case Western Reserve University he joined the faculty at Kentucky. He has been a visiting professor at the Seoul National University and the Hebrew University of Jerusalem. He saw his article into print from France, where he is currently Fulbright-Hays Senior Research Scholar at the Université Louis Pasteur in Strasbourg.

ROBERT E. M. HEDGES and JOHN A. J. GOWLETT ("Radiocarbon Dating by Accelerator Mass Spectrometry") approach their common subject from quite different starting points. Hedges began his scholarly career as a physical chemist; that is the discipline in which he received his Ph.D. from the University of Cambridge in 1968. Thereafter he spent four years doing work in chemical physics before taking up his present job as a member of the Research Laboratory for Archaeology and the History of Art at the University of Oxford. In his 13 years there he has concentrated on applying the findings of physics and chemistry to archaeology. Gowlett is senior archaeologist at the Oxford Radiocarbon Accelerator Unit. His training began as an undergraduate at Cambridge, when he helped to dig at late-prehistoric sites and Roman sites in Britain. After his graduation in 1972 he undertook field work at ancient sites in Kenya. That work formed the basis of his doctoral dissertation; he got his Ph.D. in archaeology from Cambridge in 1979 and joined the accelerator unit in 1980.

ANDREW R. BLAUSTEIN and RICHARD K. O'HARA ("Kin Recognition in Tadpoles") are respectively associate professor and research associate in the department of zoology at Oregon State University. After receiving his B.A. in 1971 from Southampton College in New York, Blaustein went west to pursue the study of behavioral ecology at the graduate level. The University of Nevada at Reno provided a good place to study the behavior of desert rodents; he earned his M.S. in zoology there in 1973. In 1978 the University of California at Santa Barbara awarded him his Ph.D. in biology for studies of the behavioral aspects of competition in small mammals. He moved to Oregon State the same year. O'Hara received both his B.S. (1972) and his M.S. in zoology (1974) at Michigan State University. He then went to Oregon State, where he earned his Ph.D. in zoology (1981) and stayed on to investigate the development of behavior in individual organisms and its evolution in species.

DIANE B. PAUL, who reviews *In the Name of Eugenics: Genetics and the Uses of Human Heredity,* by Daniel J. Kevles, is associate professor of political science at the University of Massachusetts at Boston and this year a visiting scholar at Harvard University's Museum of Comparative Zoology.

The Great Scot Parka, the cargo jean, and other "clothing that works" from Lands' End.

At Lands' End, we have a special affection for "clothing that works". Clothing that puts function first, fashion second. The kind we wear ourselves.

So we're forever adding our two cents' worth of suggestions on how various items can be improved in design and manufacture.

Every now and then, this practice really pays off, as it has in the development of our Great Scot Parka—arguably the most functional piece of outerwear you can buy.

Great Scot Parka $69.50

We give it all the features you expect in a mountain parka: weather-resistant shell, drawstring hood, drawstring waist, adjustable cuffs, storm flap. Features that allow you to seal out wet and cold. And then, in our usual Lands' End manner, we give it an extra feature you may not expect: a warm wool lining. A lining that extends even up into the hood, to keep your ears warm. All that done, we then give our Great Scot a price that represents significant value: just $69.50 ($64.50 for children).

"Tough as the game" Rugby Shirts.

Not your flimsy sport shirts, but authentic, tough 100% cotton jerseys tested in rough and tumble action. Gussetted under the arms for easy movement, and equipped with rubber buttons that won't gouge you. Best feature? They're preshrunk, so they'll fit for as long as you wear them. Just $36.50 (and wait till you see the colors)!

The backpack of jeans.

Cargo room. That's the story with our 6-pocket Cargo Jean. You can fill the pockets with just about everything you need for a day's hiking or sailing: snacks, compass, blues harp, Proust. Even strap gear to the sturdy tunnel belt loops. This jean won't let you down. And it has the full, easy-moving cut that all Lands' End garments give you. The price? Just $23.50.

6-Pocket Cargo Jean $23.50

First quality, then price.

Along with "clothing that works", our catalog also offers all manner of

"Tough as the game" Rugby Shirt $36.50

traditional dress clothing, soft luggage, and other quality merchandise from around the world. Everything is direct-to-you, at prices that give new meaning to the word "value".

We offer you first, quality. Then, price. And always, always service.

Plus a guarantee so unconditional, we state it proudly in just two words:

GUARANTEED. PERIOD.

If we're new to you, we don't ask that you trust us. Simply try us. Call toll-free 24 hours a day at 800-356-4444.

Or, fill in the coupon and let us send you a free copy of our latest Lands' End catalog.

LANDS' END® DIRECT MERCHANTS

of fine wool and cotton sweaters, Oxford button-down shirts, traditional dress clothing, snow wear, deck wear, original Lands' End soft luggage and a multitude of other quality goods from around the world.

Please send free catalog.
Lands' End Dept. Q-04
Dodgeville, WI 53595

Name ____________________

Address ____________________

City ____________________

State __________ Zip __________

Or call Toll-free:
800-356-4444

COMPUTER RECREATIONS

How close encounters with star clusters are achieved with a computer telescope

by A. K. Dewdney

Deep in space, a star cluster performs a cosmic dance to the tune of gravity. During a human lifetime the stars barely move; over a longer span, in which years are equivalent to seconds, they trace out a tangled figure of orbits. Occasionally a single star encounters a neighbor in a pas de deux that hurls it out into space. If such escapes are more than occasional, the cluster gradually shrinks and the core begins to collapse.

A powerful telescope can reveal the structure of some clusters in our galaxy but it cannot compress years into seconds—only a computer is able to do that. A computer can also be programmed to serve as a kind of telescope for viewing hypothetical clusters. At cosmic speed one can watch the movement of the members of a cluster as a succession of snapshots in which each star leaves a dotted trail that weaves through the cluster [*see illustration on opposite page*].

Do gravitational forces alone account for the evolution that astronomers infer from observed clusters? Computers help to find answers to this question and related ones. A conference of simulators and theoreticians met at Princeton University in May, 1984, to discuss the consistency of hypothetical and actual star clusters. It was the 113th symposium of the International Astronomical Union; the entire symposium was devoted to the dynamics of star clusters.

It is relatively easy to choreograph a cosmic ballet. In principle the stellar interactions within a cluster are classically simple: both members of a pair of stars experience a gravitational force that is proportional to the inverse square of the distance between them. The force is also proportional to the product of the two stellar masses. Such a formula is easy to compute: multiply the masses together; then multiply the product by a constant of proportionality and divide by the square of the distance between the two stars. The sum total of all such paired forces acting through time presumably determines the pattern of movement within the cluster. A program, called CLUSTER, computes the sum of the forces for each star and moves the sum from its present position to a new one nearby. It does this repeatedly during centuries of simulated time.

A certain tedium attends typing in the coordinates and velocities of many stars, but once this is done an armchair universe unfolds on the display screen. Stars at the center of the cluster follow wobbly, erratic courses; those at the periphery drift away, stop and then glide back. The most interesting events include close encounters and escapes.

When two stars approach each other closely, they impart a tremendous gravitational boost to each other and speed apart. Escapes are usually the result of one or more close encounters. When a star speeds away from its cluster, there are only two possibilities: either the star returns or it does not. An astronomical body has an escape velocity that depends on its mass and on the mass of the body or object from which it escapes. If the velocity is attained by a star moving outward from its cluster, it will never return. Inexperienced cluster buffs are likely to witness frequent escapes from the configurations they design. In fact, a common initial experience is to see one's hoped-for dance disintegrate. It is wise to practice by building a system of two or three stars.

The structure of the CLUSTER program is simple. It consists of an initialization loop followed by a double loop. Within the double loop the acceleration, velocity and position of each star are updated according to the summed attractions of the other stars. I shall describe a particularly simple version of the program in which the time increment, force constant and stellar masses are all built in. In spite of its simplicity, however, this version of CLUSTER seems capable of simulating almost the entire range of cluster behavior. Three sets of arrays are used. The first set keeps track of the accelerations currently experienced by the stars in each of three coordinate directions. The arrays are called *ax*, *ay* and *az*. Thus $ax(i)$, $ay(i)$ and $az(i)$ indicate the x, y and z components of the ith star's acceleration. The contents of the three arrays alone do not need to be initialized at the start of the program. The second set of arrays, *vx*, *vy* and *vz*, define velocities: $vx(i)$, $vy(i)$ and $vz(i)$ register the x, y and z components of the ith star's velocity. The third set of arrays record positions: $x(i)$, $y(i)$ and $z(i)$ are the ith star's x, y and z coordinates of position. The starting values for the arrays x, y, z and *vx*, *vy*, *vz* must be initialized at the head of the program.

The main body of the CLUSTER program follows the initialization segment. The double loop can be entered and reentered endlessly, or the programmer can establish the specific conditions that control reentry. The outer loop considers each star in turn and sets the acceleration components to zero. After this has been done the inner loop computes the forces produced on each star by its companions in the cluster.

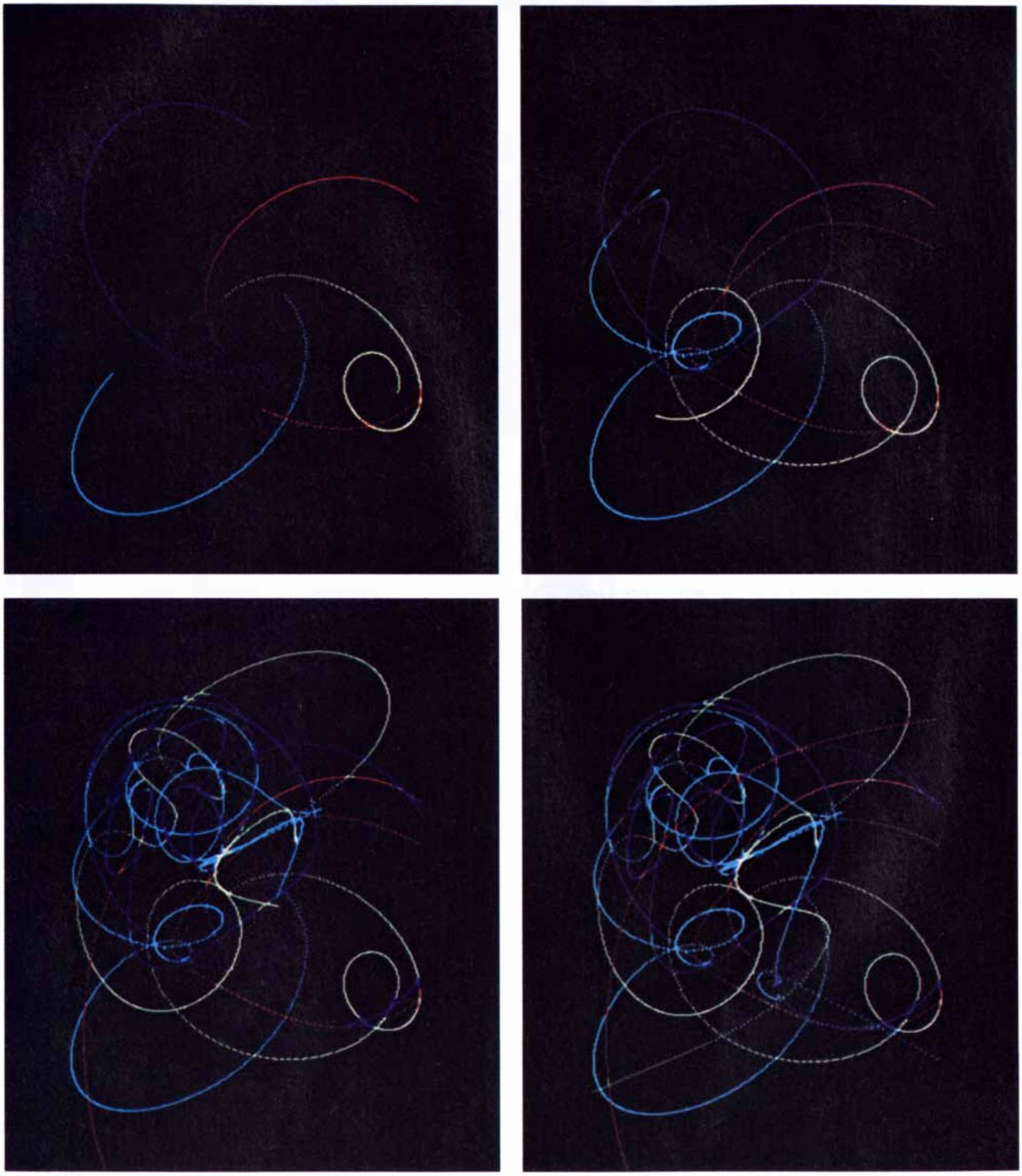
For example, let us assume that the index of the outer loop is i and that the inner-loop index is j. The inner loop first checks to determine whether i is equal to j. If it is, the program does not invoke the force computation: a star does not attract itself. In any event, to compute force under the circumstance would cause the computer to attempt division by zero. (This is the only situation that can actually make me feel sorry for a computer.) When i and j are not equal, CLUSTER uses Euclid's formula for distance between the stars: the differences of the x, y and z coordinates are squared and added together. The result, of course, is the square of the distance. Next, the inner loop tests whether this number is 0. If it is, an alarm of some kind should be raised because the computer is about to be asked to divide by zero. My version of the program prints "Collision!"

If nothing is amiss, the inner loop computes the distance between the stars by taking the square root d of the squared distance computed earlier. It then divides 1,000 by the square of the distance, a calculation that yields the force. The final task to be performed within the inner loop is to determine the acceleration components of the *i*th star. This value is obtained by adding together the force contributions from the other stars. For example, the *x* component of acceleration can be written generically as follows:

$$ax(i) \leftarrow ax(i) + f \times (x(j) - x(i))/d$$

Here *f* and *d* represent the force and distance. The ratio of the *x* distance between the *i*th and *j*th stars to the total distance is precisely the fraction of the force that acts on the *i*th star in the *x* direction. The *y* and *z* components of acceleration are computed by analogous formulas.

Two more loops, one following the other, complete the program. The first updates velocity and the second updates position. There is a subtle point here, first brought to my attention by John H. Hubbard, the Cornell University mathematician whose advice on computing the Mandelbrot set was eminently useful [see "Computer Recreations," August, 1985]. It is indeed possible to compute position before

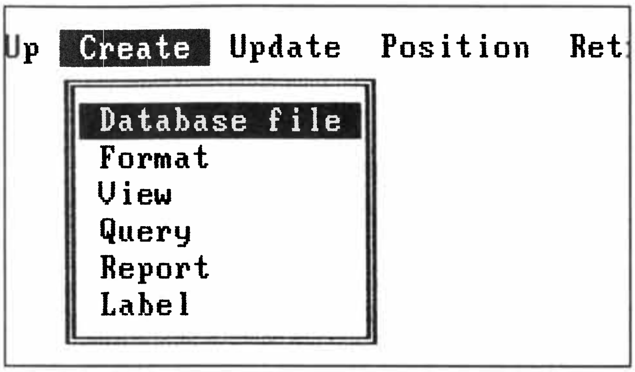


Four stars put on a cosmic ballet for a few years and then leave the stage

The database used now be used

Introducing dBASE III® PLUS.

The PLUS stands for all the improvements we've made to the world's number one selling database management software.



The Assistant helps beginning users accomplish day-to-day data management tasks without programming.

Mind you, dBASE III PLUS still has the powerful dBASE programming language, dot prompt, and all the features that have made dBASE III the standard of the industry.

We've simply raised the standard.

And just as dBASE III introduced more power to the people, our new dBASE III PLUS introduces more people to the power.

People who aren't all that crazy about programming, for example.

The Assistant feature in dBASE III PLUS now provides them with new easy-to-use pull-down menus for creating, using and modifying multiple databases.

So now anyone who can manage a simple cursor can manage day-to-day data management tasks. Without programming.

And by using our new Screen Painter, anyone can create custom screens. Without programming.

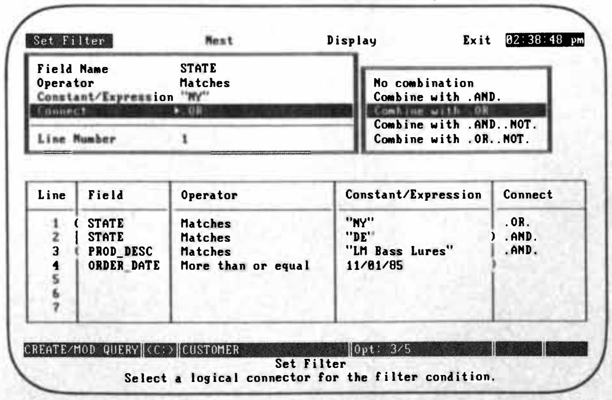
Or using View, access related information in several databases at one time. Without programming.

With Advanced Query System, another new non-programming feature, any user can build complex queries just by selecting from the dBASE III PLUS pull-down menus.

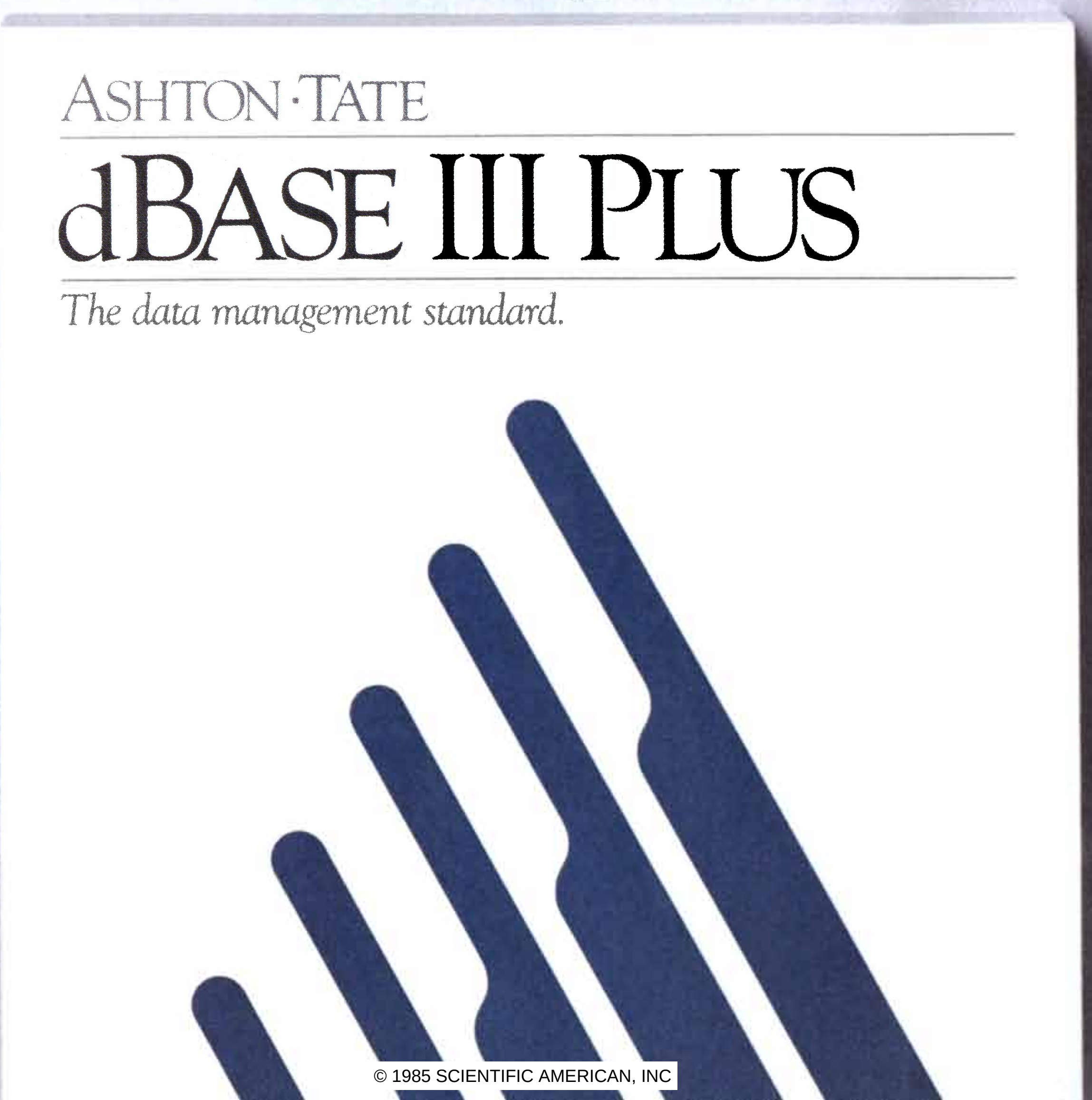
For rapidly creating entire programs, there's even a new Applications Generator.

And for all those who wish to learn to program, the Assistant can be of further assistance. By teaching you programming commands as you go along. Without disrupting your work flow.

These are only a few of the dBASE III PLUS features that can help new users quickly get up to speed. And experienced users quickly increase their speed. (Sorting, for example, is up to two times faster and indexing up to ten times faster than dBASE III.)



Advanced Query System lets you set up and answer complex queries without programming.



ASHTON·TATE
dBASE III PLUS
The data management standard.

INTRODUCING THE NEW

Take the Eldorado hallmarks: sporty elegance and legendary luxury. Add aerodynamic design, exhilarating performance and microchip technology. The result is an Eldorado that sets a new standard for personal luxury cars.

AERODYNAMIC DESIGN.

From smoothly curved fenders to tailored door-into-roof design, this is the most aerodynamic Eldorado ever. That means wind noise is reduced. Road stability is remarkable. And the styling is timeless.

GM

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

computing velocity without producing strange-looking results. Yet the motions of the stars would in time become strangely wrong, because such an operation would violate the law of energy conservation.

The velocity-updating loop merely adds acceleration to velocity, according to the following formula:

$$vx(i) \leftarrow vx(i) + ax(i)$$

Here it is assumed that the time increment equals the time unit in which velocity is expressed. The same kind of formula is used to calculate vy and vz. The position calculations done in the final loop are equally simple:

$$x(i) \leftarrow x(i) + vx(i)$$

The entries of the y and z arrays are similarly updated. Drawing on the information from the final loop, CLUSTER places each point on the two-dimensional surface of the display screen. It does so by plotting the first two position coordinates while suppressing the third. The natural result of this arrangement is that z represents depth; it is easy to imagine that one is looking into space behind the screen. The numbers produced by the cluster-simulation program are sometimes very large and sometimes very small. For this reason it is advisable to utilize double-precision arithmetic so that all relevant numbers are not inadvertently rounded.

The time taken by CLUSTER to finish one cycle of computation depends on the number of stars in one's system. As few as 10 stars will produce aesthetic intricacy; 100 or even 1,000 stars are needed to produce realistic complexity. Unfortunately the number of steps in the basic computational cycle increases as the square of the number of stars in the cluster. Although stellar simulators have found a neat method that dodges this particular limitation, other problems still arise.

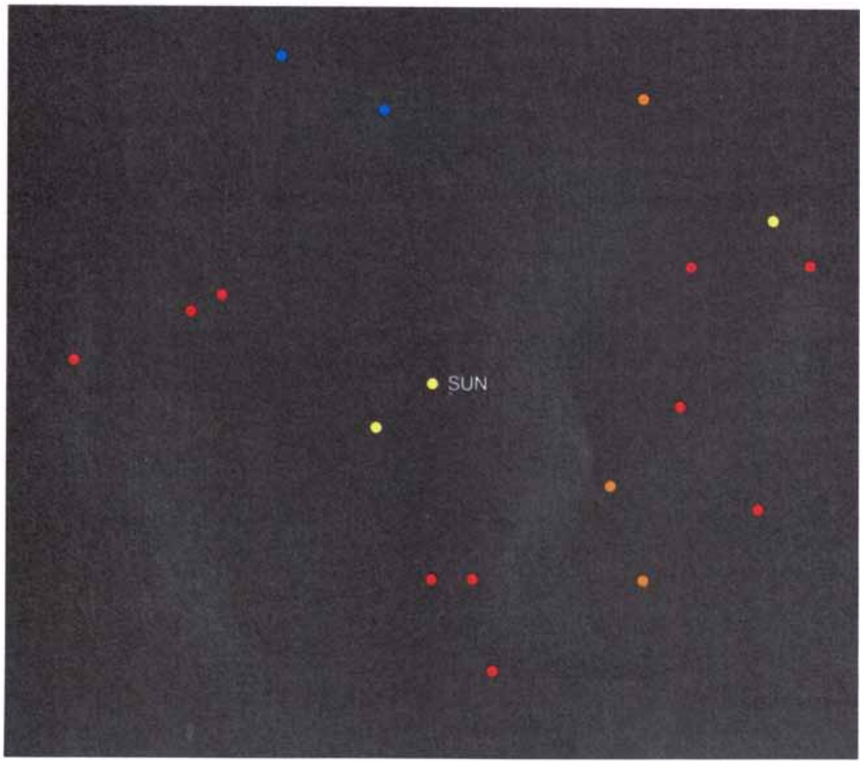
The worst problem emerges from the fact that the program is a discrete system attempting to mimic a continuous one. Continuous orbits are approximated by a sequence of jumps that depart increasingly from a star's true path through a cluster. The inaccuracy might be corrected to some extent by the presence of statistical regularities, but in close encounters between stars the system unnaturally and disastrously magnifies the slingshot effect.

For example, if the computational cycle puts one star (Stella) close to another star (Aster), a powerful gravitational pull magnifies the acceleration components of both stars. The magnification percolates through the computation to the velocity components and thence to the position coordinates. The next iteration finds Stella already widely separated from Aster and unable to repay the gravitational loan. A fiction of excess kinetic energy has been created. Artificial clusters afflicted by this problem evaporate even faster than real ones. There are two ways around the difficulty; one is hard, the other is easy. The hard alternative requires computation of a Keplerian orbit for the pair. The orbit is maintained as long as the two stars are in proximity. Theorists regard this as the method of choice because the orbital formula is perfectly accurate. An easy but occasionally inaccurate way to handle close encounters is to subdivide the time steps in the basic computational cycle. Readers may want to add this particular maneuver to the advanced version of CLUSTER that I shall now describe.

A program called SUPERCLUSTER can be derived from CLUSTER by a series of simple modifications. First, SUPERCLUSTER incorporates stars of different masses in its ballet. This is easily done at the start by entering the masses in an array called m. The force computation becomes somewhat more complicated: force is no longer proportional to $1/d^2$ but to the product of the masses divided by d^2. Next, SUPERCLUSTER incorporates spectral types. As in the case of mass, an array (called *spec*) must be filled in before the run. It is used, however, only during the display phase of the basic cycle. The colors range from blue for O-type stars to red for M types. Green is omitted. The third enhancement of CLUSTER makes arbitrary time steps possible in either version of the program.

SUPERCLUSTER uses a time-step variable called *delta*. Specified at the beginning of a run, *delta* determines the amount of simulated time between successive cycles. Naturally this time element must affect the updating formulas for both velocity and position: in the velocity formulas it multiplies acceleration and in the position formulas it multiplies velocity.

The easy way to handle close encounters can now be described. First a definition of "close" must be established. Then a test for such closeness can be inserted into the program just after the point at which the distance between two stars is calculated. If a close encounter is taking place, SUPERCLUSTER replaces *delta* by one-tenth of its value—at least until no pair of stars is that close again. This expedient certainly helps to cushion the sudden lurches of discrete gravity. It creates even worse problems when encounters are really close, however. An approach that is 10 times closer now results in a gravitational force that is 100



Would our galactic neighborhood form a cluster?

times greater! Fortunately close encounters of the worst kind are rather rare. The time-subdivision technique has been standard in cluster-simulation programs traditionally employed by professionals.

If SUPERCLUSTER is to be an astronomically meaningful program, units for distance, mass and other aspects of physical reality are needed. A convenient measure of distance is the astronomical unit (AU), which is equal to the earth's mean distance from the sun. Mass can be measured in solar masses and time is best measured in years. Under these conventions the universal constant of gravitation has the approximate value of 39. SUPERCLUSTER uses this constant instead of 1,000 in the force calculation.

All is now in readiness for putting either program to work. A preliminary exercise for CLUSTER involves four stars. Place them at the corners of a square that is an inch or two wide on the screen. It is only fair to give each star a nonzero z coordinate as well as the x and y coordinates that were mentioned above. If motion is confined to the plane of the screen, close encounters are that much commoner. Velocity components should be small (on the order of -5 to $+5$) and should specify a clockwise direction, as though the four stars were on a wheel.

SUPERCLUSTER can be tried on the system of stars shown on page 22. This is the earth's galactic neighborhood. What would happen if the sun and its neighboring stars were cut loose from our galaxy and allowed to dance endlessly in space? Would a cluster form? The question may or may not have scientific relevance, but it is fun to answer. Besides, these are the only stars for which positions and velocities are known accurately [*see illustration below*].

Clusters of stars are either open or globular. Open clusters consist of 1,000 or so stars, whereas globular clusters may consist of millions. So far investigators such as J. Garrett Jernigan at the University of California at Berkeley Space Sciences Laboratory have been able to handle only small clusters. Globular clusters are currently intractable. Even so, Jernigan and pioneering colleagues such as Sverre J. Aarseth of the University of California at Berkeley have been observing collapses of computer clusters for decades. The extent of collapse is measured by considering a spherical volume that is centered within a cluster and contains 10 percent of its mass. The radius of this volume is known as the 10 percent radius. Collapse is under way when the 10 percent radius decreases as time passes. Inexorably the core of a simulated cluster becomes ever denser. Since the simulated stars are mathematical points, nothing terrible ever happens to such clusters. No black hole comes into being at the center. This at least has been the experience of cluster theorists. But we seem able to find little evidence of extreme collapse in the clusters overhead. Something is preventing collapse out there.

Both traditional and modern simulation experiments may provide a key. On various occasions a small number of binary star systems at the center of a simulated cluster have brought the collapse of core regions virtually to a halt. In one of Jernigan's experiments a single binary seemed to be responsible. How is it possible? According to Jernigan's graduate student David Porter, it may be that "very tight binaries whizz around each other very quickly and kick wandering stars energetically around the core or even back out to a looser collection of stars around the core called the halo. This could be a mechanism for preventing the core from getting too crowded."

Jernigan used to be an observer of X-ray stars. As research focused on the search for X-ray sources in clusters, he grew increasingly interested in clusters as astronomical objects in their own right. Simulation seemed an effective way to investigate them.

Self-described as "the new kid on the block," Jernigan has discovered an important new efficiency in simulation efforts. In CLUSTER and similar programs a single computational cycle for n stars requires roughly n^2 steps. Jernigan's cycle needs only $n \times \log(n)$ steps. He organizes his cluster by grouping the stars into neighboring pairs. Each pair is then replaced by a fictitious mass and velocity that summarizes the behavior of the pair. The same process is now applied to the pairs as if they were the original stars. Continuing in this manner, a collection of grouped and regrouped mass nodes is built up in a data structure called a tree. The single node at its root simultaneously represents all the stars. Motions can then be calculated for the central node and

NAME OF STAR	POSITION COORDINATES			VELOCITY COORDINATES			COLOR	MASS
	X	Y	Z	VX	VY	VZ		
STRUVE 2398	68	–365	631	–5.69	4.76	3.35	RED	0.26
ROSS 248	464	–42	450	–8.75	1.13	–15.45	RED	0.17
61 CYGNI	394	–377	433	–2.78	22.03	0.02	ORANGE	0.69
LALANDE 21185	–404	107	307	7.32	–0.47	–20.11	RED	0.39
PROCYON 5	–295	658	68	2.38	0.75	–3.65	BLUE	1.29
BARNARD'S STAR	–7	–371	30	–0.87	24.20	16.78	RED	0.21
EPSILON ERIDANI	408	534	–114	4.60	0.69	–0.50	ORANGE	0.74
WOLF 359	–462	136	62	–0.82	9.86	–5.94	RED	0.10
SIRIUS	–98	514	–157	1.89	–2.21	–2.59	BLUE	2.96
LUYTEN 726–8	487	219	–175	2.08	10.80	–0.41	RED	0.19
ROSS 128	–683	44	13	2.51	–2.32	–4.09	RED	0.21
SUN	0	0	0	0.00	0.00	0.00	YELLOW	1.00
TAU CETI	646	307	–208	0.52	–6.62	3.92	YELLOW	0.85
ALPHA CENTAURI	–106	–86	–243	–1.95	4.68	4.51	YELLOW	1.03
LUYTEN 789–6	608	–235	–182	–6.75	10.81	10.56	RED	0.13
LUYTEN 725–32	718	227	–233	4.70	6.16	0.51	RED	0.21
ROSS 154	111	–536	–241	1.79	1.36	–0.11	RED	0.24
EPSILON INDI	334	–194	–594	–3.54	17.71	2.28	ORANGE	0.69

A table listing all but three stars in the neighborhood of our solar system for all its branches out to the individual stars.

Is this the technique of the future? It certainly helps to speed things up, according to Jernigan. Yet subsequent generations of cluster programs are more likely to resemble the hybrid variety used by Alan P. Lightman of the Center for Astrophysics of the Harvard College Observatory and the Smithsonian Astrophysical Observatory and Stephen L. W. McMillan of the University of Illinois at Urbana-Champaign: stars in the core are handled by the direct simulation methods described above; stars outside the core are modeled statistically as if they form a gas.

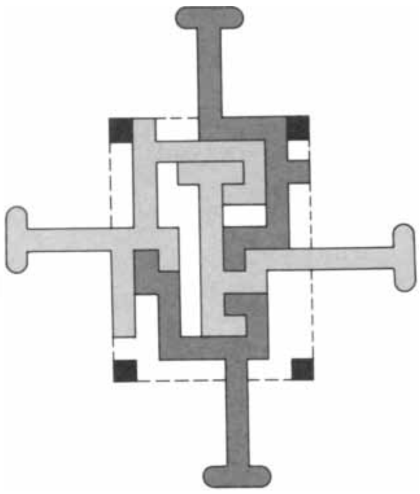
For readers proficient in the language called APL there is an interesting new publication by Gregory J. Chaitin of the IBM Thomas J. Watson Research Center in Yorktown Heights, N.Y. It is called *An APL2 Gallery of Mathematical Physics* and is a 56-page booklet containing explanations of five major physical theories, including those that describe both the Newtonian and the relativistic motion of satellites in space. APL listings are given for computer programs that illustrate each theory. Chaitin will be happy to send a copy to any reader who writes to him at the Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, N.Y. 10598.

In last October's column I described three puzzles: Bill's baffling burr, Coffin's cornucopia and Engel's enigma. Hundreds of readers have tackled the puzzles. While some seek the magic combination of moves that disassemble the burr, others scratch their head over the placement of polyominoes in a tray. Members of this group will have to get by without help from their friends: each puzzle is unique. Still other readers keep rotating the wheels of Engel's enigma in a vain attempt to unscramble it. Some of the devotees are succeeding, at least on an abstract plane: claims of solutions to the enigma have started to come in.

A call for two-dimensional burrs brought in a number of designs. The most charming design received so far is shown at the top of this page. The problem is to remove the four pieces from the tray symbolized by the rectangular outline. The pieces can only be moved in four directions confined to the plane of the page: up, down, left and right. The four corner squares are regarded as immovable. Which piece must be moved first? Jeffrey R. Carter of Littleton, Colo., designed this two-dimensional tour de force. Our three-dimensionality confers the advantage of visualizing the whole; a two-dimensional solver would push and pull at the sides of a mysterious box.

An algorithmic solution to Engel's enigma is claimed by P. Clavier of Dallas, Tex. Clavier says that his program, written in BASIC and running on a Texas Instruments CC-40 portable computer, solves typical scrambles in from 300 to 700 moves. The solution implements six fundamental exchange operations on the stones and bones. Readers who used the sequence representation I suggested may have cast their net too widely; solutions of the numerical sequence are not always solutions of the enigma. In framing the suggestion I was aware that bones were excluded from the representation. "Well," said I at the time, "take care of the stones and the bones will take care of themselves." Not so. The stones should be interleaved with the symbols that represent the bones.

The ultimate scramble-unscramble puzzle appears to have been invented by Robert Carlson of Los Altos, Calif. It is so complicated to make that he must be content with the view of it on his monitor. The puzzle is an icosahedron, the Platonic solid that consists of 20 triangular faces. Each vertex is the site of a possible scrambling operation.



Jeffrey R. Carter's two-dimensional burr

When a vertex is rotated, the five incident triangles are rotated as well. Each triangle has three colors. In unscrambled form the colors adjacent to each vertex are the same. Carlson has prepared a version of his computer puzzle for the IBM PC. In addition to colors it features a musical note for each move. Interested readers can obtain a disk by writing to Carlson at 319 Lunada Court, Los Altos, Calif. 95030.

Globular cluster Messier 13 in the constellation Hercules

You've driven in a sports car. Remember the way it hugged the road? The way it took the turns? The way it accelerated? The way you felt?

(You could accept the negatives, such as comfort which approximated that of the front seat on a roller coaster.)

You've also driven in a luxury car. Remember the legroom? The storage space? The relaxing seats? The way you felt?

(You could also accept the negatives, such as the fact it drove like the Queen Mary.)

Saab asks you not to accept negatives. Rather, add up the positives of both aspects of the new Saab 9000.

On the sports car side, there's performance in the form of a 16-valve, intercooled, turbocharged engine that takes a car from 0 to 60 in hardly any seconds and maintains speed and fuel efficiently for hours on end.

A suspension with McPherson struts that sits a Saab 9000 on a road as if car and asphalt or car and macadam or car and dirt were one.

The steering, of course, is rack-and-pinion, so the driver can't help feeling and knowing what's happening between the tires and the road.

The brakes are large-diameter discs, power-assisted and, with a dual-circuit system, almost "fail-safe."

On the luxury sedan side, there's Automatic Climate Control. You tell your Saab 9000 what temperature makes you feel the most chipper and it, through microprocessors, keeps you happy.

Size: People in government (the EPA) charged with such things have declared the Saab 9000 a "large" car.

Considering the legroom, the elbowroom, and the carrying space (up to 56.5 cu. ft.) that may even be a mite bit of an understatement.

The touches: Seats and a steering wheel that adjust to even the most extraordinary anatomy.

Instruments? Within your reach and so well-thought-out that your eyes, so used to being

assaulted on the road, will come to appreciate the soothing green illumination.

The exterior? Well, check the pictures on this page for a few seconds. Then see a new Saab 9000 for yourself at a Saab dealer, where a perfect balance always exceeds the sum of its parts.

SAAB

The most intelligent cars ever built.

Saab 9000 5-speed: 21 EPA adjusted city mpg. Use adjusted mpg for comparison only. Mileage varies with speed, trip length and weather

BOOKS

A history of the eugenics movement and of its multiple effects on public policy

by Diane B. Paul

IN THE NAME OF EUGENICS: GENETICS AND THE USES OF HUMAN HEREDITY, by Daniel J. Kevles. Alfred A. Knopf, Inc. ($22.95).

In October of 1955 Frederick Osborn, secretary of the American Eugenics Society, invited H. J. Muller, a prominent geneticist, to join the organization's board of directors. The invitation was one of many that Osborn extended to distinguished scientists in his efforts to reform and rehabilitate a movement then in disrepute. Eugenics—defined by its founder, Francis Galton, as "the science of improvement of the human germ plasm through better breeding"—had once enjoyed considerable support. In the last two decades of the 19th century and the first two of this it was widely assumed that most differences in human mentality and behavior resulted from differences in genes. "Feeblemindedness," criminality, pauperism, shiftlessness and insanity were only a few of the traits ascribed to bad heredity. That society ought to foster the breeding of those who possessed favorable traits ("positive" eugenics) and discourage or prevent the breeding of those who did not ("negative" eugenics) seemed obviously to follow.

In a lucid and intricate history of eugenics Daniel J. Kevles, professor of history at the California Institute of Technology, demonstrates that notions of genetic perfectability have been deeply entrenched over a period of time that runs back through many decades. These concepts underlay even the most sharply defined fissures in the Victorian intellectual and moral landscape. Victoria Woodhull, who outraged her contemporaries by running for president of the U.S., advocating female suffrage and calling for reformation of the relation between men and women, lined up with her critics on the question of breeding—in both its literal and its figurative sense. In her 1891 book *The Rapid Multiplication of the Unfit* she wrote: "The best minds of today have accepted the fact that if superior people are desired, they must be bred; and if imbeciles, criminals, paupers, and [the] otherwise unfit are undesirable citizens they must not be bred."

Through the act of writing Woodhull was actually winning a skirmish in her main battle. If human reproduction has an importance that makes it a major social and political issue, it is inescapable that women have a role to play in the political as well as the biological process. Indeed, Kevles points out, women were major figures in the leadership of the movement and made up more than half of the membership of the British Eugenics Education Society. They were, to be sure, the right kind of women. Socially desirable traits, resulting from good heredity, were thought to explain the success of people who were white, Anglo-Saxon, Protestant and middle class.

The assumption that social position accurately reflected genetic worth was severely shaken by the Great Depression. Biological defect no longer appeared a convincing explanation for poverty and other social ills. The result was a movement, led by geneticists, to purge eugenics of its evident class bias and place it on a firm scientific foundation. This effort was hardly under way before it was itself threatened by revelations of Nazi policies, which produced a backlash against eugenics of any kind. By the late 1940's eugenics—whether reform or establishment—was out of fashion.

Osborn hoped to reverse this trend by effecting radical changes in both the policies and the personalities associated with the movement. He wished particularly to attract to his board geneticists such as Muller and Theodosius Dobzhansky, who had been publicly critical of the simplistic scientific assumptions and racial and class bias of the old Eugenics Society. Osborn's efforts were generally successful, but he could not convince Muller to become a director, or even a member.

Muller's refusal was not prompted by opposition to the society's aims. Indeed, he was more ardent in his eugenic convictions than many who served on the society's board. He recognized, however, as they perhaps did not, that a successful eugenics required not only a new content but also a new name. As early as 1950 he chose an apparently neutral title, "Our Load of Mutations," for what is in fact a eugenic plea. One of Muller's critics has called the move, which produced a revolution in the vocabulary of population genetics, a "stroke of genius," for while "eugenics" was out, "genetic load" and its associated concepts, "mutational load," "balanced load" and the "cost of selection," were in. As Muller wrote in reply to Osborn, an argument explicitly tagged with the eugenic label or its obvious equivalents would have been dismissed in advance by many whom it did, in fact, influence. Muller did not care what terms they used, as long as people could be induced to think eugenically.

By the 1960's even Osborn was convinced that the word had outlived its usefulness. "Eugenic goals are most likely to be attained under a name other than eugenics," he wrote in his 1968 book *The Future of Human Heredity.* The following year the society's journal, *The Eugenics Quarterly,* became *The Journal of Social Biology.* As early as 1954 its British counterpart, *Annals of Eugenics,* had been renamed *Annals of Human Genetics;* the American Genetics Association, once devoted to "the improvement of plants, animals, and human racial stocks," decided in 1965 to pursue "the improvement of plants, animals, and human welfare" instead. By the late 1960's most journals and organizations with an explicit eugenic orientation had changed their name or their motto.

This is not to suggest that the changes reflected only a concern with public relations—the pouring of old wine into new bottles or the minting of new lyrics to accompany old tunes. Lionel Penrose, Galton Professor of Eugenics at University College London and a hero of Kevles' story, was not being cynical when he prompted a change in the title of both his position and his laboratory's journal. Penrose's wholesale rejection of everything associated with the movement was unusual, however. Most prominent geneticists in the immediate postwar period remained committed in principle to some form of eugenics. But at a time when a stigma attached to both word and concept, they searched for less obvious ways to pursue their aims. They turned mainly to population control and to medical genetics.

The latter raises some important and disturbing questions. The field of medical genetics was in large part developed by geneticists seeking a socially acceptable way to improve the human genome. Their efforts produced the field of genetic counseling, as well as new techniques such as genetic screening (the detection of the carriers of specific genetic disorders such as sickle-cell anemia and Tay-Sachs disease), forms of prenatal diagnosis, such as amniocentesis, and gene-therapy research. These developments Kevles collectively terms the "new eugenics." He does not mean thereby to condemn them (as he does "Jensenism" and some aspects of human sociobiology, which in his view are other and more problematic heirs of the eugenics movement). Indeed, he ultimately asserts a rather sharp distinction between the old eugenics, both establishment and reform, and the new, as reflected in genetic counseling and contemporary biomedical techniques and therapies. Kevles traces a historical continuity, in terms of intellectual leaders, sources of institutional support and the motivations of some early researchers in the field. Yet he believes the ethos of medical genetics today differs fundamentally from that of even reform eugenics. Whether this is so depends on how one characterizes the aims not only of the "new eugenicists" but also of the old. Hence it is necessary to say something more about eugenics in the period just before World War II.

In his history Kevles demonstrates that prewar eugenics enjoyed surprisingly varied support. This point has been made by others, but not in nearly such detail and therefore not as convincingly. He shows that in the first decades of this century eugenics was supported by many social radicals such as Havelock Ellis, Beatrice and Sidney Webb and George Bernard Shaw, as well as by social conservatives such as Galton, Henry Fairfield Osborn (the uncle of Frederick), Leonard Darwin and Charles B. Davenport. (The radicals used eugenic arguments to denounce restrictions on birth control, divorce and educational and occupational opportunities for women; conservatives, to defend them.) Although they differed in their attitudes toward sexuality and the "woman question," they agreed that unrestrained breeding by the lower orders was a threat to the future of the race.

The leaders of eugenic societies, in both Britain and the U.S., were mostly professionals: doctors, social workers, academics, who shared a common attitude toward the "submerged tenth" (the chronically poor, who were considered responsible for most crime and other deviant behavior). Eugenicists also joined in a commitment to the scientific understanding and control of social problems. They believed poverty and virtually all forms of deviant behavior resulted from defects in intelligence or character whose source was biological, and whose results were therefore fixed. These conditions were not the fault of the criminal or the pauper: one could no more be responsible for inheriting genes giving rise to crime or poverty than one could be responsible for inheriting genes specifying red hair (which explains why some conservatives were, and remain, hostile to any form of hereditarian thinking; it erodes traditional notions of moral responsibility). The problem of "racial deterioration" was thus viewed in new and fashionably scientific terms rather than in traditionally moral ones. Eugenicists sought the key to understanding degeneracy in the new science of genetics, and its solution in the practical application of genetic principles. Both understanding and control were thought to be matters of great urgency, since the problem of racial degeneration appeared to be rapidly growing worse.

In the past, it was supposed, the numbers of physical and mental defectives would have been kept in check by the action of natural selection, but this process had practically ceased in civilized societies. Medicine, public hygiene and charity now kept alive the physically and mentally weak. Worse, the unfit were reproducing at a much faster rate than their betters.

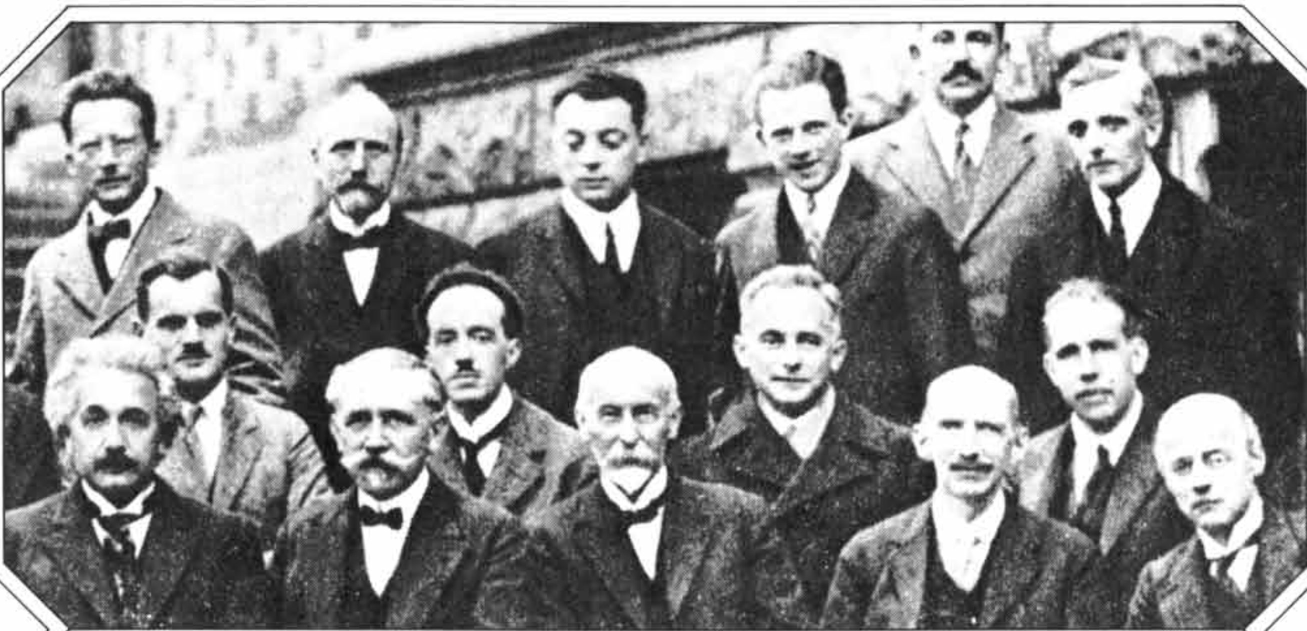
The differential birthrate was the obsessive concern and driving force of the early eugenics movement. Charles Darwin himself lamented the multiplication of the unfit, whom he identified with the lower orders. Kevles quotes a report by Alfred Russel Wallace of a conversation with Darwin: "He expressed himself very gloomily on the future of humanity, on the ground that in our modern civilization natural selection had no play, and the fittest did not survive. Those who succeed in the race for wealth are by no means the best or the most intelligent, and it is notorious that our population is more largely renewed in each generation from the lower than from the middle and upper classes."

In the U.S. these anxieties were inflamed by the "new immigration": the massive influx of southern and eastern Europeans that began in the mid-1890's. For many Americans of older stock the rapid increase in immigrants from eastern Europe, Russia, the Balkans and Italy generated intense anxiety about assimilation, biological as well as social. And it was in the U.S. that eugenic proposals were most extreme—and most effective.

In the first two decades of the 20th century American eugenicists such as Harry H. Laughlin, Davenport, Henry H. Goddard and Madison Grant published studies ostensibly demonstrating that the poor in general, and immigrants in particular, suffered disproportionately from a variety of inherited defects. Kevles cites the geneticist Davenport, director of the experimental station at Cold Spring Harbor, who thought the immigrants would produce a population "darker in pigmentation, smaller in stature, more mercurial... more given to crimes of larceny, kidnapping, assault, murder, rape, and sex-immorality."

Eugenicists were particularly con-

Winners of an American Eugenics Society "Fitter Families" contest, Arkansas State Fair, 1927

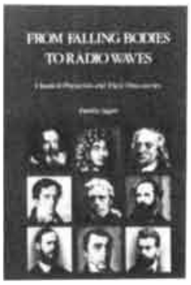


Discover the human side of physics

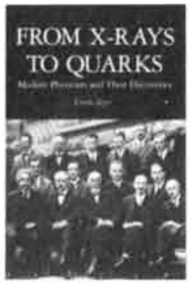
EMILIO SEGRÈ

His new book on the roots of physical thought
FROM FALLING BODIES TO RADIO WAVES

His critically acclaimed history of modern physics
FROM X-RAYS TO QUARKS



In these companion volumes, Nobel Laureate Emilio Segrè offers a warm and close-up look at the human underpinnings of scientific experiment and discovery. In examining the work of both classical and modern physicists, Segrè emphasizes the fascinating and complex personalities that are too often treated only in footnotes.



Just published, **FROM FALLING BODIES TO RADIO WAVES** traces the development of scientific thought from the works of the "founding fathers" to the more recent discoveries of Maxwell, Boltzmann, and Gibbs. In Segrè's whimsical prelude, he asks us to imagine him conversing with these scientific giants, moving within the times in which they lived. We peer through Galileo's telescope, call on an inhospitable Isaac Newton, and discuss the enigma of magnetism with Michael Faraday.

In **FROM X-RAYS TO QUARKS**, Segrè describes the evolution of atomic and subatomic physics, interwoven with the personal histories of the leading researchers, and set against the political and cultural backgrounds of those various times and places. As *The American Journal of Physics* said, "Segrè illuminates this vast and cluttered field with mastery of both detail and perspective. This excellent book is one every physicist will want. It will also be a fine gift for non-physicist friends with strong scientific interests."

In his engaging and highly readable style, Segrè communicates the inspiration, creative effort, and true drama of scientific discovery.

SPECIAL LIMITED TIME OFFER Use the accompanying coupon to order the two-volume set and save more than 25%!

Available at fine bookstores.

☐ **YES!** I would like to discover the human side of physics. Please send me:

QUAN.	ISBN	TITLE	PRICE
	1482	From Falling Bodies to Radio Waves (paper)	$13.95
	1147	From X-Rays to Quarks (paper)	$12.95
		SPECIAL TWO-VOLUME SET PRICE	$19.95
		Postage and Handling	$ 1.50
		New York, California, & Utah residents add sales tax.	$
		TOTAL	$

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

☐ MasterCard ☐ VISA ________ Account #

___/___ ________
Exp. Date Signature (All credit card orders must be signed)

Name ________

Address ________

City ________ State ________ Zip ________

Mail to: **W. H. FREEMAN AND COMPANY**
4419 West 1980 South, Salt Lake City, Utah 84104

0221

cerned with the apparent increase in "feeblemindedness": an ill-defined category incorporating a wide range of mental disabilities and socially deviant behaviors. Goddard believed the trait was generally inherited, "a condition of mind or brain which is transmitted as regularly and surely as color of hair or eyes," determined by a single (recessive) gene. He believed the defect to be widespread—afflicting from 1 to 3 percent of the population. A much larger number, he said, were carriers.

The incidence of mental deficiency was thought to be particularly high among recent immigrants. Goddard had administered a form of the Binet I.Q. test to immigrants at Ellis Island, and he concluded that about 40 percent were mentally defective. (A few years later results of I.Q. tests administered to Army draftees would be cited to support similar conclusions about blacks.)

Americans were thus even more disturbed by the prospect of racial degeneration than the British. They responded with proposals designed, in the approving words of the geneticist Edward M. East, "to cut off this defective germ-plasm." English eugenicists rarely favored compulsory measures: they stressed the need for education, financial incentives and various forms of moral suasion. The Americans were much more willing to invoke the powers of the state. "Mistaken regard for what are believed to be divine laws and a sentimental belief in the sanctity of human life tend to prevent both the elimination of defective infants and the sterilization of such adults as are themselves of no value to the community," wrote the American eugenicist Grant. "The laws of nature require the obliteration of the unfit, and human life is valuable only when it is of use to the community or race." The law of man also provided harbingers of the little corporal. By 1930, 24 states had passed laws allowing or requiring sterilization of the "feebleminded."

Eugenicists also campaigned for restrictions on immigration. In his role as "Expert Eugenics Agent" of the House Committee on Immigration and Naturalization, Laughlin conducted studies purporting to prove that the incidence of mental deficiency was much greater among the new immigrants than among the native-born. In 1924 Congress passed, and President Calvin Coolidge signed into law, a statute restricting immigration from any European country to 3 percent of the number of foreign-born of the same national origin recorded in the 1890 census, that is, before the influx of southern and eastern Europeans. (Coolidge had earlier declared that "America must be kept American. Biological laws show... that Nordics deteriorate when mixed with other races.") Only one geneticist—Herbert S. Jennings—could be found to testify against the bill. The others either agreed with its provisions or refused to take a public stand.

In the 1930's a number of distinguished geneticists—including Muller, Julian Huxley, Lancelot Hogben, J. B. S. Haldane, L. C. Dunn and Dobzhansky—repudiated what Kevles calls the "mainline" tradition. Economic depression combined with the rise of fascism to radicalize many geneticists (among other kinds of scientist) and induce at least a skeptical spirit in others. In the view of these critics, establishment eugenics traded on simpleminded science to rationalize the social and economic status quo.

This group did not abjure eugenics. Indeed, use of state power to control reproduction seemed appropriate to many geneticists of the left, particularly the Fabians and the Marxists. ("The belief in the sacred right of every individual to be a parent is a grossly individualistic doctrine surviving from the days when we accepted the right of parents to decide whether their children should be washed or schooled," asserted Hogben in his 1931 book *Genetic Principles in Medicine and Social Science.*) But neither socialists such as Muller, Haldane and Hogben nor liberals such as Dobzhansky and Jennings any longer saw the existing distribution of privileges as reflecting genetic worth. In their view it was only in a society providing equal opportunity to all its members that genetic merit could be distinguished from environmental good fortune—a distinction they considered necessary for a rational eugenics.

Just as the eugenics of the 1910's and 1920's had been used to support both traditional and unconventional views of women's roles (and both pacifism and war), so the eugenics of the 1930's and early 1940's was used to both defend and condemn the existing economic order; it has been a remarkably plastic doctrine.

As noted above the reform movement was ultimately swamped by the radical shift in social climate that followed World War II. At least, eugenics could no longer be pursued under that rubric. Some geneticists therefore undertook to search for socially acceptable ways—under socially acceptable names—to pursue their interests. And so we return to the question posed earlier: Do the new techniques and therapies that ultimately resulted (or are being developed) represent a "new eugenics"? That of course depends on what eugenics is. To define eugenics—to identify its essence, as it were—is no easy task.

If there is a core eugenic doctrine, it is the goal of genetic improvement based on some form of selective breeding; at least all the conventional definitions (for example those propounded by Galton and Davenport) imply this much. Such policies as genetic screening and abortion resulting from amniocentesis are clearly eugenic by this definition. Perhaps to avoid an unpalatable conclusion, the definition is often qualified—explicitly or implicitly—in one of two ways. Sometimes the eugenic label is applied only to policies involving some form of coercion. State laws mandating sterilization of the feebleminded would meet this criterion (as might those that only permitted sterilization, depending on how they were applied in practice; coercion can be subtle).

The problem with this standard, however, is that it excludes a larger number of policies and people ordinarily associated with eugenics. Many eugenicists, particularly in Britain, stressed the voluntary character of their proposals, holding with Havelock Ellis that "the only compulsion we can apply in eugenics is the compulsion that comes from within." Indeed, virtually all "positive" eugenics, for example Muller's scheme to inseminate women artificially with the sperm of particularly successful men (which he called "germinal choice") would be excluded. So would William Shockley's similar program, as well as his proposal to pay people of low I.Q. who agree to be sterilized.

A somewhat more sophisticated approach has been to distinguish between programs on the basis of their rationale. Policies are sometimes considered eugenic if their intent is to further a social or public purpose, such as reducing the costs borne by the sociomedical system, or even sparing future generations unnecessary suffering. For example, genetic counseling or support for biomedical research motivated by concern for the quality of the "gene pool" would be eugenic; the same practices motivated by the desire to increase the choices available to individuals would not be.

This criterion is also problematic. It makes the determination of what is eugenic depend on an assessment of motivation—a notoriously difficult task. And to the extent that we can ascertain the motives of earlier eugenicists, they seem to have often been mixed. One of the goals of the reform eugenicists at least—people such as Jennings and Hogben—was in fact to save individuals and families unnecessary hardship.

It is therefore fruitless to debate

whether the eugenic label "really" applies. Kevles' book establishes that a wide variety of policies have been advanced by groups having divergent interests and aims "in the name of eugenics." A definition broad enough to fit them all will include current technologies and therapies. A definition that broad does not necessarily condemn. In characterizing these developments as a "new eugenics" Kevles has drawn attention to the fact that medical genetics and genetic engineering have roots in the reform-eugenic program. Whether they have incorporated what was problematic in that program is not a question history can answer. (On the whole, Kevles thinks not.)

History does provide a warning, however. It reminds us how often and easily genetics has served corrupt social ends, and it alerts us to the ends it may still or in the future serve. As Kevles writes in his preface: "I am under no delusion that a history of eugenics will provide any detailed moral or political map to follow in the uncharted territory of human genetic engineering. What I do expect from such an exploration is at least some assistance in disentangling the benefits we might aim for from the pitfalls we might legitimately fear. I hope that this historical journey will suggest to the reader—as it has to me—how one might think about the human genetic future, and how one might thread a path into it of good sense, reason, and decency."

Kevles' book would be worth reading for its analysis of the benefits and pitfalls alone. There are now many studies in the history of eugenics, but his is the first to ask what we might learn from that history as we confront the social consequences of individual reproductive decisions. Kevles provides no definite answers. He implies, however, that in balancing "social obligations as against individual rights, and reproductive freedom and privacy as against the requirements of public health and welfare," we should favor private and individual interests over those that are public and social. One might question whether an individualistic ethic—which we have abandoned in most spheres of life—will in fact prove defensible in this one. How to choose between or compromise these "ancient antinomies" is the important question. It is one of the virtues of the book to have made it explicit.

Kevles is also the first historian to explore seriously the relation between eugenics and human genetics (which should make the book particularly interesting to scientists). He is also the first to examine closely the relation between eugenics and the women's movement. Kevles notes that in both Britain and America many middle-class women were involved with eugenics. Its appeal reflected eugenics' concerns with traditional women's issues; it also provided a respectable entrée to the worlds of science and public affairs, from which women were otherwise excluded.

Kevles writes engagingly. He has organized the book around a series of portraits of leading eugenicists and their critics. Since the lives of the former were often unconventional and their personalities were eccentric, the biographical approach makes for lively reading. It is nonetheless an approach that has certain drawbacks. The psychic needs of a Galton, a Pearson or a Davenport may explain their turn to eugenics. These needs cannot, however, explain its acceptance (and at times rejection) by many quite ordinary people. The explanation of these diverse reactions lies in political and social history, not in individual psychology. Kevles treats these dimensions, but he gives them rather short shrift. For example, the forces that produced the Immigration Restriction Act—perhaps the greatest triumph of the early eugenics movement—are summarized in about three pages.

It is a style that also lends itself to treatment in terms of heroes and villains. The book is fun to read in part because there are obvious good guys and bad (none of the central characters is a woman). Kevles' heart is clearly with the critics—Jennings, Haldane, Dobzhansky, Penrose—who are consistently portrayed not only as better-intentioned than their opponents but also as better scientists. For example, Jennings appears to defeat the mathematical geneticist R. A. Fisher in their dispute over the efficacy of policies aimed at reducing the incidence of mental defect. At issue was the effectiveness of segregation and sterilization of the "feebleminded" given recent evidence that the trait was transmitted recessively. Thus, the argument ran, there would be many more heterozygous carriers of the defective gene(s) than individuals actually affected. The "real menace" of the feebleminded, in the geneticist East's phrase, was constituted by this huge, invisible heterozygotic reserve. Measures preventing the feebleminded themselves from breeding, Jennings argued, could not reach this group. Hence the question: How effective were such policies?

Fisher argued that they were very effective, Jennings that they were not effective. (Neither disputed the value of feeblemindedness as a category.) Their divergent conclusions in part reflected different scientific assumptions. Fisher thought that the pattern of mental defect suggested polygenic inheritance (although why he thought this would speed selection is not clear) and that those categorized as feebleminded would tend to mate "assortatively," that is, with each other. Jennings, on the other hand, assumed that feeblemindedness was a single-gene defect and that those affected would mate at random. On these issues Fisher's assumptions were the more plausible, but the implication of Kevles' account is that he was scientifically bested by his (politically progressive and personally appealing) opponent.

To take another example: Kevles sides with Dobzhansky in his dispute with Muller over the value of genetic diversity. Muller believed selection is a purifying force, acting to produce a "best type." From this assumption it seemed to follow that individuals would be homozygous at all but a handful of loci, and that populations would be essentially monomorphic. Dobzhansky, on the other hand, believed genetic diversity is adaptive and hence is actively maintained by some form of selection. He therefore assumed individuals would be heterozygous at many loci and, within a population, a large number of alternative alleles would segregate at each locus.

The question of the extent of genetic variability has long been settled in Dobzhansky's favor. There is indeed a lot more heterozygosity and polymorphism among humans (and other species) than Muller predicted. The meaning of all this genetic variation, however, is far from settled. How much is a response to forces of selection and how much simply the accumulation of neutral mutations is not yet known. Kevles cites "the resistance to malaria conferred by the sickle-cell gene in its heterozygous state" as a substantiation of Dobzhansky's position; this tells us little. The question was, and remains, whether heterozygote superiority (or other forms of "balancing" selection) are common or exceptional. Dobzhansky may have been a more attractive person, with more appealing politics, than Muller, but he has not yet been proved right (or more right) on this issue. In a well-constructed world good people, politics and science would invariably be linked. In the real one, alas, it is not always thus.

These criticisms should not obscure Kevles' real accomplishments. The history of eugenics has long been the preserve of specialists. Yet an understanding of that history is central to informed debate on issues affecting the public in general and scientists in particular. Kevles' lively and informative book makes that debate possible. It deserves a wide audience.

Established 1845 SCIENTIFIC AMERICAN January 1986 Volume 254 Number 1

Space Science, Space Technology and the Space Station

The space-station program will seriously diminish the opportunities for advancing space science and technology if it proceeds as planned. Most national goals in space are better realized by robot spacecraft

by James A. Van Allen

There is something about the topic of outer space that induces hyperbolic expectations. With no difficulty at all I can think of a billion-dollar space mission before breakfast any day of the week and a multibillion-dollar mission on Sunday. Ordinarily I do not inflict such visions on my fellow citizens, but I note that proposals of comparable or lesser merit and of much greater cost receive public attention, and some are influential in high circles of government. I submit that the proposed permanently manned space station is in this category.

A National Commission on Space, mandated by Congress and appointed by the president, has bravely undertaken to foresee the course of the U.S. space effort over the next 50 years. The commission's final report, to be released in March, will take it for granted that the space station will be operating in orbit within a decade, as President Reagan announced in his 1984 State of the Union message. According to the timetable of the National Aeronautics and Space Administration, the initial operations capability of the space station is to be achieved by 1993. Official estimates set its development costs at $8 billion in constant 1984 dollars, but the true costs will probably be many times that preliminary figure. There have been no announcements about the costs of operating and maintaining the station in orbit or about the costs of the equipment needed to make the station a useful facility for scientific and technical purpose.

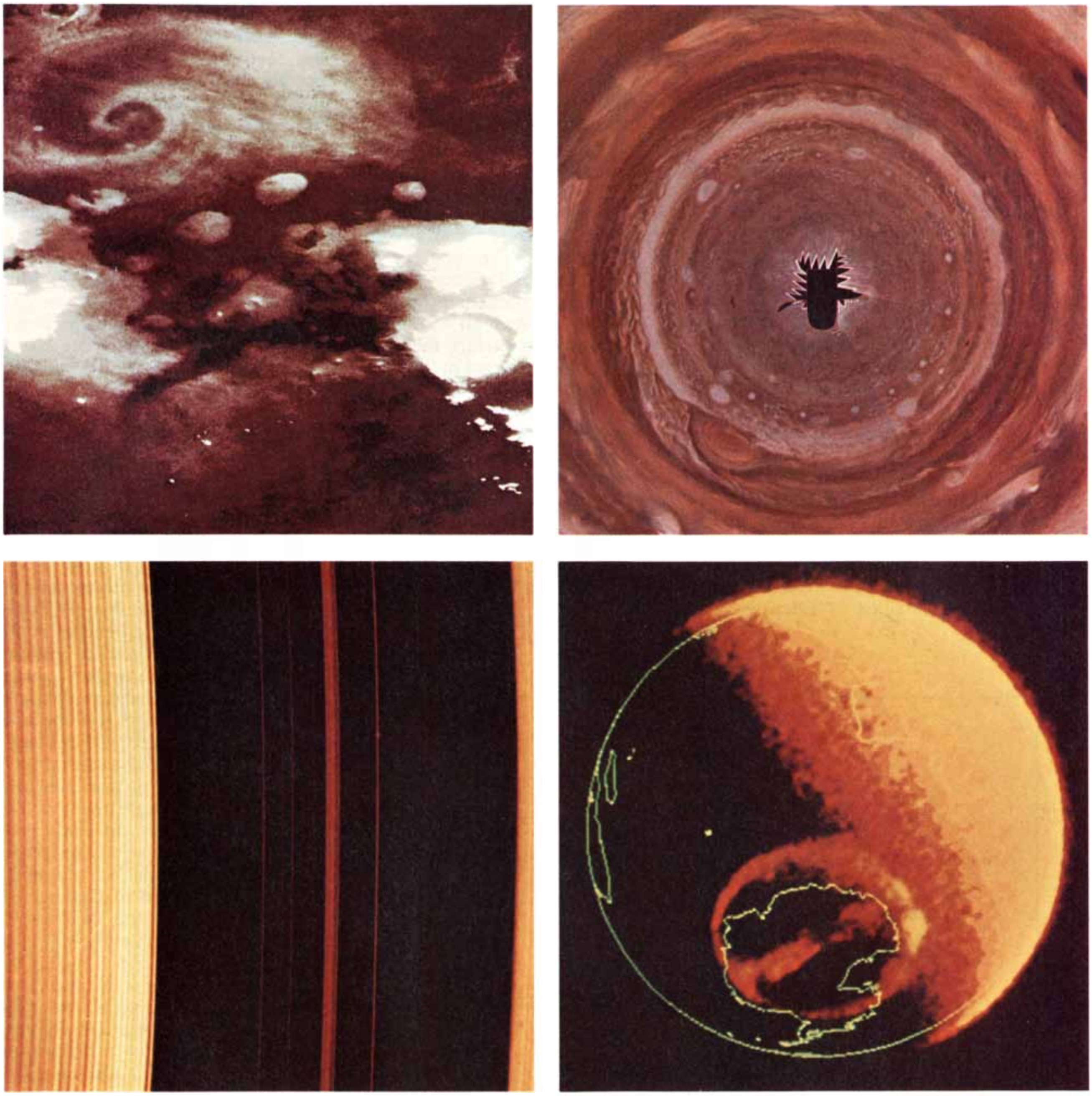
With the space station in place, the National Commission on Space envisions a number of options for building what it calls the "infrastructure required for the initial exploration and occupation of the inner solar system." The options include the construction of three more space stations, one in high earth orbit, one in lunar orbit and one in orbit around the planet Mars; the deployment of additional space stations in orbits around the earth-moon system or the earth-Mars system, to serve as long-range "buses" for earth-moon or interplanetary transport, and the construction of several vehicles to shuttle astronauts among the various space stations, moon and planets. The concept of a joint U.S.-U.S.S.R. manned mission to land on Mars has been endorsed by many officials both in and out of NASA. The presence of people living and working in space, with necessarily elaborate provisions made for their health and well-being, is common to all the major options being considered for recommendation to the president.

The acceptance of such grandiose proposals by otherwise rational individuals stems from the mystique of space flight, as nurtured over many centuries by early writers of science fiction and their present-day counterparts. Indeed, to the ordinary person space flight is synonymous with the flight of human beings. The simple taste for adventure and fantasy expressed in that sentiment has been elevated in some quarters to the quasi-religious belief that space is a natural habitat of human beings. According to this belief, the real goal of the space program is to establish "man's permanent presence in space," a slogan that does not respond to the simple question: "For what purpose?" Coupled with the public acclaim for the manned Apollo missions to the moon, this kind of advocacy has committed NASA to an overriding emphasis on the development of manned space flight: roughly two-thirds of the agency's funding is allocated to that objective.

The directions embodied in NASA's budgetary policy ignore the basic history of space flight: in the more than 28 years since the launching of *Sputnik I* the overwhelming majority of scientific and utilitarian achievements in space have come from unmanned, automated and commandable spacecraft. For example, the program of unmanned planetary exploration has been brilliantly successful and has made immense contributions to human knowledge. Robot satellites in earth orbit have revolutionized global communications and navigation, and they have yielded fundamental advances in our understanding of the atmosphere, the oceans, the weather and the distribution of natural resources. Finally, they have enhanced national security by making it possible to monitor military activities abroad.

Let me make it clear that I have no hesitation in granting the technical feasibility of a space station or of a system of space stations, given adequate resources for the purpose. Furthermore, I have no doubt that significant uses of space stations can be identified. The issue is not a technical one, however; the space-station program will consume a major fraction of the resources available for our national space activities. The plans for a space station therefore raise basic questions about the economic, political and cultural objectives of the U.S. in space.

Space science and technology are now mature enough to allow a competent, well-defined and realistic selection of goals and the assignment of well-reasoned priorities among them. Then and only then is it sensible to consider the best technical means for achieving these goals, the appropriate time scales and the necessary resources. As I see it, the primary goals of the space program include strictly utilitarian objectives, whose costs and ben-



DETAILED IMAGES of the distant planets betoken the accomplishments of the unmanned, scientific space program. The image at the upper left shows a storm on Mars; it was transmitted by the *Viking Orbiter 1* spacecraft and processed in false colors to highlight the details of the storm. At the upper right is an image of Jupiter, which has been constructed by a computer from data transmitted by the *Voyager 1* spacecraft to show the planet as it would appear from directly above its south pole; no spacecraft has ever made a real photograph of Jupiter from that vantage. There is no photographic data from the black, irregular region at the pole. The bright red band in the false-color image at the lower left is a thin ring in the Encke division in the outer main ring (ring *A*) of Saturn; the data for the computer-generated image were gathered by a photopolarimeter aboard the *Voyager 2* spacecraft, which recorded the occultation of starlight passing through the rings. In the image at the lower right the theta-aurora of the earth is shown as a yellow ring and crossbar, on which the outline of Antarctica has been superposed. The image was transmitted by the *Dynamics Explorer 1* satellite. The first three images were prepared by the Jet Propulsion Laboratory and are shown courtesy of the IBM Gallery of Science and Art in New York City. The image of aurora over Antarctica is shown courtesy of Louis A. Frank of the University of Iowa.

efits are relatively easy to determine, and cultural objectives, whose costs and benefits are harder to calculate.

One category of utilitarian objectives is the set of military applications that are deemed to be in the national interest. A second category includes civil applications of space technology that either are in the national interest as public services or are capable of paying for themselves in the marketplace. As for the cultural objectives, it seems reasonable to grant that there is value to the shared, vicarious sense of adventure that was generated by the Apollo program and similar efforts. Such a social sense can therefore probably be counted as a cultural objective. By the same token, one must grant that the conduct of scientific observations and experiments in space, without any guarantee that they will pay off in useful technology, is a legitimate cultural objective. Of course, purely scientific activity almost always yields practical applications, some of consummate importance, and so there is no implied assumption in classifying science as a cultural objective that it will not turn out to have quantifiable, utilitarian benefits as well.

Because the space program was primarily military in its inception, it seems appropriate to begin with this set of utilitarian objectives. The military applications of the space program can be further classified as defensive and offensive. Up to now, I am happy to say, the defensive applications have dominated, thanks in no small part to a succession of treaties and United Nations resolutions on the peaceful uses of outer space. Such defensive functions include worldwide reconnaissance and surveillance, oceanography, geodesy, communications, meteorology and navigation.

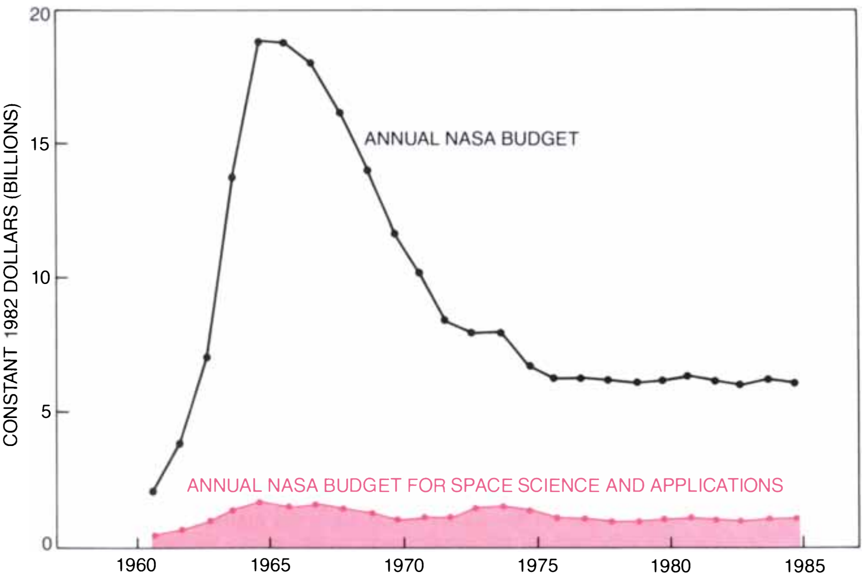
There is some persuasion to the argument that high-quality, reciprocal reconnaissance by all potential adversaries diminishes world tension: by providing advance notice of military deployments it reduces the element of surprise and buys time for intensified negotiation. The logical extension of this line of thought is that the U.S., the U.S.S.R. and the People's Republic of China should operate a joint reconnaissance program so that all observations and their interpretation would be shared. Such an arrangement would make the entire matter an academic exercise and give to warfare the aura of futility it richly deserves in the contemporary world. Military activities in space have been carried out almost exclusively by unmanned satellites, and there is every reason to think this will continue to be the case.

Yet advocates of manned missions in space argue that only a manned spacecraft makes it possible to repair robot satellites in orbit or to replace all or some of their parts. The argument attacks a straw man. Many unmanned spacecraft now deployed have expected operating lifetimes of 10 years or more, and they incorporate automatic or commandable redundancy to help ensure their longevity. Moreover, the technical obsolescence of most flight equipment over a period of 10 years makes repair or refurbishment in orbit a capability that has little or no economic justification.

Offensive military deployments in space, such as antisatellite weapons, pose deep questions of national purpose that lie mostly outside the scope of this article. Whether or not one shares my belief that the calls for a military buildup in space are to be deplored, one may see in such a buildup a potential for growth that would remove the basis for my concerns about science and the space station. Thus, goes the argument, however regrettable you may find the military buildup, the tradeoff is that under the comparative largesse of the military umbrella you can have a manned space station and a vigorous, unmanned scientific space program as well.

At first glance the point has considerable force. For those of us who remember the national trauma following the successful launching of the first Soviet satellite in October, 1957, there is little doubt that the military uses of space have provided the most powerful incentives for our subsequent effort. Indeed, President Lyndon Johnson once said that the benefits of the U.S. system of satellites for military reconnaissance had more than paid for the entire national program in space. Nevertheless, citing such statements in the present context ignores the changes in military-funding policies that have been mandated by Congress since the early 1970's. The Department of Defense must now adhere to a relatively narrow definition of what constitutes its mission: much of the loss in support from the Department of Defense for the basic sciences in the early 1970's can be traced to this evolution of policy. In justifying its expenditures the Department of Defense is unlikely to squander its credibility before Congress by supporting huge undertakings that are not manifestly in defense interests. I suspect that neither the space station nor many of the scientific interests with which it competes will receive any significant subsidies from the Defense Department budget.

One is left, therefore, to consider the objectives in space that are not overtly military in nature. The history of the civil space program in the U.S. shows that following the peak in funding generated by the Apollo program in the mid-1960's, appropriations fell by



ANNUAL BUDGET for the National Aeronautics and Space Administration in constant 1982 dollars is plotted on the graph in black; the budget for space science and applications is superposed on the graph in color. NASA's greatest spending took place in the mid-1960's, during the development of the manned Apollo missions to the moon. The growth of space science and applications in that period did not keep pace with the growth in the manned space program. Since then the overall NASA budget has fallen to about a third of its peak value, and about 15 percent of the total has been allocated to science and applications.

a factor of three in constant dollars. Since that drop more than a decade ago the funding has remained essentially constant. One may wish that it were otherwise, as I do, but the present level of Federal support has been established by our complex social and political processes, and it is difficult, if not impossible, to responsibly foresee any sizable increase in it in real dollars in the next decade. Conversely, it is reasonable to expect that the funding level in constant dollars will not shrink significantly in the near term.

Thus it appears that the U.S. has achieved an approximate equilibrium between advocates and skeptics as to the proper overall level of our national civil space effort. I shall therefore adopt the assumption of an essentially constant level of such funding for the next decade as basic to my discussion. What this means is that establishing the national priorities in space in the civil sector is a zero-sum game: any increase in one element of the NASA budget must inevitably result in an equal decrease somewhere else.

A second major category of national objectives in space is the development of space technology, including the space station and the other "infrastructure" referred to by the National Commission on Space. Advocates of the manned space station often act pained and perplexed when budgetary constraints are invoked. Do we—that is, I and those of my colleagues who are members of the "loyal opposition"—not realize that once the space station is in place the costs and effort required for commercial and scientific objectives will be reduced dramatically? Are we not aware of the so-called coattail effect, whereby the manned space program allegedly builds up enough momentum in the national space program to carry along all the other projects? Have we become so enamored of the capabilities of commandable spacecraft that we have ignored the fact that a man in space can carry out these tasks more efficiently and with less effort?

To answer these questions the history of the space program, and particularly that of the Space Transportation System, would seem to be a more reliable guide than the promises and forecasts made by interested parties. The present Space Transportation System includes a fleet of four manned, orbiting space shuttles, each of which is in essence a high-velocity aircraft and spacecraft that is launched by rockets, flies in low-altitude orbit about the earth, reenters the earth's atmosphere on command and lands on a very long airstrip. The development and initial operation of the Space Transportation System has cost American taxpayers about $30 billion to date, with much smaller but still substantial contributions from European nations through the European Space Agency. The four shuttles in the current U.S. fleet were and still are conceived as service vehicles for the space station, and so it is appropriate to consider the shuttle as a key element in the U.S. manned space program for the next 20 years.

The space shuttle represents the natural aspiration of aeronautical engineers to push the state of their art to its limits. Although I heartily applaud its impressive technical successes, I find the economic justification for building it to be quite unpersuasive, and I have so testified to the Office of Technology Assessment and to a succession of congressional committees beginning in 1971. Those of us who were on the losing side of the debate in the early 1970's as to the wisdom of developing the shuttle have no difficulty remembering the claims then being made. In brief, our opponents argued that the

Fully commercial applications
Worldwide network of satellite relays in synchronous orbits for transmission of television broadcasts, telephone and telegraphic messages and data. Operated by COMSAT, INTELSAT and private corporations
Military applications
Worldwide network of telecommunication satellites in synchronous and intermediate-altitude orbits
Worldwide network of Transit and Global Positioning System satellites for navigational purposes. Current accuracy to within 30 meters at any point on or in the vicinity of the earth. Potential accuracy to within one centimeter. Lower accuracy system also available for civil purposes
Networks of reconnaissance and surveillance satellites
Networks of meteorological satellites
Partly commercial and partly Governmental civil applications
Meteorological satellites for surveying and forecasting current global weather
Landsat and other satellites for survey of mineral resources, vegetation, icebergs, snow cover, water resources, water pollution, health of crops and geological features and for mapping
Scientific investigations and achievements
Electromagnetic and corpuscular classes of radiation from the sun and their effects on the earth
Dynamics of the solar atmosphere
In situ measurements of charged-particle populations and magnetic and electric fields in the ionospheres, the radiation belts and the magnetospheres of the earth, Mercury, Venus, Mars, Jupiter and Saturn
Plasma physical effects associated with natural and artificial comets
Geological surveys of the moon, the earth, Mercury, Venus, Mars, the satellites of Mars, Jupiter and Saturn
Closeup study of the rings of Jupiter and Saturn
Precise characterization of external magnetic fields of the moon, the earth, Mercury, Venus, Mars, Jupiter and Saturn
Detailed study of the structure, composition and dynamics of the earth's atmosphere and exploratory study of the atmospheres of Venus, Mars, Jupiter, Io, Saturn and Titan
Precise characterization of the external gravitational fields of the moon and the earth
Comprehensive observation of the solar wind and of shock waves, energetic solar particles and galactic cosmic rays in interplanetary space out to a distance of 3.4 billion miles from the sun and continuing outwards
Comprehensive surveys of stellar and planetary sources of gamma rays, X rays and ultraviolet, infrared and radio-frequency radiation and the detailed spectral study of selected sources
Marked advances in understanding the origin and evolution of the solar system and of stars and galaxies
Significant contributions to fundamental plasma physics and its role in planetary and astrophysical systems
Study of ocean currents and the global dynamics of the oceans
Negative evidence on the past or present existence of living organisms on the surface of Mars

MAJOR ACCOMPLISHMENTS of the unmanned space program are summarized in the table. They include commercial and military applications of space technology, civil applications that are partly public and partly commercial and many scientific accomplishments.

shuttle would supplant all expendable launch vehicles, such as the Scout, Delta, Atlas and Titan rockets, and that by the early 1980's there would be 50 shuttle flights per year. Each flight would deliver 50,000 pounds into low earth orbit at a cost of $100 per pound. Of the 50 annual flights at least four would carry spacecraft for the exploration of other planets.

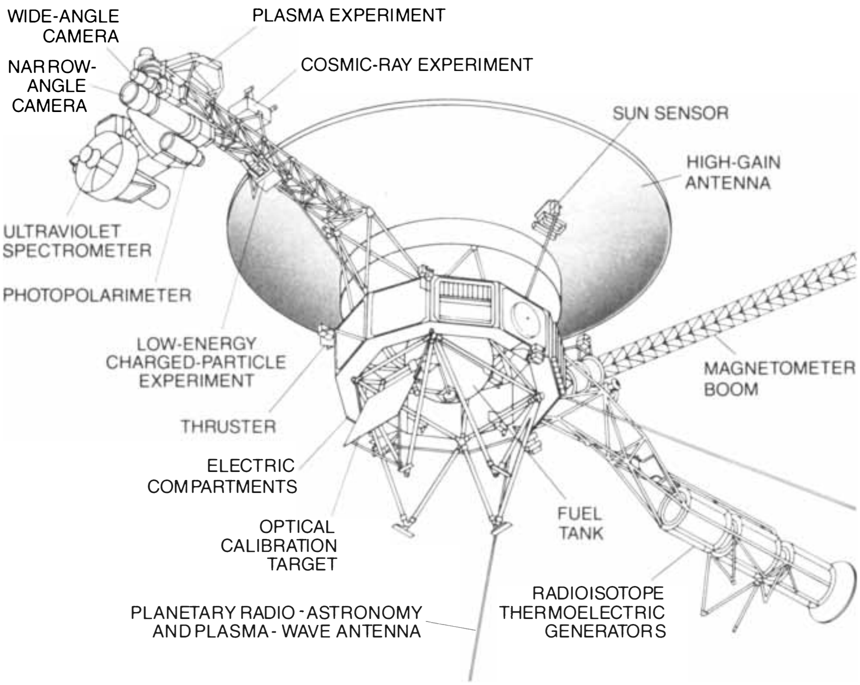
There is a striking disparity between those claims and the present situation. In 1985 only 10 shuttle flights were carried out at a true launching cost of at least $5,000 per pound, or about $2,000 per pound in 1971 dollars, a figure 20 times greater than the original estimate. No planetary spacecraft has been launched in the four years of shuttle operations.

The source of the disparity between promise and realization can be traced to NASA's gross underestimate of developmental costs and its gross overestimate of the space traffic that could reasonably be expected aboard the shuttle. As a result NASA made a wildly overoptimistic estimate of the cost-effectiveness of the shuttle compared with that of the existing expendable launch vehicles or their evolutionary descendants. I see no reason to be any more confident about NASA's economic forecasts for the space station.

There is another reason to doubt NASA's assurances that the space station will make it easier to carry out other national objectives in space. In the summer of 1981, faced with serious delays and major cost overruns on the shuttle, NASA decided that development of the shuttle must proceed, come what might to other ongoing projects. The result was a "slaughter of the innocent": massive cuts, postponements and cancellations of dozens of programs, many of which were already in advanced stages.

For example, the shuttle forced the cancellation of the U.S. member of a pair of complementary spacecraft for the International Solar Polar Mission. The surviving member of the pair, now known as *Ulysses,* was developed by the European Space Agency with the participation of some U.S. scientists and will be launched in May after a delay of approximately two years. Well-developed plans for a U.S. mission to encounter Comet Halley and subsequently to rendezvous with Comet Tempel II were also abandoned because of the shuttle. The major mission to the planet Jupiter known as *Galileo* was canceled for a time because of shuttle funding allocations, and although the mission was later reinstated, the shuttle is largely responsible for its three-year delay.

NASA's single-minded devotion to the space shuttle went unchecked for the first eight months of the Reagan Administration, and when the president finally appointed his own NASA administrator, the cuts were so deep that many of them had to be rescinded. Among other projects NASA had threatened were the worldwide Deep Space Network for tracking and receiving data from planetary missions; the ongoing reception of data from the planetary probes *Pioneer 10* and *Pioneer 11;* the reception of data from the *Pioneer Venus 1;* NASA's Infrared Observatory at Mauna Kea in Hawaii; the reception of data from the deep-space missions of *Voyager 1* and *Voyager 2;* the reception of data from the earth-orbiting satellites *IMP-7* and *IMP-8,* and the plans for further missions to Venus and Mars. In addition the shuttle caused a slowdown in the development of a gamma-ray observatory, substantial reductions in the funding of supportive space science and technology in the universities, the elimination of the office for solar-terrestrial physics programs at NASA's headquarters, the indefinite postponement of new solar-terrestrial and atmospheric research satellites in earth orbit and the indefinite postponement of the development of advanced communications technology. Finally, the provisions for developing significant scientific payloads to be flown on the shuttle are meager.



VOYAGER 2 **spacecraft is shown schematically. It has already probed Jupiter, Saturn and their systems of satellites and rings; this month it will be the first spacecraft to encounter the planet Uranus. The author served as chairman of the committee that developed the mission.**

Two more arguments are sometimes advanced by the proponents of the manned space program and these must be addressed. One argument is a peculiar reading of the history of the space program that I referred to above as the coattail effect. According to this view, the entire program of space science for the past three decades would have either been nonexistent or run on a very small scale had it not been for the manned program. This assertion is impossible to either prove or disprove; one cannot rerun history with different boundary conditions. Nevertheless, I can offer several reasons for doubting the assertion.

In 1946 the U.S. began a vigorous and successful program of high-altitude rocket flights carrying scientific instruments. The work was given much impetus during the International Geophysical Year in 1957–58, and it served as the technical and scientific basis for all the subsequent advances in the scientific and utilitarian use of earth satellites and interplanetary and planetary spacecraft. The major period of growth in these fields coincided with the Apollo program, but in my view neither set of activities depended to any important extent on the other.

Both the scientific activities and the Apollo program took place in an optimistic, expansionist epoch in national and international history, as did all kinds of other scientific activities unrelated to space. Many other major research agencies of the Federal Government have grown to a sustained level of support comparable to that of the research component of NASA without the benefit of huge, public spectaculars; examples include the National Institutes of Health, the National Science Foundation, the U.S. Geological Survey, the Department of Energy and the National Oceanic and Atmospheric Administration.

The second argument often put forward in favor of a manned space program is that a person in a spacecraft is superior to any conceivable machine because of judgment, resourcefulness, flexibility and the like. If one considers the complexity and sophistication of modern space equipment and the ready control of such equipment by command from earthbound stations, such an argument has very limited validity.

My own feelings about this issue are aptly expressed by a story from the early development of large balloons and manned balloon flight. At that time, about 30 years ago, there were advocates of the idea that a large network of manned balloons should be maintained and continually replenished for the purpose of observing both natural and artificial activities on the earth. The classic comment on ideas of this nature was made by Edward P. Ney of the University of Minnesota, who was one of the pioneers in the use of balloons for scientific purposes. Ney had given a public lecture on some of his work in the late 1950's. In the ensuing discussion period a member of the audience stood up to ask a question. "Professor Ney, please tell me: Is there anything a man can do in a balloon gondola that an instrument cannot?" Ney's answer, after only a moment's hesitation, was, "Yes, there is. But why would anyone wish to do it at such a high altitude?"

INTERNATIONAL SOLAR POLAR MISSION (U.S. SATELLITE OF PROPOSED PAIR)	CANCELED
U.S. MISSION TO COMET HALLEY	CANCELED
GALILEO PROBE TO JUPITER	CANCELED (LATER RESCINDED)
DEEP SPACE NETWORK FOR TRACKING PLANETARY MISSIONS	THREATENED CLOSING
DATA RECEPTION FROM *PIONEER 10* AND *11*	TERMINATED (LATER PARTLY RESCINDED)
DATA RECEPTION FROM *PIONEER VENUS 1*	TERMINATED (LATER PARTLY RESCINDED)
INFRARED OBSERVATORY AT MAUNA KEA, HAWAII	CLOSED (LATER RESCINDED)
DATA RECEPTION FROM *VOYAGER 1* AND *2*	CUT BACK (LATER PARTLY RESCINDED)
DATA RECEPTION FROM *IMP-7* AND *IMP-8*	TERMINATED (LATER PARTLY RESCINDED)
LANDSAT PROGRAM	CUT BACK
GAMMA-RAY OBSERVATORY	DELAYED SEVERAL YEARS
PLANNED MISSIONS TO VENUS AND MARS	CANCELED (REVIVED IN REDUCED FORM AFTER DELAYS OF SEVERAL YEARS)
SUPPORTING UNIVERSITY RESEARCH	CUT BACK
NASA OFFICE FOR SOLAR-TERRESTRIAL PHYSICS PROGRAMS	CLOSED
PLANNED SOLAR-TERRESTRIAL AND ATMOSPHERIC RESEARCH SATELLITES	INDEFINITELY POSTPONED
SCIENTIFIC PAYLOADS ABOARD SPACE SHUTTLE	INADEQUATELY PROVIDED FOR
ADVANCED COMMUNICATION TECHNOLOGY	INDEFINITELY POSTPONED

"SLAUGHTER OF THE INNOCENT" was the result of the decision made by NASA in 1981 to proceed with the development of the space shuttle over all other projects. The table summarizes the effects of the decision. Some of the program cuts have since been rescinded, but the effect has been a severe chill on scientific and other civilian activities in space.

The burden of experience is that, apart from serving the spirit of adventure, there is little reason for sending people into space. On the contrary, there are strong reasons for keeping operating personnel on the earth. The life-support systems and the overriding concern for the safety of personnel in any manned space mission are extremely costly and restrictive. Moreover, most space missions of scientific or utilitarian importance require high earth orbits, lunar orbits, interplanetary orbits or planetary orbits that involve months or years of in-flight operation. Such missions will be inaccessible to manned spacecraft for many years to come.

Some experiments one would like to carry out in space require highly stable platforms and the accurate aiming of scientific instruments, and so they must be free of vibrations and accelerations. An astronaut's sneeze could wreck a sensitive experiment in a microgravitational field; clouds of gas or droplets from thrusters of the spacecraft or from dumps of water or urine ruin the local vacuum and optical observing conditions, and complex magnetic and electric fields associated with manned spacecraft preclude certain kinds of radio observations.

The simplest repair and refurbishment of equipment in space requires heroic measures, even if the equipment is accessible. The high cost of such "space rescues" casts grave doubt on their economic viability. Moreover, it is much harder and more expensive to design and build space equipment in such a way that it can be repaired and refurbished in space than it is to build equipment that need not meet such specifications.

Inside a spacecraft the working conditions for people are extremely restrictive and the resources available for experimental work are limited. Simple functions that can be carried out by a skilled technician are all that can be expected, whereas all the real sophistication and resourcefulness of an in-flight experiment must be exercised by radio command or built into the equipment before the flight, just as they are in a robot spacecraft. Nearly all investigations can be monitored and controlled much more effectively by people on the ground, who are working under far more comfortable and efficient conditions and with easy access to all the resources available there. Finally, the apparatus in an unmanned spacecraft does not get tired, it is free of human contamination and it is not subject to the kind of human error that can result from onboard manipulation.

All the foregoing leads one to conclude that the development of advanced technology for launching and maintaining people in space is a goal largely independent of other legitimate national objectives in outer space. There is a large and diverse body of other civil applications of space technology that deserve consideration on their own merits. Foremost among such applications is worldwide telecommunications by satellite relays. More than half of all transoceanic communications go by way of satellite relays, and this capability is being continually expanded. Furthermore, domestic communications in far-flung countries such as Canada and Indonesia have been revolutionized by satellite methods.

Some 20 years ago I was among those who expressed great hope that satellite communications would be employed in worldwide educational efforts, particularly within developing countries. The hope was based on the recognition that substantial benefits to mankind can result even from the spread of simple literacy and a knowledge of basic arithmetic. In 1974 an Advanced Technology Satellite in synchronous orbit was assigned to deliver elementary educational materials in India on an experimental basis. In all technical respects the experiment was an unqualified success, but there are still many thorny cultural, sociological and political issues to resolve.

Telecommunications is the only application of space technology that has achieved economic viability, in the sense that the direct beneficiaries ask for certain services and both voluntarily and consciously pay their full costs. Thus I distinguish between the market support of a commercial service and the taxpayer support of a government service in the public interest. The future growth rate of satellite communications will be determined by market forces, at least in the short term, although eventually there will be technical limits to that growth.

Some planners envision a gradual transfer of most domestic communications within the U.S. to satellite systems. At the same time there are immensely promising developments in the transmission of information by modulated beams of laser light carried by optical fibers. Tens of thousands of miles of optical fibers are already installed between cities in the U.S., and a transatlantic cable of optical fibers is under construction. Optical-fiber carriers may therefore come to dominate high-traffic communications between fixed points in the next 20 years, and so they may limit or slow the growth of corresponding techniques in space. Of course, optical fibers cannot be used for communications to or from mobile stations, such as aircraft in flight and ships at sea. All these matters are under continuous engineering study by many private corporations and government agencies in Europe, Japan, the U.S. and, undoubtedly, the U.S.S.R. A proper role for NASA in this field is to conduct advanced research and contribute to the development of hybrid communications systems.

The other principal civil application of space technology comes under the generic term remote sensing. Remote sensing includes not only ordinary photoreconnaissance, including what have become routine forecasts of the weather on a worldwide scale, but also the imaging of the earth's surface and atmosphere over a broad range of electromagnetic frequencies. There have been exquisite instrumental developments in the field, and it is now possible to choose well-defined frequency bands of radiation in the radio, infrared, visible, ultraviolet and X-ray portions of the electromagnetic spectrum. For example, the two automated satellites *Landsat 4* and *Landsat 5* carry instruments called thematic mappers, which map radiation emissions from the earth's surface in several frequency bands that are important to geologists in their search for worldwide mineral resources.

Such applications of remote sensing yield substantial public and private benefits, but they still have not met the crucial test of full commercial success. The Landsat program, for example, operates under the Department of Commerce as a data-service agency to industry and to other Government agencies, but it is heavily subsidized by the Government. Virtually the entire field of remote sensing, as well as the many other useful applications of space technology such as surveying and aircraft and marine navigation, remains in the realm of Government services. As such, they are all exposed to budget cuts caused by reallocations of funding to the manned space program; indeed, the Landsat program has suffered severely for precisely this reason.

Ironically, far more tenuous proposals are put forward as justifications for building the shuttle and the space station. One example is the processing of materials in space; for example, it has been widely advertised that the microgravitational environment of space can be exploited to grow large crystals of ultrahigh purity or to refine pharmaceuticals on a commercially viable scale. Objective studies of this subject by the National Research Council and other agencies do not support such sanguine expectations. The studies review the relatively meager results in the field to date and endorse the validity of further exploratory investigation. Nevertheless, they conclude that the prospects for viable commercial applications have as yet no convincing foundation commensurate with the costs of space flight.

Another proposal that has been given much public exposure is the solar-power satellite. The satellite is envisioned as a solar-power collector, some 20,000 acres in area, that would be assembled in earth orbit. Microwave beams would transmit the solar energy to receivers at stations on the earth, which would deliver the energy over conventional power lines. I am gratified to learn that the voice of sanity has placed this proposal in limbo. Former senators James G. Abourezk (D—S.D.) and Floyd K. Haskell (D—Colo.) have pointed out that the estimated cost of one such satellite would be equivalent to the cost of providing every U.S. household with a simple solar-energy collector that would meet 65 percent of its energy needs.

These two examples, space manufacturing and the solar-power satellite, are leading elements in forecasts of an explosive rate of growth in space traffic. I am not so foolish as to suggest that such undertakings are totally out of the question at some time in the remote future. Not one of them, though, withstands critical scrutiny in the context of the 20th century, and their ratio of cost to benefit may never be less than unity.

The two major cultural objectives of the U.S. in space demand quite different consideration. The first is the realization of a kind of collective human adventure. Popular interest in real—as opposed to fictional—space activity was highest during the first manned landing on the moon in July, 1969. In subsequent years the role of the space program in creating vicarious adventure has dwindled markedly, having been supplanted to a considerable extent by the romance of motion pictures depicting far more dramatic exploits. The American public has now spent more than $200 million to see *Star Wars* and hundreds of millions of dollars more to see its derivative successors; the total is about the sum needed to carry out a major planetary mission. I draw no moral from these facts, but I do consider them a point of reference concerning the public motivations for manned space flight.

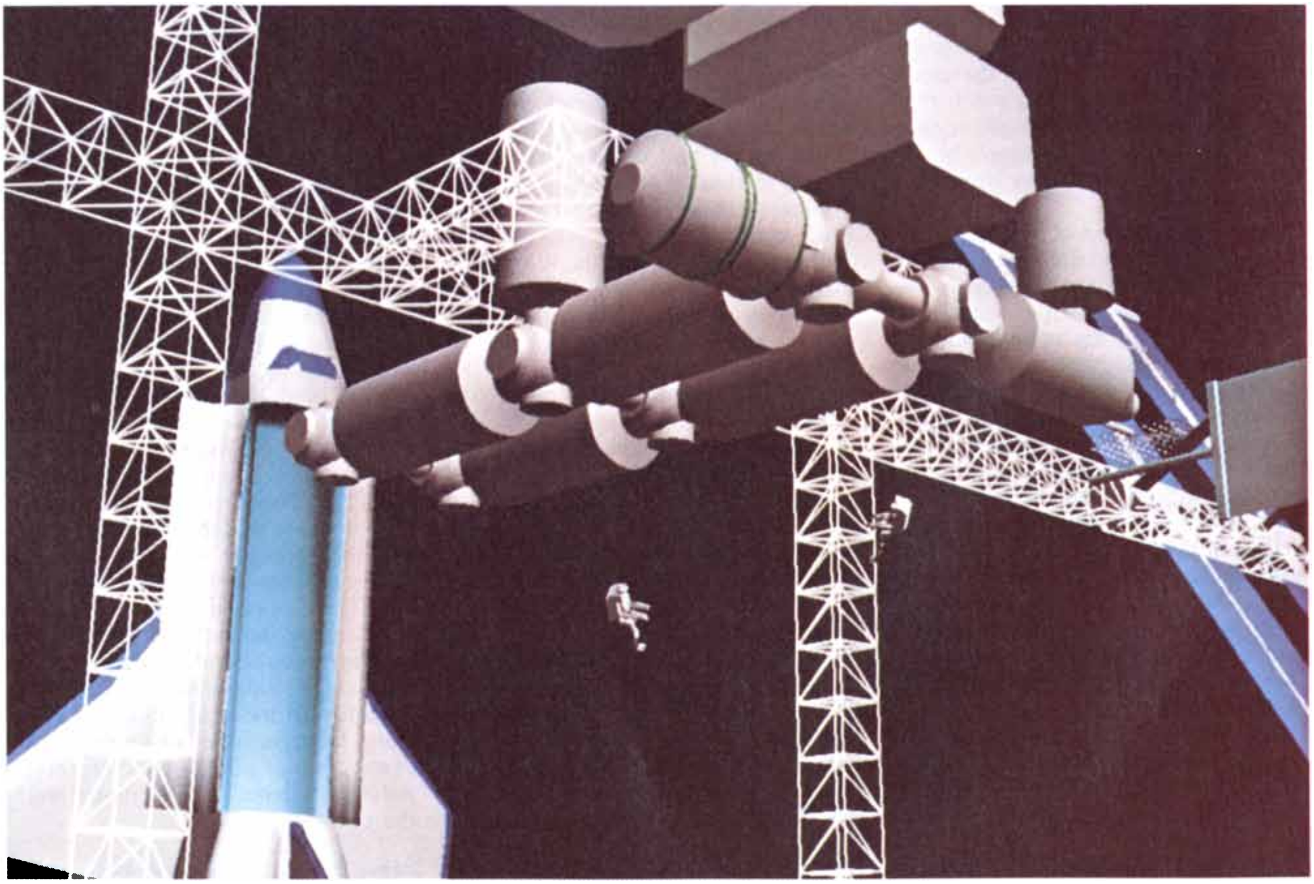
The second cultural objective in space is the conduct of space science. A possible definition of the term is the investigation of natural phenomena that take place above the surface of the earth. By this definition astronomy qualifies as the most ancient of the space sciences. A somewhat different definition, which is the one usually intended by contemporary practitioners, is the investigation of phenomena, both terrestrial and extraterrestrial and both natural and artificial, by means of apparatus carried aloft in rocket-propelled vehicles. Thus space science is not a clearly delineated scientific discipline in the usual sense of the term; instead its common element is a shared set of basic techniques. The substance of space science is best thought of as a sophisticated, and expensive, mixture of the traditional disciplines of astronomy, geology, geophysics and oceanography.

In the decades since the first satellites there have been tremendous advances in observing and understanding the oceans, the atmosphere, the ionosphere and the magnetosphere of the earth, the many types of radiation from the sun and their effects on the earth, and the nature and evolutionary history of the moon and planets. There have been many discoveries of basic importance to stellar astronomy. All the objects of the solar system as well as the interplanetary medium are now accessible to closeup study. Probes have been dispatched to the planets Mercury, Venus, Mars, Jupiter and Saturn, and some of them will also transmit data from Uranus and Neptune. *Pioneer 10,* one of my favorite spacecraft, has been operating in flight for nearly 14 years and is now the remotest manmade object in the universe. It is still functioning well into the outer heliosphere and is farther from the sun than Pluto is. A few months ago the *International Comet Explorer* flew through the coma of Comet Giacobini-Zinner, and this month *Voyager 2* will be the first spacecraft to make an encounter with Uranus.

The scientific community has developed a great variety of superb instruments that can withstand the rigors of launching, and space science is teeming with fresh discoveries and well-formulated plans for the future. Support is now moderately secure for the ongoing missions of the *International Ultraviolet Explorer,* the *Dynamics Explorer,* the *International Sun-Earth Explorer* and the Pioneer and Voyager spacecraft. Yet the number of new opportunities for flight has been reduced markedly in the 1980's by cancellations and prolonged delays.

The major emphasis in recent years in space science has been on billion-dollar missions, such as the Space Telescope, the *Galileo* mission to Jupiter, the Viking landers on Mars and the Voyager probes. This trend also accounts in part for the reduction in scientific payloads; indeed, the Space Telescope and *Galileo* are the only major U.S. scientific spacecraft that have been or will be scheduled for launching in the years from 1983 through 1988. Such missions represent a tendency within space science toward ever greater complexity and sophistication, and they do have high merit. Unfortunately, however, like the large, manned space projects, they tend to squeeze out more flexible and much less expensive undertakings that historically have been highly productive. Smaller projects nurture space science on a broad, national basis and continue to have a potentially important role in our national program, but they are now nearly extinct.

In the meantime the European Space Agency, Japan and the U.S.S.R. are forging ahead with important scientific missions. The progressive loss of U.S. leadership in space science can be attributed, I believe, largely to our excessive emphasis on manned space flight and on vaguely perceived, poorly founded goals of a highly speculative nature. Given the current budgetary climate and a roughly constant level of public support for civil space ventures, the development of a space station, if pursued as now projected, will seriously reduce the opportunities for advances in space science and in important applications of space technology in the coming decade.



SPACE STATION, if it were constructed, could resemble the design shown here, but fundamental decisions about the design are still pending. The latest version of the basic structural design differs from a previous one in having two main "towers" instead of one. In this computer-generated image of the current version only the shuttle-docking area of the space station is shown, and the scale of the structure is indicated by the human figures near one of the towers. At this stage the two designs and, indeed, several others can be quickly interchanged on the color monitor of a computer-aided design system, such as the one responsible for the image shown, made by the McDonnell Douglas Corporation. Nevertheless, the current NASA schedule for development calls for the electronic image to be translated into a real device in earth orbit by the year 1993. The cost of that effort may be as high as $30 billion in constant 1984 dollars.

Growth, Differentiation and the Reversal of Malignancy

Specific proteins regulate the growth of normal white blood cells and their differentiation into nondividing forms. Leukemic cells can also be made to differentiate, suggesting new approaches to cancer treatment

by Leo Sachs

The cells in the body descend from precursors known as stem cells. Stem cells can multiply rapidly; their progeny, after maturing and differentiating into specialized forms, generally stop growing. During normal infancy and adulthood the processes of multiplication and differentiation are in harmony: the growth of stem cells provides new tissue and replaces dying cells, while the cessation of growth after cells reach their final form keeps cell multiplication in check. In cancer the harmony breaks down: there are too many immature, multiplying cells.

An understanding of how the processes of growth and differentiation are regulated in normal cells makes it possible to answer several questions about cancer. Have all the cellular mechanisms controlling growth and differentiation gone awry in malignant cells, or do some of the controls still operate? If some of the mechanisms are intact, can they be reactivated to make cancer cells differentiate and stop growing? I have concluded that malignant cells can retain the genetic bases for differentiation. Appropriately stimulated, they can complete the normal sequence of growth, differentiation and cessation of growth. These findings have opened new possibilities for cancer therapy.

I studied the processes of growth and differentiation in normal and leukemic blood cells. In hematopoiesis, the process by which the body generates new blood cells, a single kind of precursor (the multipotential stem cells in the bone marrow) develops into a number of more specialized precursor cells, among them myeloid stem cells and lymphocyte precursors. These more specialized stem cells in turn mature and differentiate into all the various white blood cells and the red blood cells. As a model system I studied myeloid stem cells, the mature white blood cells to which myeloid stem cells give rise and myeloid leukemic cells. The cells in myeloid leukemia appear to be frozen in their precursor form. Ordinarily they do not differentiate into mature cells, and so they continue to multiply.

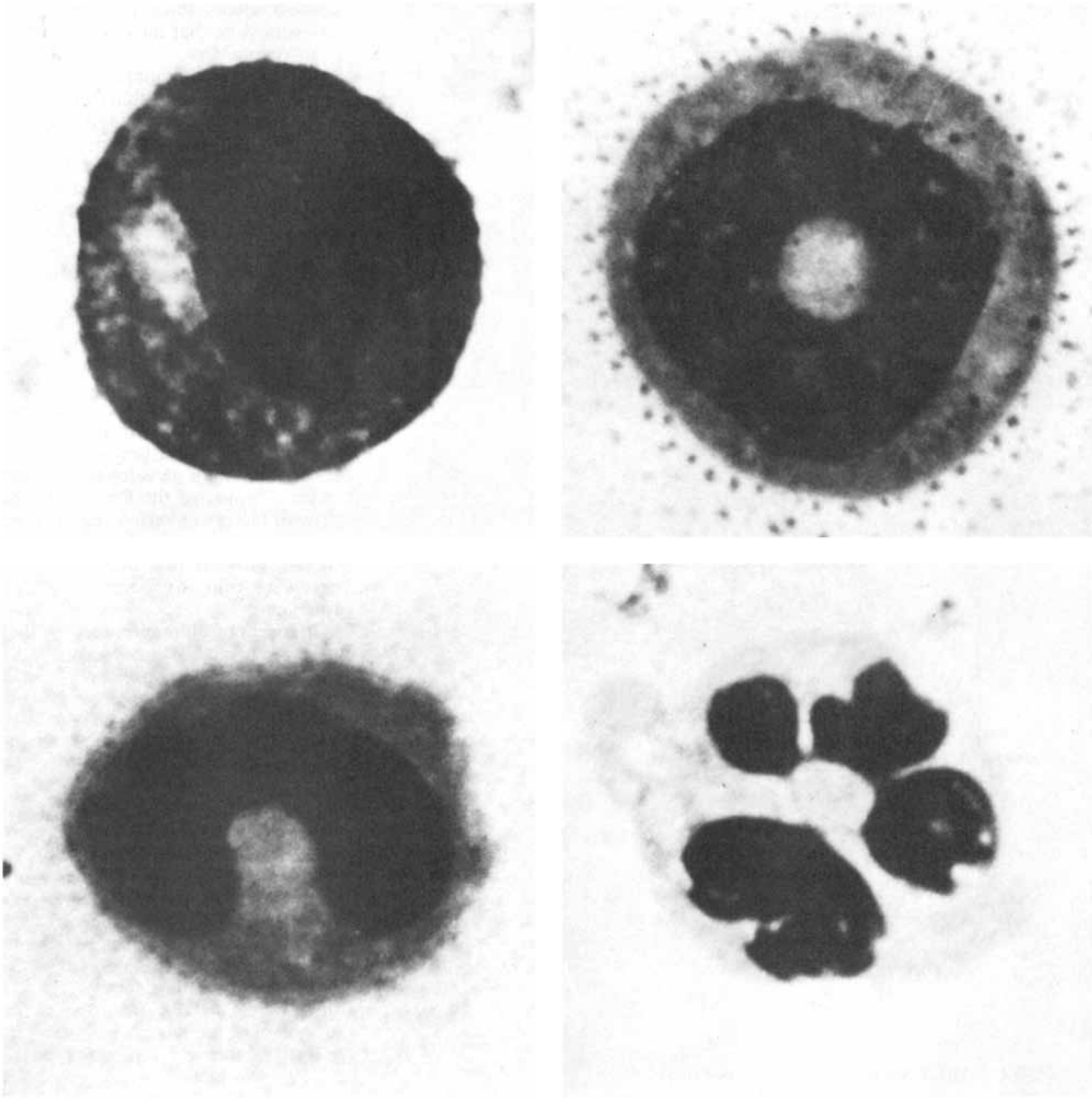
In trying to establish general principles governing cell growth and differentiation and the reversibility of malignancy I had to establish a cell-culture system that made it possible to observe these processes in vitro. In 1963, working with my graduate student Haim Ginsburg at the Weizmann Institute of Science in Israel, I found that when normal blood-cell precursors are placed in a liquid culture medium containing fibroblasts, or connective-tissue cells, the precursors multiply into clones (descendants of a single cell) and ultimately differentiate. In 1965, working with another graduate student, Dov H. Pluznik, I showed that clones can also develop and differentiate in a semisolid medium containing agar. The gel inhibits the mobility of the cells, making it easier to distinguish separate clones. In 1966 Thomas R. Bradley and Donald Metcalf of the University of Melbourne and the Walter and Eliza Hall Institute of Medical Research in Melbourne also cloned precursor cells in agar. The first normal precursors for which the method was used were myeloid stem cells, which gave rise to macrophages and granulocytes.

Macrophages and granulocytes are the body's phagocytic cells: they engulf and dispose of foreign material, including invading microorganisms. Macrophages also display molecules from the engulfed foreign material on their cell membrane; there the foreign molecules are "recognized" by other kinds of white blood cells, which then mount an immune response to the intruder. James E. Till and Ernest A. McCulloch of the Ontario Cancer Treatment and Research Foundation have shown that in the body macrophages and granulocytes develop from a common precursor form, the myeloid stem cells. The cells in our culture therefore followed the same pattern of development as they do in the body.

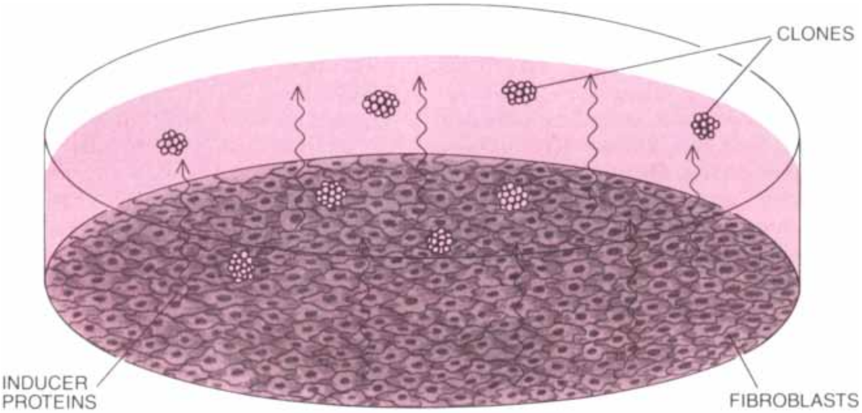
What was the role of the fibroblasts, which had to be present in the culture as a "feeder layer" if the myeloid cells were to grow and differentiate? I had shown in 1965 that the cells of the feeder layer secrete substances inducing growth and differentiation. The following year, working with my students Pluznik and Yasuo Ichikawa, I showed that the inducers are detectable in the culture medium; moreover, the inducers that acted on clones containing different types of cells seemed to differ. The inducers were purified and shown to be proteins, either glycoproteins (proteins with attached sugars) or proteins without detectable sugars. The sugars do not seem to be necessary to the inducers' activity. Using the same cell-culture system, other laboratories have isolated inducer proteins for all the many kinds of blood cells, including the various lymphocytes. A range of cell types can secrete inducer proteins, both in culture and in the body.

The development of a clone of mature macrophages or granulocytes requires the multiplication of myeloid precursors, followed by their differentiation. It seemed plausible that a specialized protein induces each of the processes. Eitan Fibach and I found that mature granulocytes in culture secrete a protein that induces myeloid stem cells to differentiate but does not cause them to multiply into clones. Joseph Lotem and I then tested liquid taken from cultures of various cell types for the ability to induce growth or differentiation in myeloid cells and showed that the processes are indeed triggered by separate factors. In contrast to the differentiation inducers, the growth inducers do not cause precursors to mature but are crucial for cell multiplication and also for cell viability; without them the precursors die. Other workers, including Ichikawa, now at the Chest Disease Research Institute in Japan, Motoo Hozumi of the Saitama Cancer Center Research Institute, Inge L. Olsson of Lund Hospital in Sweden and Robert C. Gallo of the National Cancer Institute, have obtained further evidence for the distinct nature of the factors.

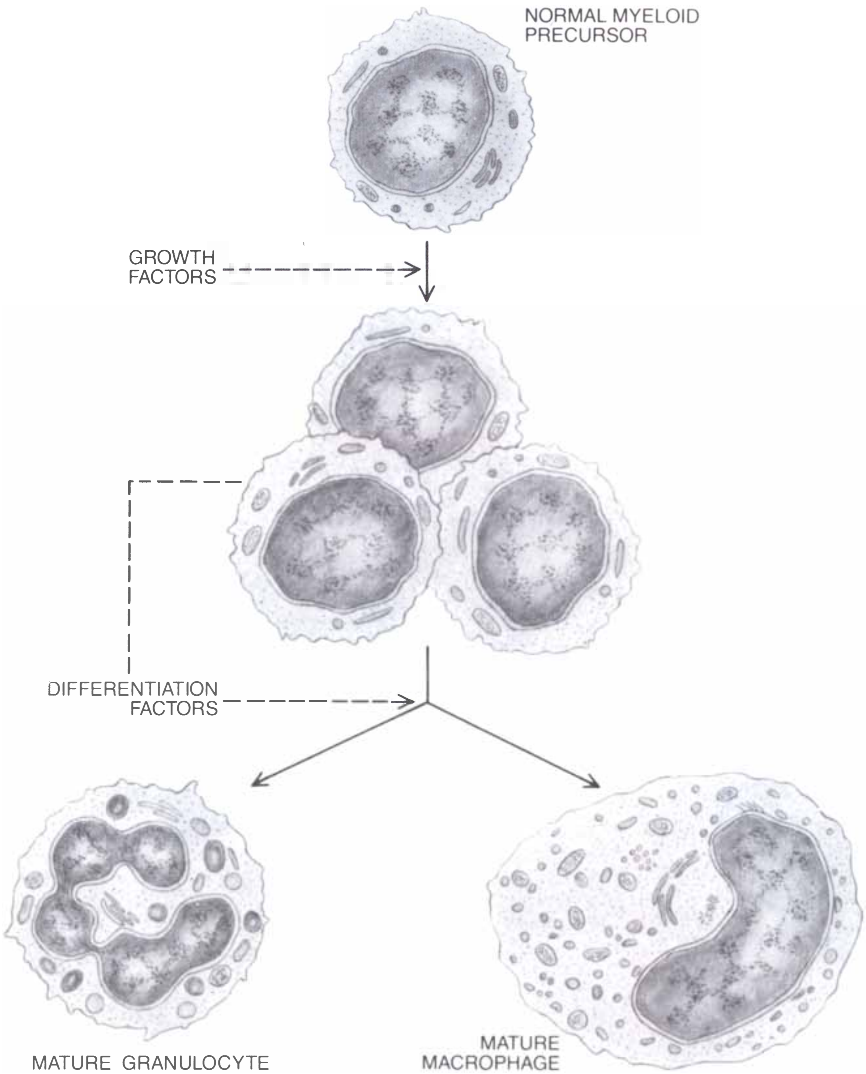
The two kinds of inducer seem to achieve their effects differently. Gary Weisinger and I found that differentiation inducer can bind directly to the DNA of a precursor cell, presumably activating the genes that are expressed as the cell matures. Growth inducer, in contrast, does not bind to DNA. It must stimulate the multiplication of



MYELOID LEUKEMIC CELLS are shown in stages of development from an undifferentiated, rapidly growing form (*top left*) to a mature granulocyte, which does not divide (*bottom right*). Normal myeloid precursor cells, a variety of immature, undifferentiated blood cells, pass through these stages as they mature into granulocytes, one of the cell types they can give rise to. Leukemic cells normally do not differentiate, but in this case the cells were induced to do so: they were incubated with differentiation factor, a protein made by normal cells. Once the leukemic cells had assumed a mature form they stopped growing and hence were not malignant.



CULTURE SYSTEM devised by the author and his students enabled them to observe the growth and differentiation of myeloid cells and identify the regulatory molecules affecting both processes. A layer of fibroblasts, or connective-tissue cells, had to be present in the culture for the myeloid cells to multiply into clones and differentiate into mature forms, either macrophages or granulocytes. Purification of the culture medium showed that fibroblasts secrete proteins that are essential for growth and differentiation of myeloid cells.



COUPLING OF CELL GROWTH AND DIFFERENTIATION in myeloid cells results from an interaction between regulatory proteins. Growth factors from an external source (such as fibroblasts) induce precursor cells to multiply. The growth factors also cause the cells to produce their own differentiation factors. Eventually the supply of differentiation factors is sufficient to induce the cells to assume their final forms. Growth then stops. Thus cell differentiation (and the cessation of growth) is efficiently linked to cell multiplication.

precursor cells by some other means.

It is now clear that each kind of factor includes a multiplicity of proteins. Studies in my laboratory and by other investigators elsewhere, including Metcalf and Anthony Burgess at the Walter and Eliza Hall Institute, Richard Stanley at the Albert Einstein College of Medicine, James N. Ihle at the Frederick Cancer Research Facility in Maryland and T. M. Dexter at the Victoria University of Manchester, have shown that four different growth-inducing proteins affect myeloid precursors alone. Each protein acts on myeloid cells that have a specific developmental fate.

One inducer stimulates the growth of myeloid stem cells that can later differentiate into a wide range of progeny: macrophages, granulocytes, red blood cells, megakaryocytes (the precursors of platelets), and eosinophils and mast cells (two kinds of cells that take part in inflammatory reactions). Another acts on myeloid cells that can be made to differentiate into macrophages and granulocytes only, and the third and fourth growth inducers affect only the precursors of macrophages and granulocytes respectively. It appears that the growth inducers make up a hierarchy of specificity: different growth inducers act on myeloid cells as they mature and become more restricted in their developmental program. Cloning of the DNA for these growth factors in various laboratories has shown that each of the growth-inducing proteins that act on myeloid precursor cells is encoded by a separate gene.

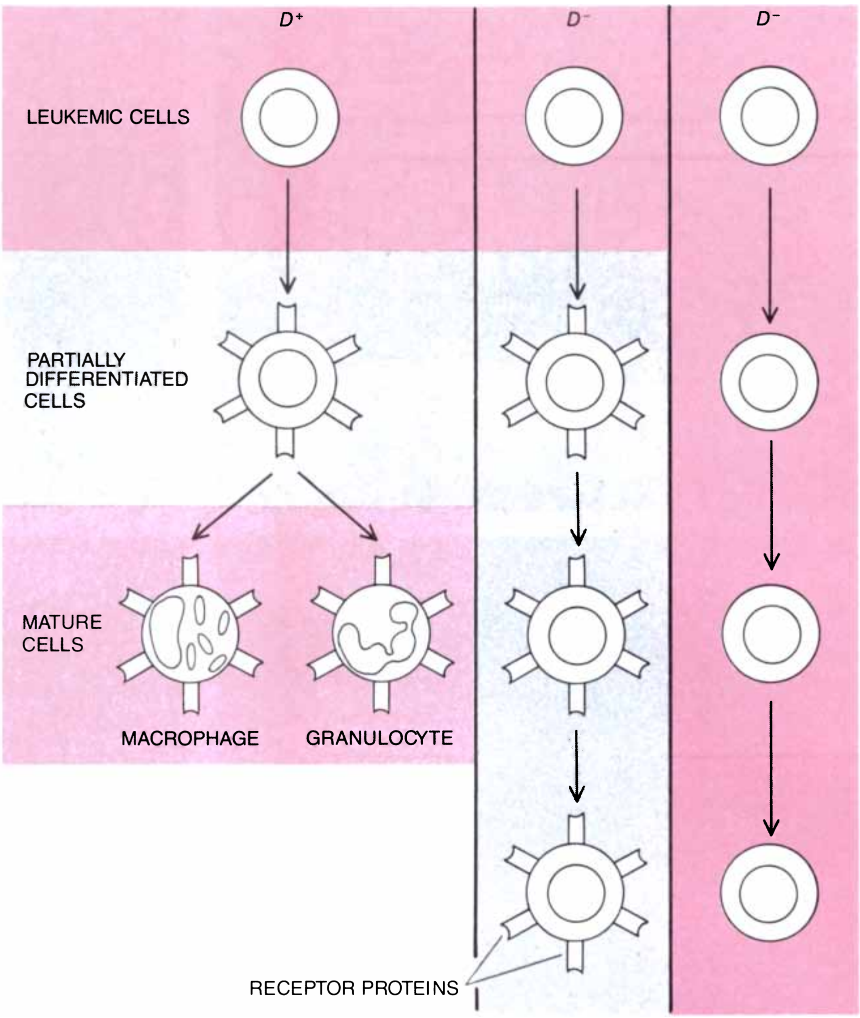
Inducers of differentiation, for their part, are probably as numerous as the cell types whose maturation they induce. Cloning of the DNA for differentiation inducers will also clarify the relations among the genes that code for these proteins.

The inducers of growth and differentiation have been given a variety of names. As a collective term for the inducers I first used *mashran gm,* from the biblical Hebrew word *sharo,* meaning "to send forth," with the initials for granulocytes and macrophages. Later other terms were adopted, including "colony stimulating factors" for all the proteins, MGI-1 (macrophage and granulocyte inducers type 1) for the growth factors and MGI-2 for the differentiation factors.

In normal cells the action of growth and differentiation inducers is elegantly coupled. Lotem and I found that when normal myeloid precursors are incubated with growth factor purified from a culture of other cell types, the precursors multiplied and then differentiated, even though we had not added differentiation factor. We showed that the source of the differentiation inducer was the myeloid cells themselves. Growth factor, then, has two effects: it causes the precursor cells to grow and it induces in them the production of differentiation factor. When the precursors have multiplied into a large enough population, their combined production of differentiation factor is sufficient to cause them to differentiate. This normal relation between growth and differentiation can be short-circuited, however, by certain compounds that cause cells other than the precursors themselves to produce and secrete differentiation factor.

The production of differentiation factor may take place later in the growth of some precursors than it does in others, allowing a larger population of cells to form before differentiation occurs. Specific growth factors may induce precursor cells to make specific differentiation factors, thereby leading to particular kinds of mature cell. This possibility might explain why each of the four myeloid growth factors is associated with a different set of mature cells; it offers a possible alternative to the explanation that different growth factors act at different stages of cell development.

The normal coupling between cell growth and differentiation is accompanied by a second linkage, in the cells of the blood and other tissues: between differentiation and the cessation of growth. Why do mature cells stop multiplying? Mature red blood cells in human beings and other mammals eliminate their nucleus and so are unable to divide. Other mature cells retain their nucleus but do not grow; it has been suggested that such terminally differentiated cells produce growth-inhibiting compounds, which block the cells' own multiplication.



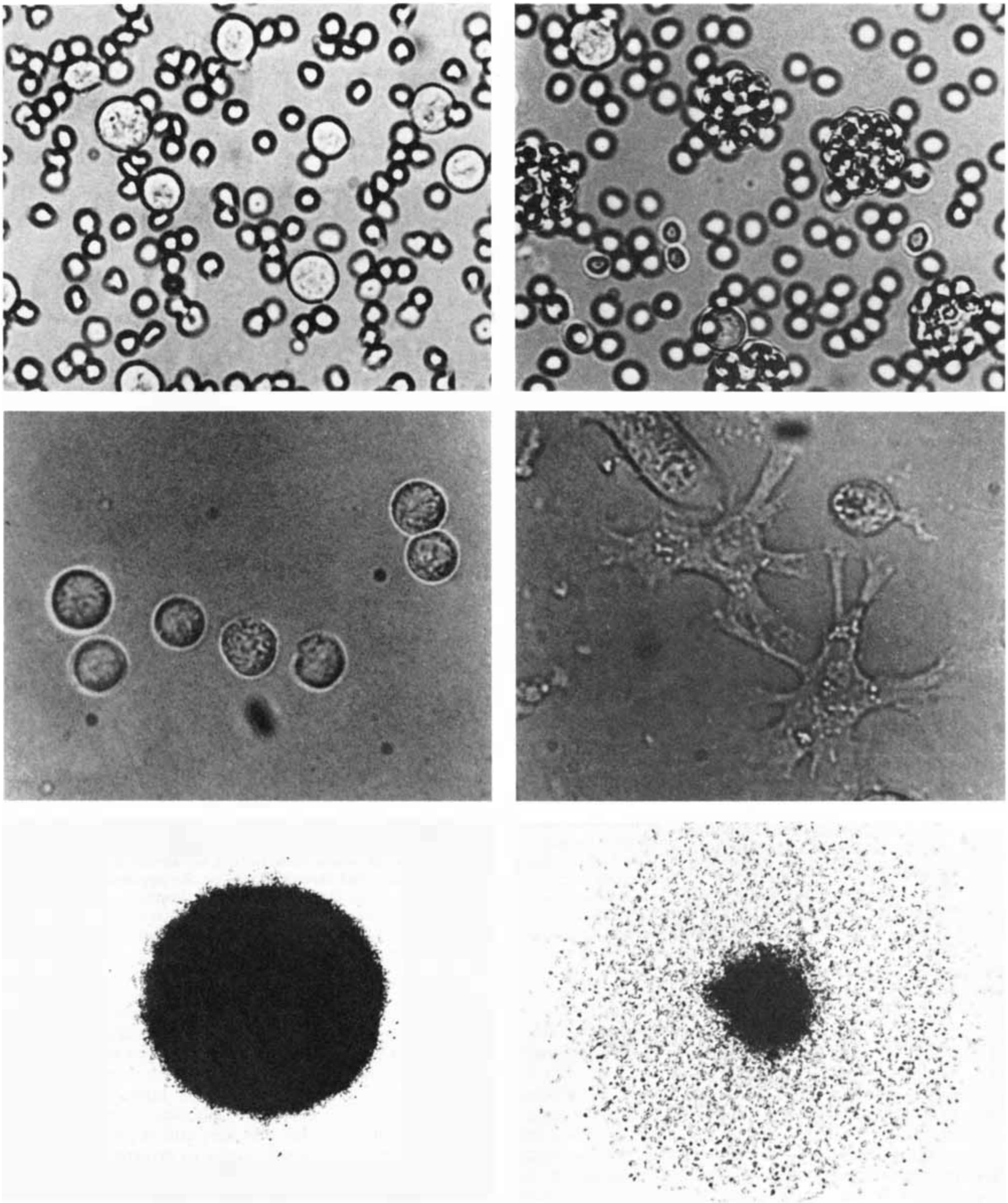
RESPONSE TO DIFFERENTIATION FACTORS varies among clones of leukemic cells. Some clones are differentiation positive (D^+): the addition of differentiation factors isolated from cultures of normal cells causes them to develop normally. First they assume an intermediate stage of differentiation, in which their surface membrane displays receptor proteins characteristic of mature cells, and then they take on the appearance of mature, nondividing macrophages or granulocytes. Other clones are differentiation defective (D^-). Under the influence of differentiation factors they may reach an intermediate stage of differentiation, in which their growth is merely slowed, or they may be entirely unaffected and continue to multiply as before. Other compounds can make D^- cells differentiate, however.

Leukemic cells have escaped from the constraints governing the growth of normal myeloid cells. One limitation on normal growth is the supply of growth factor. Normal myeloid precursors must depend on other cells to produce the factor, and the supply can be sporadic. During an infection, for example, fibroblasts and other kinds of cells secrete large quantities of the factor, thereby increasing the body's population of white cells, but at other times the level of growth factor declines. Does the continued growth of leukemic cells indicate that they have escaped from the normal requirement for growth factor?

To answer the question we had to obtain cloned lines of malignant cells that could be grown in culture and were not contaminated with normal cells. In my laboratory and in others several such lines of myeloid leukemic cells have been isolated. Their study revealed that the cells circumvent the growth limitation ordinarily imposed by the supply of growth factor in two ways. Certain strains of leukemic cells simply need less growth factor than normal cells, and the amount they require decreases as the cells grow in culture until they have no need of any inducer. Another strain produces its own growth factor.

Either alteration makes the cells capable of multiplying steadily rather than intermittently, as normal cells do. These changes can also explain how cancer cells can metastasize to sites where the growth factors their normal counterparts need to survive and grow are lacking. Conversely, the preference of some kinds of metastatic cells for particular organs may indicate that they still require a small external supply of growth factor, which their preferred organs supply.

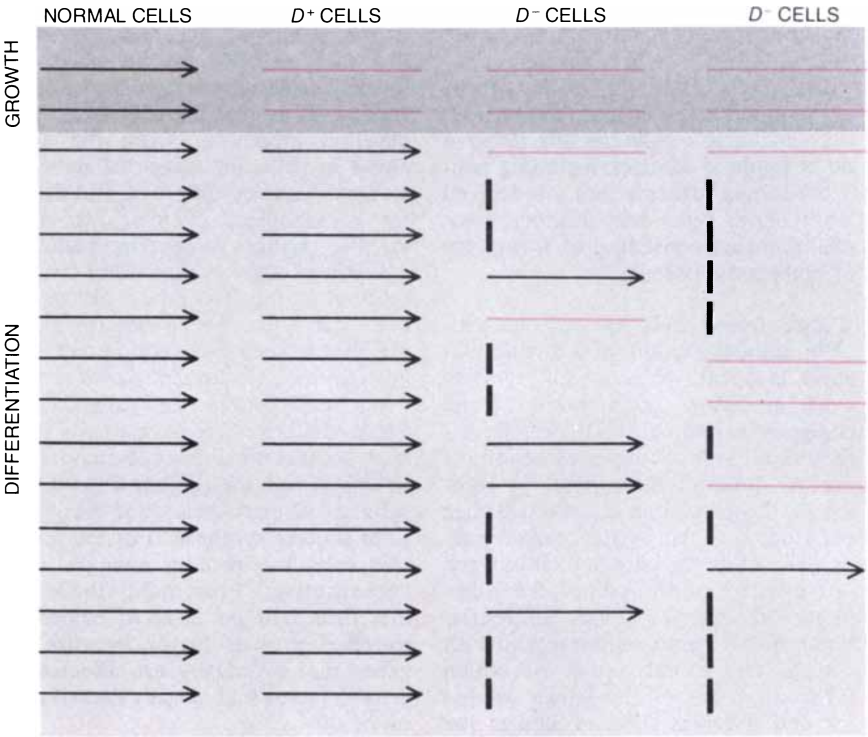
To qualify as malignant a cell must also escape the second constraint limiting the growth of normal cells: the coupling of growth and differentiation, which regulates the balance between multiplying cells and cells that have stopped growing. If growth factor induced the production of differentiation factor in leukemic cells, then malignant cells producing growth factor on their own or exposed to it from oth-



HALLMARKS OF MATURITY are evident in leukemic cells that have been made to differentiate. One way to detect certain surface receptor proteins characteristic of mature white blood cells is to mix the cells with the red blood cells of sheep. The red cells gather on the surface of the mature cells, forming "rosettes." Sheep red cells do not cluster on myeloid leukemic cells (*larger cells at top left*), but the rosettes do form on leukemic cells that have been exposed to a differentiation inducer (*top right*). When leukemic cells are placed on a surface, they ordinarily remain round and unattached (*middle left*). After a specific differentiation factor has caused them to differentiate into macrophages, the cells display the ability of normal macrophages to spread out on a surface and move across it (*middle right*). Mature macrophages and granulocytes are capable of chemotaxis: movement toward a chemical stimulus. A clone of leukemic cells exposed to a chemical stimulus remains compact because the cells do not migrate (*bottom left*). In a clone of cells that have been made to differentiate the stimulus causes the edge of the clone to grow diffuse as cells migrate (*bottom right*).

er cells would differentiate and stop growing. We added purified growth factor to cultures of leukemic cells that multiplied without growth factor and found that it does not trigger the production of differentiation factor. Thus the development of leukemia reflects two sets of genetic changes in myeloid precursor cells: one set that reduces or eliminates the cells' need for an external supply of growth factor and another that uncouples cell multiplication from differentiation.

Although leukemic cells in the body generally do not produce differentiation factor, we wondered whether differentiation factor supplied artificially might cause them to mature and stop growing. Certain clones of myeloid leukemic cells, we found, do differentiate when they are incubated with normal differentiation factors. When such cells are induced to differentiate, they take on the distinguishing characteristics of mature macrophages and granulocytes. They form the same protein receptors on their membrane that are found on mature normal cells and develop the same ability to move in the direction of certain chemical stimuli. The leukemic cells incubated with a differentiation factor that induces the development of macrophages also become able to spread out and crawl on surfaces to which they are attached. Work in other laboratories has confirmed that when myeloid leukemic cells are induced to differentiate, they develop the traits of mature macrophages and granulocytes. Once they have differentiated, the leukemic cells take on a further characteristic of mature normal cells: they stop growing.



VARYING PATTERNS OF PROTEIN SYNTHESIS are observed among normal cells, differentiation-positive (D^+) leukemic cells and differentiation-defective (D^-) cells. Normal myeloid cells undergo an array of changes (*arrows*) in the synthesis of cellular proteins when they are exposed to growth and differentiation factors; the cells begin making some proteins and stop making others. In leukemic cells the protein changes that growth factor ordinarily induces are constitutive, that is, they have already taken place (*color*). The constitutive changes can explain why the cells are able to multiply without growth factor. In D^+ cells the protein changes normally triggered by differentiation factor remain to be induced, and so D^+ cells respond normally to the factor. In D^- clones many of the changes either are constitutive or cannot take place (*bars*). The clones that are the most resistant to differentiation factor display the largest number of constitutive changes. Normal differentiation seems to require that the entire array of changes occur in synchrony; constitutive protein changes produce asynchrony and prevent D^- cells from responding to differentiation factor.

Some myeloid cell lines, then, become leukemic by losing their normal requirement for growth factor but remain capable of following the normal pathway of differentiation into macrophages or granulocytes. In contrast to these differentiation-positive (D^+) clones, my colleagues and I also isolated clones of leukemic cells that were differentiation-defective (D^-). Differentiation factor caused some of these clones to develop to an intermediate stage of differentiation, in which cell multiplication was slowed, whereas other clones could not be induced to differentiate even to this intermediate stage. The D^- trait occurs both among leukemic cells producing their own growth factor and among those that do not need growth factor. I have suggested that D^+ clones represent an early stage of malignancy and D^- clones correspond to a more advanced stage, in which the malignant cells have undergone more extensive changes in their genetic makeup.

Do leukemic cells lose all the genes for differentiation as they develop to the D^- stage? We found that differentiation can be triggered in D^+ clones by means other than normal differentiation factor. Certain steroid hormones, small amounts of X rays and low doses of chemicals used at much higher doses in conventional cancer therapy, such as cytosine arabinoside, adriamycin and methotrexate, can cause such leukemic cells to mature. Other compounds that can induce differentiation include insulin, some vitamins, bacterial lipopolysaccharide (a component of the bacterial cell wall) and certain compounds made by plants: lectins and some phorbol esters.

Such compounds, administered singly, in combination or together with normal differentiation factor, can also induce differentiation in D^- clones, although we found the effectiveness of a given compound often varied among the clones. This finding indicates D^- leukemic cells retain genes that are active in normal differentiation. It is probable that an appropriate combination of compounds can induce any myeloid leukemic cell no longer susceptible to normal differentiation factor to mature and stop growing.

Malignant cells of other lineages can also be artificially induced to differentiate. Charlotte Friend, now at the Mount Sinai School of Medicine, found that the compound dimethylsulfoxide causes erythroleukemic cells (leukemic red blood cells) to differentiate, and Paul A. Marks and Richard A. Rifkind, both of whom are now at the Memorial Sloan-Kettering Cancer Center, found that other chemicals have the same effect. The erythroleukemic strains the investigators studied did not differentiate in response to erythropoietin, a protein that normally induces the production of hemoglobin (a hallmark of mature red blood cells) in developing cells. The leukemic cells therefore resembled our D^- clones in that they did not respond to a normal differentiation factor. Workers elsewhere have found that some of the compounds that stimulate differentiation in leukemic cells can also do so in other kinds of cancers, including neuroblastomas (cancers that are derived from nerve cells) and teratocarcinomas (cancers consisting of a mixture of embryonic tissues).

We found that the various substances capable of inducing myeloid leukemic cells to differentiate work in diverse ways. Some of the compounds that cause D^+ cells to differentiate induce the production of normal differentiation factor by turning on the genes that encode it. Other compounds, such as the steroid hormones, act on the genome without the intermediary of normal differentiation factor. In some D^- clones differentiation requires combined treatment with several compounds, each of which turns on some of the genes needed for cell differentiation. Together the compounds activate a combination of genes that is sufficient to result in differentiation.

A study of cellular proteins in normal and leukemic cells revealed some of the genetic bases for the varying responses of leukemic strains to differentiation-inducing substances. Dan Liebermann, Barbara Hoffman-Liebermann and I used the technique of two-dimensional gel electrophoresis to resolve cellular proteins according to their molecular weight and electric charge. By comparing gels made from different clones or from the same clone in different stages of development and noting differences in the array of separated proteins, we could identify changes in protein synthesis.

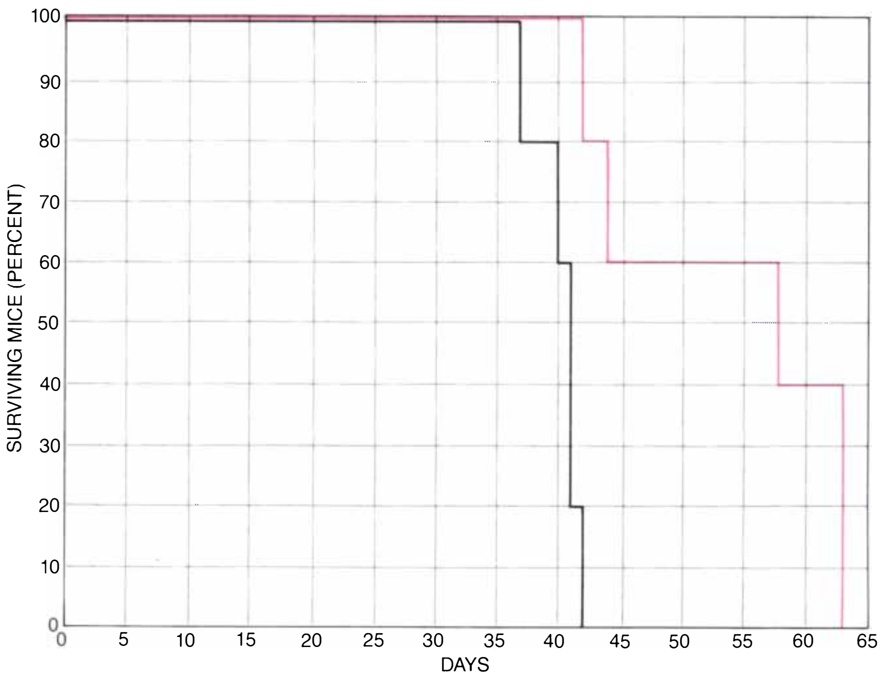
Some of the changes bore a clear relation to the leukemic cells' malignant character. We found, for example, that in normal myeloid precursors growth factor induces the synthesis of a number of proteins and stops the synthesis of others. In leukemic cells, however, these protein changes do not need to be induced: the gels showed that the cells had already undergone the changes in protein synthesis. For the leukemic cells the protein changes were "constitutive." Presumably the leukemic cells had no need of externally supplied growth factor because the genes that ordinarily are affected by growth factor had already been turned on or off.

We also identified some of the protein changes induced in normal cells by differentiation factor. D^+ leukemic cells, like normal cells, only showed the protein changes after exposure to differentiation factor: the changes had to be induced. Many of the changes in protein synthesis were constitutive in D^- clones, however. The more resistant a clone was to the induction of differentiation by normal differentiation factor, the larger was the number of constitutive protein changes the cells displayed.

Our results suggest that the normal process of differentiation requires the synchronous expression of an array of genes; interactions among the proteins encoded by the genes presumably are crucial to differentiation. When the expression of a gene is constitutive rather than induced, the necessary coordination of gene expression fails. The asynchrony can block the normal differentiation program, resulting in a cell that never differentiates fully and that therefore continues to multiply. Geoffrey Symonds and I found that certain serums cause constitutive protein changes that are characteristic of D^- leukemic cells to revert, so that their expression has to be induced. After treatment with these serums the cells became capable of maturing under the influence of normal differentiation factor. We had in effect unblocked the genetic program for differentiation.

The finding that various compounds can stimulate D^- cells to mature in spite of constitutive protein changes that make the cells resistant to normal differentiation factor indicates that there are several genetic programs for differentiation. Lydia Cohen and I used two-dimensional gel electrophoresis to examine the cellular proteins of two kinds of myeloid leukemic cells: those that could be induced to differentiate by normal differentiation factor but not by the steroid hormone dexamethasone and those that responded to dexamethasone but not to differentiation factor. The clones displayed different sets of constitutive protein changes. It appears that normal differentiation factor and some other inducers of differentiation activate distinct sets of genes, although the net effect of the genes—cell maturation—is the same. It is likely that cells in other cancers can also follow a variety of genetic pathways to differentiation.

My colleagues and I have found that none of the cells in our myeloid leukemic clones have a normal array of chromosomes. The abnormalities we observed include changes in the number of chromosomes as well as rearrangements and deletions of chromosome segments. We noted consistent differences in the clones that can be induced to mature by differentiation factor and those that cannot. The chromosome changes that can be observed in leukemic cells presumably are responsible for the genetic abnormalities in the control of cell growth



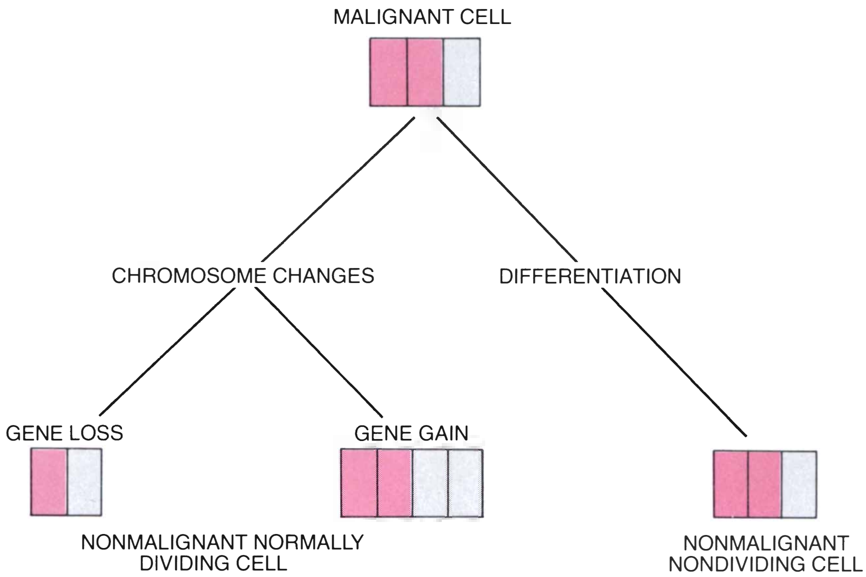
SURVIVAL OF MICE injected with myeloid leukemic cells was longer (*color*) when they were given differentiation factor than when they were not treated (*black*). The leukemic cells were from D^+ clones, which mature and stop growing in culture when they are exposed to differentiation factor. The extended survival of the mice that were treated with differentiation factor shows that it inhibited the development of leukemia in the animals.

and differentiation that make such cells malignant.

One way that malignancy can be reversed is by restoring the chromosomes of malignant cells to a more normal pattern. In 1968 Zelig Rabinowitz and I showed that clones of sarcoma cells (malignant cells arising in connective tissue) cultured under certain conditions often yielded cells that had reverted to a nonmalignant state. In sarcomas that Carmia Borek and I induced in cultured normal fibroblasts by means of chemical carcinogens or X rays, the revertant cells had the limited life span of normal fibroblasts. They also had a chromosome composition different from that of their malignant progenitors: one more closely resembling the chromosomes of normal cells.

The chromosomes whose alteration was crucial to the reversion of malignant cells carried genes that at the time I called *E* genes (genes for the expression of malignancy) and *S* genes (genes for the suppression of malignancy). Chromosome changes that acted to remove *E* genes from the influence of *S* genes caused cancer; those that restored the gene balance led to reversion. The *E* genes were later isolated in various laboratories and named oncogenes; the *S* genes can be called soncogenes or anti-oncogenes.

We asked whether a similar restoration of a more normal chromosome makeup takes place when leukemic cells are made to differentiate and stop growing. Are the genetic abnormalities corrected or does the reversion of malignancy by differentiation bypass the genetic changes that made the cells malignant in the first place? We examined the chromosomes of myeloid leukemic cells that had been made to differentiate and found their chromosome composition was still abnormal. It appears that arresting multiplication by inducing differentiation bypasses the genetic abnormalities that originally disrupted the normal sequence of growth and differentiation.



REVERSAL OF MALIGNANCY can be achieved in two ways. Cells become malignant as a result of genetic changes disturbing the normal balance between genes that can cause malignancy (*color*), generally known as oncogenes, and genes that suppress oncogene expression (*gray*). In culture malignant cells have been observed to resume normal growth following chromosome changes that restored the balance between oncogenes and their suppressors (*left*). Arresting malignancy through cell differentiation, in contrast, bypasses the genetic changes (*right*). A malignant cell's genetic makeup remains abnormal, but after differentiating the cell assumes a mature form, stops dividing and is no longer malignant.

The finding that malignancy can be reversed by causing cancerous cells to differentiate and stop growing opens new possibilities for therapy. The reversal of the cancer-causing genetic abnormalities themselves may figure in cancer treatment sometime in the future. The therapeutic value of bypassing the genetic changes by inducing cell differentiation can already be tested. My group at the Weizmann Institute has demonstrated that the development of leukemia in mice injected with D^+ leukemic cells is slowed when the mice are given differentiation factor or a compound that causes normal cells in the body to increase the production of differentiation factor. It may be that similar effects can be achieved in human beings, providing an alternative to the cytotoxic drugs now employed in cancer chemotherapy, which kill many normal cells as well as cancer cells.

In another therapeutic use, normal macrophage and granulocyte inducers—both growth and differentiation factors—might be administered to patients undergoing conventional cytotoxic therapy for other cancers. The factors might bolster their macrophage and granulocyte populations, which chemotherapy depletes. The factors might also be given to alleviate nonmalignant deficiencies in macrophages or granulocytes. Factors that act on other cell types could be put to similar clinical uses.

The treatment of leukemias in which the cells, like our D^- clones, no longer respond to normal differentiation factor could employ other differentiation-inducing substances. Such compounds are already used, in high doses, in cancer chemotherapy. Our work suggests that their efficacy may derive not only from the killing of cells but also from the induction of differentiation, either directly or by stimulating the malignant cells to produce differentiation factor. Differences in the response of patients to chemotherapy may reflect the varying susceptibility of malignant cells to the differentiation-inducing compounds.

Differentiation therapy for leukemia could use the same substances in much lower doses, perhaps in combination with normal differentiation factor. Some of the compounds could be given in small doses to induce differentiation while others are administered in larger quantities for their cytotoxic effects, thereby reducing the total number of leukemic cells. In order to choose the best combination of compounds and approaches before treatment, leukemic cells from the patient could be tested outside the body in order to determine their susceptibility to various substances.

My suggestions have already led several workers to attempt clinical trials of differentiation inducers in myeloid leukemia. Laurent Degos of the Hôpital St. Louis in Paris is among those who have had encouraging results using small doses of cytosine arabinoside, one of the chemicals with which we induced differentiation of myeloid leukemic cells in vitro. It is likely that the findings with myeloid cells will prove to be generalizable to other kinds of cells in which other specific factors control growth and differentiation. If they are, reversing malignancy by inducing cell differentiation could eventually figure in the treatment of a range of cancers.

The Structure of Comet Tails

The plasma tail forms and disconnects from the comet in response to the solar wind and its magnetic field. Observations of comets Giacobini-Zinner and Halley may help to clarify such phenomena

by John C. Brandt and Malcolm B. Niedner, Jr.

The years 1985–86 will one day be regarded as a golden age for cometary astronomy. Indeed, if we had been allowed to choose two years in which to be active as cometary scientists, these would have been our clear choices. Two important comets, Giacobini-Zinner and Halley, have approached within range of observation as they orbit around the sun. Giacobini-Zinner has already yielded a bounty of information as a result of being the first comet to be visited by a spacecraft, and astronomers have deployed an unprecedented array of resources to examine Comet Halley. Data are being gathered by observatories on the earth's surface, by spacecraft orbiting the earth, by vehicles in space and in orbit around other planets and by six spacecraft that will fly near or into the comet's atmosphere.

It is fortunate that Giacobini-Zinner and Halley—among the few known periodic comets that are sufficiently bright and are also nearly complete in their range of cometary features—have come under close scrutiny at virtually the same time. Astronomers will compare findings derived from similar observational techniques for the two rather different comets. The massive efforts organized to study Giacobini-Zinner and Halley promise to provide direct evidence for theories about the origin, composition and dynamics of comets and their tails; we also expect them to raise many new questions.

The National Aeronautics and Space Administration's *International Cometary Explorer* (ICE) sped through the tail of Giacobini-Zinner on September 11, 1985, and relayed a mass of data that are still being evaluated. Probes from the European Space Agency, Japan and the U.S.S.R. will venture to Halley in March, transmitting valuable measurements bearing on the structure, composition and physical conditions of cometary atmospheres, as well as providing the first images of a cometary nucleus. The missions will be supported by networks of ground-based observations; concurrent observations will be made by the crew of NASA's Astro 1 space shuttle mission.

The current missions extend a rich history of comet observation that has stretched across the centuries. The word "comet" comes from the terminology of Greek astronomers, who first named these solar visitors "aster kometes," or long-haired star. It is now known that comets are composed of three main parts: the atmosphere, the tail and the nucleus. The visual part of the atmosphere is often called the coma or head; it is an essentially spherical cloud of gas and dust. The atmosphere may range from less than 1,000 kilometers in diameter to several million kilometers, depending on the species of gas. One or more tails, which are directed away from the sun, extend behind the atmosphere. There are two principal types of tails: dust tails and plasma tails.

Dust tails typically consist of solid, micrometer-size particles that have been pushed away from the coma by the pressure force resulting from sunlight striking the dust grains. They are curved and fuzzy, usually displaying little or no internal structure. Plasma tails are quite different. They are made of molecules that have been ionized in the atmosphere by solar radiation, trapped on interplanetary magnetic fields generated by the sun, and wrapped around the comet so that the ions form a long, hairpin structure, which often exhibits threadlike formations, knots and large-scale disturbances. The gases in the coma and the tail fluoresce, that is, they absorb sunlight and reradiate it.

According to a generally accepted model proposed by Fred L. Whipple in 1950, the source of all cometary material is the nucleus, which is inside the atmosphere but is never observed through telescopes because it is too small. Whipple compared the nucleus to a dirty snowball; it consists of ices of water and other molecules. Dust grains, and possibly rocky material, are interspersed more or less uniformly throughout the icy matrix. The nucleus is often considered to be somewhat spherical, measuring several kilometers in diameter. How can such a minor body give rise to plasma tails that are sometimes 50 million kilometers long?

The answer lies in the dynamics of the material the nucleus releases into interplanetary space. As a comet approaches the sun the nucleus absorbs sunlight and heats up until it reaches the temperature at which the ices sublimate, or go directly from the solid phase into the gaseous one. The escaping gases leave the nucleus at an initial speed of several tenths of a kilometer per second; as they move outward they undergo many chemical reactions. In addition fleeing gas molecules collide with newly liberated dust grains and propel them outward. Many of the gas molecules absorb ultraviolet photons of sunlight and gain kinetic energy as they break into smaller molecules—a process called photodissociation. This

COMET WEST, photographed on March 9, 1976, was characterized by two types of comet tail. The dust tail is the wide, diffuse tail to the left, consisting of three broad bands with faint emissions between them. It is composed of solid particles that have been pushed away from the comet's atmosphere by the pressure force resulting from sunlight striking the dust grains. The plasma tail is the narrower one to the right, which is threaded with intricate streamers. It is composed of molecules that have been trapped by interplanetary magnetic field lines (carried by particles of the solar wind) that have been wrapped around the comet.

COMET GIACOBINI-ZINNER is seen in an electronic image made by Uwe Fink with a charge-coupled device at the Lunar and Planetary Laboratory in Tucson. The date was July 26, 1985. An extensive atmosphere surrounds the comet, somewhat elongated in a direction away from the sun. No plasma tail is seen, probably because the exposure was short.

conversion of sunlight into kinetic energy and the pressure of denser gas forcing these molecules outward cause them to accelerate to speeds that average about one kilometer per second. The gas and dust mixture expands to become the cometary atmosphere, which continues to expand and to be replenished as material constantly escapes from the nucleus.

After the gases leave the nucleus they undergo a complicated series of transformations that produce new molecules; a recent computer simulation of these activities included more than 1,200 processes and reactions. The escaping gases can also react with particles that surround the comet as they speed away from the sun. In one type of reaction, called charge exchange, a proton captures an electron from a cometary gas molecule or atom to yield a neutral hydrogen atom and a positively charged ion. The escaping gases are also modified by reactions between neutral molecules and ions, in which they exchange charge. The various reactions altering the original nuclear gases yield a concentration of ions in the coma that will become the gaseous constituents of the comet's plasma tail.

Understanding of the formation and structure of the plasma tail was greatly advanced by the introduction of photography in the late 1800's. The first high-quality photographs of comets were probably the ones obtained by Sir David Gill in 1882, but the technique was not practiced regularly until Comet Swift appeared a decade later. Edward Emerson Barnard pioneered in the intensive exploitation of photography, working with such simple equipment as a portrait-lens camera. He was the first person to observe that plasma tails are highly intricate, threaded with raylike structures and streamers often accompanied by knots, condensations and helixes. In the photographs he obtained several times in the course of each clear night, he noted that parts of the plasma tail moved at high speeds. Indeed, photographs that had been made on successive nights usually bore no resemblance to one another because of great changes in the tail. Although Barnard did not know the detailed composition of the tails he observed, he was able to distinguish them from the slower-changing tails now known to be dust tails.

Barnard's studies led to the major revelation of these early years of cometary photography: the understanding that a plasma tail disconnects and floats away into space, to be replaced by a new tail. In a 1905 paper, "The Anomalous Tails of Comets," Barnard described this mystifying cyclic phenomenon and advocated that frequent photographs be made throughout the night. He argued: "The day-to-day history of a comet has too great an interval, and the changes are not necessarily at all connected. It is the hour-to-hour history that must be studied to understand the changes taking place in the comet. In the case of a very bright comet, exposures at intervals of half an hour should be made as long and as continuously as the conditions will permit. By this means it will be possible to determine the exact value of the motion of the particles in the tails of various comets." Barnard conjectured from these studies that disconnections are caused by interactions with "currents in interplanetary space across which the tail may sweep."

Little progress was made beyond Barnard's prescient work until 1951. In that year Ludwig Biermann began to discern the nature of the currents by analyzing the motions and accelerations of structures observed in cometary plasma tails. He calculated that the pressure of the sun's light was not nearly sufficient to account for the strength of the force pushing the tail away from the sun. Biermann argued that some form of "corpuscular radiation" continuously emitted from the sun must collide with cometary ions to form the tail and produce the large observed accelerations.

The next step was to determine the nature of these hypothesized solar particles and explain how they account for the plasma tail's intricate morphology. It was known that hydrogen and helium are ionized in the corona (the tenuous outer part of the sun's atmosphere) and that the resulting protons and electrons flow away from the sun as a result of the high solar temperatures. E. N. Parker of the University of Chicago determined that these atoms must accelerate as they leave the sun's gravitation and travel through interplanetary space. He coined the term solar wind to account for the radiation's dynamic motion as it sweeps across the solar system. In 1957, a year before Parker's pioneering paper was published, Hannes Alfvén of the Royal Institute of Technology in Stockholm speculated that this corpuscular radiation probably carries the sun's magnetic field into space. His paper showed that the solar wind's magnetic field couples the solar-wind plasma with the cometary ions to form plasma tails. Alfvén's "magnetic flux tube" model is widely accepted today. Our work and the observations of Giacobini-Zinner and Halley continue to add detail to his theory.

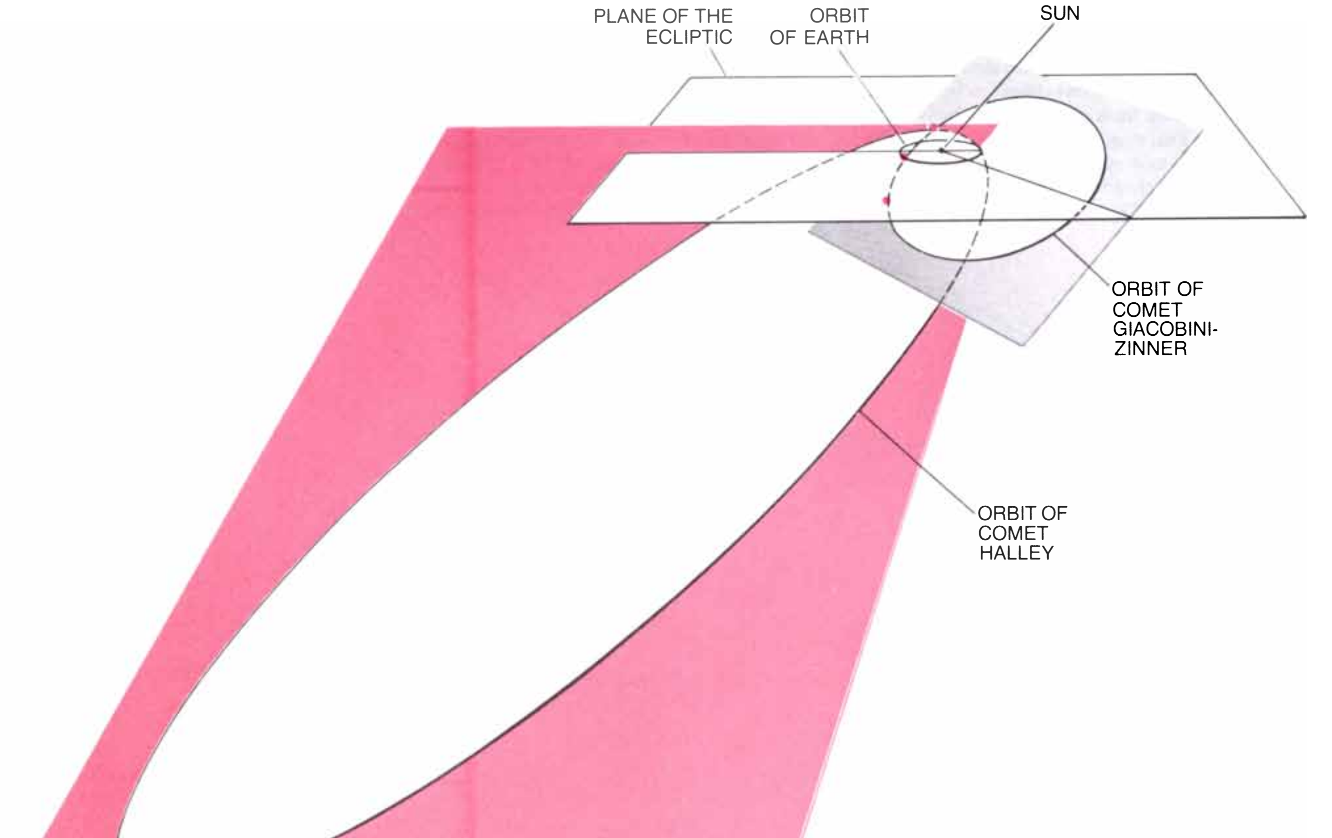
At the heart of the theory is the violent collision between the solar wind and the comet's atmospheric gases. The solar wind and its magnetic field both flow toward the comet at typical speeds of 400 kilometers per second. In contrast to the solar-wind plasma traveling away from the sun, the newly created cometary ions are traveling toward the sun at perhaps one kilometer per second. Some basic knowledge of electricity and magnetism helps to predict what will happen when the gases collide. It is significant that charged particles cannot freely cross a magnetic field, but instead perform helical orbits along its lines of force. When ions from the outer regions of the atmosphere (500,000 kilometers or more from the nucleus) are deposited into the solar wind, they are "captured" on the solar wind's magnetic lines of force and consequently travel back toward the comet in the same direction as the solar wind.

Because the solar wind has had mass added to it in the form of cometary ions, it must slow down to conserve momentum. This deceleration process continues as progressively more ions are captured during the solar wind's flight toward the inner atmosphere. Eventually a point is reached where so many ions have been captured, and the flow has been so decelerated, that the outward pressures from the ions and other gases closer to the nucleus are balanced by the inward pressures exerted by the solar wind carrying the captured ions. At this point the flow of the solar wind stops; it is said to stagnate. The magnetic fields it carries, which have been continuously compressed, form a magnetic barrier that is also at rest. This takes place far inside the atmosphere; in the case of a bright comet such as Halley it may occur between 1,000 and 10,000 kilometers from the nucleus.

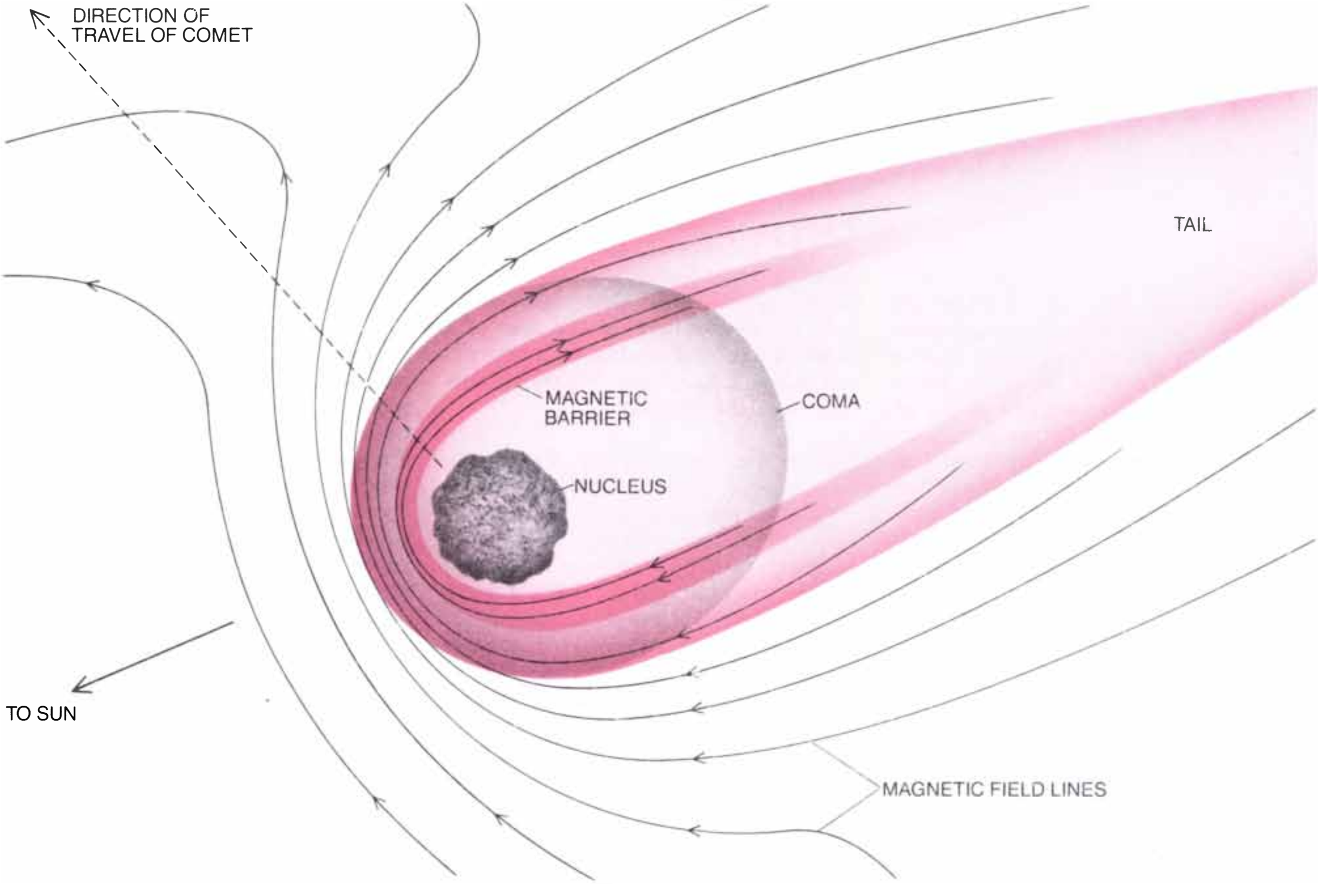
Since the comet is an obstacle to the solar wind, a "bow shock" is formed between perhaps 50,000 and 100,000 kilometers from the nucleus. It is similar to the bow wave made by a ship moving through the water. Far off to the sides of the comet smaller numbers of ions are captured, and consequently the solar wind is not significantly impeded. The magnetic fields in these areas, which are connected to fields in the barrier, wrap behind the comet, forming two lobes of opposite polarity. This magnetic tail can be observed because it channels fluorescing cometary ions such as those of carbon monoxide and water vapor.

The magnetic field embedded in the solar wind also interacts periodically with the magnetic field of the comet's tail to cause the kind of detachment Barnard observed. Our study of Comet Kohoutek in 1973–74 shows how the process works. The comet was photographed extensively at the Joint Observatory for Cometary Research near Socorro, N.M. Its images revealed a wealth of structural detail in the plasma tail. The photographs made on one night indicated that the tail extended outward to a point some distance from the head, stopped and then seemed to start again. We studied Barnard's writings on what are now called disconnection events carefully as we sought to determine whether the Kohoutek photographs were indeed demonstrating a property of comets or whether we were simply misinterpreting faint images on the film. His photographs provided the support we needed to view disconnection events as general properties of comets.

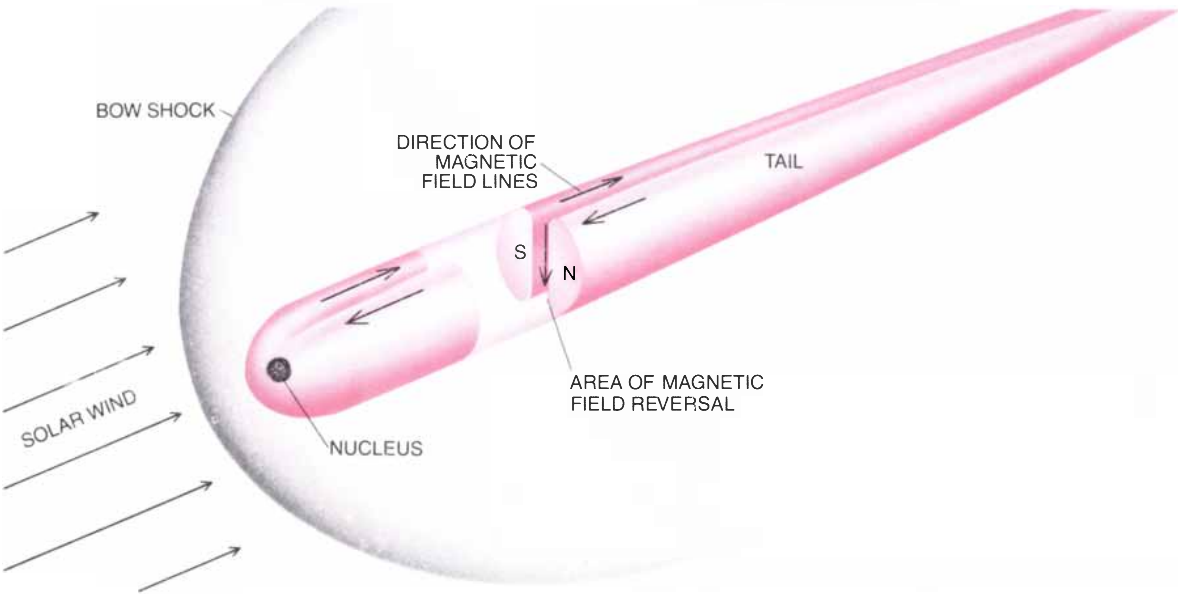
We next returned to data from Kohoutek, seeking to understand how



GIACOBINI-ZINNER AND HALLEY are periodic comets, which visit the inner solar system in the course of their regular journeys around the sun. Halley's comet is observable from the earth approximately once every 76 years and Giacobini-Zinner approaches the earth every 6.5 years. The colored dots on the three orbits indicate the position of the two comets and the earth as of January 15.

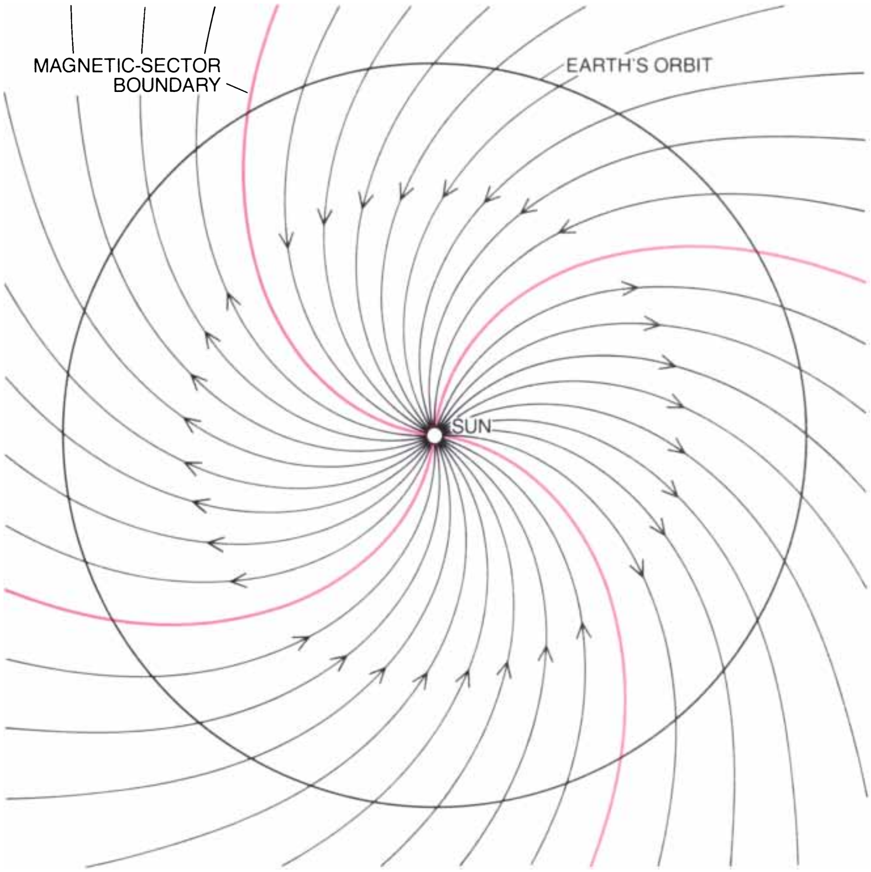


MAGNETIC FIELD LINES from the incident solar wind are compressed in the comet's atmosphere, whose ions become captured by the magnetic field lines as the solar wind streams into the cometary atmosphere. The compression is caused by the increasing mass of the solar wind as a result of picking up ions; conservation of momentum dictates that the flow must slow down. At some point deep inside the atmosphere, where the concentration of captured ions has become very large, the pressure of outwardly flowing ions matches the pressure of the inwardly flowing solar wind. This balance makes it impossible for the solar-wind plasma, and the lines of force it carries, to penetrate any farther into the atmosphere; the result is a magnetic-free region in the comet. Off to the sides the magnetic field lines wrap around and behind the comet. These lines drag ions in the atmosphere along with the solar wind, away from the sun.



IONS IN COMET TAIL are channeled along the magnetic field lines of the solar wind as it wraps around the comet. A sheet of electric current (*vertical arrow*) and a magnetic field of low strength separate lobes of opposite magnetic field polarity in the tail. When the supersonic solar wind has captured a large enough mass of cometary ions, it is abruptly slowed as it passes through a "bow shock," which is analogous to the shock wave caused by a supersonic aircraft or the bow wave caused by a boat as it moves through water.

disconnections can take place even though magnetic field lines have no beginning and no end. The Eighth Interplanetary Monitoring Platform satellite measured the solar-wind plasma and the magnetic field near the earth. By translating these data to the position of Kohoutek we found the likely cause of disconnection events. Kohoutek lost its tail when the comet passed through a magnetic sector boundary: a border between sectors of opposite polarity in the magnetic field. We concluded that a tail disconnection probably occurs every time a comet passes from one magnetic sector to another. There are plenty of opportunities for comets to lose their tails; during the sun's 25-day rotation period it forms four magnetic sectors of alternate polarities that expand constantly as they turn with the sun. When the sectors are detected in the plane of the earth's orbit, they appear in the shape of four swirling spirals.



SPIRAL SHAPE of the solar wind's magnetic field (as seen from above the plane of the earth's orbit) results from the sun's rotation as it emits plasma carrying the field. Four "field sectors" are usually observed with each rotation. Successive sectors are of opposite polarity.

According to our model, a comet drops its tail when it crosses a sector boundary because the new sector contains a magnetic field that opposes the field from which the tail developed. From a plasma-physics point of view, the opposition of the two fields creates a situation that could hardly be more unstable. The result is a process known variously as magnetic reconnection, magnetic merging or magnetic annihilation. Although the theoretical details of the phenomenon are poorly understood in spite of decades of study, it is generally accepted that the topology of the magnetic field in the cometary atmosphere changes in a fundamental way. When the old magnetic field lines of the comet are approached by the magnetic field lines of the new sector, the old field lines are cut and reconnected into the pattern of the field lines of the new sector. When the field lines of the comet are thus cut, the material they contain remains trapped in the old field lines while the comet continues to move into the new magnetic field. The tail appears to detach when the last part of the material embedded in the old field moves away from the comet. When the disconnection process is complete, the comet immediately starts to grow a new plasma tail: one whose polarity corresponds to that of the new magnetic sector.

Our model calling for disconnection events every time a comet crosses a sector boundary, or roughly every week, is supported by high correlations between observed disconnection events and crossings of magnetic sector boundaries. We have also found agreement between observed morphological changes during a disconnection event and those expected from the model. Other mechanisms have been proposed to explain disconnection events, but the explanation based on the cutting and reconnection of magnetic field lines at a sector boundary seems to be the most viable.

There are basic weaknesses in this model of disconnection events that should be solved by the extensive current observations of Halley. The first weakness is that our macroscopic model of the morphology of disconnection events and the evolution of plasma tails has been plagued by fragmentary data. The picture has been pieced together by utilizing images of different comets and analyzing data from various observatories that rely on a variety of instruments, emulsions and exposure times. We expect that the Large-Scale Phenomena Network of the International Halley Watch will help by providing a comprehensive record of Halley's journey. The network consists of approximately 100 facilities around the world, each one equipped with wide-field photographic instruments that should record a major fraction of the comet's plasma and dust tails. These observatories are tracking Halley from its passage across the northern hemisphere of the celestial sphere on its inbound journey through its return to the outer solar system across the southern hemisphere this spring.

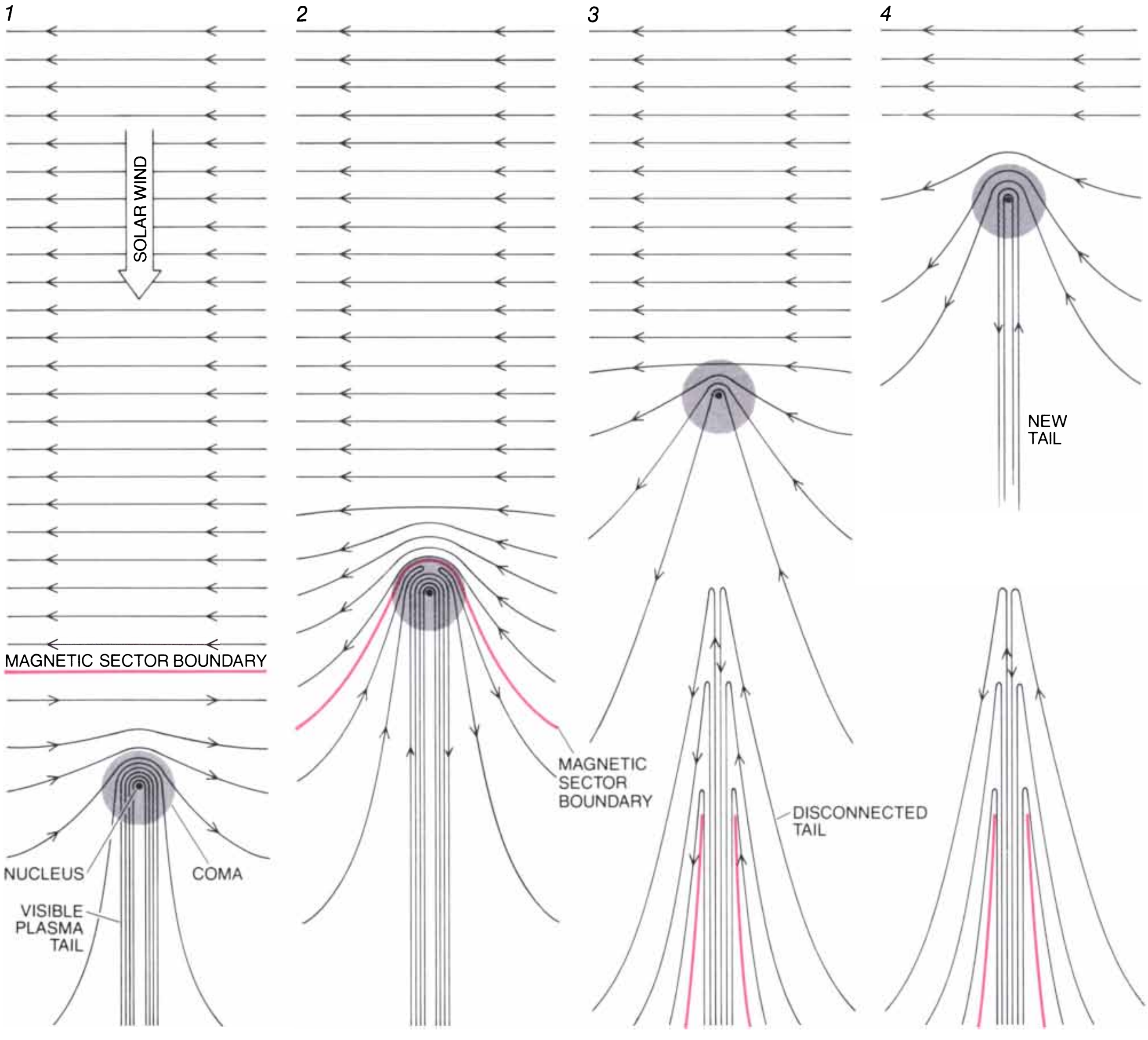
In order to achieve extended periods of frequent coverage, good weather and an even distribution of observation in both the earth's hemispheres is required. In the Southern Hemisphere, which is largely covered by oceans, four island-network sites have been established, two in the Pacific and one each in the Atlantic and Indian oceans. Portable telescopes will be set up at these sites to fill in what would otherwise be sparse coverage as the comet moves across the Southern Hemisphere in March and April.

Halley's dust tail and plasma tail should both be observable. The tails will point roughly in the same direction on the comet's journey toward the sun; on its outbound leg the tails should be separated by a wide angle. There is every reason to expect that the impressive tail lengths observed during the 1910 apparition will be repeated, although the tails will not appear as long this time owing to foreshortening and the greater distance from the earth. Analysis of photographs made in 1910 suggests that Halley's tail achieved lengths of several tenths of an astronomical unit. (A tenth of an astronomical unit is equal to 15 million kilometers.) As a result of highly favorable geometry and projection factors, the angular tail lengths in 1910 were as great as 50 degrees. In 1985–86 these apparent lengths should be in the range of from 10 to 15 degrees.

The coverage by the Large-Scale Phenomena Network will be supplemented by photography of Halley by the wide-field cameras on board the Astro 1 space shuttle mission. During one week in March the cameras will obtain images every six hours. They will also record the comet every hour and a half for one-day intervals when Halley is approached by two spacecraft, the U.S.S.R.'s *Vega-2* and the European Space Agency's *Giotto.*

The second weakness in our model—insufficient data on the microscopic conditions of cometary plasma and the tail's magnetic fields—is being addressed by six spacecraft. One, ICE, has already had a successful encounter with Giacobini-Zinner. The other five will encounter Halley in March of this year, and ICE will be an upstream monitor. Prior to these missions there have been no in situ measurements of any cometary parameters; we have had no accurate values for the magnetic field, temperature, density, composition or bulk velocities in any cometary plasma. Data on these parameters are necessary to determine which processes are important and how rapidly changes in the comet can be expected.

The first of the missions, ICE, passed through the plasma tail of Giacobini-Zinner some 8,000 kilometers from the nucleus on September 11, 1985. It provided significant evidence for the shape of the captured magnetic field, the two-lobed tail and the flow of electric current in the tail—features that follow from Alfvén's 1957 model. ICE measured a striking amount of turbu-



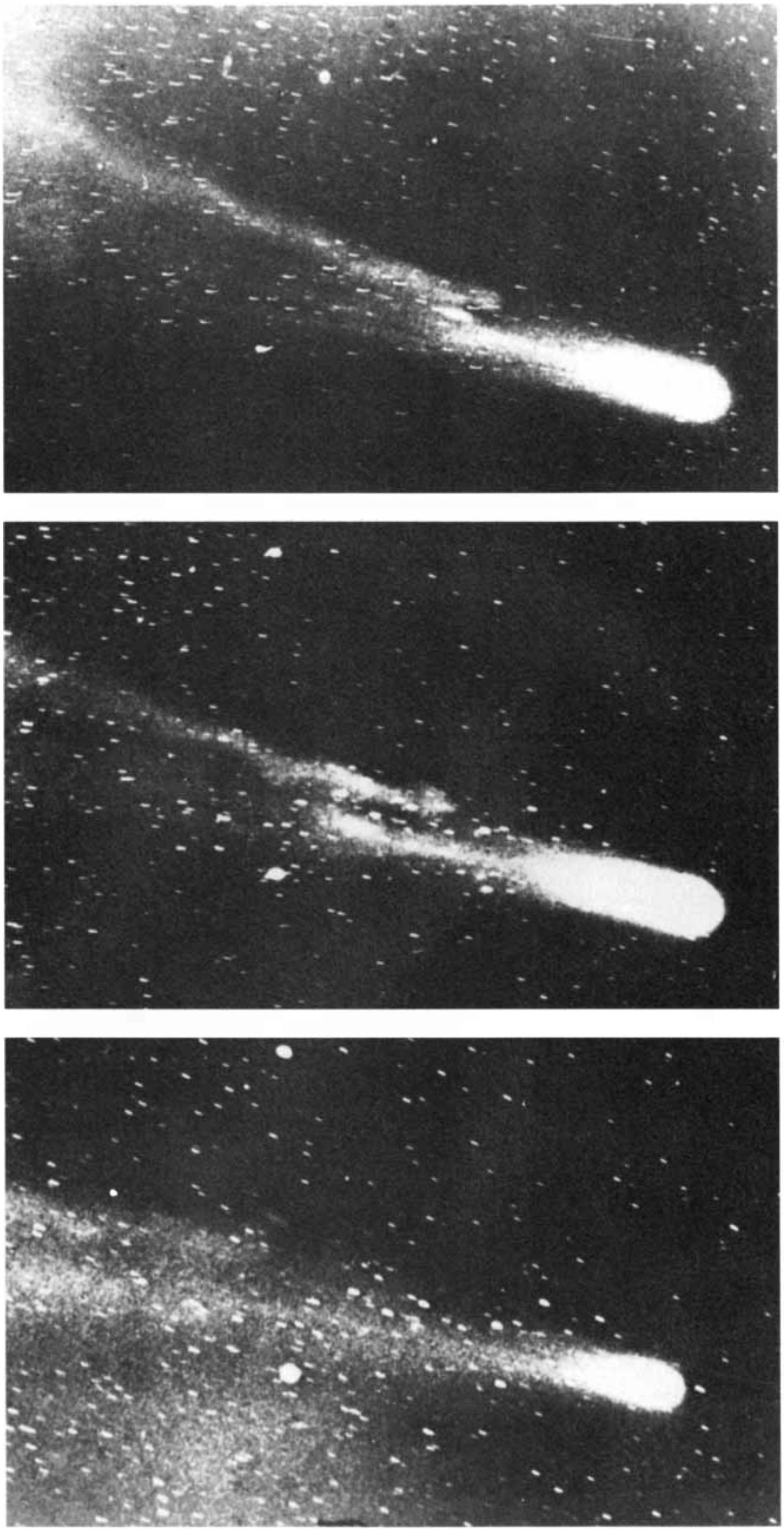
PLASMA TAIL DISCONNECTS when the comet crosses a sector boundary, moving from a sector where the magnetic field has the same polarity as the tail to a sector where the field has opposite polarity. When the new field enters the coma and is pressed into the oppositely directed old fields threading the tail, the old field lines are cut by a process known as magnetic reconnection. The old field lines retain ions they have captured, but they are no longer bound to the comet. When all the old magnetic fields have been ejected from the atmosphere, the newly created ions in the atmosphere have no magnetic connection to the plasma tail. The tail appears to disconnect from the coma. The atmosphere immediately supplies ions for a new tail having the polarity of the new magnetic field sector.

lence as it encountered the bow-shock region. The mission also produced two major surprises: not all the expected signs of bow shock were observed and some unexpected high-energy ions were detected by two experiments. The observations of the bow shock have led some workers to speculate that comet plasma may under certain conditions gently slow the solar wind instead of decelerating and shocking it more rapidly, as was predicted in the classical model. Detailed study of the ICE mission's data should help to explain these findings and others.

The launchings of the Halley-bound spacecraft are timed to minimize the launch energy necessary to reach the comet. Aiming for the comet when it is passing through the earth's orbital path takes advantage of the momentum of the earth's orbital motion, which transfers momentum to the spacecraft and thus effectively boosts their energy during the launching. The most daring of these missions will be carried out by *Giotto,* which will encounter a large concentration of cometary dust as it flies within 500 kilometers of the nucleus. The leading surface of *Giotto* is protected by a double shield. The first shield should fragment and slow the incoming particles; the second shield is intended to prevent the particles from penetrating the main body of the spacecraft.

In order to approach closely, *Giotto* will need data on the position of the nucleus. The U.S.S.R. will provide this information through an agreement called Pathfinder. Cameras on the Soviet probes *Vega-1* and *Vega-2* will record the position of the nucleus, which will be relayed under the auspices of an Inter-Agency Consultative Group composed of the four flight agencies and the International Halley Watch. *Vega-1* will fly within 10,000 kilometers of the nucleus; *Vega-2* may be targeted closer to it. The Vega and *Giotto* missions should supply the first photographs of the nucleus of a comet. The Japanese probes, *Sakigake* and *Suisei,* will provide additional information about Halley while staying far enough away from the comet to avoid the hazard of cometary debris. *Sakigake* will monitor the solar wind streaming onto the comet from a position about one million kilometers or more upstream of the comet. *Suisei* will penetrate the comet's atmosphere at a distance of 200,000 kilometers and obtain data on plasma density and velocity. It will also obtain ultraviolet images of the comet's atmosphere.

The massive efforts directed toward Halley and Giacobini-Zinner should dramatically advance understanding

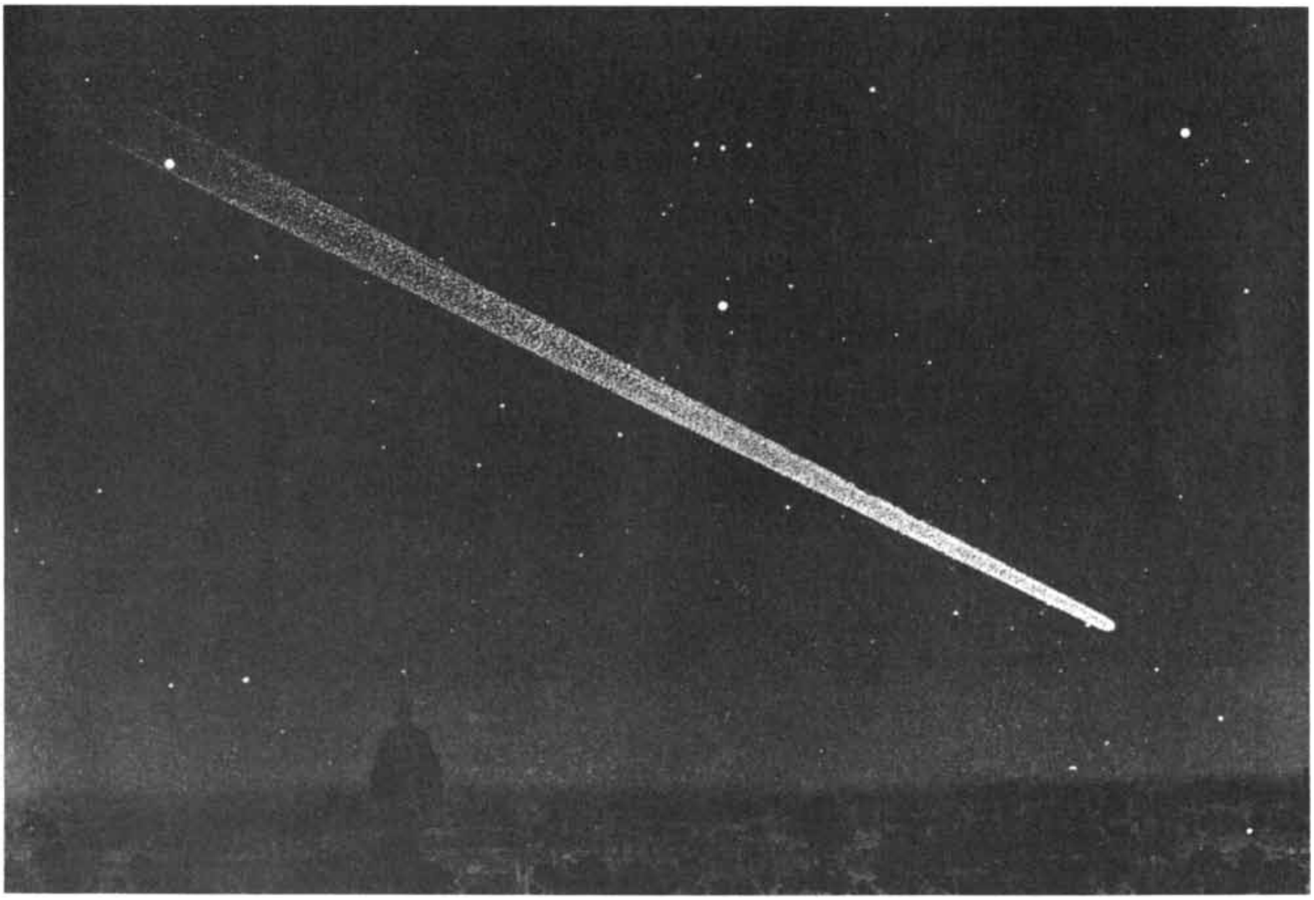


DISCONNECTION of Comet Halley's tail was photographed during its last visit, in 1910. Disconnections were first reported by Edward Emerson Barnard, who correctly hypothesized in 1899 that they were caused by some interaction with the interplanetary medium. He advocated photographing comets frequently to trace such changes. Halley's disconnection event was pieced together from different photographs made in 1910 on June 6 at the Yerkes Observatory (*top*) and at Honolulu (*middle*) and on June 7 at Beirut (*bottom*).

	GIACOBINI-ZINNER	HALLEY				
	NASA	U.S.S.R.		EUROPEAN SPACE AGENCY	JAPAN	
VEHICLE	ICE	*VEGA*-1	*VEGA*-2	*GIOTTO*	*SAKIGAKE*	*SUISEI*
LAUNCH DATA	AUGUST 1978 (DECEMBER 1983)	DECEMBER 1984	DECEMBER 1984	JULY 1985	JANUARY 1985	AUGUST 1985
ENCOUNTER: DATE	SEPTEMBER 11, 1985	MARCH 6, 1986	MARCH 9, 1986	MARCH 13, 1986	MARCH 8, 1986	MARCH 7, 1986
DISTANCE TO NUCLEUS (KILOMETERS)	8,000 (TAILWARD)	10,000 (SUNWARD)	3,000 (SUNWARD)	500 (SUNWARD)	4×10^6 (SUNWARD)	2×10^5 (SUNWARD)
FLYBY SPEED (KILOMETERS)	21	80	77	69	74	75
DISTANCE FROM SUN (KILOMETERS)	1.55×10^8	1.185×10^8	1.245×10^8	1.335×10^8	1.215×10^8	1.20×10^8
DISTANCE FROM EARTH (KILOMETERS)	7.05×10^7	1.74×10^8	1.635×10^8	1.470×10^8	1.665×10^8	1.71×10^8
COMET EXPERIMENTS	8	13	13	10	3	1

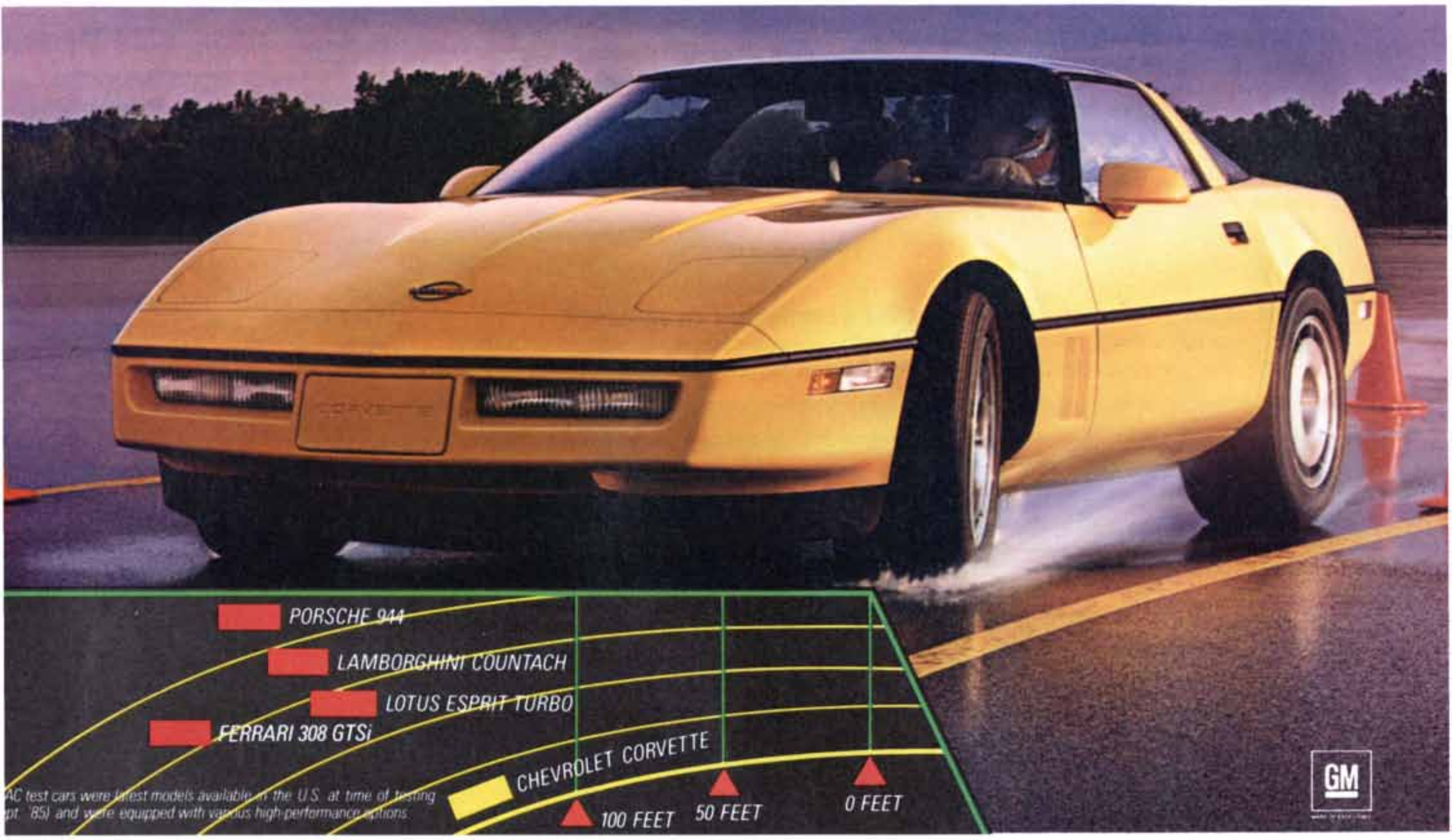
SPACECRAFT ENCOUNTERS with two important comets are listed in this table. The encounter with Giacobini-Zinner provided valuable data on the structure of comet tails when the National Aeronautics and Space Administration's *International Cometary Explorer* (ICE) passed through the plasma tail last September 11. (ICE was launched into orbit between the earth and the sun in August, 1978, and was retargeted in December, 1983, for the encounter with Giacobini-Zinner.) The missions that will intercept Halley in March should provide a more detailed understanding of the properties of comet tails by supplying information from the sunward side.

of cometary physics, but many questions, including those that emerge from the new data, will remain. The direct exploratory missions provide information based on a series of snapshots taken along single trajectory lines. Global data recording cometary changes over time will be needed to deepen understanding of comet tails. NASA's Comet Rendezvous and Asteroid Flyby mission, planned for launching in the early 1990's, should provide this important information. The vehicle is expected to approach Comet Wild II in 1995 and record valuable data as it travels with the comet along its orbital path for approximately two and a half years. If the mission is successful, it will mark the next logical step in attempts to explore and understand the nature of comets.



GREAT COMET OF 1843 dominated the northern skies. It was depicted in this lithograph as it was seen over Paris on March 19. The comet, perhaps the brightest one of the past two centuries, has not been seen since the 1843 appearance. The tail was about 300 million kilometers long, or longer than the distance from the sun to the orbit of Mars. The artist recorded a bright tail that is probably of the dust variety. Although there must have been a plasma tail, it was probably embedded in the dust tail and was too faint to be observed.

Anti-lock braking power you can grade on a curve.



You're driving 55 MPH on a rain-slick curve. Suddenly the unexpected: You stand on the brake pedal and steer to stay in your lane. You might expect Europe's most exotic cars to handle such a crisis effortlessly. Yet for all its awesome straight-line braking ability, Ferrari 308 GTSi failed to negotiate a 150-foot radius curve at maximum braking in USAC-certified testing. Lamborghini Countach failed. Lotus Esprit Turbo failed. Porsche 944 failed. Only the 1986 Corvette demonstrated the ability to steer and stop in these conditions at the same time. Only Corvette made the turn while coming to a controlled stop. When conditions turn foul, Corvette's new computerized Bosch ABS II anti-lock braking system is designed to help improve a driver's ability to simultaneously brake and steer out of trouble.

Why does the Corvette feature the world's most advanced braking technology? Because a world-class champion should give you the edge in an emergency.

Corvette. A world-class champion.

TODAY'S CHEVROLET Live it!

Artificial Intelligence:

Summary:

GTE research in Artificial Intelligence has produced exciting results in several areas of knowledge-based systems. In addition, research is under way to teach computers to learn by themselves, much as humans do.

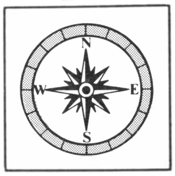
It's extremely tedious and difficult to teach a computer to respond to specific problems in an intelligent way. Despite this, GTE has created several workable systems, which are in the field now.

But training a computer to respond to analogous or unexpected situations —teaching it to *learn*—is a very different challenge. And this is one of our long-range programs in AI research.

The ultimate brain-picking.

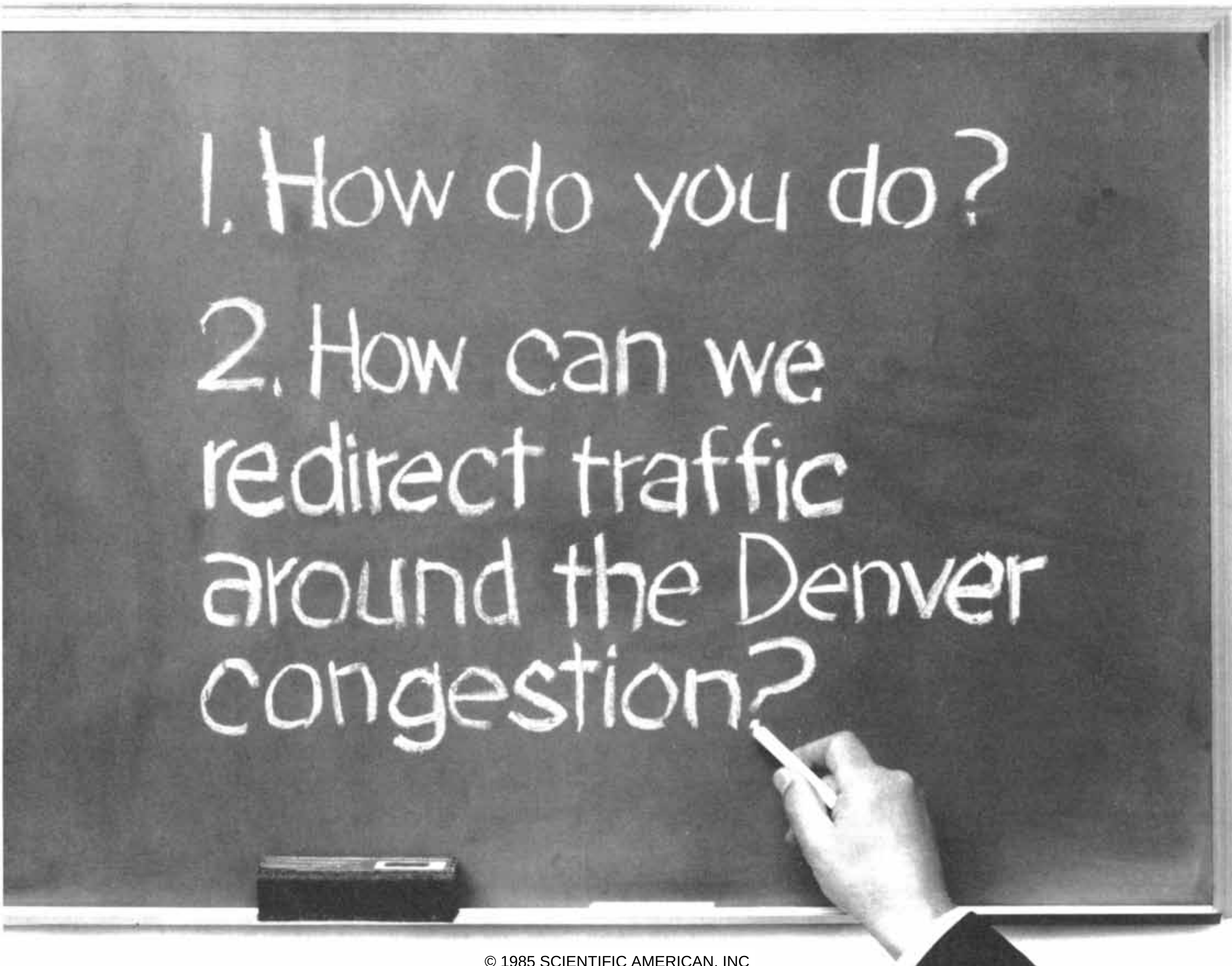
The Expert-Systems version of AI is literally the result of programming the experiences of experts into a computer.

Once these human reasoning processes have been codified, the computer has the information it needs to mimic the experts' responses to an immense variety of problems.



COMPASS (Central Office Maintenance Printout Analysis and Suggestion System) is an Expert System we devised for telecommunications. It is being phased into field use to monitor switch per-

Which of these questions is easier for a computer to answer?
The apparently simple greeting is loaded with semantic traps. On the other hand, the complex question relating to traffic redirection can be tackled by Expert Systems.



reality and promise.

formance, diagnose problems and recommend corrective actions in large communications networks.

Say hello to FRED.

The proliferation of databases and their integration in a large information system is increasing computer uses. Increasing user friendliness is becoming all the more necessary for computers to be used by less skilled operators.

GTE has developed FRED (Front End for Databases), which enables operators to frame information requests from multiple databases, in plain English. FRED untangles the request, breaks it into segments the computer understands—and provides the data, in plain English.

For its next evolution, we are teaching FRED to approach several databases at once (rather than one at a time), and put all relevant data into a single reply.

The nature of thought.

Another of our AI research directions is basic, long-range research into ways of teaching computers to learn for themselves, through experience and/or inference.

This involves research into such an area as the way children learn, as well as deep studies into the nature of decision-making itself.

Much remains to be discovered, of course—but the promise of true machine learning is perhaps the most exciting in the entire computer field.

The outcome of these projects—some near-term, some more in the future—will be to make the computer a far more useful and friendly tool for an immense variety of industrial and human problems.

The box lists some of the pertinent papers GTE personnel have published on various aspects of Artificial Intelligence. For any of these, you are invited to write to GTE Marketing Services Center, Department AI, 70 Empire Drive, West Seneca, NY 14224. Or call 1-800-828-7280 (in New York State 1-800-462-1075).

Pertinent Papers

"COMPASS: An Expert System for Telephone Switch Maintenance," S.K. Goyal, D.S. Prerau, A.V. Lemmon, A.S. Gunderson and R.E. Reinke, Expert Systems: The International Journal of Knowledge Engineering, *Vol. 2, No. 3, August 1985. pp 112-126.*

"Selection of an Appropriate Domain for an Expert System," D.S. Prerau, AI Magazine, *Vol. 4, No. 2, Summer 1985; pp 26-30.*

"A Natural Language Interface for Medical Information Retrieval," G. Jakobson, C. LaFond, E. Nyberg and V. Shaked. Third AASMI Joint National Congress on Computer Applications in Medicine, May 1984, San Francisco, California. pp 405-409.

Computer Experience and Cognitive Development, *R.W. Lawler. Ellis Horwood Limited, Chichester, U.K. (1985). (Summary of book.)*

"The Learning of World Models by Connectionist Networks," R.S. Sutton and B. Pinette. Proceedings of the Seventh Annual Conference on Cognitive Science Society, 54 (August 1985).

"Training and Tracking in Robotics," O.G. Selfridge, R.S. Sutton, A.G. Barto. Proceedings of the Ninth International Joint Conference on Artificial Intelligence, 670 (August 1985).

GTE

Signing Off

Managers of the Strategic Defense Initiative (SDI) program (popularly known as "Star Wars") have recognized from the start that developing a shield against enemy ballistic missiles must rely heavily on fundamental science done in university laboratories. They have established an Office of Innovative Science and Technology that plans to spend some $600 million on research over the next five years, much of it at universities. The funds would support what the office's director, James A. Ionson, calls goal-oriented basic research in fields relevant to the program, such as lasers, computer science, materials science and plasma physics.

The office's invitation for research proposals, announced last March, drew a flood of submissions. By June they numbered about 3,000, and workers at about 60 universities had already been given research contracts funded by the SDI. Yet many of the people whose skills could be important to the Star Wars program are publicly shunning it. By November, on some 90 campuses a total of more than 1,600 faculty members and almost 1,200 graduate students in relevant disciplines had signed statements condemning Star Wars as unworkable and strategically unwise. The signers pledged that they would neither solicit nor accept money from the program.

According to John B. Kogut, professor of physics at the University of Illinois at Urbana-Champaign and a coordinator of the boycott there, a total of 58 percent of the faculty members in 14 top physics departments have signed such pledges; the institutions include Princeton University (where some 75 of the physics faculty signed), the University of Illinois (74 percent), Cornell University (69 percent), the California Institute of Technology (60 percent), Harvard University (49 percent) and the Massachusetts Institute of Technology (38 percent). Although the campaign, initiated last spring at Illinois and Cornell, first took hold among physicists, it has gained significant support in other disciplines. Zellman Warhaft, associate professor of engineering at Cornell, says the pledge circulating there carries the signatures of 75 percent of the astronomy faculty, 46 percent of the chemistry faculty and 44 percent of the engineering faculty.

The signers are motivated not only by an aversion to Star Wars itself but also by concerns about the effects of SDI funding on academic freedom. Ionson maintains that SDI funds for unclassified research at universities come with no strings attached. Ambiguities in recent policy statements by the SDI and the White House worry the boycott organizers, however. They are afraid that classification might be imposed on an initially unclassified research project as the work progresses toward militarily significant results. Even if the research were to remain unclassified, organizers think the Government might seek to limit the communication of results by such means as export controls, which would restrict access to the information by foreign nationals. Ionson calls such fears groundless, saying his office will try to ensure that any unclassified work funded by the SDI on campuses is not subject to further controls.

Another objection stems from the view that the SDI money, together with other recent increases in funding of research by the Department of Defense, could make much of university science dependent on a single source of funding—to its peril, should the political climate change. Signers of the pledges also express concern at the prospect of a vast diversion of talent away from research they say the U.S. needs to pursue if it is to maintain its commercial and scientific strength.

The boycott is meant to be a frank political statement: its organizers want to show the public and Congress that much of the scientific community is opposed to the SDI and doubts its feasibility. They fear that the infusion of SDI research money into universities could raise the level of support for Star Wars on campuses by creating a dependent, and therefore silent or supportive, constituency for the program. Kogut points out that several of the fields, such as optics, for which ample support will be available are ones that were starved for funds until recently.

Whatever its broad political effects turn out to be, the campaign may already have succeeded in reducing the number of applicants for SDI funds. Warhaft thinks that at Cornell, where faculty have submitted only about a dozen preliminary proposals ("white papers") to the Office of Innovative Science and Technology, the "discussion and heated debate" engendered by the drive has caused many investigators to reconsider their views of the program. "In my opinion," Warhaft says, "the very good schools are not going to take part" in SDI research.

Ionson insists he has more good proposals than he can fund; he admits the boycott will reduce scientific input to the SDI, but only to a negligible degree. His chief worry, he maintains, is that such boycotts of fundamental, unclassified research are themselves an infringement on academic freedom because they generate political pressure against taking part in research.

High-Tech Big Brother

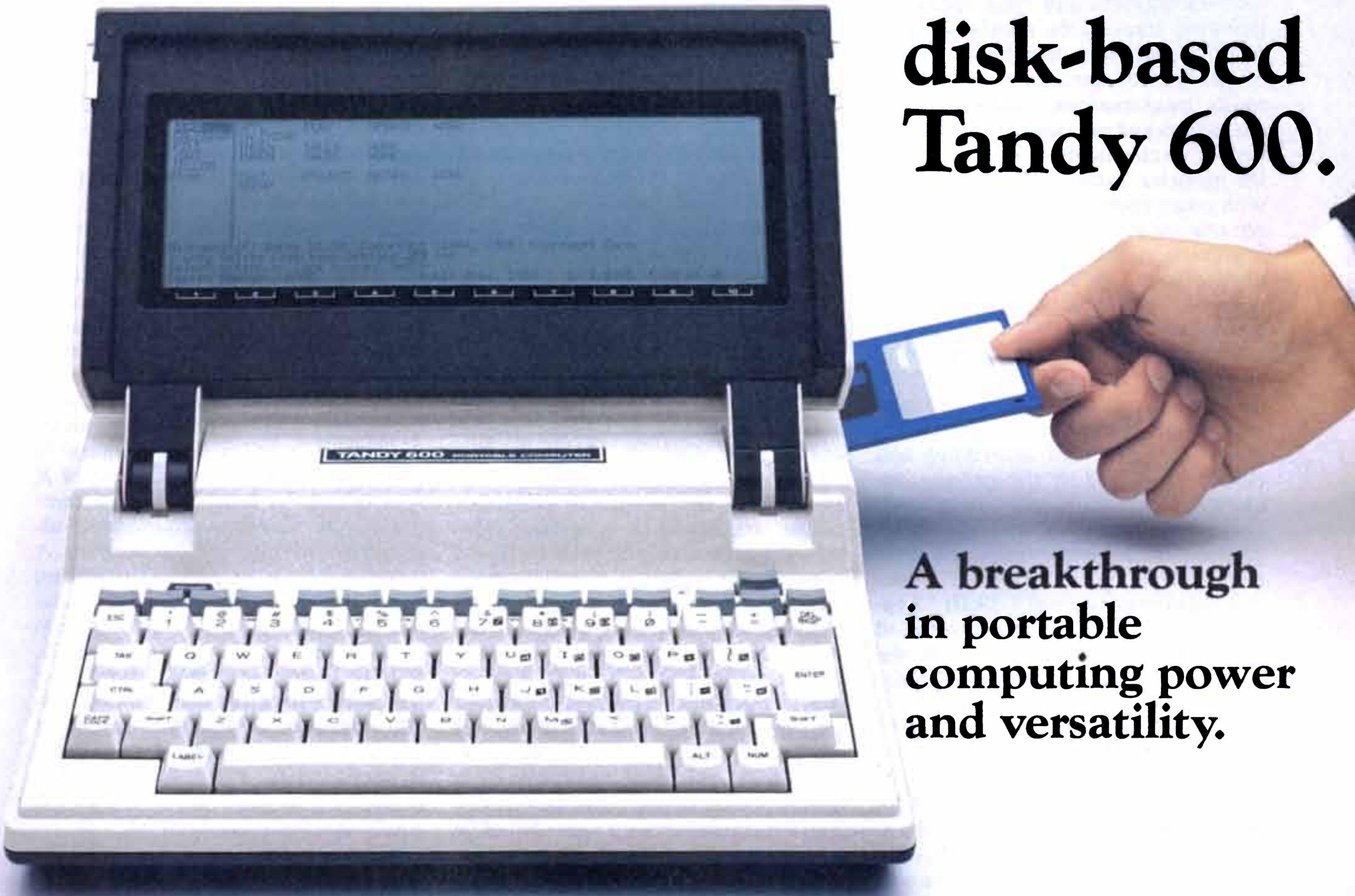
The swift advance of communication technology has outpaced the laws protecting individual privacy, according to a report issued by the congressional Office of Technology Assessment. Congress should consider the situation, the OTA said, with a view to balancing the individual's right to privacy and the Government's responsibility for maintaining public safety.

As recently as 20 years ago, the report points out, electronic surveillance was limited to the tapping of telephones and the installation of that staple of espionage and police fiction, the concealed microphone. Since then "there has been a virtual revolution in the technology relevant to electronic surveillance." Now devices are available that will keep track of a person's movements, actions, written messages and speech—even if the person is at home with the doors locked and the curtains drawn. For example, a sensitive microphone pointed at the outside of a window can pick up a conversation in the room and record it; moreover, the investigator does not have to listen to the entire recording but can put it through a computer programmed to pick out certain topics or names.

With other techniques a person's activities at home or at work can be recorded on videotape; his movements outdoors at night can be tracked by cameras that function in dim light; his travel by automobile can be followed if a "beeper" emitting an electronic signal has been attached to the car; the numbers he dials on his telephone can be recorded, and his bank and credit records stored in computers can be examined electronically.

The OTA found that aside from the Central Intelligence Agency, the Defense Intelligence Agency and the National Security Agency (the study did not encompass classified information), 35 out of 142 Federal agencies surveyed are engaged in or planning to engage in electronic surveillance. Fa-

The new disk-based Tandy 600.



A breakthrough in portable computing power and versatility.

Finally—a portable that matches the performance of a desktop computer. Our new Tandy 600 features a 16-bit microprocessor, an 80-character by 16-line display, a built-in 3½" disk drive that stores 360K of data and 32K RAM (expandable to 224K*).

Five Resident Applications

With Tandy 600's larger display and memory, Multiplan spreadsheet analysis can hold more information. Word processing is as easy as using MS-Word. You'll have quicker access to documents and better storage with the built-in disk drive. File is an electronic data base for names and addresses, expenses, client billing, inventory and more. And you can keep a large number of different files on the pocket size diskettes. With Telecom and the Tandy 600's built-in modem, you're able to communicate with other computers over phone lines. Telecom will even dial the phone number of anyone listed in the File program. Calendar helps you keep track of daily tasks and activities.

The easy-to-learn resident System Manager lets you run each application, exchange information between applications and manage the files created. You can even set it to turn itself on for unattended operations. And you can add BASIC language ROM (26-3904, $129.95) to write your own programs.

Tandy...Clearly Superior™

If you've been looking at portable computers, compare them with the Tandy 600 (26-3901, $1599). You'll be amazed at what it can do! Visit a Radio Shack Computer Center today.

Available at over 1200 Radio Shack Computer Centers and at participating Radio Shack stores and dealers.

Radio Shack®
COMPUTER CENTERS
A DIVISION OF TANDY CORPORATION

Send Me a Free Tandy 600 Brochure.

Radio Shack, Dept. 86-A-569
300 One Tandy Center
Fort Worth, TX 76102

Name ______
Company ______
______
Address ______
City ______
State ______
Zip ______
Phone ______

*RAM upgrades are in banks of 96K each (26-3910, $399.95). Prices apply at Radio Shack Computer Centers and participating stores and dealers. Multiplan and MS-Word/TM Microsoft Corp.

vorite devices include closed-circuit television, night-vision systems, miniature transmitters and "pen registers" that keep track of the numbers dialed from a particular telephone. Procedures include interception of cellular-radio transmissions, monitoring of computers and monitoring or interception of electronic mail. In many cases the agencies initiate surveillance only with court approval, but the law does not require that. The major applicable law, Title III of the Omnibus Crime Control and Safe Streets Act of 1968, protects only oral communications transmitted by wire.

The OTA said Congress has several choices. One is to do nothing, in which case policy would be made by administrators and court decisions. The effect would be "continued uncertainty and confusion regarding the privacy accorded phone calls, electronic mail, data communication and the like." Another option is to "bring new electronic technologies and services clearly within the purview of Title III." Legislation to do that has been introduced in the Senate and the House as the Electronic Communication Privacy Act of 1985.

State of Controversy

For a little more than a year crystallographers and solid-state physicists have been faced with the possibility that there is a previously unknown "quasicrystalline" form of ordered matter. Unlike a crystal, such matter cannot be described in terms of a repeated lattice. Unlike an amorphous solid, on the other hand, it does have a long-range orientational order.

The patterns of speculation over the nature of this new state of matter have been disturbed by Linus Pauling. He says the data on which the hypothesis is based can be explained, without proposing such a quasicrystalline state, by a well-known crystallographic phenomenon called directed twinning, in which several identical crystals grow along shared faces.

The original evidence for the existence of quasicrystals was an electron-diffraction pattern published in November, 1984, in *Physical Review Letters* by Dany Shechtman of the Israel Institute of Technology–Technion and colleagues at the Technion, at the French National Center for Scientific Research and at the U.S. National Bureau of Standards (NBS). The pattern, which was made by aiming a beam of electrons at a rapidly cooled alloy of aluminum and manganese, had fivefold rotational symmetry; in other words, when the pattern was rotated by one-fifth of a full circle, it looked the same as an unrotated pattern. This result was surprising because an elementary theorem of crystallography says that, just as it is impossible to tile a plane with five-sided figures, so a crystal, which consists of one elementary unit repeated indefinitely, cannot have fivefold rotational symmetry. The discovery of a material having such an unusual property led a large number of investigators to search for other quasicrystals and to propose structures that would make such materials possible.

The atoms of a normal crystal are arranged regularly on parallel planes. The distances between adjacent planes are determined by simple periodic functions. Through each atom an infinite number of different planes can be drawn, each of which belongs to a different infinite set of parallel planes. In most of the proposed quasicrystalline structures, as in normal crystals, the atoms can be seen to lie on parallel planes. Sets of parallel planes intersect each other at angles of 108 degrees, which is the angle between adjacent sides of a pentagon, and so the planes themselves fall into a pattern that has fivefold symmetry: if the quasicrystal is rotated by one-fifth of a full circle, each set of planes takes on an orientation formerly occupied by a different set. For such symmetry to be possible, however, the atoms cannot be arranged on the planes in a regular way, and the distances between adjacent parallel planes cannot vary periodically. It is therefore not clear how to describe a quasicrystal in a way that indicates the positions of every atom.

John W. Cahn of the NBS, one of Shechtman's coauthors, points out that "in crystallography you can say 'Here is what the unit cell looks like; repeat that and you have the crystal.' But how much must one say about a quasicrystalline structure in order to define it uniquely?"

Pauling contended, in a talk delivered in August to a session on "crystallography applied and misapplied" at a meeting of the American Crystallographic Association, that it is not necessary to consider such exotic structures. In his talk, a revised version of which was later published in *Nature,* he suggested that the same fivefold symmetry could be caused by twinning: by a growing together of 20 identical crystals from a single seed.

In Pauling's model each crystal grows as a compressed trigonal pyramid: a squat pyramid with a triangular base. The 20 pyramids grow with their tips pointing in toward the seed. Each crystal fits together with the neighboring pyramids, and so as they grow their bases fit together as the faces of an icosahedron (a 20-faced polyhedron each of whose faces is a triangle). The icosahedron has inherent fivefold symmetry: five triangular faces meet at each of its 12 vertexes. Twenty crystals twinned in such a way would indeed produce a diffraction pattern that has fivefold symmetry: each of the individual crystals would give rise to its own diffraction pattern and the superposition of the patterns would have the same symmetries as an icosahedron.

Pauling has determined the lattice structure of an aluminum-manganese crystal that could grow in such a conformation. The crystal, which resembles crystal structures first discovered in the early 1950's, has a cubic unit cell with an edge length of 26.7 angstrom units. He checked his model against an X-ray powder pattern (a pattern made by passing a beam of X rays through a finely powdered sample of the aluminum-manganese alloy) sent to him by Shechtman. A powder pattern provides information about the lattice structure of a crystal but not about the relative orientation of different crystals within a sample. Pauling is convinced that he has found the correct structure because his model, constructed without reference to the pattern, nonetheless matched it very closely. "The probability of chance agreement," he estimates, "is surely less than one in 1,000."

Investigators trying to determine the possible structure of quasicrystals contend, however, that the hypothesis of twinning does not fit all the available data. Indeed, Cahn recalls that "when Shechtman first showed me the diffraction pattern, I said, 'Go away, Dany, that's just twinning.' We then spent more than two years, before we published, ruling out just such conventional crystallographic explanations." Cahn and other solid-state physicists, including David R. Nelson of Harvard University, insist that results obtained by the techniques called microdiffraction and high-resolution microscopy eliminate Pauling's hypothesis from consideration.

In microdiffraction a diffraction pattern is made of an area as small as 100 angstroms in diameter. Such patterns have been made from many locations on the same sample of aluminum-manganese alloy; every microdiffraction pattern made by an electron beam aimed at the appropriate angle, regardless of position within the sample, shows fivefold symmetry. Cahn argues that if the material were composed of twins, not all the microdiffraction patterns would include every one of the twinned crystals; some would therefore consist of only the diffraction patterns due to a few of the crystals and would not display fivefold symmetry.

High-resolution microscopy can provide an image of a feature as small as two angstroms across. A feature such as Pauling's unit cube, with a characteristic length of 26.7 angstroms, should surely be recognizable. None of the high-resolution micrographs made of the alloy shows either unit cells or the pattern of growth that is characteristic of twinned crystals. Pauling is readying for publication his own analysis of the high-resolution micrographs.

Universal Transition

Could a tabletop experiment shed light on the birth of galaxies? It very well might, according to Wojciech H. Zurek, a theoretical physicist working at the Los Alamos National Laboratory.

Zurek bases his bold proposal on the universal behavior of matter near phase transitions. A phase transition takes place when a material changes from one state to another. Perhaps the best-known phase transition is the melting of ice to water. Remarkably, the laws of physics that describe this common example also apply to phase transitions taking place at temperatures in a range from near absolute zero (zero degrees Kelvin, or minus 273 degrees Celsius) to 10^{28} degrees K. The extremes correspond respectively to the approximate temperatures at which helium liquefies and at which the universe was born.

Writing in *Nature,* Zurek proposes that phase transitions in liquid helium can serve as models of phase transitions in the universe soon after its birth. Specifically, Zurek's proposal focuses on the transition from normal liquid helium to superfluid. The properties of the superfluid state are exotic indeed. As a superfluid, for instance, liquid helium seems to flow without friction. In addition its thermal conductivity, which is a million times greater than that of the normal fluid, is better than the conductivity of any metal. Moreover, quantized vortex lines, or long-lived whirlpools, are known to exist in the superfluid.

Such vortex lines are analogous to cosmic strings, or loops of very dense material, that correspond in a mathematical sense to stable solutions of certain fundamental equations of high-energy physics. Phase transitions in the early universe are thought to have resulted in the copious production of these postulated but as yet undetected entities. Cosmic strings may be as long as a million trillion (10^{18}) miles. (They should not be confused with superstrings, which, it has been proposed, may make up elementary particles such as quarks and electrons.)

About a decade ago Yakov B. Zel'dovich of Moscow University suggested that loops of cosmic strings having lengths on the order of a few hundred trillion miles could have precipitated the gravitational condensation of galaxies. The Zel'dovich proposal has been developed by a number of other investigators. In particular, T. W. B. Kibble of the Imperial College of Science and Technology in London has proposed scenarios for string formation in cosmological phase transitions that enabled him to estimate the initial density of strings. Unfortunately his prediction cannot be verified directly.

What Zurek offers is a way out: an analog experiment. By determining the density of vortex lines formed during a phase transition to superfluid helium one might infer the density of cosmic strings formed in the early universe. The experiment he puts forward is simple. A ringlike tube containing liquid helium in its normal state at about two degrees K. is subjected to high pressure. By rapidly removing the pressure a transition to the superfluid state is induced. If Kibble's scenario is correct, the forming of vortex lines—the analogues of cosmic strings—should accompany the transition to the bulk superfluid. According to the same scenario, the superfluid should also begin to circulate around the tube. Both the density of vortex lines and the rate of flow in the tube are readily measurable; either would provide a quantitative test of the string-formation scenario.

Cosmological considerations aside, Zurek's experiment is interesting in itself as a study of superfluids. Although quantized vortex lines are not new entities, the possibility of inducing vorticity by a rapid phase transition is novel. Several research groups are considering plans to carry out the Zurek experiment.

Landslide

A thick shell of the earth, perhaps even including part of its very core, has slipped erratically around the earth's center in the course of millions of years. The evidence for this proposition is presented in *Journal of Geophysical Research* by Jean A. Andrews of Columbia University's Lamont-Doherty Geological Observatory. Andrews' findings, startling in their own right, also provide an answer to the puzzling question of the earth's wandering poles.

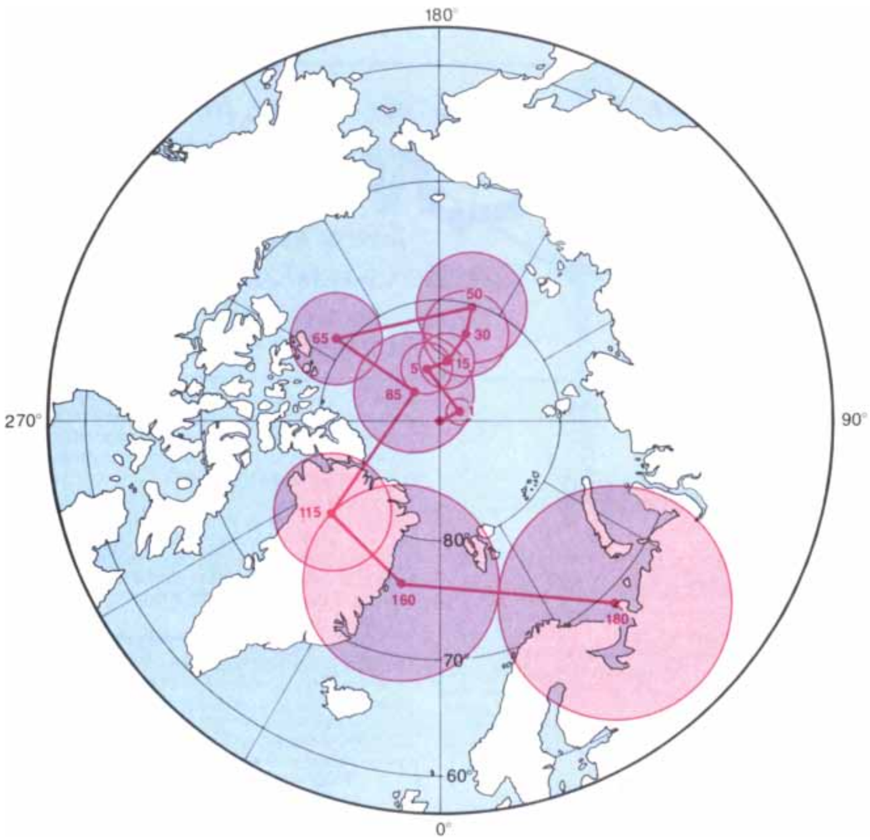
For many years it has been known that the geomagnetic poles appear to have wandered over the surface of the earth. The evidence for this seeming motion was accumulated through studies of the vestigial paleomagnetism exhibited by certain rocks: if the age and original position of such a rock is known, the orientation of the geomagnetic field lines when the rock was formed can be reconstructed and the old position of the poles can be deduced.

The observed changes in the positions of the ancient magnetic poles were originally thought to be the result of sporadic geomagnetic instabilities that caused the magnetic poles to fluctuate about the fixed geographic poles (whose locations coincide with the stable spin axis of the earth). It soon became clear, however, that the distances traveled by the paleomagnetic poles were greater than could be allowed by the prevailing theory of the earth's magnetism. Moreover, the apparent paths traced by the wandering magnetic poles varied according to the continent where the data were collected.

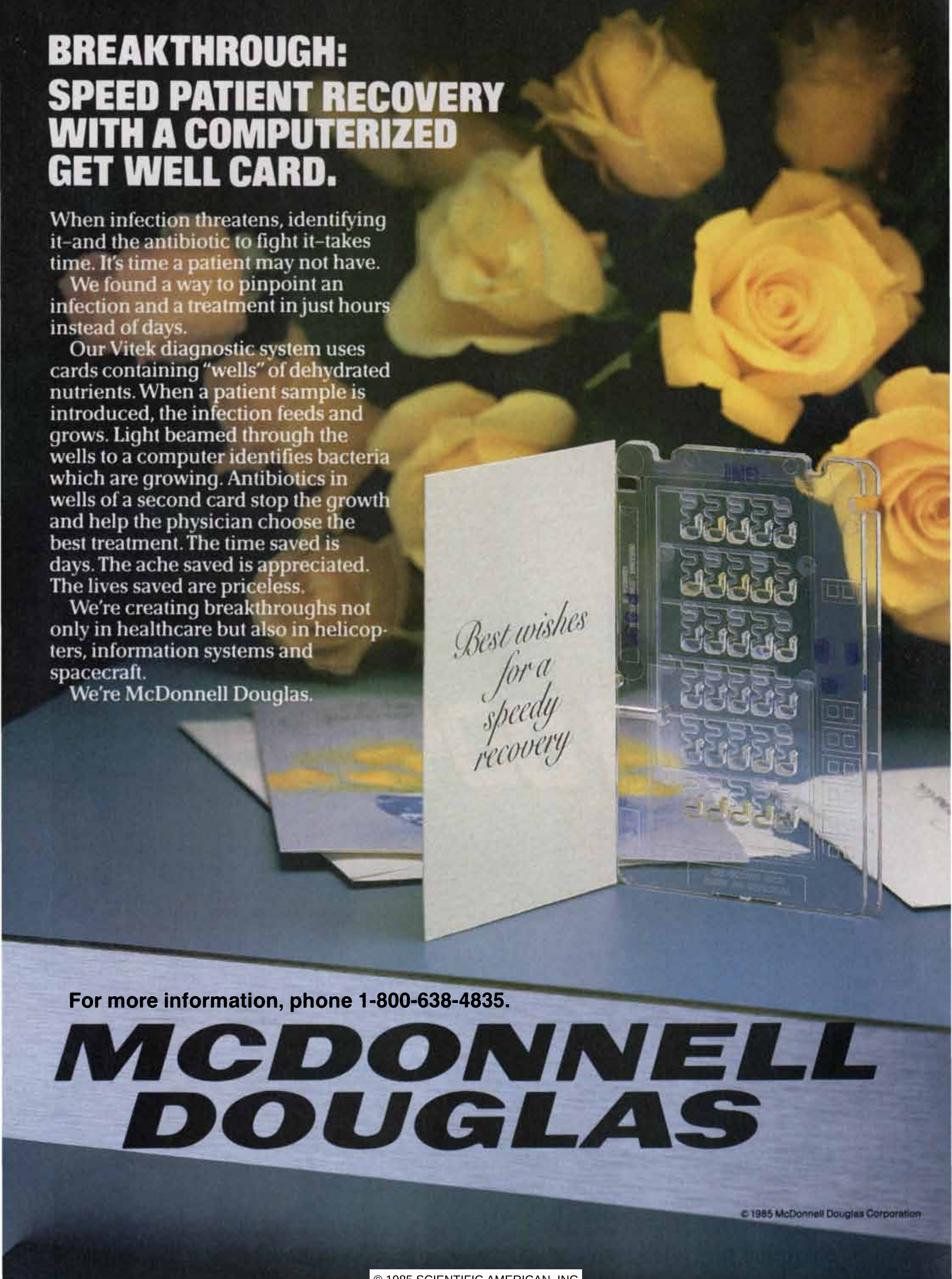
For some time these discrepancies seemed to be explained by the theory of plate tectonics: the drift of continents on great plates of the earth's solid outer layer, the lithosphere, had jumbled the paleomagnetic data by rearranging the positions of the rocks from which the data had been collected. The apparently excessive distances between past magnetic and present geographic poles could therefore be viewed as the combined result of plate motions and geomagnetic instabilities. Yet calculations of ancient pole locations that took these two factors into account still left a substantial amount of the polar wandering unexplained.

Because the geomagnetic poles are linked to the geographic ones and could not have wandered freely over the surface of the earth, Andrews reasoned that the entire surface of the earth must have wandered over the poles instead: a major part of the earth's thick plastic mantle, which underlies the lithosphere, must have shifted about the center of the earth, carrying the entire lithosphere with it. Such a large-scale motion of the mantle and lithosphere, when added to the relative motions of the lithospheric plates and the vagrant motion of the geomagnetic poles, could fully explain polar wandering.

To prove this hypothesis Andrews had to select well-defined points in the earth's mantle from which she could measure any change in position with



APPARENT WANDERING of the North Pole is due to slippage of the mantle and lithosphere around the center of the earth; the pole has actually remained stationary in space. The endpoints of the line segments designate positions of the pole on the surface of the earth at particular times (millions of years ago) in the past. Each point is encircled by its range of locational error. Landmasses are shown at their present latitudes and longitudes.


BREAKTHROUGH: SPEED PATIENT RECOVERY WITH A COMPUTERIZED GET WELL CARD.
When infection threatens, identifying it–and the antibiotic to fight it–takes time. It's time a patient may not have.
We found a way to pinpoint an infection and a treatment in just hours instead of days.
Our Vitek diagnostic system uses cards containing "wells" of dehydrated nutrients. When a patient sample is introduced, the infection feeds and grows. Light beamed through the wells to a computer identifies bacteria which are growing. Antibiotics in wells of a second card stop the growth and help the physician choose the best treatment. The time saved is days. The ache saved is appreciated. The lives saved are priceless.
We're creating breakthroughs not only in healthcare but also in helicopters, information systems and spacecraft.
We're McDonnell Douglas.
Best wishes for a speedy recovery
For more information, phone 1-800-638-4835.
MCDONNELL DOUGLAS
© 1985 McDonnell Douglas Corporation

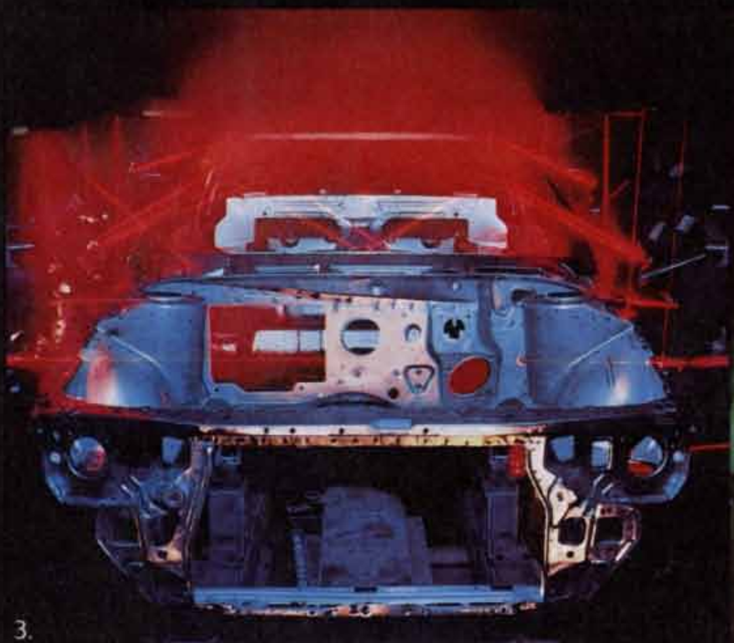
DRIVING AMERICA DOWN ROADS WE'VE NEVER BEEN BEFORE.

General Motors is on a journey to a far-off place just around the corner, the 21st Century. With the help of its thousands of scientists, designers and engineers, GM is embarking on an odyssey into the unknown. Roads paved with scientific and technological wonders that might seem like science fiction. But at GM, they're reality. Now. Lasers that probe body structure as no man's eyes ever could. Electronic navigation systems that tell you where you are and how to get where you're going. But the road doesn't stop there. The 21st Century lies dead ahead. GM is leading the way there by adding new dimensions of electronic wizardry and computer expertise to its vast worldwide resources. Today's GM is evolving into a GM the future will demand. Bolder. Daring. More responsive. Speaking a new language. Venturing in new directions.


1.

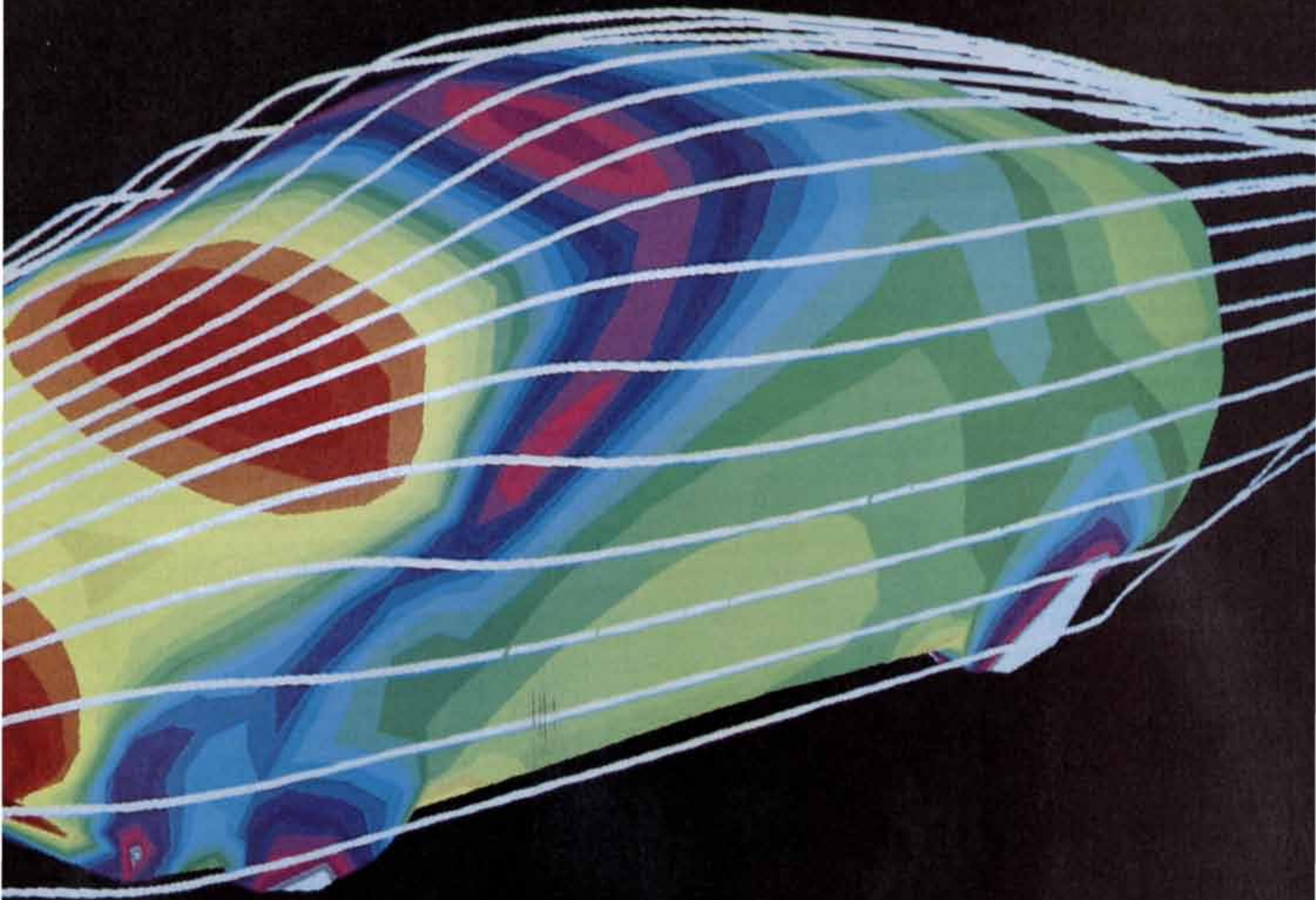

2.

THE GM ODYSSEY: SCI


3.


4.


5.



6.

respect to an "average" paleomagnetic pole, corrected for the random geomagnetic variations. Such points exist. They are hot spots: upwellings of molten rock that originate in fixed regions of the mantle. Their current and historic positions can be identified by geologic features such as the Hawaiian Islands, which are permanently impressed on the surface of the earth as a crustal plate drifts over a hot spot.

The mean positions of past magnetic poles, when they are measured against past hot-spot positions, indicate that hot spots, and therefore the earth's mantle and lithosphere, have shifted in unison: about 180 million years ago the North Pole was displaced some 22 degrees in latitude from its present position.

What could cause such slippage of the earth's outer layers? Andrews believes the likeliest explanation is one that Thomas Gold of Cornell University proposed 30 years ago. He pointed out that if the mantle of the earth is sufficiently plastic, the presence of a region of high density in the mantle could cause just such large-scale motion of the surface of the earth. Applying the equations that govern rotating bodies, he showed that an extraneous mass placed on a rotating, deformable sphere will cause the sphere to redistribute its mass in such a way as to achieve the lowest possible energy state. For a plastic sphere the minimum-energy state is reached by shifting the entire surface so that the extraneous mass ultimately finds itself on the sphere's equator. The mantle could similarly shift, as a whole, around the center of the earth in order to move a region of high density toward the Equator.

Regardless of its cause, the identification of this shifting of the earth's mass about its center is likely to have an impact in many areas of geology. A number of conclusions in paleogeography and paleoclimatology, which depend critically on the correct determination of ancient latitudes, may have to be revised.

Tender Sediments

The construction of dams for irrigation, hydroelectric power and flood control has greatly diminished the amount of sediment being delivered to the oceans by the major rivers of the U.S., according to the U.S. Geological Survey. Dams tend to trap sediment. Contrary to what one might think, the effects are mostly undesirable. Shorelines in the Mississippi delta are receding rapidly because ocean waves and tides can easily carry off the reduced amount of sediment deposited by the river. Rivers downstream from dams have a greater capacity for carrying solids and therefore tend to erode the land they flow through.

Robert H. Meade and Randolph S. Parker of the Geological Survey discuss the phenomenon in the second annual *National Water Summary.* They report that the Mississippi now discharges into the Gulf of Mexico less than half the sediment it discharged 35 years ago. In recent decades the Rio Grande's delivery of sediment to the Gulf of Mexico has dropped from 20 million tons per year to less than one million. Since the Hoover Dam began impounding water from the Colorado River in 1935, the amount of sediment delivered to the Gulf of Mexico has fallen from 150 million tons per year to 100,000 tons.

Philip Cohen, chief hydrologist of the Geological Survey, discussed the implications of the findings. "Just because certain rivers are carrying less sediment does not mean that we can relax efforts to control soil erosion," he said. "The amount of sediment delivered to the oceans by rivers is only 10 percent of the total soil eroded." Hence 90 percent of the eroded soil is not going immediately into streams and being carried to the sea. Instead it is being stored somewhere between the eroded site and the sea, mostly on hill slopes, behind dams, in stream valleys and on floodplains. "The water-quality implications of this large amount of sediment in storage are enormous. Because many of the toxic materials that travel in streams are attached tightly onto sediment particles, any accurate prediction of the fate of toxic substances in a stream will require an understanding of what is happening to the sediment."

Hazards

The tragic events in Colombia and Mexico have lent unexpected timeliness to a paper published recently in *Geology* by Joseph V. Smith of the University of Chicago. Smith argues that it is time to stop thinking of volcanic eruptions, earthquakes or even the impact of an asteroid or comet as being forever beyond human control. The suggestion is not new; for example, proposals for destroying an incoming asteroid before it can strike the earth have been advanced at least since the 1960's. Smith thinks there are two reasons the suggestion should now be taken more seriously. First, schemes for protecting people against natural disasters, although still highly speculative, no longer seem technically fantastic. Second, the stakes have been raised: recent research indicates that an asteroid impact or a massive volcanic eruption might cause not only local destruction but also deleterious changes in the global climate.

Smith proposes some specific protective measures. Most of the world's 700 or so dangerous volcanoes are not monitored systematically; Smith thinks that mass-produced seismometers, volcanic-gas detectors and tiltmeters (which could detect bulges that might precede an eruption) should be deployed on all of them. The data could be transmitted by satellite to research centers. With every volcano watched as closely as Mount St. Helens, scientists could issue warnings with greater confidence and precision, and government authorities would be more likely to order evacuations. Ultimately the goal would be to prevent eruptions or reduce their intensity. Smith considers two ideas worthy of further study: drilling into the magma chamber under a volcano in the hope of allowing the lava to escape harmlessly, and draining lakes situated in volcanic calderas to prevent the explosive interaction of underground lava with water.

The worst natural disaster would be the impact of a large asteroid like the one thought by many workers to have precipitated the extinction of the dinosaurs 65 million years ago. Such impacts are obviously rare, but Smith holds that an impact of a much smaller but still dangerous object should be expected in a populated area every 1,000 to 10,000 years. Hence he proposes a program to find and track comets and asteroids whose orbits intersect that of the earth. The program would entail the construction of at least 10 new optical telescopes, the launching of orbiting infrared and optical telescopes and regular flyby missions like those now under way to Halley's comet. On the basis of data from these missions workers would frame plans for deflecting or destroying an object headed for the earth, perhaps by detonating small nuclear bombs in its vicinity.

Some of the technology necessary to ward off asteroids would be similar to that envisioned in the Reagan Administration's Strategic Defense Initiative. Like the "Star Wars" plan, Smith's proposals (he also suggests stepped-up research on earthquake prediction and prevention) would be extremely expensive. The money would come from a worldwide agreement to reduce expenditures on nuclear weapons, initially by a few percent per year. It would be disbursed by international agencies. In Smith's view the cooperation needed to protect people against natural hazards could in itself be "a small but important catalyst toward attainment of a peaceful world."

Tailored Photosynthesis

A novel composite material that is part biological and part metallic shunts photosynthesis from its usual role, the synthesis of carbohydrates, to a role it never played in the natural world: the production of hydrogen, which in principle could supersede fossil fuels. The biological part of the composite material is thylakoid membrane, the structure in chloroplasts that captures the energy of sunlight for photosynthesis; the metallic part is platinum. Elias Greenbaum, who created the new composite at the Oak Ridge National Laboratory, describes his work as the extraction of a functional organization from a plant and the tailoring of that organization to serve human needs.

Greenbaum's novel composite preserves essentially intact the initial step in photosynthesis, a step that requires light. In this step the energy of an arriving photon, or quantum of light, serves to liberate an electron in photosystem I, a part of the thylakoid membrane. The electron, relayed by a sequence of molecules inside the membrane, then takes part in the reduction of the electron-transporting molecule nicotinamide adenine dinucleotide phosphate ($NADP^+$), which is thereby converted into NADPH. (Reduction signifies the donation of electrons.) The next photosynthetic step is also undisturbed. Here the energy of an arriving photon enables the part of the thylakoid membrane called photosystem II to pull an electron from a molecule of water and supply it to photosystem I, thus readying the membrane for another reaction cycle.

Ordinarily in photosynthesis the NADPH would fuel further reactions in which carbon atoms from carbon dioxide are incorporated into carbohydrates, notably hexoses, or six-carbon sugars. In Greenbaum's composite the light-dependent photosynthetic reactions are decoupled from these so-called dark reactions. What Greenbaum has done is to precipitate platinum onto the thylakoid membrane at the places where electrons are donated to $NADP^+$. The electrons go instead to the platinum, which acts as a catalyst for a chemical reaction quite different from the dark reactions of natural photosynthesis. The light-dependent reactions in photosynthesis can be summarized as the extraction of electrons from water, so that what remains of the water is oxygen atoms and protons, or hydrogen nuclei: hydrogen atoms from which the electron has been stripped. The reaction catalyzed by platinum is the reconstitution of hydrogen from these protons and electrons. In sum, then, the platinized chloroplasts redirect photosynthesis so that the energy of light breaks water into molecular oxygen and molecular hydrogen.

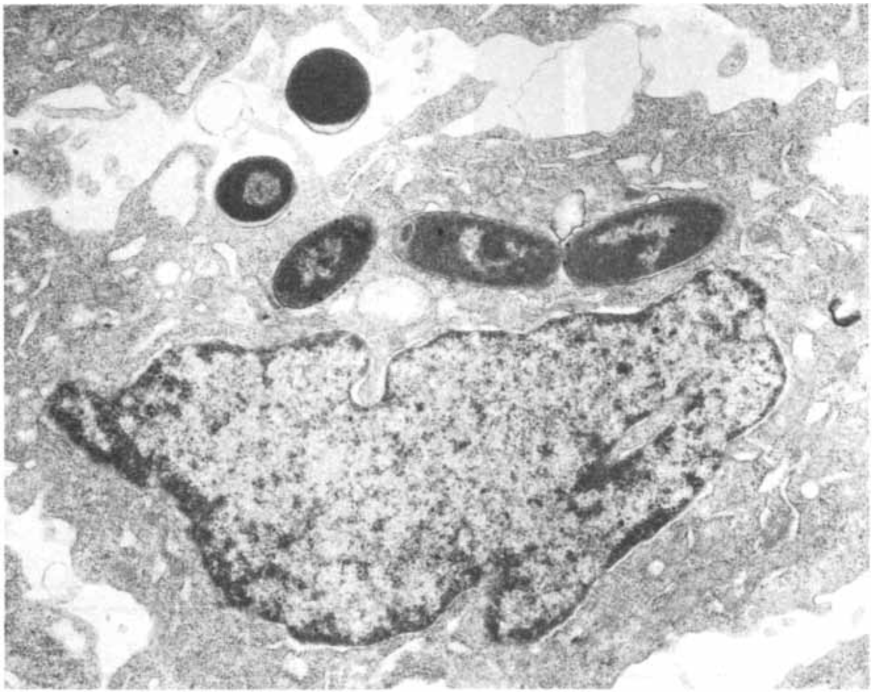
Greenbaum's composite emerged from a series of experiments in which chloroplasts extracted from spinach leaves were ruptured, exposing the thylakoid membrane. Next a suspension of thylakoid membrane was combined with a solution of platinum ions. In the presence of a reducing agent (an electron donor, typically molecular hydrogen) metallic platinum would precipitate out of solution and onto the membrane fragments. A solution of tetraamineplatinum ions, in which platinum atoms are surrounded by amine (NH_3) groups, failed to yield a composite material with photosynthetic activity. Each tetraamineplatinum ion has a net electric charge of +2.

A solution of hexachloroplatinate ions did yield photoactivity. (That is, the thylakoid-platinum composite, trapped on filter paper, responded to light by simultaneously generating hydrogen and oxygen from water.) Each hexachloroplatinate ion, in which a platinum atom is surrounded by six chlorine atoms, has a net charge of −2. Greenbaum proposes that the successful precipitation results from the electrostatic attraction of opposite charges: evidently the reducing site of photosystem I (the site where the thylakoid membrane donates electrons to $NADP^+$) has a net positive charge. In a further experiment Greenbaum dispensed with the addition of a reducing agent to the platinum-plating solution and relied instead on the reducing site of photosystem I. Exposing the membrane to light caused electrons to become available at the site, and metallic platinum precipitated there. Again a photoactive composite resulted.

It is a bit early to proclaim that platinized chloroplasts may generate hydrogen for the energy economy of the future. Still, the advantages of such a fuel cycle are plain. The making of hydrogen by means of tailored photosynthesis would consume only water, and the combustion of the fuel would regenerate the water.

Inside Job

Some bacteria, including those responsible for typhoid fever and tuberculosis, invade animal cells and survive there; other infectious bacteria are content to make their living outside the cell. Now two investigators have isolated a gene responsible for the invasion process. Understanding of such genes could lead to the development of alternatives to antibiotics for combating harmful invasive bacteria.



BACTERIAL INVADERS of a mammalian cell are the five dark bodies toward the top of this electron micrograph of the interior of a lung-carcinoma cell. The bacteria are *Escherichia coli,* which normally cannot invade; a single gene, designated *inv,* was transferred to the *E. coli* to change them into invaders. Each bacterium is about two micrometers long. The micrograph was made by Stanley Falkow of the Stanford University Medical Center.

Ralph R. Isberg and Stanley Falkow of the Stanford University Medical Center worked with *Yersinia pseudotuberculosis* and *Escherichia coli. Y. pseudotuberculosis* is an invader of cells: ingested in contaminated food, it enters the body by penetrating the cells that line the intestine. Ultimately it causes guinea-pig plague, a veterinary disease resembling typhoid fever. *E. coli* is not an invader of cells. It harmlessly inhabits the inner face of the intestinal tract of man and other mammals. Isberg and Falkow report in *Nature* that the transfer of a single gene from *Y. pseudotuberculosis* into *E. coli* converts the latter into an invader (although not a cause of disease).

Isberg and Falkow reasoned that some limited region of the genetic material in *Y. pseudotuberculosis* must underlie the ability to invade, so that the ability should be transferable. To find the region they first broke down the *Yersinia* DNA into short fragments. They packaged the fragments into phage lambda, a virus that infects bacteria, and proceeded to parcel out the total genetic "library" of *Y. pseudotuberculosis* in different *E. coli* cells.

The *E. coli* were applied to a monolayer of some 10 million *HEp*-2 cells, a laboratory cell line derived from a human lung carcinoma. After three hours the monolayer was washed repeatedly. The only *E. coli* remaining would be those that somehow had linked themselves to the cells of the monolayer. Some might have got inside the cells. The survivors of the washing were cloned (that is, allowed to generate colonies of identical offspring) and the clones were given a second opportunity to enter *HEp*-2 cells. Then the *HEp*-2 cells were bathed in the antibiotic gentamycin, which cannot penetrate cells. The surviving bacteria would have to be inside the cells. Twelve clones survived. The cells were examined under the electron microscope. *E. coli* indeed were inside them.

The genetic material from *Y. pseudotuberculosis* that confers the ability to invade proves to be a DNA strand no longer than 3,200 bases; Isberg and Falkow call it the *inv* locus. It is in fact a single gene, which specifies the structure of a single large protein with a molecular weight of 108,000. In the case of *Y. pseudotuberculosis,* at least, the genetic basis of the ability to invade cells is strikingly uncomplicated.

Isberg and Falkow observe that their work "may be a first step towards defining a previously undefined class of proteins encoded by many invasive disease-producing microorganisms as part of their strategy of pathogenesis." Just what this protein does to promote invasion remains to be learned. When it is, methods may emerge by which to undermine the invasion and so block infection by those bacteria that cause harm by invading cells. Harmless bacteria, on the other hand, might conceivably be transformed into living ferries that carry proteins or DNA into animal cells for research and perhaps for therapeutic purposes.

Weed Out

In an ecologically conscious era biological methods for the control of pests and weeds have strong appeal. The release of sterile adult insects, for instance, may diminish screwworm and mosquito populations, thereby avoiding the need for chemical pesticides. Workers at Montana State University and Cornell University have now proposed that similar tactics can be exploited to replace synthetic chemical herbicides. They have taken the first steps toward utilizing a fungal toxin to control Bermuda grass.

Although some varieties of Bermuda grass can prevent soil erosion and make excellent lawns, in more than 80 countries it is considered a troublesome weed that interferes with the farming of at least 40 different crops, notably sugarcane. (It is also a significant cause of hay fever.) Fumio Sugawara and Gary A. Strobel of Montana State and Larry E. Fisher, G. D. Van Duyne and Jon Clardy of Cornell write in *Proceedings of the National Academy of Sciences* that they have successfully isolated and characterized a fungal product known as bipolaroxin that is toxic to Bermuda grass. (Some other plant species such as corn and wild oats are also susceptible to the toxin, but only at much greater concentrations.) The investigators believe bipolaroxin is the first such host-selective toxin from a weed pathogen to be identified.

Sugawara and Strobel worked with laboratory cultures of the fungus *Bipolaris cynodontis,* a pathogen that causes leaf blight in the plant. To winnow out the toxin, they applied successively purified extracts of the fungus to puncture wounds on Bermuda grass samples and noted whether or not spots or lesions developed. Having isolated bipolaroxin, Sugawara and Strobel crystallized the molecule and Fisher, Van Duyne and Clardy characterized it by X-ray-diffraction analysis.

Although the workers are far from deploying bipolaroxin in the field, the investigation has already increased understanding of the molecular structure and biological activity of the toxin. It could ultimately provide chemical clues to the development of new herbicides from natural sources.

Announcement of William C. Foster Fellows Program for 1986-87

The U.S. Arms Control and Disarmament Agency (ACDA) has announced that it is accepting applications for visiting scholars, under the William C. Foster Fellows Program for 1986-87. This program is designed to give specialists in the physical sciences and other disciplines relevant to ACDA's activities an opportunity to participate actively in the arms control and disarmament activities of the Agency and to give the Agency the perspective and expertise such persons can offer.

William C. Foster Fellows for 1986-87 will be appointed for twelve months beginning in the summer or early fall of 1986. They will be compensated in accordance with the Intergovernmental Personnel Act which allows the Agency to reimburse a university for the services of its employees. Fellows must be citizens or nationals of the United States and on the faculty of a recognized institution of higher learning. Prior to appointment they will be subject to a full-field background security and loyalty investigation for a top secret clearance.

Applications should be made in the form of a letter indicating the perspective and expertise which the applicant offers accompanied by a curriculum vitae and any other materials such as letters of reference and samples of published articles which the applicant believes should be considered in the selection process. The deadline date for applications is January 31, 1986. Applications and requests for information on available assignments should be sent to William C. Foster Fellow Program, Attention: Personnel Officer, room 5722, U.S. Arms Control and Disarmament Agency, 320 21st Street, N.W., Washington, D.C. 20451, telephone (202) 632-2034.

UNITED STATES ARMS CONTROL AND DISARMAMENT AGENCY
Washington, D.C. 20451

How Exxon broke complex fluids to solve a

It costs about $7,000 per day to drill for oil or gas on land. Offshore, comparable costs range upward to $100,000 daily. Even short interruptions are expensive.

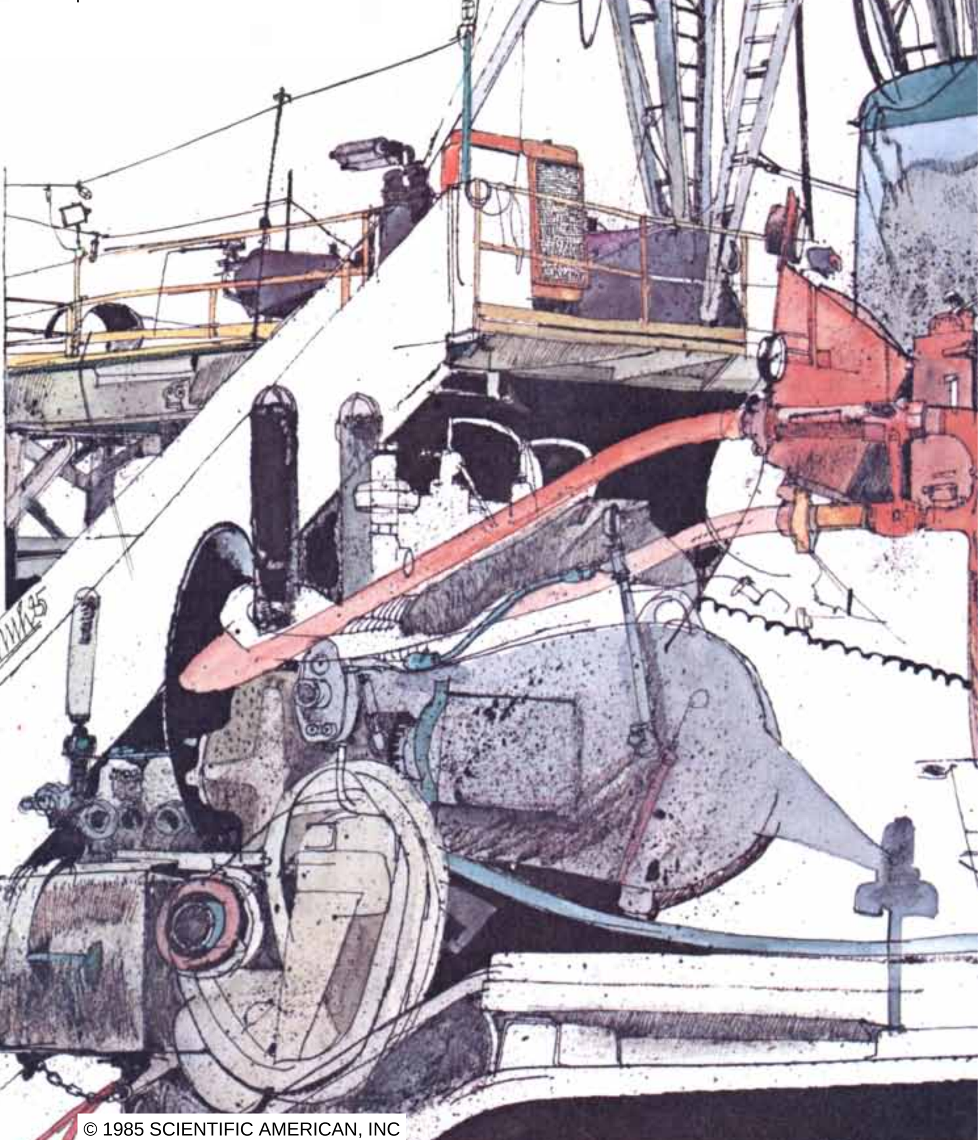
Delays can occur when drilling fluid, which is used to carry rock chips to the surface and to control well pressures, is lost into underground formations. When this happens, drilling must often be stopped until the problem is corrected.

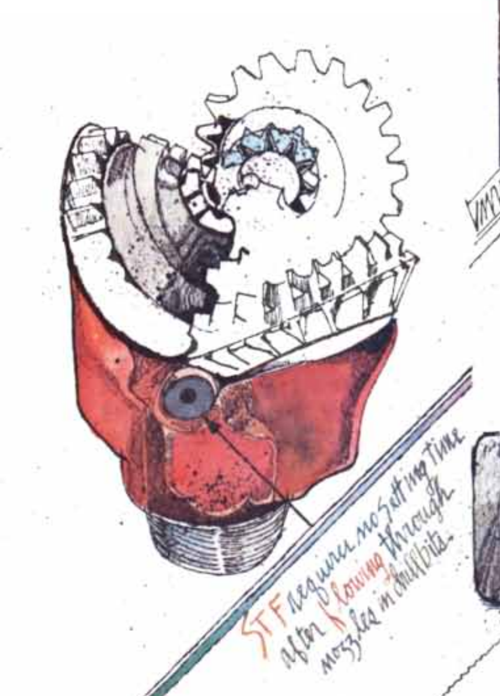
Traditional methods of correcting drilling fluid loss are not always effective. Granular materials pumped downhole may not plug the leaks. Cement may need too much time to harden. Other plugging substances are unreliable because they depend on mixing two or more components downhole. Where plugging fails, the only effective remedy has been to isolate the loss zone by placing an extra steel casing over it, often at considerable cost.

Complex Fluid Development

To solve the problem, Exxon sought to create a unique fluid that could be pumped down the drill pipe to the bottom of the well, where it would then thicken into a viscous paste.

The fluid Exxon developed consists of water droplets and granules of water-swellable clay finely dispersed in a continuous oil phase. Each water droplet is surrounded by a tough film created by a water-soluble polymer that reacts with an oil-soluble surfactant. This film is ruptured by high shear forces when the fluid flows through the nozzles in a drill bit. The water is released and interacts rapidly with the clay. As the shear-thickening fluid (STF) emerges from the nozzles, it becomes a viscous paste with the consistency of modeling clay. The paste can seal the loss zone.





new ground in serious drilling problem.

Interactive Research

Development of STF was the result of a team effort involving scientists and engineers in several Exxon units.

Exxon Production Research Company (EPR) knew the oil field environment with its problems, complexities and engineering requirements. Exxon Research and Engineering Company (ER&E) had expertise in the chemistry of surfactants and polymers. Together, they refined the formulation through a series of tests in full-scale surface equipment, then in wells provided by Exxon Company, U.S.A., using materials supplied by Exxon Chemical Company. The final result was successfully field-tested, solving a number of drilling fluid loss problems in wells where other approaches had failed.

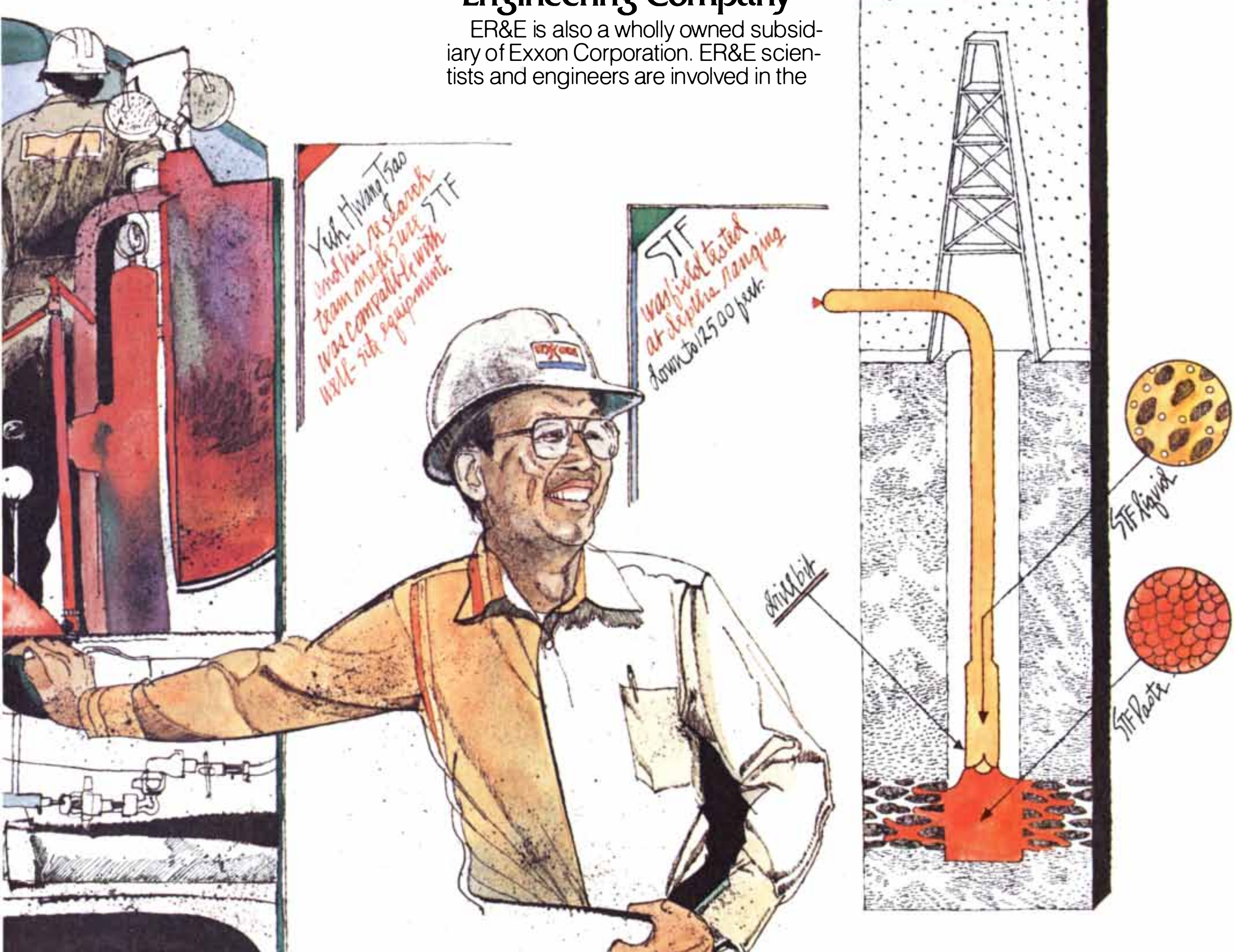
Exxon Production Research Company

EPR, a wholly owned subsidiary of Exxon Corporation, is involved in research and development on methods of finding and developing oil and gas reserves around the world. For more information on shear-thickening fluid, write Exxon Production Research Company, P.O. Box 2189, Houston, TX 77252-2189.

Exxon Research and Engineering Company

ER&E is also a wholly owned subsidiary of Exxon Corporation. ER&E scientists and engineers are involved in the research and development of petroleum processing and products, as well as pioneering science, and the engineering required to develop and apply new technology in the manufacture of fuels and other products. For more information, write Exxon Research and Engineering Company, Room 201, Box 101, Florham Park, NJ 07932.

EXXON



Applications of Optical Phase Conjugation

"Time-reversed" light waves can be used to improve laser-beam quality, compensate for atmospheric turbulence, track a moving satellite, encode and decode messages and compare image patterns

by David M. Pepper

Boris Ya. Zel'dovich and his colleagues observed a curious phenomenon as they did an experiment at the P. N. Lebedev Physical Institute in Moscow in 1972. The investigators intentionally distorted an intense beam of red light from a pulsed ruby laser by directing it through a frosted glass plate. Then they aimed the smeared beam down a long tube containing high-pressure methane gas. In accordance with a well-known effect called stimulated Brillouin scattering the beam interacted with the molecules of the gas and was reflected backward; the gas acted as if it were a mirror, but a very unusual one. What surprised the investigators was that after the reflected wave passed back through the same piece of frosted glass a nearly perfect, undistorted optical beam emerged. In other words, the distortions introduced during the first passage through the glass had been undone. (Reflection from a conventional flat mirror, in contrast, would have increased the distortions.) The backward-traveling wave could therefore be loosely thought of as the "time-reversed" replica of the incident wave.

The phrase "time-reversed" is intended to imply that the beam reflected back by the gas faithfully carried all the distortions introduced by the frosted glass plate, but in a reversed sense. When the beam returned through the glass, the distorting properties of the plate therefore canceled the effects they had originally produced. In other words, if it were possible to make a motion picture of the incident beam, the "time-reversed" beam would be portrayed by playing the same film backward. "Time-reversed" waves are more accurately known as phase-conjugate waves, or wave-front-reversed replicas. The technology by which they are generated is known as optical phase conjugation [see "Optical Phase Conjugation," by Vladimir V. Shkunov and Boris Ya. Zel'dovich; SCIENTIFIC AMERICAN, December, 1985].

Phase-conjugate-wave technology has many intriguing applications. A high-quality optical beam can, for instance, be transmitted through a turbulent atmosphere and, after generation of its phase-conjugate beam, be made to retrace its path exactly. When the beam returns to its point of origin, it will therefore be free of degradation. Such beams can be used in the pointing and tracking of moving objects, the processing of images, optical computing, interferometry, laser gyroscopes, fiber and satellite communication systems, laser weapon systems and photolithography. The "mirrors" that generate phase-conjugate beams also make possible novel laser resonators.

Generating "Time-reversed" Light

Two standard methods of producing phase-conjugate waves are stimulated Brillouin scattering and four-wave mixing. In the years following the landmark experiment by Zel'dovich investigators have found that phase conjugation by stimulated Brillouin scattering can occur in a variety of substances other than compressed gas. Such substances are called nonlinear mediums. The term is used here in a special sense: the optical properties of a nonlinear medium, in contrast to those of a linear medium, are affected by light. Examples of nonlinear mediums include semiconductors, crystals, liquids, plasmas, liquid crystals, aerosols and atomic vapors.

Nonlinear mediums are also utilized in four-wave mixing, a scheme that was proposed in 1977 by Robert W. Hellwarth of the University of Southern California, followed by studies by Amnon Yariv of the California Institute of Technology and me (then at Caltech), and by David M. Bloom, Gary C. Bjorklund and Paul F. Liao of the AT&T Bell Laboratories. The technique involves the interaction in a nonlinear medium of four optical beams: three input beams and one output beam. The three input beams consist of a probe beam, whose "time-reversed" replica is sought, and two counterpropagating pump beams that are necessary to "sensitize" the nonlinear medium. The steps leading to the production of the fourth beam, the phase-conjugate output beam, are analogous to those of conventional holography. In holography a photographic emulsion is illuminated with light from an object and a reference beam. The reference beam and the light from the object interact to produce a hologram, or three-dimensional interference pattern, in the emulsion. After the film is developed the fixed, or static, hologram can be read by illuminating it with the same reference beam. The result is a three-dimensional reconstruction of the object's image.

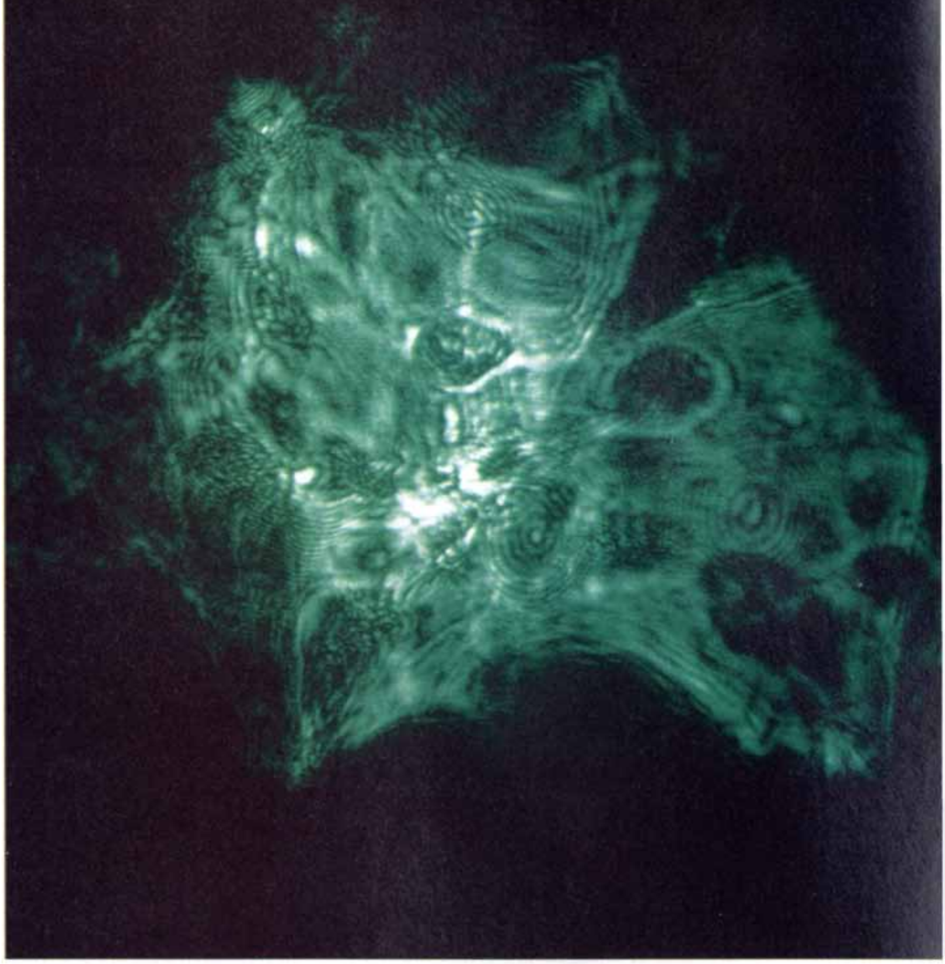
In four-wave mixing the nonlinear medium acts as the photographic emulsion. The interaction of the probe, or object, beam with one of the pump, or reference, beams produces a wave pattern of reinforcement and cancellation (constructive and destructive interference) in the medium that amounts to a real-time, or dynamic, hologram. The phase-conjugate output beam is generated when the other counterpropagating pump beam is reflected from the hologram. Actually two holograms are created: each pump beam, in conjunction with the probe beam, yields a hologram from which the other reference beam is reflected.

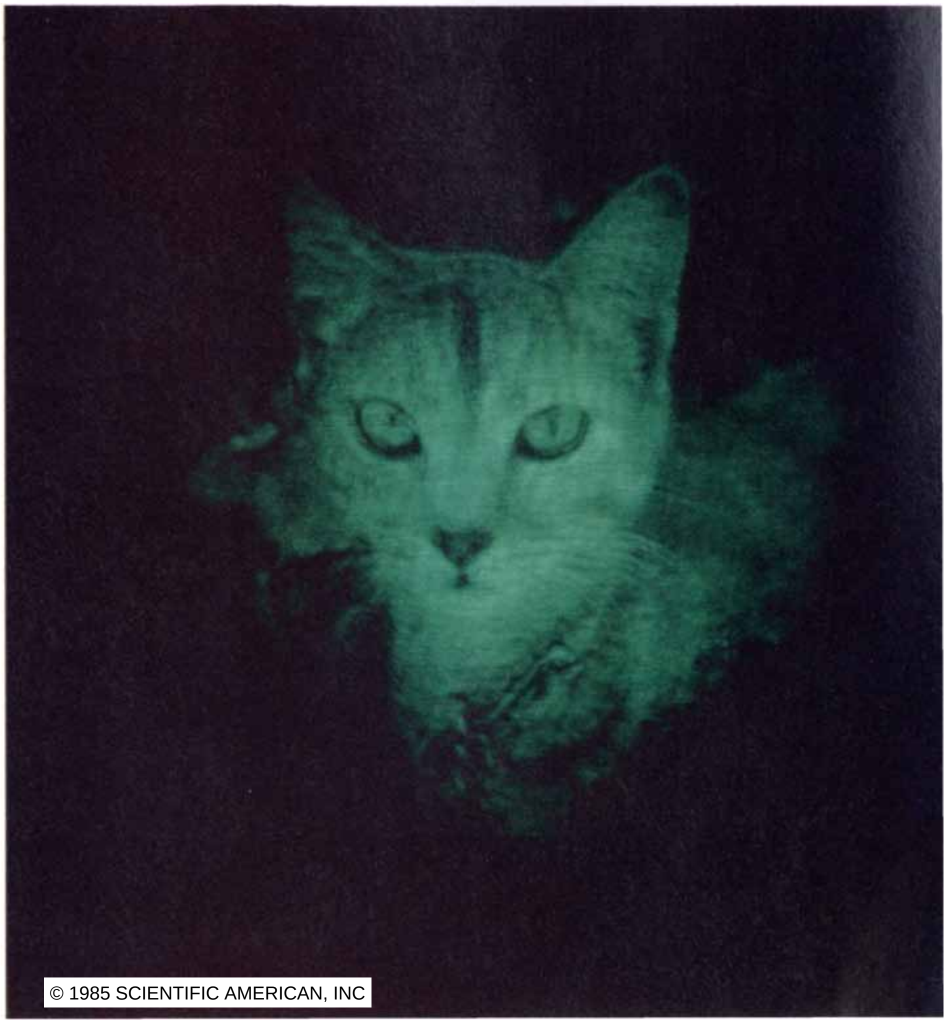
Although each method—stimulated Brillouin scattering and four-wave mixing—has advantages and disadvantages, the end result of both is a phase-conjugate mirror, a peculiar mirror that exists in three dimensions and returns a "reflection" of the initial beam that exactly retraces the path of the initial beam. In both cases the dynamic aspect of the nonlinear medium enables the phase-conjugate mirror to respond to time-varying input beams. The reader is challenged to imagine while reading the article what he or she might see while gazing into such a mirror. (The answer is given at the conclusion of this article.)

Adaptive Optics

Optical phase conjugation has many applications because there is often a need to compensate for static and dynamic distortions encountered in optical systems. High-power lasers, tracking systems, atmospheric communication networks and photolithographic systems are all hampered by such noise, which can also hamper the effectiveness of a weapon system. Either four-wave mixing or stimulated Brillouin scattering can help.

Many of these applications depend on a so-called double-pass geometry: an arrangement in which a laser beam is phase-conjugated and reflected in such a way that it passes through the same medium twice. Such a reflected beam exactly retraces the path of the incident beam in a "time-reversed" sense. In one example of such a geometry a low-power seed laser injects a highly directed light beam into an amplifier that is typically made up of a solid or a gas of highly excited atoms or molecules. As the beam passes through the amplifier it "tickles" the molecules into releasing their energy as radiation. A powerful beam is produced, but at the cost of directivity: inhomogeneities in the amplifying medium distort the beam so that it diverg-





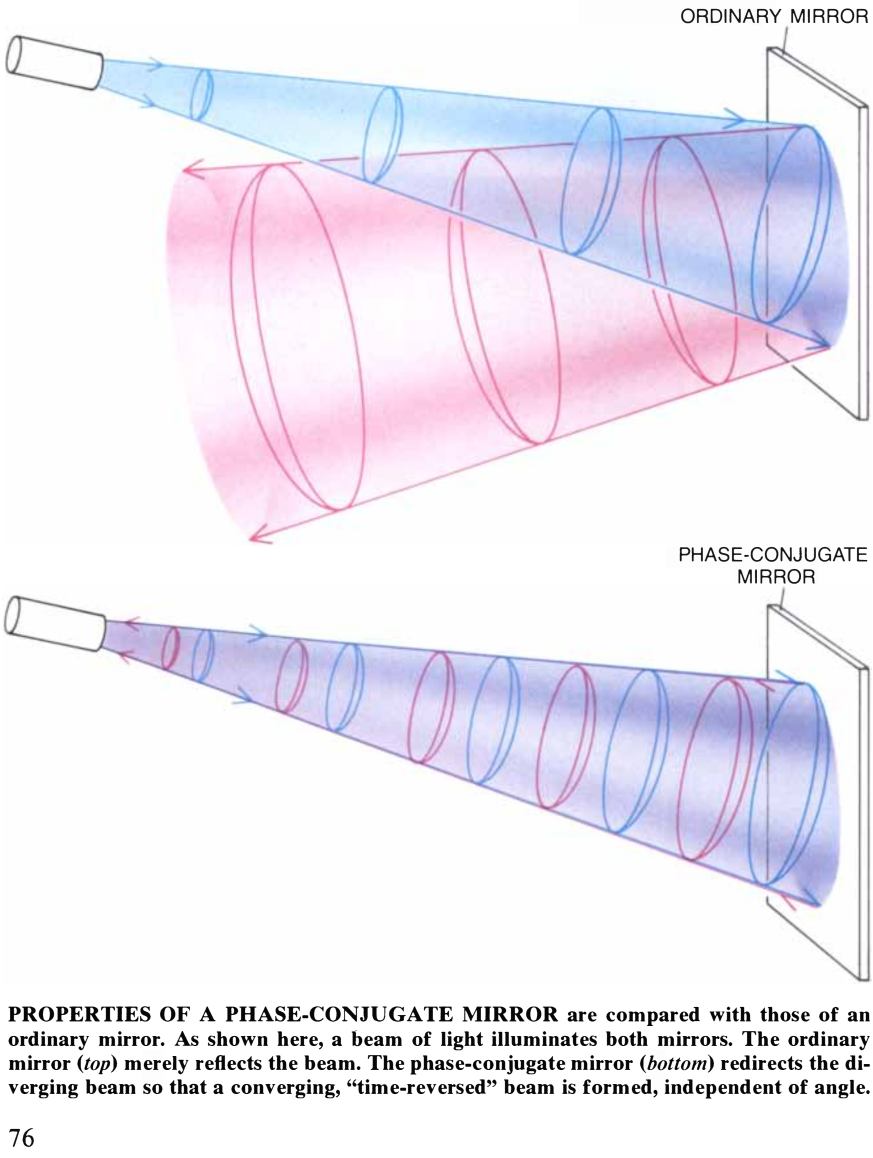
PHASE-CONJUGATE MIRROR is able to compensate for distortions imposed on an image of a cat. In both photographs the image was distorted by transmitting it through a piece of frosted glass. Reflection of the image back through the same piece of glass by an ordinary mirror yielded an unrecognizable image (*top photograph*). Reflection of the image back through the frosted glass by a phase-conjugate mirror, on the other hand, corrected the distorted image (*bottom photograph*). It did so because a phase-conjugate mirror produces a beam that propagates back through the distorting glass in a "time-reversed" sense: its trajectory retraces that of the original beam and thereby undoes the distortions. Jack Feinberg of the University of Southern California performed the experiment using an argon laser.

es. The distortions can be removed if a phase-conjugate mirror placed at the end of the amplifier receives the beam. When the phase-conjugate beam from the phase-conjugate mirror travels back through the amplifier, its motion, which is the reverse of the original beam, undoes the degrading effects of the medium. The double-passed beam that emerges is both powerful and highly directed.

As a consequence of fundamental and practical considerations there are limits on the physical size of a single amplifier and hence on the output power it can produce. By using a set of amplifiers in a parallel arrangement and a phase-conjugate mirror one can both compensate for optical distortions within each amplifier and "phase up," or synchronize, an entire ensemble of amplifiers. Since the ensemble acts as an optical phased array of amplifying elements, the peak intensity of the output beam is proportional to the square of the number of amplifiers. The peak intensity produced by an array of amplifiers in which the phase of the radiation is random is merely proportional to the number of amplifiers. The basic scheme, proposed by Nikolay G. Basov and his colleagues at the Lebedev Physical Institute and by Thomas R. O'Meara and his colleagues at the Hughes Research Laboratories, has been recently demonstrated by David A. Rockwell and Concetto R. Giuliano, also at Hughes.

Another example of a double-pass geometry is the self-targeting of radiation, which also relies on a beam from a low-power seed laser. In this case, however, the seed laser directly illuminates a target, such as a fusion pellet containing a mixture of deuterium and tritium (isotopes of hydrogen). Some of the light scattered by the pellet passes through a nearby amplifier. As the beam passes through the amplifier its power is increased. A phase-conjugate mirror at the end of the amplifier produces an intense "time-reversed" beam that is directed back to the target. If the beam were powerful enough, the nuclei in the pellet would fuse, releasing useful energy.

Self-targeting can be applied to other situations as well: the fusion pellet can be replaced by an orbiting satellite. On illumination any scattered light returning from the satellite and collected by a phase-conjugate laser system can be amplified and sent back to the satellite. To the extent that the intervening atmosphere and satellite position do not appreciably change during the round-trip optical transit time, the phase-conjugate mirror not only will compensate for atmospheric turbulence (the phenomenon that causes the "twinkling" of stars) but also will make it possible to track the satellite by maintaining a continuous beam of illumination on it. The satellite could subsequently direct the amplified laser beam to a missile and, if the radiation were sufficiently intense, destroy it. Such schemes have been proposed and extensively evaluated for laser weapon systems since the early 1970's.



PROPERTIES OF A PHASE-CONJUGATE MIRROR are compared with those of an ordinary mirror. As shown here, a beam of light illuminates both mirrors. The ordinary mirror (*top*) merely reflects the beam. The phase-conjugate mirror (*bottom*) redirects the diverging beam so that a converging, "time-reversed" beam is formed, independent of angle.

Lensless Imaging

Many optical systems rely on large numbers of lenses that are difficult to control and adjust. Phase-conjugate schemes have been proposed that could eliminate such difficulty. They could, for instance, be used to transfer a given two-dimensional pattern from one plane in space to another. That is precisely the goal of photolithography, the technique by which a mask pattern containing a microelectronic circuit layout is transferred to a semiconducting chip that is coated with a photographic emulsion. (Direct physical contact of the mask with the substrate may be undesirable.)

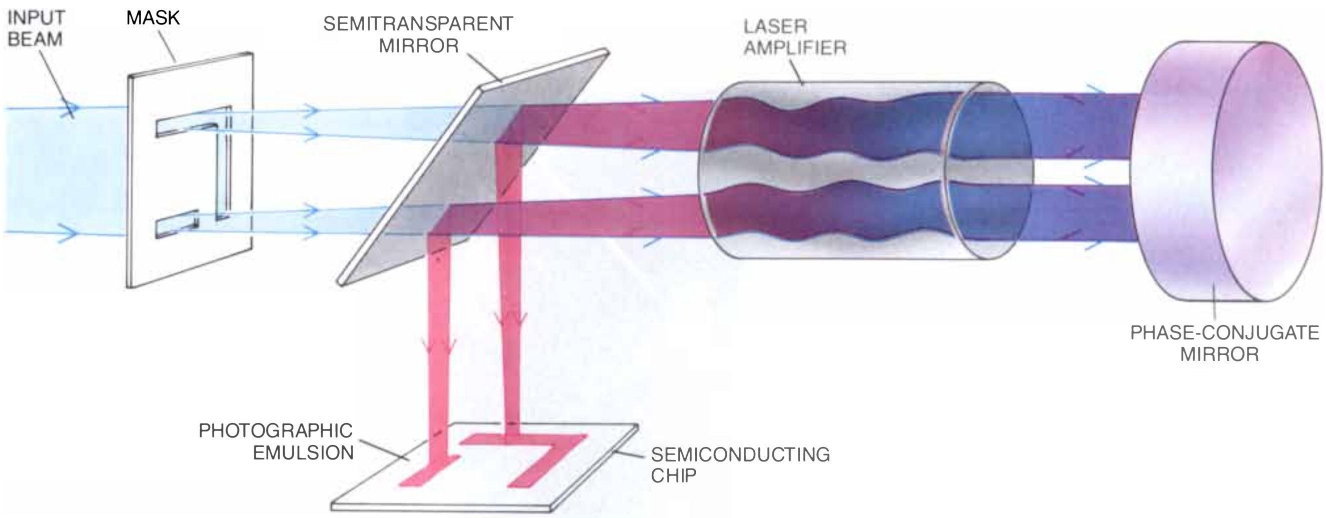
The feasibility of such image transfer has been demonstrated by Marc D. Levenson and his colleagues at the IBM Research Laboratory in San Jose and more recently by Malcolm C. Gower of the Rutherford Appleton Laboratory in England. Light from a laser passes through the mask pattern, a semitransparent mirror and then an amplifier. The intensity of the beam increases at the expense of introducing distortions into the beam. The resulting image is sent back through the amplifier by a phase-conjugate mirror. The "time-reversed" beam is both powerful and free of distortions, so that when it reflects from the semitransparent mirror, it exposes the emulsion with the pattern. The phase-conjugate system has advantages in relation to conventional methods such as compensating for optical aberrations.

Lensless-imaging schemes involving optical phase conjugation are useful in other ways. One is the removal of distortions introduced by fiber-optic cables. Such distortions arise as follows. Imagine sending a three-dimensional image (say of a tree) through a plate-glass window. The window does not degrade to any serious degree the quality of the image. If the same glass window is replaced by a long optical fiber, the image will be unrecognizable after traveling only a few centimeters. The reason is that the image of the tree travels in many optical modes, each of which corresponds to a given ray that zigzags down the fiber. Since all the modes traverse different paths, they are out of step by the time they reach the end of the fiber. (They no longer have the same phase relation to each other as they did at the beginning of the fiber.) In other words, just as a piece of frosted glass scrambles an image, so does an optical fiber.

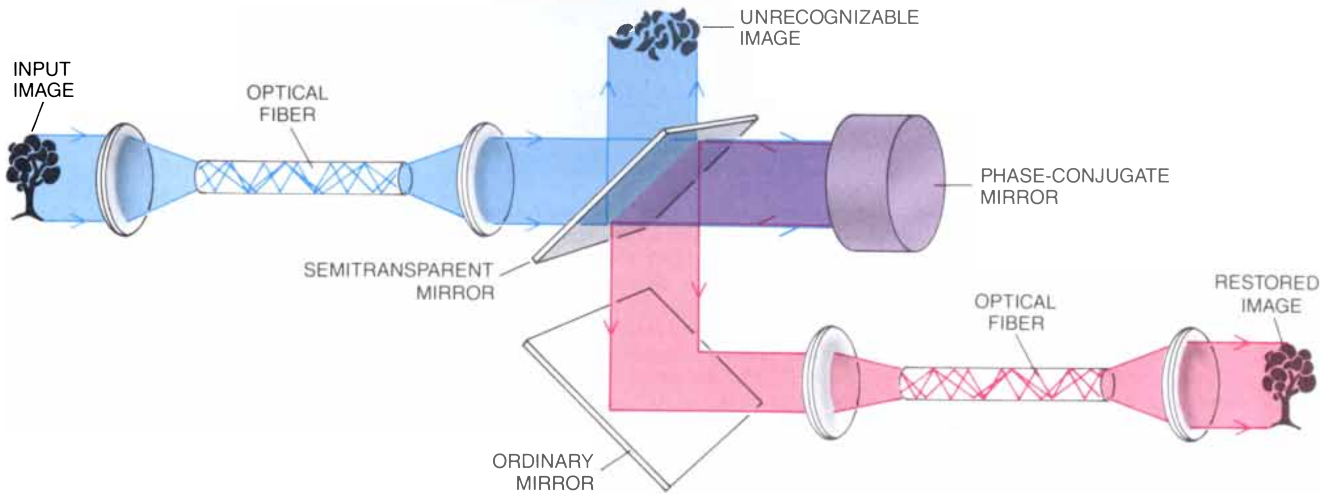
The image can be unscrambled if it is phase-conjugated and sent back through the same fiber, or through another, identical fiber link. This concept, initially pointed out by Yariv, has been successfully demonstrated by Gilmore J. Dunning and Richard C. Lind of Hughes, and more recently by Baruch Fischer and Shmule Sternklar of the Technion–Israel Institute of Technology in Israel. Even though they worked with two-dimensional images, the scheme can in principle also allow for the transmission of three-dimensional images through fibers.

Phase-Conjugate Resonators

How might the performance of a laser change were one or both of its cavity mirrors to be replaced by a phase-conjugate mirror, forming a so-called phase-conjugate resonator? A resonator is a hollow cavity for reinforcing a sound wave or an electromagnetic wave. Typically it consists of a long tube filled with an amplifying medium (a solid or a gas), capped at each end with a highly polished semitransparent



PHOTOLITHOGRAPHY, the transfer of a two-dimensional pattern from one plane in space to another, can be done without optical lenses by employing a phase-conjugate mirror. The process can be used to transfer a mask pattern containing a microelectronic circuit layout onto a semiconducting chip coated with a photographic emulsion. A beam of light (*blue*) from a laser passes through the mask, a semitransparent mirror and then a laser amplifier. Although the amplifier increases the power of the beam, inhomogeneities in the amplifying medium distort its quality. After the beam is phase-conjugated (*red*) it propagates back through the amplifier, undoing the distortions and emerging with the original beam information restored. On reflection from the semitransparent mirror, an intensified image of the mask pattern exposes the emulsion. The system compensates for optical aberrations, has high resolution over a large field of view, eliminates laser speckle, minimizes beam spreading and avoids physical contact with the substrate.



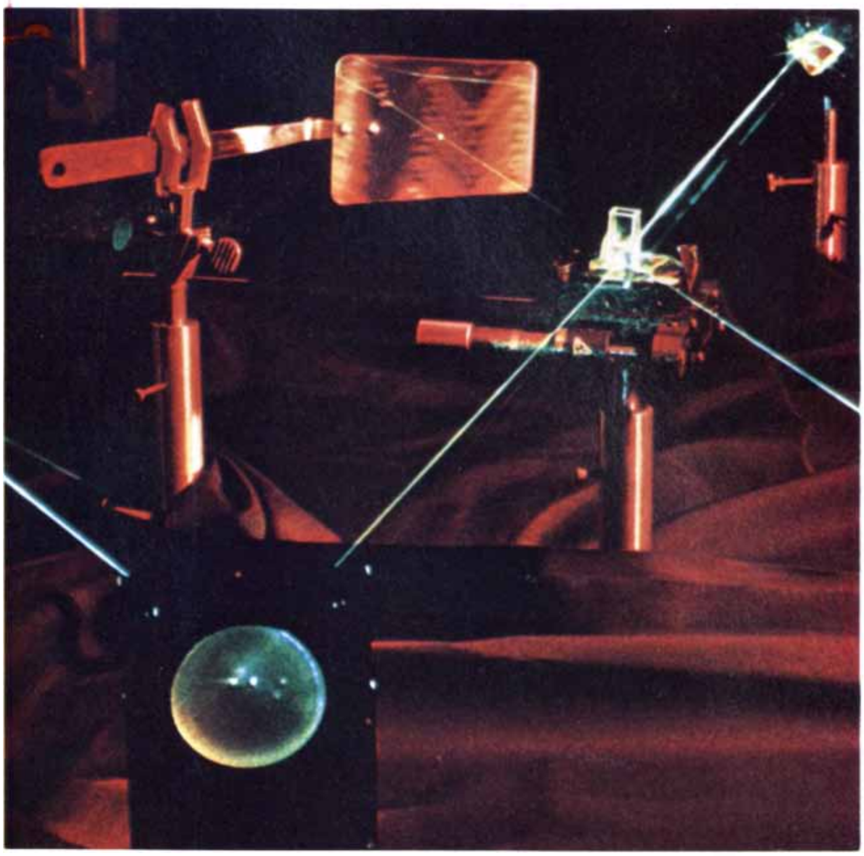
DISTORTIONS DUE TO FIBER-OPTIC CABLES can be removed by a phase-conjugate mirror. The distortions arise because a spatial image (*left*) sent down a cable travels in many optical modes: each mode corresponds to a given ray that zigzags down the fiber (*blue*). Since all the modes traverse different paths, by the time they reach the end of the fiber they are out of step, producing a scrambled image (*center*). The image can be unscrambled if it is phase-conjugated (*red*) and sent down an identical cable (*right*).

mirror. Laser output results when light passes through one of the mirrors.

In 1978 a team at Caltech headed by Yariv that included John C. AuYeung, Dan Fekete and me explored this question. In the course of our work we successfully developed the first pulsed phase-conjugate laser. In our experiment a ruby rod served as the amplifying medium. We placed a conventional mirror at one end of the rod and a phase-conjugate mirror at the other end. The phase-conjugate mirror, in this case a four-wave mixer, consisted of a cell filled with carbon disulfide; the pump beams were obtained from another pulsed ruby laser.

We found that a phase-conjugate resonator has unique properties. Perhaps the most obvious property is that the phase-conjugate mirror can compensate for static and dynamic aberrations in the cavity due to, for instance, imperfect optical components, as well as dynamic thermal and mechanical perturbations. A phase-conjugate resonator is therefore highly efficient at extracting the optical energy stored in the laser medium. A more subtle effect relates to what is called "resonator stability." The ability of a conventional resonator to store energy is dependent on the relation between the radiation, the cavity length and the mirror curvature. A phase-conjugate resonator is free from such constraints.

The properties of the phase-conjugate resonator occupy the attention of a number of research groups including those at Hughes, Caltech and Stanford University and in the Soviet Union, France, Great Britain and Canada. Devices based on four-wave mixing and stimulated Brillouin scattering have both been employed.



"SPATULA LASER" consists of a reflecting mirror (*spatula*) and a phase-conjugate mirror (*crystal at beam intersection*). A conventional laser consists of a long tube capped at each end with a highly polished mirror. The tube is typically filled with an amplifying medium (a highly excited solid or gas). The spatula laser requires no amplifying medium because the phase-conjugate mirror provides the necessary gain. The amplification is achieved by pumping, or exciting, the rectangular crystal with two laser beams known as pump, or reference, beams. One of the pump beams enters the photograph from the corner at the lower left and is reflected by a mirror (*green disk*) to the phase-conjugate mirror. Some of the laser beam is absorbed by the crystal; most of the beam continues on to the wedge-shaped mirror in the upper-right corner. The beam reflected by that mirror becomes the other pump beam. After the crystal has absorbed sufficient energy from the two counterpropagating pump beams it will begin to lase with any shiny surface in its vicinity, as if the region between the crystal and the shiny surface were filled with an amplifying medium. The laser beam lying along the line between the spatula and the crystal is thereby generated. If the spatula were moved, the phase-conjugate mirror would track it by maintaining the beam on its surface. The laser beam emerges through the crystal (*right*). The experiment was done by Feinberg and Robert W. Hellwarth of the University of Southern California.

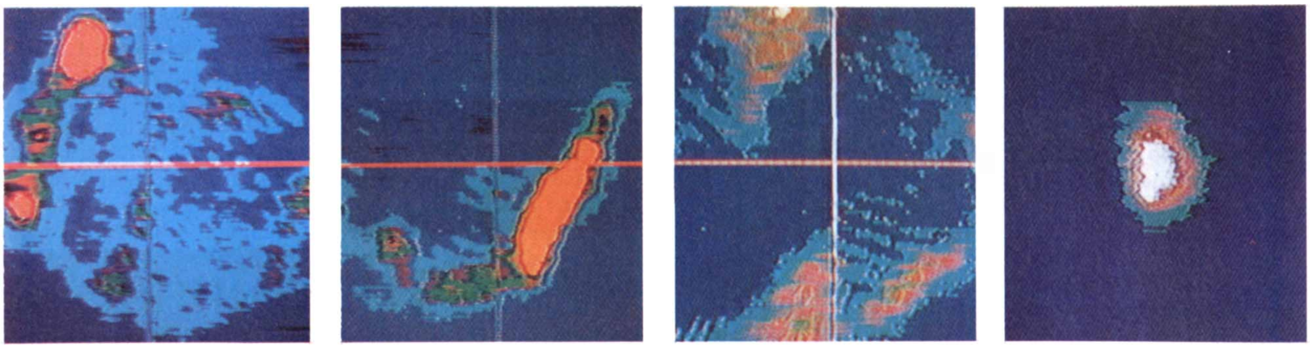
Lind and Duncan G. Steel (then at Hughes) have, for instance, recently constructed the first continuously lasing phase-conjugate resonator. The phase-conjugate mirror they used consisted of a one-centimeter cell filled with atomic sodium vapor, pumped by a dye laser. Inserting a frosted glass sheet within the laser cavity had virtually no effect on the system: it continued to produce a reasonably good beam, even as the plate was moved into the cavity. The same frosted plate prevented a conventional version of the same dye laser from producing any beam at all.

A phase-conjugate resonator of particular interest can be built using a four-wave mixer. Since two pump beams are employed, the intensity of the output phase-conjugate beam can exceed the intensity of the input probe beam. In other words, a four-wave mixer can be made to have optical gain. Such a four-wave mixer is said to be an amplifying phase-conjugate mirror. Even an "empty" laser resonator, consisting of only an amplifying phase-conjugate mirror and a conventional mirror, can therefore be made to lase; no internal amplifying medium is necessary. Since the phase-conjugate mirror also reflects any light striking it back to the point of origin, lasing will still occur even if the conventional mirror is moved about. Yariv and I proposed such a device and with the assistance of Fekete have tested it; it has also been demonstrated by Bloom, Liao and N. P. Economou of the AT&T Bell Laboratories.

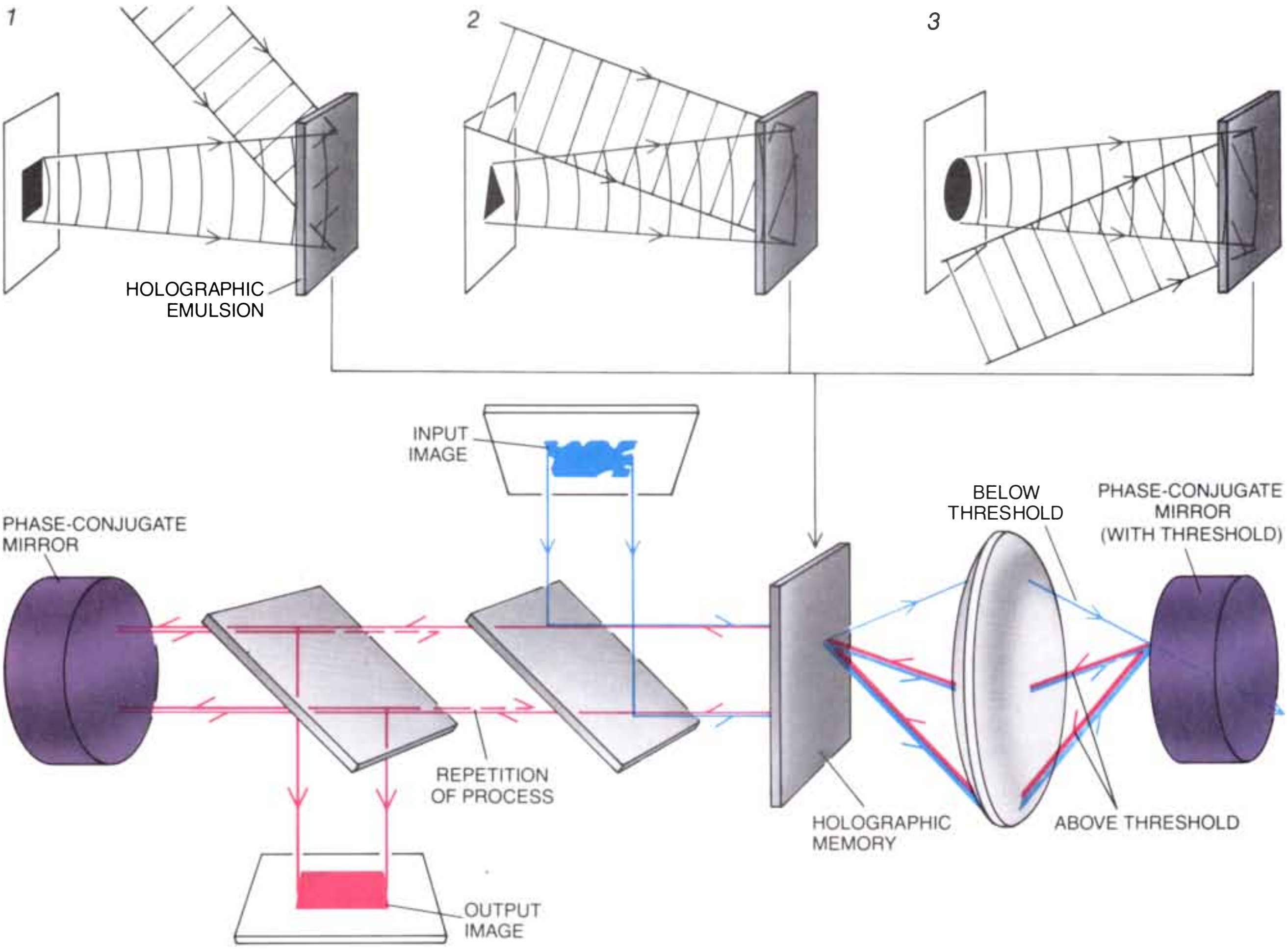
Perhaps the most novel laser resonator was built by Jack Feinberg of the University of Southern California and Hellwarth. They merely positioned a kitchen spatula in the vicinity of a phase-conjugate mirror (a barium titanate crystal pumped by a laser). The amplifying properties of the phase-conjugate mirror created a phase-conjugate resonator, with the crystal and the reflective surface of the spatula forming the cavity mirrors. An intense beam of light was then seen between the spatula and the crystal. The phase-conjugate mirror tracked any motion of the spatula by maintaining the beam on its surface. Moral: Do not wear anything reflective near an amplifying phase-conjugate mirror!

Image Processing

The common element in all four-wave-mixing processors is the optical



ATMOSPHERIC TURBULENCE is compensated for by optical phase conjugation. Laser light that traversed a 100-meter outdoor range produced highly distorted beam spots (*sequence of three photographs at left*) that vary in space and time. (The turbulence that produced the distortions is also responsible for the twinkling of stars and the apparent undulation of objects viewed over a heater or on a paved road on a hot afternoon.) When the beam was phase-conjugated and sent back through the atmosphere, the near-perfect quality of the beam spot was restored (*right*). The video-processed, color-encoded photographs were made at an exposure of 1/2,000 second. Gilmore J. Dunning and Richard C. Lind of the Hughes Research Laboratories did the experiment with a tunable dye laser.



ASSOCIATIVE MEMORY can recognize two images that share common features with each other. A memory-storage element is formed by illuminating a photographic emulsion with a reference beam and an image of a square (*top left*), creating a hologram, or an interference pattern, in the film. The process is repeated on the same emulsion for a triangle and a circle. The incident angle of the reference beam is changed in each case so that the images can be distinguished. The photographic emulsion—now a holographic memory element—is developed and placed between two phase-conjugate mirrors (*bottom*). The image to be categorized (*imperfect square shown here*) is reflected by a semitransparent mirror onto the emulsion (*blue*). Every time a ray of light from the imperfect square matches a ray from one of the holograms stored in the film a strong output ray is produced (*heavy line*); if the rays do not match, a weaker output is produced (*medium and thin lines*). Strong and medium output rays are reflected by the phase-conjugate mirror at the right (*red*). Weak output rays, however, do not have enough energy to cause conjugation to occur and therefore pass through the phase-conjugate mirror (*blue ray at right*). In this way the qualities of "squareness" are selected. After the strong and medium output rays are conjugated a second time by the phase-conjugate mirror at the left the process is repeated. Successive iterations yield a square as an output. In other words, the system "recognized" that the input image was an imperfect square, not an imperfect circle or triangle.

interaction of the beams. On a fundamental level the conjugate beam can be thought of as an algebraic product of the three input optical beams (the probe beam and the two counterpropagating pump beams); the nonlinear medium in the four-wave-mixing cell provides for the multiplicative effect through the coupling of the beams. Hence if any or all of the input beams are encoded either in a spatial fashion (by passing the beams through various slides or other optical transparencies) or in a temporal manner (by pulsing them), the conjugate output beam will be encoded in a way proportional to the product of all the information of the interacting beams.

Research groups led by Yariv at Caltech and by Osamu Ikeda at the Tokyo Institute of Technology have shown that by spatially encoding the pump beams, instead of the probe beam, in a four-wave mixer one can transmit images in a "one way" manner, even through a distorting medium. (Remember that the lensless-imaging techniques discussed above involve the double-pass transmission of an image.) Suppose a projectionist wants to send an image or a motion picture (say of a sportswoman diving off a springboard) to a friend some distance away. Owing to atmospheric turbulence, the diver would be unrecognizable.

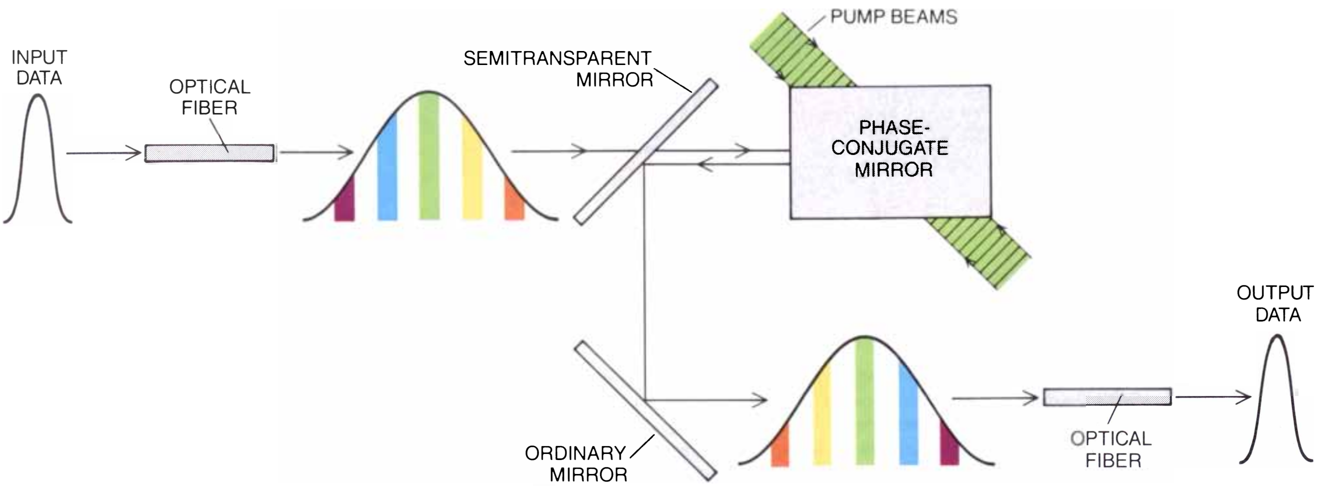
The image can be transmitted by replacing the projector with a four-wave mixer and equipping the friend with a beacon laser. The sequence of events is as follows. The friend aims his beacon laser toward the projectionist. The distorted beam, as seen by the projectionist, in essence samples, or probes, the atmospheric turbulence. If the projectionist allows the light to strike a conventional four-wave mixer, the "time-reversed" beam will undo the atmospheric distortions and make its way back to his friend with no information. By placing a transparency of the diver in the path of one of the pump beams, however, the phase-conjugate mirror can be spatially encoded with the wanted information. In this way the phase-conjugate mirror reflects a predistorted image of the diver. The result is that a unique "time-reversed" beam bearing a distortion-free image reaches the friend at the beacon laser.

Image processing also plays an important role in a variety of fields such as criminology (the comparing of fingerprints), biological sciences and medicine (classification and identification of various classes of cells and their mutations), artificial intelligence, robotics and automation systems. Sophisticated optical systems could facilitate these and other comparison techniques. Conventional holography has, for instance, been employed for some time to perform image-processing functions. The theory was developed in 1963 by Anthony B. Vander Lugt of the University of Michigan, who also demonstrated various classes of image filters and pattern-recognition schemes by means of conventional fixed holograms recorded on film emulsions. Four-wave-mixing technology, since it is a form of real-time, or dynamic, holography, can enhance the power of the technique. Jean-Pierre Huignard of the Thomson-CSF Laboratories in France, Feinberg and Y. H. Ya of the Telecom Research Laboratories in Australia have made significant contributions to this field.

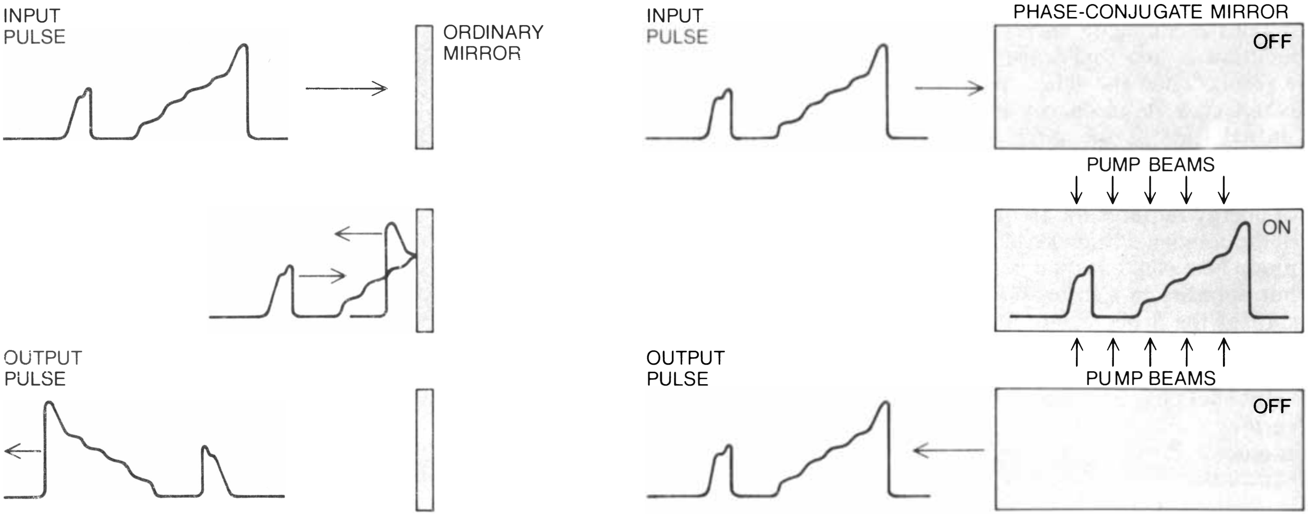
One potential application of four-wave mixing to image processing is a real-time pattern-recognition device. It was proposed by several colleagues and me at Caltech in 1978 and demonstrated by Jeffrey O. White (then at Caltech) and Yariv in 1980 and independently by S. G. Odulov and M. S. Soskin of the Institute of Physics of the Academy of Sciences of the Ukrainian S.S.R. during the same year. The scheme involves a four-wave mixer. The three input beams to the mixer (that is, the two pump beams and the probe beam) pass through respective transparencies whose spatial patterns are to be compared. The pattern formed by the resulting conjugate output beam reveals the overlap, if any, between the patterns of the transparencies. In White and Yariv's experiment, for instance, one slide bore the character string *C-A-L-T-E-C-H* and another slide bore the letter *C*. The output of the system was two bright spots, each of which "pointed" to one of the two occurrences of the letter in the character string.

Associative Memory

Implicit in the design of an ideal pattern-recognition device is the assumption that the test letter perfectly matches the key letter in the character string. What if the test letter differs somewhat from the key letter? Suppose, for instance, that the letters are



TEMPORALLY BROADENED PULSE can be reconstituted by a phase-conjugate mirror. The broadening occurs because the various frequency components (*shown here as different colors*) that constitute a pulse propagate at slightly different speeds through space. A four-wave-mixing phase-conjugate mirror (pumped at the central frequency component of the pulse) placed midway along a fiber-optic cable transmitting the pulse will rearrange the pulse in such a way that the slower traveling components enter the second half of the cable first. As the conjugated pulse propagates through the remainder of the cable the faster traveling components will therefore catch up with the leading (that is, slower) part of the pulse: the pulse is renarrowed and the cable can transmit more information.



TEMPORAL REVERSAL of a pulse sequence is achieved by a phase-conjugate mirror. An ordinary mirror (*left*) preserves the temporal ordering of a pulse sequence on reflection. A phase-conjugate mirror (*right*) will reverse the temporal ordering if the mirror is turned on when the entire pulse is inside the mirror. The mirror is actuated by applying two reference, or pump, beams to the mirror.

set in different typefaces (Gothic and Roman, say) or that one or both are blurred or even have missing areas. These examples are of particular interest because they correspond to daily experience: one recalls an incomplete image through a complicated process known as associative memory.

Teuvo K. Kohonen of the University of Helsinki and more recently John J. Hopfield of Caltech have developed models of associative memory, variations of which can be realized by means of a system containing phase-conjugate mirrors. The model might provide insight into how neural networks within the brain function. The basic idea consists of a mathematical algorithm that, in essence, compares a given input data string with one in memory, generates a new version of the data string and then iterates, or repeats, the procedure with the new version as the input. The comparison, selection and feedback process repeats until the input and output data strings converge, or agree. Many research groups, including those led by H. John Caulfield of the University of Alabama, Demetri Psaltis of Caltech, A. D. Fisher of the U.S. Naval Research Laboratories and A. A. Vedenov of the I. V. Kurchatov Institute of Atomic Energy in the U.S.S.R., have been exploiting electronic techniques (computers) and electro-optical techniques to simulate the process.

More recently Bernard H. Soffer, Dunning, Yuri Owechko and Emanuel Marom, all at Hughes, and, independently, Dana Z. Anderson of the Joint Institute for Laboratory Astrophysics in Colorado are developing architectures for all-optical associative memories. The Hughes scheme uses a conventional fixed hologram to store the memory information and phase-conjugate mirrors to provide the necessary image feedback [*see bottom illustration on page 79*]. Physically the device is simply a resonator capped with a phase-conjugate mirror at each end. The holographic memory element is positioned within the resonator cavity. One of the phase-conjugate mirrors has a threshold characteristic that enables it to select the features stored in memory that have the most in common with those of an input image for the next iteration.

Although digital computers can execute image processing and associative-memory functions, they do the job in discrete steps. The optical approach offers an advantage: image information can be processed in a parallel manner. In other words, the entire picture field is processed at the same time.

Frequency Effects

Throughout the discussion of the applications of four-wave mixing the implicit assumption has been made that the three input beams and the one output beam have the same frequency, or optical wavelength. Such a scenario is known as the "degenerate" case. In 1978 Richard L. Abrams of Hughes and I analyzed the case of "nearly degenerate" four-wave mixing, an example of which occurs when the frequency of the incident probe is slightly shifted with respect to the frequency of the two identical pump beams.

There are several situations in which nearly degenerate four-wave mixing occurs: the frequency of a probe beam reflected from a moving satellite would be Doppler-shifted with respect to the pump beams in a ground-based phase-conjugate mirror; two separate lasers (one to provide the pump beams and the other to provide the probe beam) that have slightly different frequencies may be involved in a given system; a pulse of light, which consists of a continuous band of frequencies, could be incident on a phase-conjugate mirror.

As Yariv, Fekete and I suggested in 1979, a phase-conjugate mirror could possibly be exploited to compensate for pulse-spreading effects in fiber-optic cables. Such effects limit the information-handling rate of a cable since densely spaced pulses that spread appreciably can spill into one another, rendering the data string unrecognizable. The spreading occurs because the frequency components of a pulse travel at slightly different speeds through a given material. If a four-wave-mixing phase-conjugate mirror is placed at the midpoint of a fiber-optic cable that is transmitting a pulse, the frequency components will be reordered so that the slower-propagating frequencies leave the mirror before the faster-propagating frequencies do. The faster-propagating ones will therefore catch up with slower ones on transit through the second half of the cable, restoring the quality of the pulse by renarrowing it to its original state.

How does a phase-conjugate mirror reorder the frequency components of a pulse? In the degenerate case the probe beam and one of the two counter-propagating pump beams generate a stationary optical interference pattern within the medium of the four-wave mixer. The other pump beam is reflected from the pattern and produces the phase-conjugate beam. In the nearly degenerate case the frequency shift causes the interference pattern to move through space, creating a moving phase-conjugate mirror within the medium of the four-wave mixer. As a result, when the other pump beam is reflected, its frequency is Doppler-shifted (just as the pitch of a train whistle changes when the train passes a stationary observer). Conservation of energy requires the frequency shift of the conjugate beam in relation to the pump beams to be equal in magnitude but opposite in sign to the frequency shift of the probe beam; the frequency distribution of the probe wave is "flipped" symmetrically.

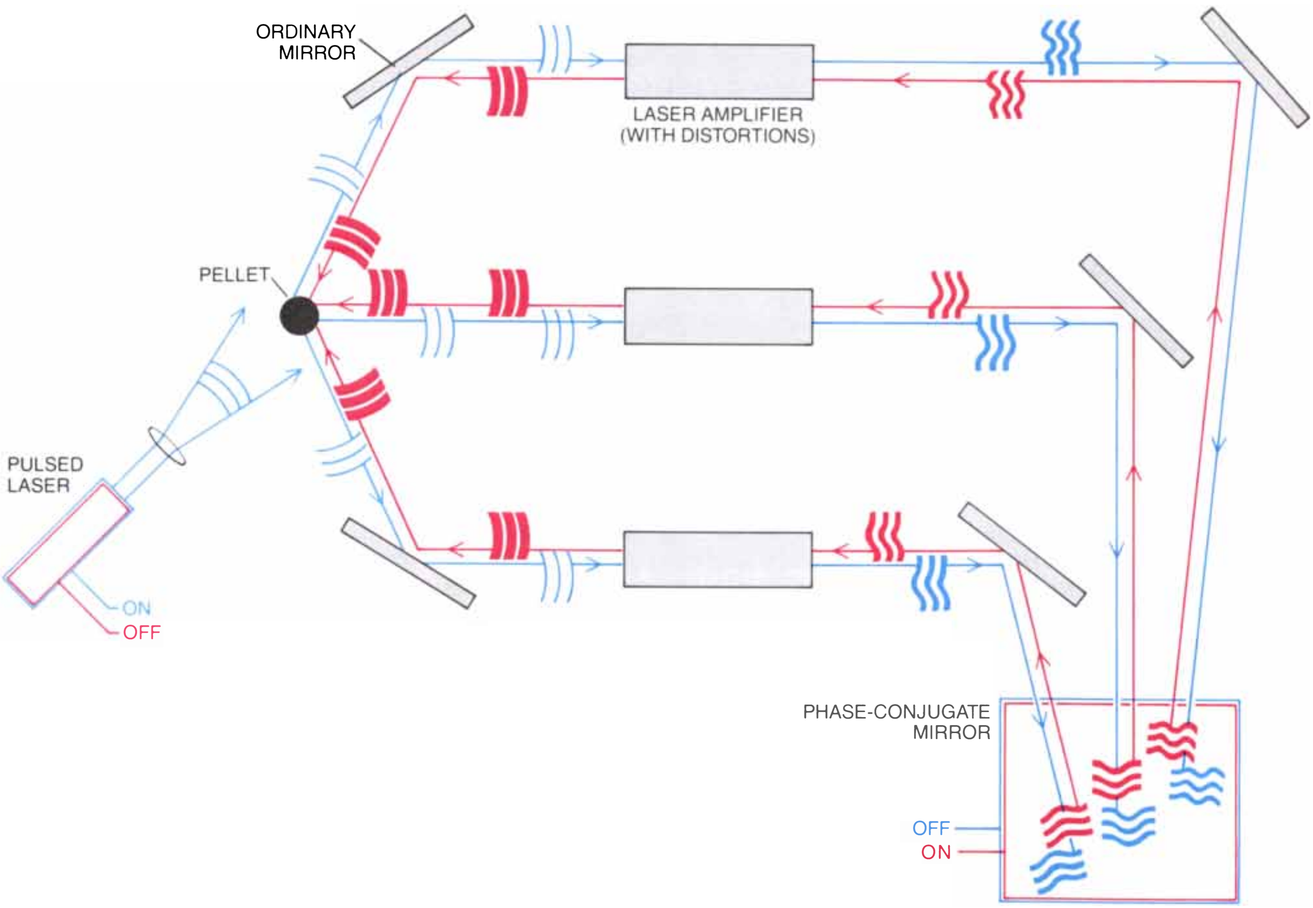
By changing the frequencies of the various beams that interact in a phase-conjugate mirror a so-called narrow-band reflective optical filter can also be realized. Typically such a filter will block (that is, not conjugate or reflect) all frequencies except those within a very narrow band that can be one ten-millionth of the optical frequency. In contrast, a simple piece of colored glass has a relatively coarse band-pass filter: it will pass light having a frequency that is within a band equal to about a hundredth of the optical frequency. The actual properties of the filter depend on the geometry of the phase-conjugate mirror and the specific nonlinear medium employed. The filter will not function, for example, unless the atoms of the medium can respond fast enough to the moving interference pattern. First proposed in 1978 by Abrams and me, phase-conjugate narrow-band filters have now been constructed in various laboratories, using a host of nonlinear mediums and laser sources. Although in the future the filters might replace conventional models in satellite and other communication systems, at present they are merely intriguing devices.

Time-Domain Applications

Another major class of information-processing techniques involves pulsing one or more of the interacting optical beams of a four-wave mixer to encode messages. Such a system could be used to effectively communicate with a satellite. A reflected wave can be generated by a phase-conjugate mirror only when the probe beam as well as both pump beams are illuminating the nonlinear medium. Hence by turning on or off one or more beams the reflectivity of the mirror is controlled and a message can be encoded in the output conjugate beam.

Suppose a satellite is equipped with a low-power beacon laser. Assume now that this probe laser is directed to a ground station equipped with a laser amplifier and a four-wave-mixing phase-conjugate mirror. If the conjugator is pulse-encoded, it not only will compensate for atmospheric turbulence and amplifier distortions but also will direct an encoded message back to the satellite. Moreover, under the proper conditions the system can track the spacecraft's motion.

A phase-conjugate mirror can also reverse a pulse sequence in time so that the last part of the sequence to enter the device is the first to leave. Such a pulse-sequence reversal device was proposed independently by David A. B. Miller, then at Heriot-Watt University in Edinburgh, O'Meara and Yariv.



PARALLEL AMPLIFYING SYSTEM could be used to initiate fusion. A fusion pellet is illuminated by a pulsed low-power laser (*blue*). Pulses scattered from the pellet pass through the three parallel laser amplifiers shown here. The intensity of the individual pulses is increased, but at the expense of introducing distortions. The pulses are then directed to a four-wave-mixing phase-conjugate mirror. By the time they reach the mirror they are out of step because they have traveled different distances. When all the pulses are inside the mirror, it is turned on, conjugating each pulse and reversing their temporal order with respect to one another (*red*). On the return trip the distortions are removed and the pulses become synchronized, so that an intense pulse of radiation hits the pellet.

It could have important consequences in laser-fusion research.

Imagine a phase-conjugate mirror so long that a pulse sequence can spatially fit in the nonlinear medium. (A pulse that has a duration of a billionth of a second would, for example, require a foot-long phase-conjugate mirror.) When all the pulse sequence is contained within the medium, one immediately turns on the phase-conjugate mirror by activating the two counterpropagating pump waves. Once the pump waves are activated, the last part of the pulse sequence to enter the medium will be the first part to leave it.

O'Meara has combined the features of the pulse-sequence reversal device with a double-pass compensation scheme. In this arrangement a phase-conjugate beam propagates backward through a medium to compensate for distortions. He has also proposed a system that could simultaneously synchronize pulses from a chain of laser amplifiers arranged in a parallel configuration. The system could illuminate a fusion pellet and compensate for optical distortions introduced in the process, thereby optimizing performance.

The freedom to encode temporally the various beams in an independent manner can be exploited to realize yet another class of all-optical processors known as time-domain convolvers and correlators. The correlator, as an example (proposed by O'Meara and Yariv), is in essence a temporal analogue to the spatial-domain, real-time pattern-recognition device and so can recognize a given pulse sequence. Other variations of such a processor can rapidly encode and decode temporal messages. These time-domain schemes could be employed in signal processors, radars and optical devices.

Govind P. Agrawal and Christos Flytzanis of QUANTEL in France have suggested that phase conjugation could also be used to sense the state of the optical logic elements forming the heart of potential optical computers. An optical logic element might consist of counterpropagating beams in a nonlinear medium, an arrangement resembling the configuration of the pump beams in a four-wave mixer. A probe beam properly directed into such a logic element would therefore be conjugated and could be read to ascertain the logic state. The probe would not affect the operation or modify the logic state of the memory element.

Fundamental Studies

The emergence of optical phase conjugation has unified many areas of applied and fundamental optical physics. Spectroscopy, the study of the interaction of matter and radiation, has particularly benefited. In this application the experimental picture is turned around: the three input beams (the probe beam and the two pump beams) are now used to investigate the properties of the nonlinear medium itself. One can, for example, explore the physics and spectroscopy of four-wave-mixing processes by measuring the intensity of the phase-conjugate beam as a function of such parameters as the polarization, frequency, intensity and angle of incidence of the pump and probe beams, the pressure of a buffer gas (in the case of a gaseous medium) and the intensity of applied electric and magnetic fields.

The application of these techniques to such nonlinear mediums as atomic vapors, aerosols, photorefractive crystals and various organic species provides another source of detailed information about fundamental atomic and quantum-mechanical processes. Even though the field of laser spectroscopy has existed for many years, the phase-conjugate geometry offers yet another tool for gaining insight into the interaction of light with matter.

In addition the quantum-optical properties of interactions producing phase-conjugate beams may be useful in realizing novel classes of ultrasensitive detectors, which might be sensitive enough to respond to gravitational waves or to tap an optical-fiber telephone line. Although the development is still in its infancy, research is being carried out on an international level.

The concepts, techniques and basic applications of optical phase conjugation can in principle be applied to most other areas of the electromagnetic spectrum. This may open the door to microwave phase conjugation, for instance, with obvious applications in the fields of radar, millimeter-wave imaging systems and high-frequency temporal signal processing, as well as microwave spectroscopy.

Groups in the U.S.S.R. and the U.S. are also considering such candidates as phonons (sound waves) for phase-conjugation experiments. Acoustic-wave signal-processing devices and sonar are two areas that could benefit from such research. The technology might ultimately allow for the detection and tracking of submarines.

Answers and Questions

Now that the reader has been exposed to many of the potential applications of optical phase conjugation, I shall provide the solution to the question posed earlier: What do you see when you look at a phase-conjugate mirror? In an ordinary mirror the viewer obviously sees his or her face. Some of the light scattered from various parts of the face will be reflected by the mirror and find its way into one's eye. The phase-conjugate mirror, on the other hand, would direct any light scattered from the face right back to its point of origin. Light scattered from the nose will therefore be reflected and sent back to the nose. The only light seen by the viewer would be the illumination light that initially scatters from the cornea, strikes the phase-conjugate mirror and returns to the eye. In other words, the viewer would see a uniformly lighted mirror surface with no obvious details. A second person looking at the viewer directly (that is, not by way of the mirror) would, however, see the viewer's face illuminated by the action of the phase-conjugate mirror.

As a rather speculative note on which to conclude, I shall point to an intriguing curiosity. The particle-wave duality of light and matter raises the question of whether the propagation of matter can be reversed. The nonlinear optical interactions leading to phase conjugation involve stimulated effects and are therefore restricted by a law in physics, the Pauli exclusion principle, to particles known as bosons. All bosons are integer-spin particles. (Spin is a parameter of quantum theory.) A photon, or quantum of light, for instance, has a spin of 1 and is therefore a boson. Electrons, protons, neutrons and certain unstable elementary particles, on the other hand, are fermions, or half-integer spin particles, and as such are forbidden by the Pauli exclusion principle to take part in stimulated interactions.

Many classes of bosons such as helium-4 nuclei (composed of two protons and two neutrons), Cooper pairs (electron-positron quasi particles existing in superconductors) and certain short-lived elementary particles (pi and K mesons, for example) can in principle take part in "time-reversed" matter-wave interactions. Superfluids and superconductors may be suitable mediums for such interactions. The actual realization, however, seems a faint possibility because of the weakness of most particle-particle and particle-photon interactions, the requirement of an intense, undistorted beam for observing the effects, and the short lifetime of many bosons. Yet before the invention of the laser the field of nonlinear optics (which is the backbone of optical phase conjugation), although known in principle, was itself a *gedanken,* or thought, experiment and was therefore ruled out as a viable experimental possibility.

Mineral Deposits from Sea-Floor Hot Springs

Seawater circulating through fractured volcanic rock above sources of heat participates in chemical exchanges with the rock. A major result is significant deposits of metal, some now uplifted onto land

by Peter A. Rona

When the theory of plate tectonics gained wide acceptance some 15 years ago, it brought good news and what seemed at the time to be bad news. The good news was that the theory explained the way the earth works: Continents are continually in motion and ocean basins open and close as new oceanic crust is created by sea-floor spreading and old oceanic crust is destroyed by reassimilation into the earth's interior. The apparent bad news was that the young age and homogeneous composition of the newly created oceanic crust would seem to preclude the formation of significant deposits of metals in ocean basins. Recent discoveries of metal-rich hot springs associated with mineral deposits in ocean basins have changed the picture. The finding may eventually lead to mining in the sea on a commercial scale. It also elucidates the origin of some of the major mineral deposits on land.

The older model of a static earth envisioned the ocean basins as passive sinks for the material weathered from continents and transported to the oceans, primarily by rivers. This model accounted for the only deposits of minerals on the sea floor then known. They consist of gold, tin, titanium, diamonds and other heavy materials mechanically eroded from exposed rocks on land and concentrated in pockets on the sea floor by flowing water, and also of phosphorite and manganese chemically precipitated in the form of nodules and crusts from materials dissolved in seawater.

As the dynamic-earth model of plate tectonics developed, it became evident that the ocean basins were themselves a source of heat and metals. The metals result from an exchange process between the ocean and the oceanic crust that is capable of concentrating metallic mineral deposits.

Such deposits could not form were it not that the oceanic crust is fractured. The crust fractures because it contracts as it cools from its original molten state to a solid. Movement of the crust as it forms at spreading centers and fluid pressures augment the effect. Hence seawater can penetrate through the crust, which is typically six kilometers thick, to the underlying mantle. Where heat sources exist in or under the crust, a convective circulation of seawater through permeable crust results. (In convection heat is transported by a fluid, which can be air as well as water.) A sea-floor convection system involves a downwelling of cold and dense seawater through permeable oceanic crust that is associated with a heat source. The water is heated and wells upward. Such a system, involving the convective circulation of hot and aqueous (hydrothermal) solutions, promotes both efficient thermal and chemical exchanges between the ocean and the oceanic crust.

The principal sources of heat are chambers containing magma, or molten rock. Heat also originates from chemical reactions between seawater and certain constituents of the mantle rock. An additional heat source is the geothermal gradient: the increase in the temperature of the earth with depth, primarily due to the dissipation of heat from the radioactive decay of elements in the crust and mantle.

Seawater and fractured oceanic crust are ubiquitous in ocean basins, but heat sources capable of convectively driving seawater through the crust are localized. They exist mainly at sites where molten rock separates from the mantle and rises into the crust to form magma chambers. The chambers are distributed along the boundaries that segment the earth's lithosphere, or rigid outer shell, into 10 major plates and numerous minor ones. Magma chambers are also found under volcanic sites in plate interiors away from plate boundaries, an example being the Hawaiian Islands within the Pacific plate.

The magnitude of the thermal and chemical exchange between the oceanic crust and seawater convectively circulating at plate boundaries is impressive. Thomas J. Wolery and Norman H. Sleep of Northwestern University estimated the circulation. They assumed that the difference between the calculated amount of heat delivered to sea-floor spreading centers (the boundaries where plates diverge) by the creation of oceanic crust and the average measured heat flow through sea-floor materials is attributable to cooling by the convective circulation of seawater. They estimated that the entire mass of the oceans circulates through the crust at oceanic ridges every 10 million years. On an annual basis the mass flow rate is comparable to that of the Amazon River.

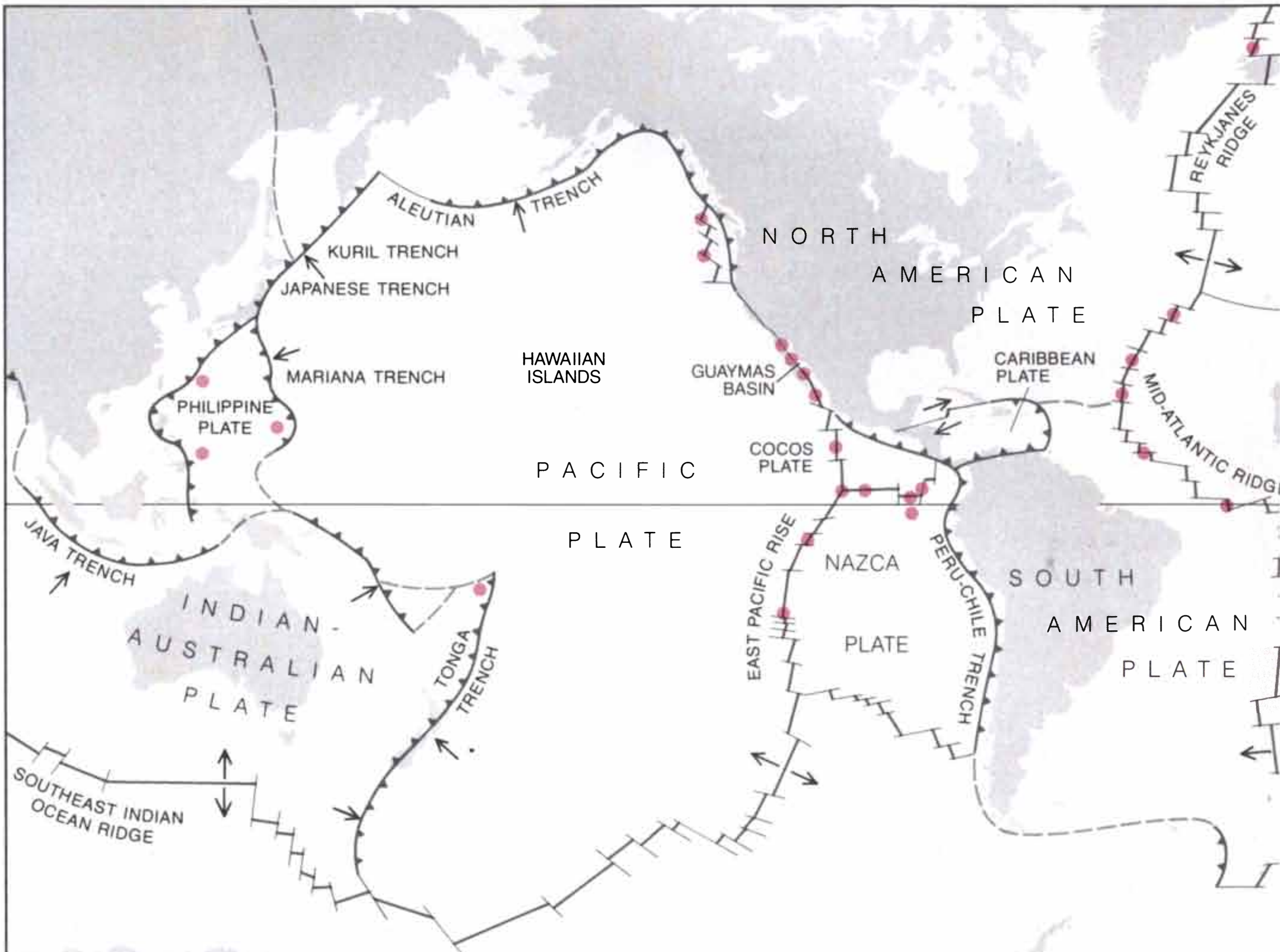
A two-way chemical exchange occurs as seawater convectively circulates through oceanic crust at high

SEA-FLOOR CHIMNEY several meters high on the East Pacific Rise vents hot water, precipitating particles of iron-copper-zinc sulfides assimilated in chemical reactions with fractured rocks in the oceanic crust. The investigators who found such chimneys in 1979 called them black smokers because they resemble smokestacks. The photograph was made by Kyung-Ryul Kim from the submersible *Alvin*. Part of the sampling equipment is visible.

temperatures. Certain elements and compounds, notably magnesium and sulfate, are removed from the water and transferred into the crust. Other elements, notably certain alkali metals such as lithium and potassium, alkaline earths such as calcium and barium and transition metals such as copper, iron, manganese and zinc, are dissolved from the crust and transferred to the circulating seawater, turning the water into a metal-bearing solution. John M. Edmond and his colleagues at the Massachusetts Institute of Technology determined that the magnitude of some of these exchanges is comparable to the input of the same elements and compounds by rivers carrying materials weathered from the continents to the oceans.

A circulating fluid is not enough: anomalous physical and chemical conditions must apply if ore-forming hydrothermal convection systems that concentrate mineral deposits are to arise. The physical conditions include (1) high thermal gradients to vigorously drive the upwelling limb of convectively circulating fluid; (2) a combination of low permeability in the zone of heat exchange (to maintain the temperature of the solution at several hundred degrees Celsius) and discrete zones of high permeability (to channel the upward flow and focus the discharge through an impermeable caprock, or overlying layer, so that deposits are concentrated rather than dispersed); (3) crustal movements to create a suitable distribution of permeable zones; (4) injections of magma to replenish the magma chamber energizing the hydrothermal circulation, and (5) continuing volcanic activity or sedimentation to bury a sea-floor mineral deposit and so protect it from oxidation and disintegration.

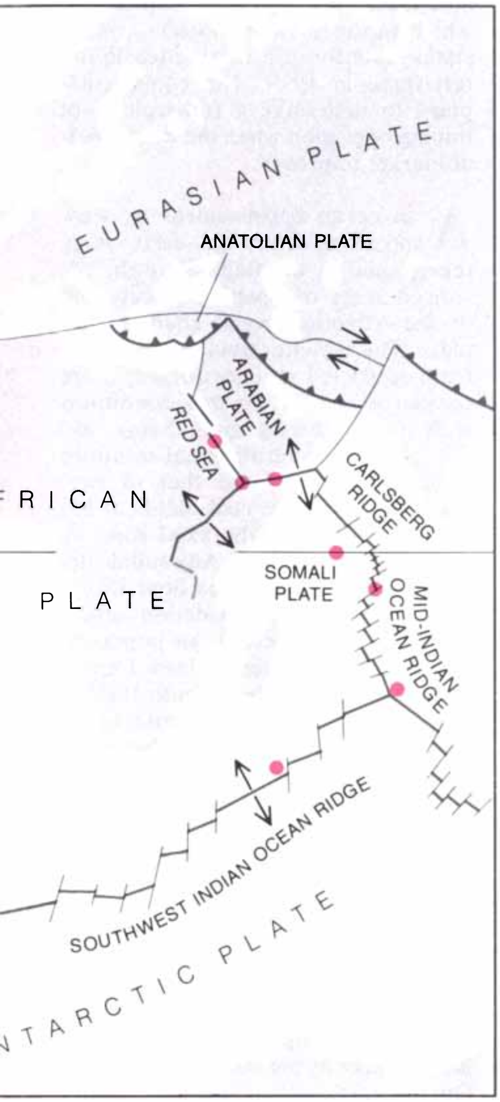
The chemical environment that facilitates the dissolution of metals from the crust and their transport, precipitation and concentration in a sea-floor deposit can be described as follows. Seawater that is cold, dense, alkaline, oxidizing and poor in metals penetrates down through permeable crust, assimilates heat by flowing near a magma chamber and rises in the upwelling limb. As the thermally expanded seawater moves up through the crust it takes part in the two-way chemical reactions with the surrounding rocks. The magnesium transferred from the water combines with hydroxide and



MINERAL DEPOSITS formed by sea-floor hot springs are identified by the colored circles on this map. The conditions necessary for the formation of such a deposit are seawater as an ore-forming fluid, fractured volcanic rock or sediment as a permeable medium and a source of metals, and a magma chamber as a source of heat. Favorable conditions occur at sites along the boundaries of the plates

silica to form mineral phases in the crust. The reaction is accompanied by the release of hydrogen ions, producing acidity in the solution. Seawater sulfate (SO_4) is reduced to sulfur by reaction with ferrous iron in volcanic rock and forms hydrogen sulfide (H_2S). An acidic solution evolves that aggressively dissolves metals existing at very low concentrations (parts per billion) in the oceanic crust.

Additional sources of metals are fluids derived from magma and volatile elements derived from the mantle (antimony, arsenic, mercury and selenium). The metals are transported, primarily as complexes with chlorine from the original seawater, at concentrations of from one to 100 parts per million in the hydrothermal solutions. Ed T. C. Spooner of the University of Toronto and William S. Fyfe of the University of Western Ontario have calculated that cooling 350 cubic kilometers of magma will heat 1,000 cubic kilometers of circulating seawater (equal to the annual runoff of the Chang Jiang, or Yangzi River) to 300 degrees C. At a metal concentration of one part per million that seawater has the capacity to deposit one million metric tons of metal.



that make up the lithosphere (the earth's rigid outer shell) and in volcanic regions within plates, as in the Hawaiian Islands.

A variety of chemical and physical mechanisms precipitate the metals from solution and concentrate them as mineral deposits, both along the upwelling zone within the crust and on the sea floor, where the solutions discharge as metal-rich hot springs. The two chief precipitation mechanisms are mixing and boiling.

Mixing of the hydrothermal solution with normal seawater may take place when the confining pressure of the overlying seawater prevents boiling, as is the case when the water depth is more than 2,000 meters and the water temperature is 350 degrees C. The result is a quenching of the solution, lowering its temperature and changing its composition from acidic to alkaline. Metallic minerals precipitate rapidly. Under reducing conditions (the absence of free oxygen) the metals combine with sulfur derived from seawater and dissolved from the crust to deposit polymetallic sulfides. Under oxidizing conditions the metals combine with oxygen, hydroxide, silica sulfate or carbonate to deposit metallic oxides, hydroxides, silicates, sulfates or carbonates.

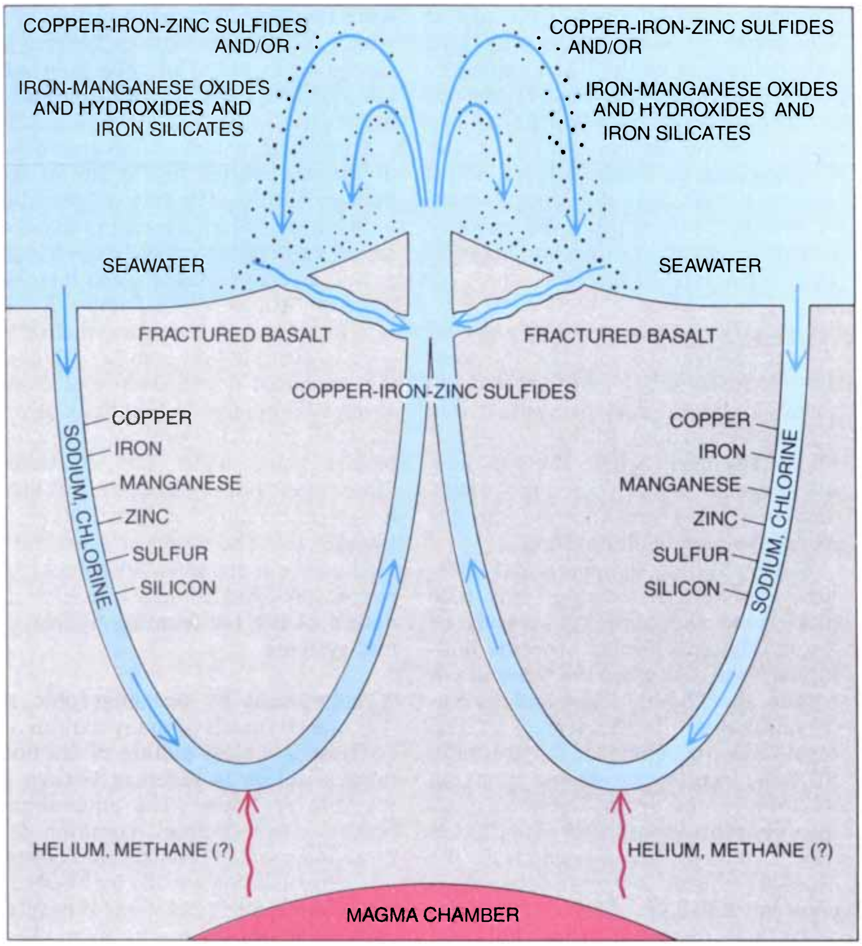
Boiling takes place when the confining pressure of the overlying seawater decreases to the critical point, which usually occurs at water depths of less than 2,000 meters for a solution at 350 degrees C. A liquid phase containing the sodium chloride separates from a vapor phase that retains the hydrogen sulfide. Boiling may result in the precipitation of polymetallic sulfide deposits under the sea floor or in the discharge of a hypersaline brine denser than normal seawater.

The anomalous physical and chemical conditions for ore-forming hydrothermal convection systems appear in sea-floor spreading centers that exhibit both common features and differences. All spreading centers have a segmented structure along their axes, consisting of linear segments some 10 kilometers long that alternate with transverse offsets at fracture zones ranging from a few kilometers to hundreds of kilometers in length. All spreading centers have an axial zone of volcanic activity about one kilometer wide, flanked by marginal zones within which the oceanic crust generated by spreading in the axial zone is broken into fault blocks by movements of the crust.

Spreading centers exhibit systematic differences related to the rate of sea-floor spreading. The rate of spreading on each side of an axis is classified as slow (two centimeters or less per year), as in the Atlantic Ocean, and intermediate or fast, as in the Pacific. As the rate of spreading increases, the difference in altitude between the axial and the marginal zones decreases (from several kilometers to hundreds of meters) and the underlying magma chambers become wider (20 kilometers rather than one kilometer) and shallower (one kilometer rather than six kilometers). The common features as well as the systematic differences between spreading centers influence the nature of the ore-forming hydrothermal systems.

Explorations by oceanographic research vessels of many nations are developing a clear picture of sea-floor mineralization in different settings of sea-floor spreading. The mineralizing process exhibits many common features and some differences that depend on the setting. Sites in the Red Sea represent mineralization in one setting: an early stage of the opening of an ocean basin, associated with a slow spreading center that bisects the Red Sea. There hot, metal-rich brines fill a number of basins along the axial zone. The brines are produced by ore-forming hydrothermal convection systems that operate in the basins. Cold, dense seawater that penetrates down through volcanic rocks in marginal zones becomes unusually salty because it passes through thick beds of rock salt (sodium chloride) buried in the crust. Such beds are characteristic of an early stage of the opening of an ocean basin, when under dry climatic conditions evaporation exceeds the replenishment of seawater because the inflow from exterior oceans is restricted by the surrounding landmasses.

As the salinity increases, salt crystals precipitate, settle and accumulate in thick beds on the sea floor. The high salinity of the circulating solutions enhances their capacity to transport dissolved metals as complexes with chloride from the salt. It also decreases their density, so that the heated solutions collect as density-stratified brines when they discharge from the floors of the basins. Metals precipitate from the hot brines as particulate mineral phases that settle in the basins and are trapped, forming layered deposits



MINERALIZATION PROCESS is shown in a cross-sectional view of two symmetrical hydrothermal convective circulation systems at a sea-floor spreading center. Seawater penetrates downward to a depth of several kilometers through fractured crust containing various metals. Heat from an underlying magma chamber causes the water to expand and to rise convectively. The heated water dissolves a number of elements, including metals, from the rock and loses a few elements. Additional metals may come from the magma chamber and the underlying mantle, along with helium and possibly methane gas. In "leaky" systems fresh seawater penetrates downward and mixes with the upwelling hot, metal-rich solutions, triggering the precipitation of metals as sulfides in the crust; metals remaining in solution are deposited on the sea floor from hot springs. In "tight" systems the solutions discharge directly into hot springs that deposit sulfides on the sea floor. The diagram is schematic.

CONSTITUENT	BASALT	SEAWATER	RIDGE/RIVER
COPPER (Cu)	–	+	
IRON (Fe)	–	+	
MANGANESE (Mn)	–	+	1
ZINC (Zn)	–	+	
POTASSIUM (K)	–	+	2/3
LITHIUM (Li)	–	+	10
RUBIDIUM (Rb)	–	+	
BARIUM (Ba)	–	+	2/3
CALCIUM (Ca)	–	+	1
SILICA (SiO_4)	–	+	2/3
MAGNESIUM (Mg)	+	–	1
SULFATE (SO_4)	+	–	1
SODIUM (Na)	+	–	VARIABLE
CHLORINE (Cl)	?	?	VARIABLE

EXCHANGE OF CHEMICALS between circulating seawater and the basaltic rock of the mantle is charted. The temperature is assumed to be between 200 and 400 degrees Celsius. A minus sign means the basalt or the seawater is a source of the element, a plus sign that it is a sink. The column at the right shows the ratio of the amount of the material introduced to the oceans by hot springs at oceanic ridges to the amount introduced by rivers flowing off the continents. With lithium, for example, ridges deliver 10 times more than rivers.

of metalliferous sediments up to 100 kilometers thick.

The metal-complexing capacity and the density stratification of the brines make the Red Sea hydrothermal convection systems exceptionally efficient in the transport and trapping of metals. The largest metal deposit known at any sea-floor spreading center is a layered, polymetallic sulfide body within the seven-kilometer diameter of the Atlantis II Deep in the axial zone of the Red Sea, in water 2,000 meters deep due west of Mecca. The deposit is estimated to contain a salt-free bulk dry weight of 100 million metric tons and to consist of 29 percent iron, 2 to 5 percent zinc, .3 to .9 percent copper, 60 parts per million of silver (6,000 metric tons at 1.8 ounces per ton) and .5 part per million of gold (50 metric tons at .015 ounce per ton).

This deposit ranks with the larger ancient sulfide deposits on land. The Saudi-Sudanese Red Sea Commission, which represents the adjacent coastal states, sponsored a mining-feasibility test there in 1979. The commission plans to undertake a full-scale pilot mining operation when the world metal market improves.

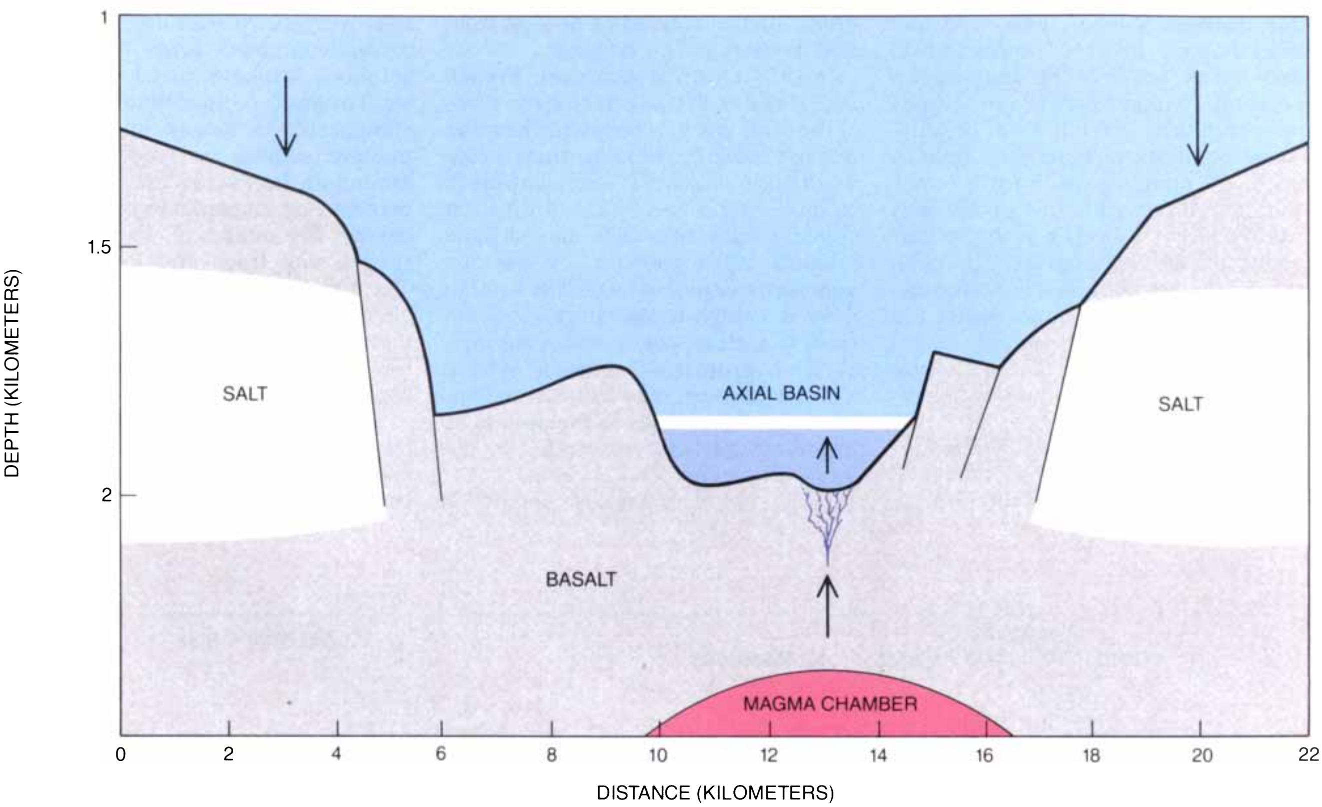
As an ocean basin widens by slow spreading from the early stage represented by the Red Sea to the advanced stage of opening represented by the Atlantic, several changes take place. The oceanic circulation changes from restricted to unrestricted; more oxygen is present, so that conditions shift from reducing to oxidizing, and the salinity of hydrothermal solutions generally decreases to that of normal seawater as the rock-salt beds are moved away from the axial zone by the spreading sea floor. Any sulfide deposits exposed on the sea floor would be decomposed by oxidation unless they were preserved by an impermeable cover of sediment or lava. Deposits such as those of the Atlantis II Deep may lie under such layers at sites along the margins of the Atlantic. Indeed, a strip of deposits may be concentrated in oceanic crust along the flow lines of sea-floor spreading from an enduring hydrothermal system such as the one at the Atlantis II Deep.

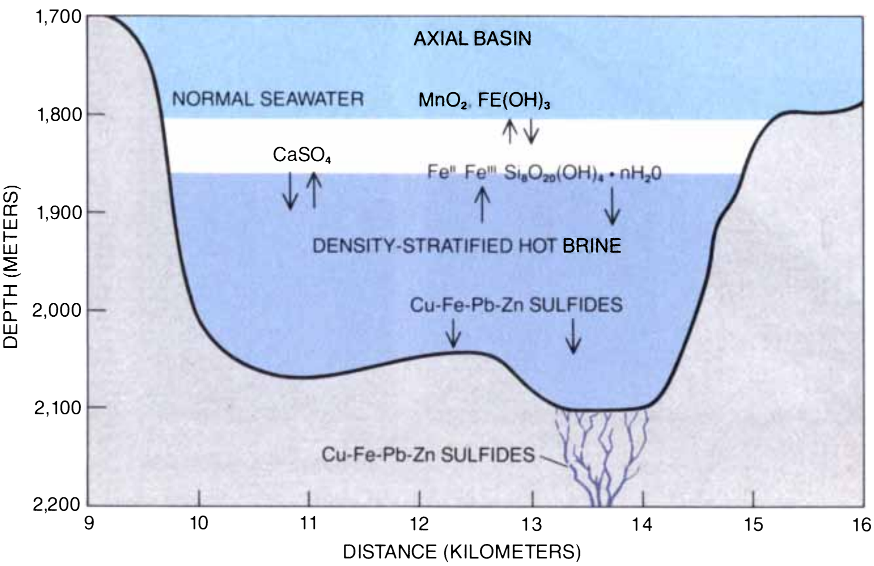
Evidence that deposits of this type exist was obtained by the ship *Discoverer* of the National Oceanic and Atmospheric Administration in 1972. The ship was making a transect to characterize the sea floor across the Mid-Atlantic Ridge at the latitude of Miami. On the east wall of the axial valley of the Mid-Atlantic Ridge the ship unexpectedly dredged a black, layered, crumbly specimen of rock. It was extremely pure manganese oxide (50 percent manganese) that had accumulated quite rapidly, indicating that it originated by precipitation from low-temperature (less than 200 degrees C.), metal-rich hot springs. The site has been named the TAG Hydrothermal Field (for the Trans-Atlantic Geotraverse project). It is still being investigated, in part through dives with the submersible *Alvin*.

The layered manganese oxide deposits produced by the discharge of low-temperature hot springs at the TAG field lie on top of the volcanic rocks of oceanic crust. The walls of large fracture zones that offset segments of the Mid-Atlantic Ridge provide partial windows deeper into the crust. In 1976 Enrico Bonatti, who was then at the University of Miami, and his colleagues dredged volcanic rocks from the walls of the equatorial Atlantic fracture zones that contained a network of copper-iron sulfide veins. Bonatti and his associates drew an analogy between these veins and a type of sulfide that underlies many ancient ore bodies, called massive sulfide bodies, on land.

The combination of such deposits below the sea floor and layered manganese oxide deposits on the sea floor could be produced by a "leaky" hydrothermal convection system. In such a system normal seawater leaks downward and mixes with upwelling hydrothermal solutions to precipitate either copper-iron sulfides, under reducing conditions within the crust, or layered manganese deposits, under oxidizing conditions where the residual solutions discharge at the sea floor.

Mineralization in the setting of an early stage of the opening of an ocean basin associated with an intermediate-to-fast spreading center is represented by sites in the axial zone





RED SEA sea-floor mineralization typifies one kind of setting in which the process takes place: an early stage of the opening of an ocean basin, associated with a slow sea-floor spreading center. The crust of the Red Sea basin contains thick beds of rock salt formed under earlier conditions. Seawater penetrating down through the crust becomes unusually salty; the salt both enhances the ability of the solution to transport metals as complexes with chloride from the salt and increases the density of the solution. The result is a density-stratified, metal-rich hot brine that forms ponds in axial basins as is shown at the left. The reactions represented are the precipitation of metals as copper-iron-lead-zinc sulfide particles under reducing (no free oxygen) conditions in the lower layers of the hot brines and as sulfate particles under oxidizing conditions where the upper layers of the brine mix with seawater. The particles settle to the sea floor to form layered deposits of metalliferous sediment.

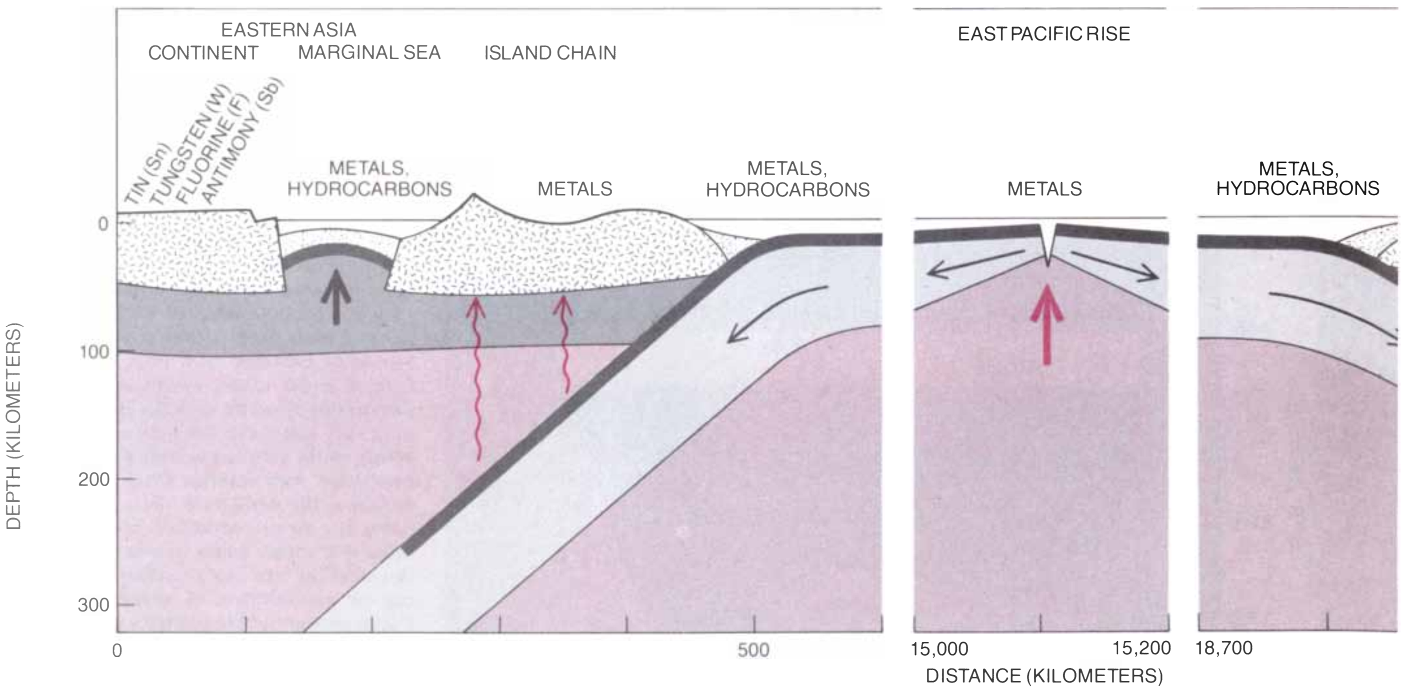
of the Gulf of California such as the Guaymas basin. This basin contains several hundred meters of sediment contributed by the surrounding rivers. The sediment is intermixed with tiny shells of calcium carbonate and silica from organisms that settle from surface waters. An ore-forming hydrothermal convection system is inferred, involving a downwelling of seawater through the sediment and its interaction with hot volcanic rocks underlying the basin to form hydrothermal solutions that dissolve metals from the rock and sediment. As the solutions well upward they become alkaline by reacting with the carbonate-rich sediment, and they precipitate metallic sulfide deposits (copper, iron, lead and zinc) that are conserved within the sea floor. At the sea floor the solutions deposit the remaining metals in various mineral phases in the form of pagoda-shaped edifices surmounting mounds up to 20 meters high. Such mounds were first observed in dives with *Alvin* led by Peter F. Lonsdale of the Scripps Institution of Oceanography in 1977. Where the hot solutions flow through sediments rich in organic matter the matter is cracked to form gas and oil that impregnate the sediment and rise from the sea floor as plumes of bubbles and droplets.

A site representative of mineralization in an advanced stage of the opening of an ocean basin lies south of the Gulf of California at 21 degrees north latitude on the East Pacific Rise. In 1978 a team of French, American and Mexican investigators, making a geologic transect across the site in the French submersible *Cyana,* happened to observe and sample curious mounds up to 10 meters high at a margin of the axial zone in water 2,600 meters deep. Some months later, workers in the IFREMER (Institut Français de Recherche pour l'Exploitation de la Mer) Laboratory in Brest determined that the material was composed of copper-iron-zinc sulfides, making the curious mounds the first massive sulfide bodies found on the sea floor. Massive sulfide bodies are economically among the most important types of ancient mineral deposits mined on land.

In 1979 a team of American, French and Mexican investigators made dives at the same site in *Alvin.* Near the relict massive sulfide mounds they made a dramatic discovery: turbulent black clouds of fluid that billowed up from chimneylike vents. They named these features black smokers by analogy with factory smokestacks. The venting fluid is a high-temperature (350 degrees C.), clear, acidic, reducing metal-rich hydrothermal solution. Mixing with the ambient seawater causes copper-iron-zinc sulfides to precipitate as fine black particles suspended in the buoyantly rising black plumes.

Individual black smokers persist for only a matter of years. Their short life is deduced from radiometric dating of chimney materials, the residence time of convectively circulating solutions, the age of associated organisms and the calculated rate of heat dissipation in excess of the heat replenishment by sea-floor spreading. The chimneylike vents rise as much as 10 meters above mounds of about the same height that are mainly composed of debris from the disintegration of the chimneys. A horizontal section through a chimney is likely to show several minerals within a radius of a few centimeters. Copper-iron sulfides precipitated from the hottest solutions line the orifice; moving outward, one finds iron-zinc sulfides intergrown with calcium sulfate precipitated from lower-temperature solutions that have mixed with seawater. The grade of the chimney metals is comparable to that of many ancient massive sulfides on land. A typical mound at the Pacific site contains 31 percent zinc, 14 percent iron, 1 percent copper, five ounces of silver per metric ton and trace amounts of gold. The quantity of metal in a typical single chimney mound (up to several thousand metric tons), however, is too small to qualify as a prospect for commercial mining.

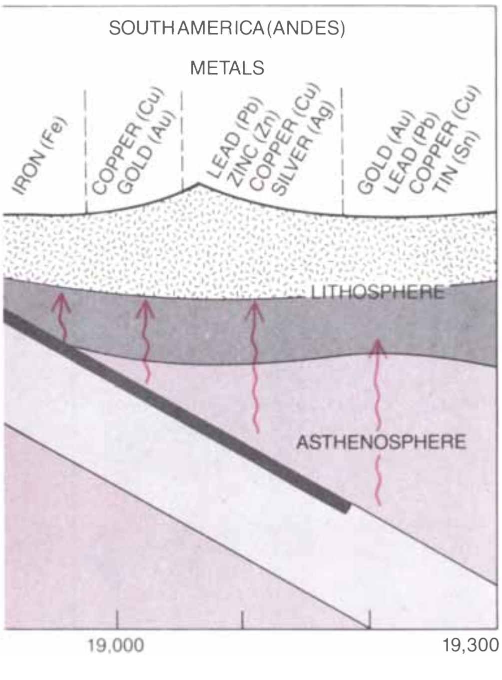
The hydrothermal convection systems that produce the black smokers are primarily "tight." Seawater penetrates one or two kilometers to a mag-



SETTINGS OF SEA-FLOOR HOT SPRINGS are shown in this schematic cross section of the Pacific Ocean. The settings are a spreading center (the East Pacific Rise), a volcanic island arc above a subduction zone at a convergent plate boundary and a marginal

ma chamber. The heated water wells up, evolves into a hydrothermal solution by chemical exchange with the volcanic rocks and discharges directly from the sea floor without significant prior mixing. Under conditions of unrestricted oceanic circulation on an oceanic ridge such as the East Pacific Rise this convection is inefficient as an ore-depositing system because nearly all the metals discharged in a black smoker literally go up in smoke and are dispersed by oceanic currents. Only small deposits remain in the form of mounds surmounted by chimneys.

Two volcanic seamounts that have summits 2,000 meters down in the water off the spreading axis of the East Pacific Rise were investigated in 1982 by Bonatti, Rodey Batiza of Washington University and Thomas Simkin of the Smithsonian Institution. In *Alvin* they dived into the summit calderas (circular depressions several kilometers in diameter formed by collapse resulting from the withdrawal of magma from an underlying chamber) of the seamounts. Various hydrothermal deposits were present in the calderas. The investigators saw to their surprise shimmering hydrothermal fluids that were vented from chimneys and actually flowed downhill. That meant the emerging hot fluid was denser than the ambient cold seawater. Bonatti and his colleagues inferred that the hydrothermal solutions had boiled. As a consequence a vapor separated from a dense liquid that had formed a pond in the caldera.



sea between the island arc and an adjacent landmass, which in this case is eastern Asia.

What are the applications of this emerging knowledge of sea-floor mineralization? Sea-floor mining may still be in the future, but the discoveries of metal-rich hot springs at sea-floor spreading centers are elucidating the processes that formed many ancient hydrothermal ore deposits now on land. Since the 19th century a few economic geologists had inferred that submarine volcanic "exhalations" involving the localized expulsion of fluids must have had a significant role in the formation of certain ore deposits. The recent discoveries fulfill that vision. Moreover, the oceanic hot springs offer natural laboratories where ore-forming processes can be observed directly.

A narrow interpretation of this application is to extend the observations to ophiolites: sections of oceanic crust and upper mantle generated at former sea-floor spreading centers and subsequently lifted onto a landmass. The hydrothermal deposits in ophiolites are regarded as direct analogues of the deposits formed in the sea-floor settings I have described. A classic example is the ore deposits of the Troodos Massif, a 2,000-square-kilometer area of ancient oceanic crust exposed on the island of Cyprus. The word copper was derived from Kypros, an old name of the island, and reflects the importance of the Troodos deposits as a source of that metal from preclassical to modern times. Some 90 massive copper-iron-zinc sulfide deposits occur as saucer-shaped bodies with dimensions of up to hundreds of meters in diameter and weights of up to 15 million metric tons. They fill depressions in volcanic lavas that erupted on the sea floor about 85 million years ago.

The Troodos Massif is part of a belt of ophiolites containing similar ore bodies that extends through southeastern Turkey, southern Iran and Oman. Other ophiolite belts of various ages containing hydrothermal mineralization extend through Tibet along the northern margin of the Himalayas; on islands of the western Pacific including the Philippines, New Guinea and New Caledonia; along the eastern and western margins of North America and the northern margin of South America; in Italy's Apennine Mountains; in western Scandinavia, and in the Ural Mountains of the U.S.S.R.

A broader application is to treat the sea-floor hot springs as natural laboratories in order to study hydrothermal ore-forming processes wherever the components of a hydrothermal convection system are present. Experimenters have caused rock to react with seawater at elevated temperatures and pressures in closed containers. They have found that reaction with diverse volcanic rock types or certain types of sediment produces acidic, reducing, metal-rich solutions similar to those discharged from sea-floor hot springs. The finding implies that sea-floor hydrothermal ore-forming processes transcend any particular setting and apply to such diverse settings as the spreading centers I have described, to volcanic arcs and to rift zones in continents.

The sea floor around volcanic arcs overlying subduction zones at convergent plate boundaries along the western Pacific encompasses the least explored and most promising settings. Among them are the marginal seas that lie between volcanic island arcs and an adjacent landmass. Examples are found near Japan, the Philippines, Indonesia and Melanesia. Sea-floor spreading centers exist in many of these marginal seas. Most of the massive sulfide deposits in ophiolites of the Cyprus type were formed at such spreading centers.

Another volcanic island-arc setting conducive to ore-forming hydrothermal convection systems is the island arc proper and the region between it and a deep-sea trench where oceanic crust enters a subduction zone. The Kuroko (meaning black ore) massive copper-iron-lead-zinc sulfides found on northern Honshu Island in Japan represent such deposits. They contain remnants of sulfide chimneys like those observed on the East Pacific Rise and are thought to have been formed by black smokers venting on the sea floor about 13 million years ago. There are similar deposits of various ages on the island of Fiji, in the Buchans mine of Newfoundland, at Avoca in eastern Ireland and at Captain's Flat and Woodlawn in eastern Australia.

In contrast to the Kuroko deposits in volcanic rocks of an island arc, the Besshi copper-iron-lead-zinc massive sulfide deposits of southern Honshu occur in sedimentary rocks interbedded with volcanic rocks. Steven D. Scott of the University of Toronto has drawn an analogy between the setting in which they were formed some 200 million years ago and the deposits forming in and on the sediments of the Guaymas basin.

In the perspective of earth history, episodes of local or global rifting of landmasses created settings favorable for sea-floor hydrothermal convection systems. The last global rifting episode took place about 200 million years ago, when the supercontinent of Pangaea began to break up. Rifts that opened (and incipient rifts that failed to open) in the Atlantic and the Mediterranean passed through stages similar to the events taking place now in the Red Sea and the Gulf of California. Hydrothermal deposits formed then that may now be buried under kilometers of sediments. When plate divergence was followed by the convergence and collision of plates, some of these deposits were uplifted and exposed on land, as they were in Cyprus.

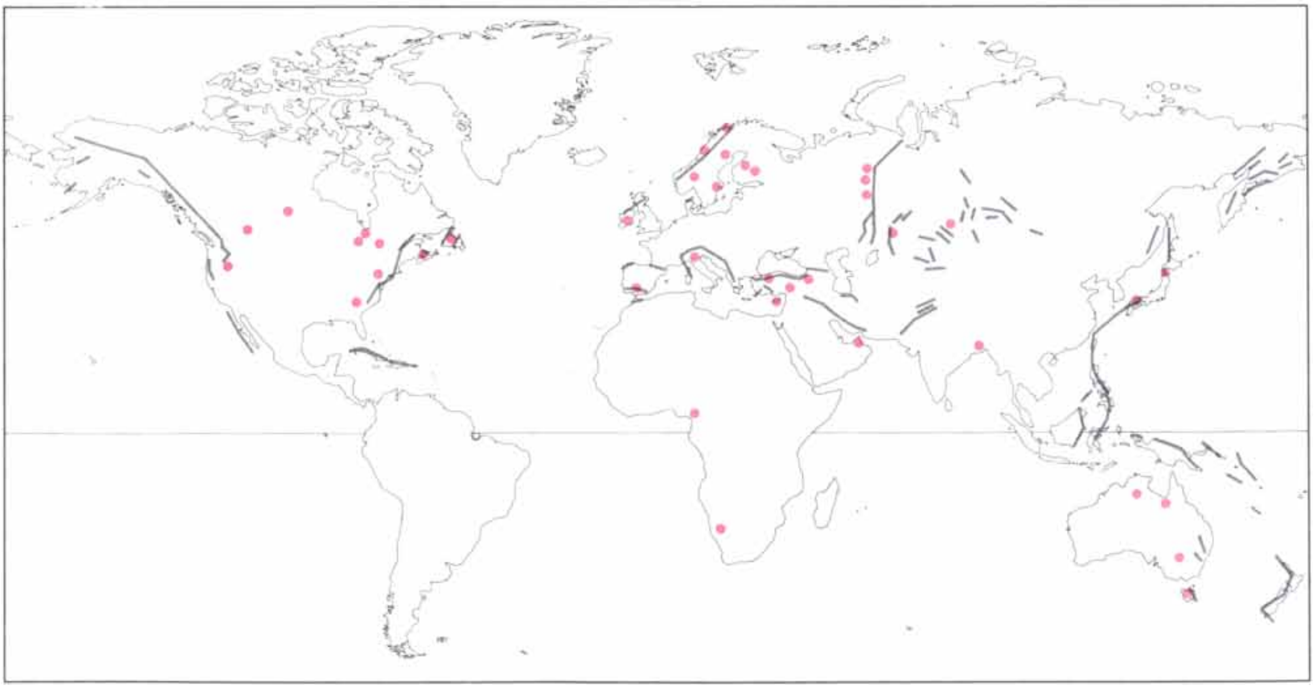
Similar episodes and sites can be identified in association with the earlier Proterozoic period. Frederick J. Sawkins of the University of Minnesota has interpreted a number of sedimentary deposits formed between 750 and 1,500 million years ago in the Proterozoic as the product of metal-rich hot springs discharging in settings like that of the Guaymas basin. The deposits include the Sullivan layered lead-zinc sulfide ore body of southeastern British Columbia, the layered silver-lead-zinc deposits of Mount Isa in Australia and the lead-zinc deposits of the McArthur River, also in Australia. Kurt Bostrom of the University of Stockholm has interpreted Proterozoic iron and manganese oxide ores of northern Sweden in terms of a rift setting.

Still farther back in geologic time, conditions conducive to sea-floor ore-forming hydrothermal convection systems prevailed between 2.6 and 2.7 billion years ago in rocks of the Archean period, which are exposed in the Canadian Shield of eastern Canada. More than 80 copper-zinc massive sulfide deposits are known to have formed during this period, including the largest one known (the 200-million-metric-ton Kidd Creek ore bodies in Timmons, Ontario) and deposits of the productive Noranda mining district of Quebec.

The analogies that have been drawn between such old ore bodies and the mineralization processes taking place in the sea today show how the thinking of geologists has reversed since the advent of plate tectonics 15 years ago. At the time it was the general geological opinion that mineralization was unlikely in a dynamic sea floor. The discoveries of mineralization at sea-floor spreading centers since then demonstrate otherwise. What oceanic crust lacks in age, diversity of composition and complex deformational history is made up for by the intensity of hydrothermal ore-forming processes wherever there are adequate magmatic heat sources at submerged plate boundaries and sites of midplate submarine volcanism.

An understanding is developing that links metal-rich hot springs in different sea-floor settings with ancient mineral deposits concentrated by such springs. The unifying process involves the exchange of heat and mass between volcanic rock and seawater circulating in sea-floor hydrothermal convection systems. The process is associated with crustal rifting and volcanic activity in the different settings of sea-floor spreading centers, volcanic island arcs and continents.

Variations of the basic process produce different kinds of deposits and may even converge to produce similar deposits. For example, similar saucer-shaped, layered massive sulfide deposits are formed by the development of excess salinity either by the dissolution of local salt beds, as in the Atlantis II Deep, or by boiling, as in the calderas of seamounts. In contrast, mound-shaped deposits may be concentrated by black smokers buoyantly venting from the sea floor. Variations in the metal content of the deposits are related to the metal content of the source rocks; for example, hydrothermal circulation derives more lead from continental rocks than it does from oceanic rocks. The interplay between the investigation of mineral deposits at sea-floor hot springs and the study of their ancient analogues is enhancing human ability to find large hydrothermal mineral deposits on land and advancing human knowledge of the effect of chemical and thermal exchanges between seawater and the oceanic crust in the sea floor.



ANCIENT MINERAL DEPOSITS originally formed by sea-floor hot springs and now on land account for a number of major sources of ore (*colored circles*). They include deposits in volcanic rock or in sediment that originally formed in a sea-floor setting. The gray lines indicate ophiolite belts: material that was generated at former sea-floor spreading sites and was later lifted onto a landmass.

Head into the future, in this Century.

Buicks are becoming more popular every year, and Century is the most popular of all Buicks.

Perhaps this is because Century embodies all of Buick's traditional virtues: comfort, luxury and convenience.

Perhaps also because these traditional virtues are engineered into a contemporary, stylish, front-wheel-drive package, with responsive performance and sophisticated road manners. Century has power-assisted rack-and-pinion steering and MacPherson strut front suspension, standard.

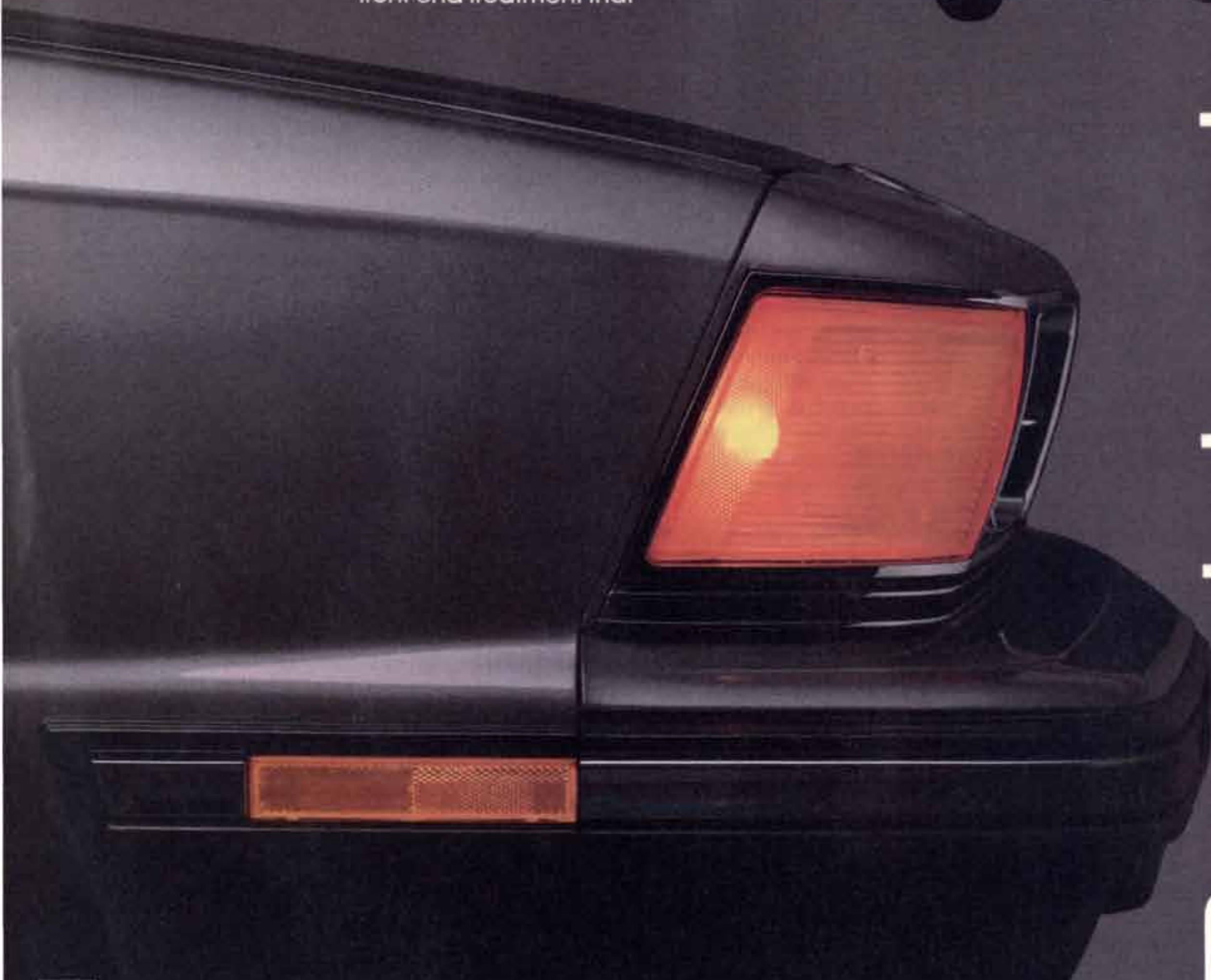
Then again, perhaps Century's popularity is due to continuing refinements, which will keep it contemporary for years to come. This year, for example, Century sports a rakish new front-end treatment that enhances both its appearance and its aerodynamics.

Or, perhaps Century's popularity is a result of all these attractions, in a uniquely satisfying combination.

Buckle up and visit your Buick dealer.

Then, head into the future — in this Century.

Wouldn't you really rather have a Buick?



BUICK

Buick Century. For more information or a test drive, call 1-800-86-BUICK (1-800-862-8425).

The Chemical Defenses of Higher Plants

Some plant-produced chemicals poison herbivores or repel them; others reduce plants' nutritive value or impede an insect's growth. Herbivores in turn have ways of exploiting these natural products

by Gerald A. Rosenthal

Even a casual consideration of the relation between higher plants and the multitude of animals that consume them makes one marvel at the plants' ability to survive. When it is attacked by a predator such as an insect, a plant can neither run away nor defend itself physically. Yet many plants have evolved subtle ways to protect themselves that are no less effective: they employ chemical defenses. Such defenses may be simple or elaborate. Some plants manufacture toxins that poison the attacking herbivore, whereas others produce complex compounds that interfere with the attacker's growth cycle or its ability to digest the plant.

Insects and other herbivores have in turn developed responses to this chemical warfare. Many herbivores have managed to adapt to the defensive mechanisms of plants, developing chemical defenses of their own. Some insects have developed ways to convert potentially harmful substances produced by plants into sources of nutrition or protection from insectivores. The study of such chemical interactions among organisms forms the basis for an emerging discipline known as chemical ecology.

The chemical ecologist studies the role of natural chemical products in the relations among organisms. One such relation is feeding: all plant-eating insects have basically the same nutritional requirements, which can be satisfied, more or less, by most higher plants. What is it, then, that determines an insect's specific feeding pattern? On what basis does it select or reject a particular plant as a food resource?

As the late Gottfried S. Fraenkel suggested, it is not a plant's primary metabolites (the substances it synthesizes that are essential for its growth and reproduction) that make it suitable or unsuitable as a source of food. Rather, the plant's suitability depends to a large degree on its secondary metabolites: metabolic compounds that are not involved in the common processes of life and that vary from plant to plant, helping to determine each plant's unique characteristics.

In 1971 the late Robert H. Whittaker and Paul P. Feeny of Cornell University added a new level of precision to Fraenkel's concept. They suggested that secondary metabolites produced by an individual of one species and able to affect the growth, health, population biology or behavior of another species should be called allelochemics. (Chemical ecologists now use the terms allelochemic and allelochemical interchangeably; I much prefer the latter term.) Among the many types of allelochemicals are attractants, repellents, allergenics and toxins. In this article I shall discuss the allelochemicals employed by certain plants to defend themselves from predation by insects and various other herbivores.

The customary way to determine the defensive capacity of a higher plant's allelochemicals is to demonstrate their toxicity toward one or more of a variety of insects that have come to be accepted as standard reference species in evaluating biological toxicity. The usual approach is to incorporate the natural allelochemicals into an artificial diet that would normally sustain the insect. J. M. Erickson, then a student of Feeny's, and Feeny, working on the black swallowtail butterfly, *Papilio polyxenes,* modified this method to provide a more natural approach. Instead of creating an artificial diet for their insects, they introduced a plant allelochemical into a plant that is part of the butterflies' natural diet.

Adult *P. polyxenes* avoid plants of the group Cruciferae (the mustards), which produce such allelochemicals as sinigrin, a compound that contains allylisothiocyanate, a toxic constituent. On the other hand, the butterflies forage avidly among the Umbelliferae, which include such plants as celery. Erickson and Feeny reared *P. polyxenes* larvae on a diet of celery leaves that had been induced to take up sinigrin. The larvae fed, and their growth was markedly inhibited. Celery containing a level of sinigrin equivalent to the level found in cruciferous vegetation was lethal to all the tested larvae. These experiments demonstrated that toxic allelochemicals could render an otherwise suitable host plant unacceptable to an insect pest.

David A. Jones and his colleagues at the University of Hull developed another experimental verification of the effectiveness of toxic allelochemicals. They studied bird's-foot trefoil and white clover, species that are capable of producing cyanogenic glycosides, compounds made of sugars bound to cyanide complexes, and storing them in their leaves. If two particular enzymes are present when the plant's leaves are damaged, the cyanogenic glycosides are broken down to release the cyanide complex, from which free cyanide is eventually liberated. Bird's-foot trefoil and white clover are "polymorphic" for cyanogenesis: only some individual plants can produce both the cyanogenic glycosides and the enzymes required to liberate cyanide. Hence not all plants of each species can defend themselves by means of cyanogenic glycosides.

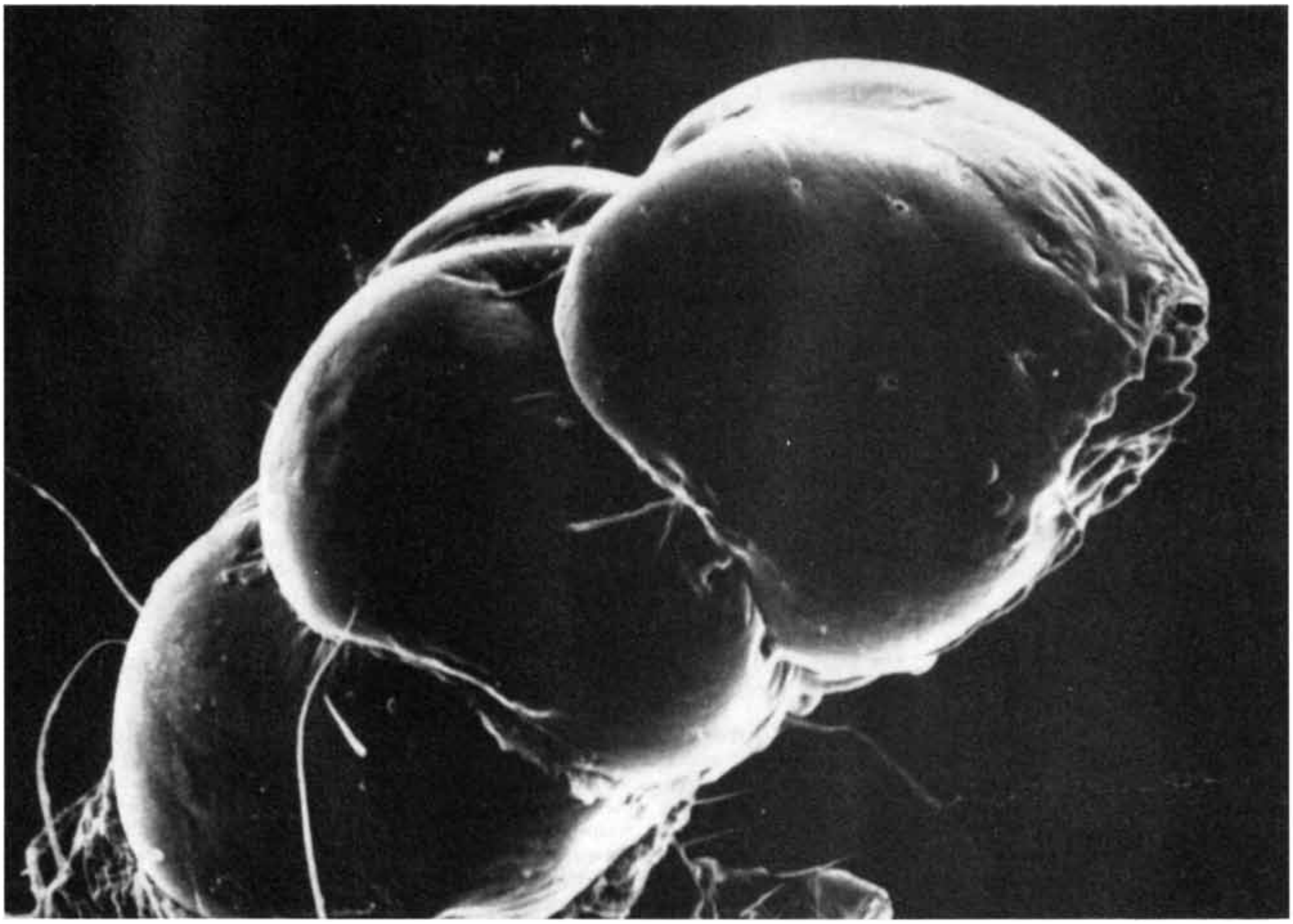
Jones exploited this peculiar property to determine how effective a stratagem cyanogenesis is for the plant population within his region of study. He examined a map, published in 1954 by Hunor Daday, then at the Welsh Plant Breeding Station in Aberystwyth, showing the geographic distribution of plants able to synthesize both cyanogenic glycosides and the appropriate enzymes and thus able to produce free cyanide. Daday had found a dramatic relation between the January mean temperature of the regions within his study area and the proportion of clover plants in those regions that could produce free cyanide. In warmer regions, such as the area around the Mediterranean Sea, about 70 to 90 percent of the collected plants exhibited cyanogenesis, whereas in colder regions, such as portions of the U.S.S.R., almost none of the analyzed clover was cyanogenic.

Jones and his colleagues observed that slugs and snails, two major predators of the bird's-foot trefoil, tend to consume acyanogenic plants rather than cyanogenic ones. The ability to produce free cyanide would therefore be more advantageous to plants in warmer habitats, where slugs and snails remain active during the winter. Cyanogenesis would be less valuable to plants in areas where winter weather controls herbivore populations.

Another kind of defense employing toxic allelochemicals has been described by Eloy Rodriguez of the University of California at Irvine. In a noteworthy study of the vegetation of Baja California and Chihuahua he has found that the trichomes, or hairs, of many desert plants are storehouses of toxic natural products. The trichomes of the desert plant *Phacelia,* for example, contain numerous poisons, insecticides and allergenic substances. The trichomes of another desert plant, the thistle *Parthenium hysterophorus,* contain certain allergenic chemicals that deter herbivores from feeding.

Some allelochemicals protect plants that produce them not by poisoning or repelling herbivores but by interfering with the predators' normal cycles of growth and development. Many insects grow in three morphologically distinct stages: larval, pupal and adult. Some go through several larval stages before pupating. Certain other insects do not go through morphologically distinct stages; the hatchlings, which resemble adults in appearance, do go through several periods of ecdysis, or molting, however, growing in size at each stage. Two types of hormones that play a key role in insect development are the juvenile hormones produced by larvae and a hormone group known collectively as ecdysteroids, which act to initiate the cycles of ecdysis that occur within the developmental sequence.

Adolf F. J. Butenandt of the Max Planck Institute for Biochemistry in Munich and Peter Karlson, who was then at the University of Tübingen, were the first investigators to isolate ecdysteroids. From nearly 1,000 kilograms of silkworms they eventually isolated 25 milligrams of an ecdysteroid called ecdysone and about a third of a milligram of another one called 20-hydroxyecdysone. Soon afterward



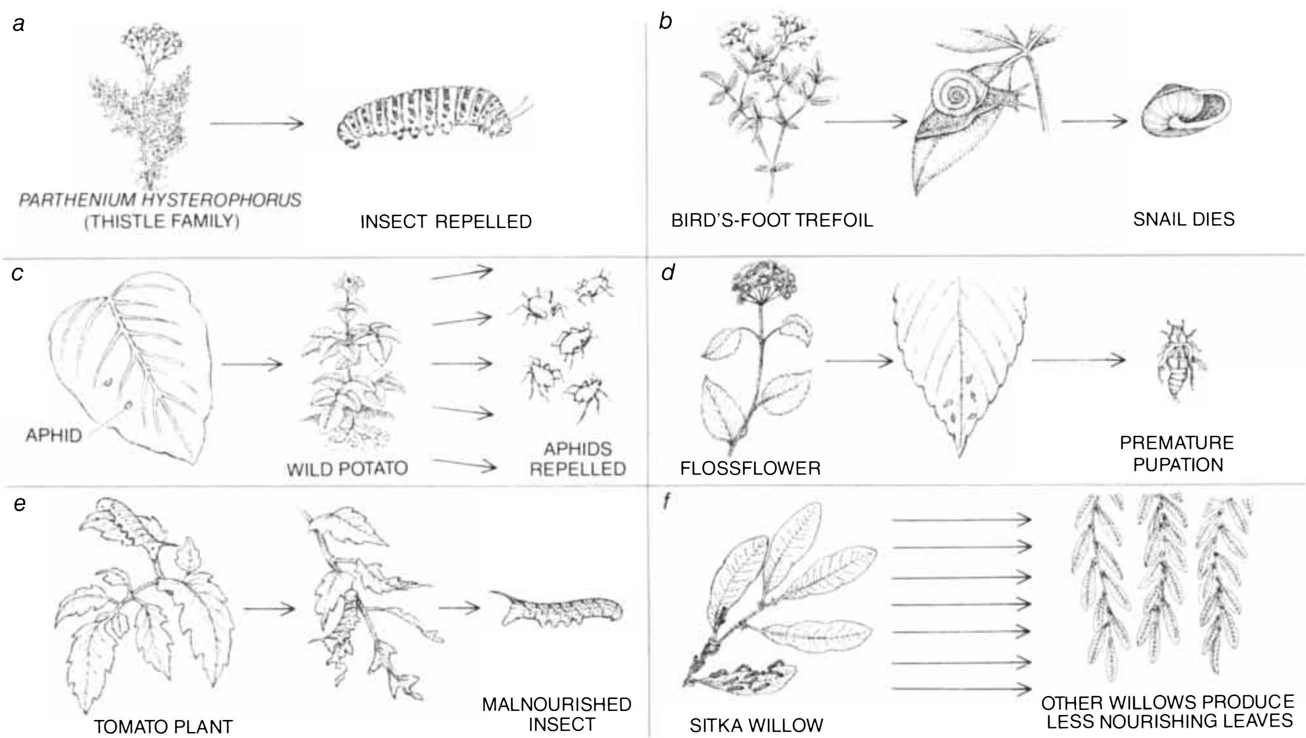
ABNORMAL PUPA with three heads develops when the larva of the fall armyworm eats an extract of *Ajuga remota,* a bugleweed. The plant produces phytoecdysones: compounds that mimic normal growth hormones, called ecdysteroids, of the larva. Phytoecdysones cause the insect to undergo the cellular events that normally precede metamorphosis several times without actually shedding the larval exoskeleton, each time developing a new head capsule. The resulting extra head capsules block the function of the mouthparts, and the larvae starve. Mimicking a predator's natural hormones is one of the more sophisticated forms of chemical defense practiced by plants. Other methods include poisoning the herbivore or making substances that can deter it from feeding or from depositing its eggs.

Koji Nakanishi, now at Columbia University, and T. Takemoto of the University of Tokushima discovered that certain plants are excellent sources of ecdysonelike materials. They were able to extract 25 milligrams of 20-hydroxyecdysone from a mere 2.5 grams of a dried rhizome (a subterranean food-storage organ) of the common fern, *Polypodium vulgare*. That inspired a search by many other investigators for other phytoecdysones (plant-produced ecdysteroids). Eventually several dozen structurally different compounds were isolated, a few of which were found to be even more potent than similar compounds that occur naturally in insects.

In recent years Isao Kubo and his colleagues at the University of California at Berkeley, working in Kenya, confirmed that phytoecdysones can act as powerful protective agents by disturbing the growth cycles of insects that prey on phytoecdysone-producing plants. These workers found that after swarms of locusts had virtually denuded broad stretches of savanna vegetation the sole surviving plant species was a bugleweed, *Ajuga remota*. When they fed extracts of *A. remota* to a number of insects, a striking developmental aberration resulted: in metamorphosing from larvae to pupae, each insect grew not one but several head capsules. The insects' extra head capsules blocked their mouthparts and they starved. Kubo established that this developmental abnormality was due to the presence in the plant of several phytoecdysteroids that had blocked the normal metamorphosis from larva to pupa. In each insect the cellular events that normally initiate metamorphosis had occurred but the actual shedding of the larval exoskeleton had been aborted.

Just as *A. remota* protects itself by producing analogues of ecdysteroids, so certain other plants employ analogues of juvenile hormones. In many insects juvenile hormone acts as a control on the process of development. As long as the growing larva produces juvenile hormone, ecdysteroids can initiate a molt only from one larval stage to the next; the insect cannot metamorphose into a pupa unless the juvenile hormone is degraded and thus unable to influence the process of pupation directed by ecdysteroids. Juvenoids, which are substances that resemble juvenile hormones but are produced by plants, can be of significant protective value to the plant if they are lethal. (Otherwise they may only prolong the insects' larval stage, which, ironically, is typically the most destructive phase of an insect's life.)

The first juvenoids were discovered in 1964, when Karel Sláma came from the Entomological Institute of the Czechoslovak Academy of Sciences to work with Carroll M. Williams



HIGHER PLANTS' CHEMICAL DEFENSE METHODS may range from the simple to the elaborate in their mode of action. Some plants, such as the thistle *Parthenium hysterophorus* (*a*), manufacture repellent chemicals that discourage predators from feeding or laying eggs. Others, such as the bird's-foot trefoil (*b*), manufacture lethal toxins; some bird's-foot trefoil plants produce cyanogenic glycosides (compounds consisting of sugars bound to cyanide complexes) and enzymes that liberate cyanide from the cyanogenic glycosides. More complicated defensive strategies consist in producing natural substances that mimic substances normally produced by a herbivore, as *A. remota* mimics natural growth hormones (*see illustration on preceding page*). Another such plant is the wild potato *Solanum berthaulthii* (*c*), which synthesizes an ingredient of the aphid's alarm pheromone (the substance an aphid can release when under attack to warn others that an insectivore is present). Aphids usually flee from a pheromone-releasing plant as they would from a predator. Other plants defend themselves by blocking the function of an insect's biochemicals. For example, the flossflower, *Ageratum houstonianum* (*d*), manufactures a substance that blocks insects' juvenile hormone, killing larvae by forcing them to molt prematurely. Many plants do not keep permanent stores of defensive chemicals but manufacture them only in response to predation. When the tomato (*e*), for example, is attacked by a chewing insect, it is stimulated to manufacture proteinase inhibitors, macromolecules that inhibit the insect's digestion of many plant proteins. A more exotic kind of inducible defense may also occur. When a Sitka willow (*f*), is attacked by certain caterpillars, the nutritional quality of the leaves of neighboring willows—even those that have not been attacked—deteriorates. Perhaps the tree emits a substance that acts like a pheromone, warning other willows to prepare for an attacking herbivore.

at Harvard University. The two investigators planned collaborative studies of *Pyrrhocoris apterus,* a bug found in Europe. *P. apterus* normally goes through five distinct stages before metamorphosing into an adult. Sláma and Williams made the surprising discovery that insects reared at Harvard did not ecdyse from the fifth stage to the adult; instead they went through a sixth and sometimes even a seventh stage. Sláma had never observed such a growth pattern in Czechoslovakia.

Eventually it became evident that the insects must have been exposed to a juvenile hormone, which was interfering with normal development. A series of tests revealed the critical factor to be the paper placed in the petri dishes that housed the bugs. The investigators tested a variety of paper products and found that many American newspapers and journals caused the same anomalous effect, whereas similar paper products from European and Japanese sources did not. (In Czechoslovakia, Sláma had naturally lined his dishes with European-made filter paper.) Sláma and Williams were able to isolate the active factor from paper towels and to determine that it was effective only against the Pyrrhocoridae family. A closely related group of bugs, the Lygaeidae, was totally unaffected by the factor.

Finally the active factor was traced to the wood pulp from which the paper had been made. Pulp derived from the balsam fir, a primary source of pulp for North American paper products, was found to be particularly active. Several years later William S. Bowers, then working at the New York State Agricultural Experimental Station, isolated and characterized the active factor, which he named juvabione. It is quite similar in structure to insect juvenile hormones.

A few years afterward, Bowers decided to determine whether other plants protect themselves by blocking, rather than imitating, an insect's normal juvenile hormone. A juvenile-hormone antagonist could kill insect larvae by causing them to molt prematurely into the adult stage. Eventually he succeeded in isolating two such substances from *Ageratum houstonianum,* a small plant that grows in temperate regions. He named the two compounds precocene I and precocene II because of their ability to elicit precocious metamorphosis by preventing the secretion of juvenile hormone. Precocene II has a number of other effects as well: it terminates the production of sex attractants by the American cockroach, causes several types of insects to lay infertile eggs and forces the Colorado potato beetle to enter a hibernational state called diapause. When insects that have been fed precocene II are later treated with juvenile hormone, the deleterious effects of the precocene are reversed.

Plants can mimic many other substances that are naturally secreted by insects. For example, the aphid *Myzus persicae,* which preys on such plants as the wild potato, secretes a substance called an alarm pheromone when it is attacked by a predator. The volatile pheromone alerts other aphids to impending danger. An important ingredient of the aphid's alarm pheromone is a compound called (*E*)-beta-farnesene. Richard Gibson and John A. Pickett of the Rothamstead Experimental Station in England have found that *Solanum berthaulthii,* a wild, tuber-bearing potato, releases (*E*)-beta-farnesene from trichomes on its leaves. Thus the plant can repel a major pest by mimicking that pest's alarm signal. The insect cannot readily counter or adapt to this type of defense, since it cannot ignore the very chemical signal that is critical to its survival.

Some insects have, however, been able to adapt to the chemical defenses of plants. In many cases adaptation has taken a remarkable turn: sometimes the insect can store and make effective use of the chemical the plant employs as a defense.

One example of such an insect is the grasshopper *Poekilocerus bufonius,* which feeds solely on plants of the Asclepiadaceae, or milkweed, family. The milkweeds manufacture a number of complex compounds known as cardenolides, toxins that can severely disrupt normal cardiac function.

When this grasshopper is attacked by a potential predator, it can defend itself by ejecting a spray from a poison gland. Analysis of this fluid reveals that it contains two major cardenolides, calactin and calotropin, both of which can also be extracted from the milkweeds on which the grasshopper feeds. When a grasshopper is maintained on a diet that includes no milkweeds, the cardenolide content of its protective spray is reduced tenfold. It seems evident that the plant itself is the source of the defensive compounds in the insect's defensive fluid.

The existence of insects, such as the grasshopper, that are able to store and utilize a plant's defensive chemicals led Whittaker and Feeny to refine the concept of an allelochemical by dividing the allelochemicals into several groups. Two groups are of particular importance: allomones, which confer an adaptive advantage on the organism that produces them, and kairomones, which confer an advantage on the organism that receives or consumes them. In the case of the grasshopper and the milkweeds, a chemical that may originally have been an allomone (the cardenolide) has become a functional kairomone, benefiting the insect that consumes the plant.

The grasshopper's adaptation to the milkweeds' defensive chemicals may be one example of coevolution. Coevolution is a reciprocal process in which the properties and characteristics of one organism evolve in response to specific properties of another organism; the two interacting species exert pressures that influence each other's genome. Some coevolutionary relations between plants and insects have become remarkably specialized. For example, certain flowers can be pollinated only by a single species of insect; the insects in turn exhibit absolute fidelity to flowers of their host plant.

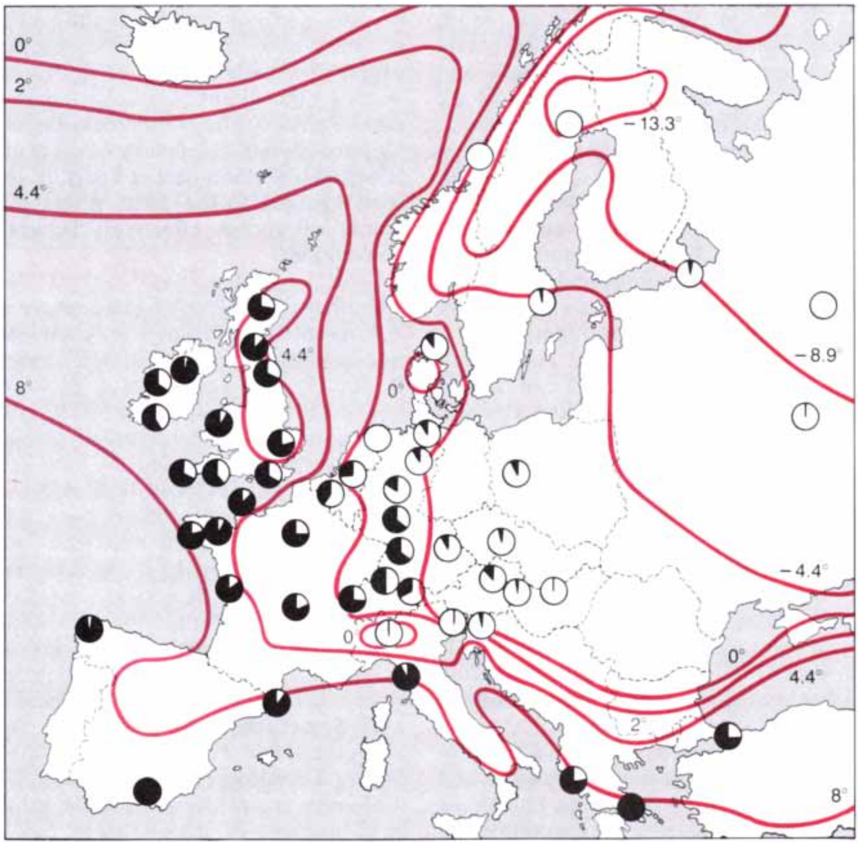
There is considerable interest in reconstructing the extent to which higher plants and their insect pests and predators may have evolved together over evolutionary time. In the case of defensive chemicals, perhaps higher plants evolved toxic natural products as part of their defensive barrier against herbivores. Certain herbivores may simply have avoided the plants, whereas others may have evolved an ability to detoxify the plants' defensive chemicals. Perhaps plants counteradapted by intensifying the efficacy of their allelochemicals, and perhaps certain insects evolved to the point where the plants' allomones effectively became kairomones.

Another example of an insect that exploits a plant's defensive chemical as a kairomone is the monarch butterfly. In 1967 Lincoln Pierson Brower, now at the University of Florida, observed that monarch butterflies grown from larvae that had fed on *Asclepias curassavica,* a milkweed that stores a good deal of cardenolides, were unacceptable as food for blue jays [see "Ecological Chemistry," by Lincoln Pierson Brower; SCIENTIFIC AMERICAN, February, 1969]. Jays that consumed the butterflies, or even certain parts of one butterfly, became violently sick. Brower's finding was consistent with observations made by other investigators that butterflies of the subfamily Danainae, which includes the monarch, are rejected on sight by a large number of insectivorous birds. Brower analyzed the compounds in adult and pupal monarchs and found about 10 different cardenolides; the total amount of toxin in the body of one butterfly is several times higher than the amount necessary to kill a cat or a small dog. The insects rarely store all the cardenolides found in the host plant, but all the cardenolides the insect does store have a counterpart in the plant. Butterflies grown from larvae maintained on a cardenolide-free diet had no harmful effect on blue jays.

The monarch butterfly is aposematic: it "advertises" its intrinsic toxicity by the bright coloration and distinctive markings of its wings. The presence of such aposematic insects as the monarch butterfly has spawned an array of mimics. These mimetic insects, such as the viceroy butterfly, do not store cardenolides, but they resemble insects that do. Thus they are placed under the cardenolide umbrella of protection without the need to produce, regulate, store or utilize the toxic chemicals.

Another interesting dimension has been added to investigators' knowledge of these interactions by Brower's field work in central Mexico, the wintering ground for a vast number of monarch butterflies. In this habitat, assaults by black-backed orioles and black-headed grosbeaks account for more than 60 percent of the butterflies' mortality. Detailed observation of the feeding activity of these insectivores revealed that the two species of birds feed quite differently. The orioles, which are sensitive to cardenolides, pick the butterflies apart; they selectively remove the thoracic muscles and abdominal contents without eating the cardenolide-laden cuticle (the outer covering of the body) or wings. Grosbeaks, on the other hand, are much less sensitive to the toxins, and they feed randomly, voraciously eating the intact abdomen. No defense is inviolate, and, like the milkweeds, monarch butterflies are not completely protected by their chemical defenses.

One of the most striking examples of an insect's adapting to the point where a plant's allomone has become a kairomone comes from the study of *Caryedes brasiliensis,* a beetle of the Bruchidae family. It is the sole predator of the seeds of *Dioclea megacarpa,* a vinelike legume that grows in the deciduous forests of Costa Rica. About 13 percent of the dry matter of the *Dioclea* seed is made up of L-canavanine, an amino acid that normally is not incorporated into proteins and that can exhibit potent insecticidal properties.



ABILITY TO PRODUCE FREE CYANIDE varies among individual bird's-foot trefoil plants. A large percentage of plants in warmer regions manufacture both cyanogenic glycosides and the enzymes necessary to liberate cyanide from them. (The percentage of plants in a region that are cyanogenic is indicated by the darkened part of the circle in that region.) In areas with colder winters (contour lines are labeled with the January mean temperature in degrees Celsius) far fewer plants are cyanogenic. Plants in cold regions, where winter weather is an effective predator control, have less need to defend themselves chemically.

Daniel H. Janzen of the University of Pennsylvania and I have found a number of unique biochemical adaptations achieved by the beetle that enable it to thrive on the potentially toxic seed for its entire larval life [see "A Seed-eating Beetle's Adaptations to a Poisonous Seed," by Gerald A. Rosenthal; SCIENTIFIC AMERICAN, November, 1983]. The insect has developed the necessary enzymes for converting canavanine into urea and then converting urea into ammonia. It uses the nitrogen of the ammonia, which had originally been part of the canavanine molecule, to support the production of virtually every amino acid it manufactures. In this way a potentially highly poisonous allomone of the plant has been manipulated so that it functions as a kairomone, wholly supporting the dietary nitrogen needs of the developing larvae.

There is yet another way in which insects can exploit a plant's chemical defenses. Vincent G. Dethier of the University of Massachusetts at Amherst and Louis Schoonhoven of the University of Wageningen in the Netherlands, along with other investigators, has found that many insects are equipped with taste receptors sensitive to the allelochemicals of plants on which they feed. For example, larvae that consume plants of the family Rosaceae have receptors for sorbitol, a sugar alcohol that is stored by such plants. Other insects usually lack such receptors. Insects equipped with these receptors can rely on their olfactory ability in finding and identifying chemically defended plants to which they are immune or resistant.

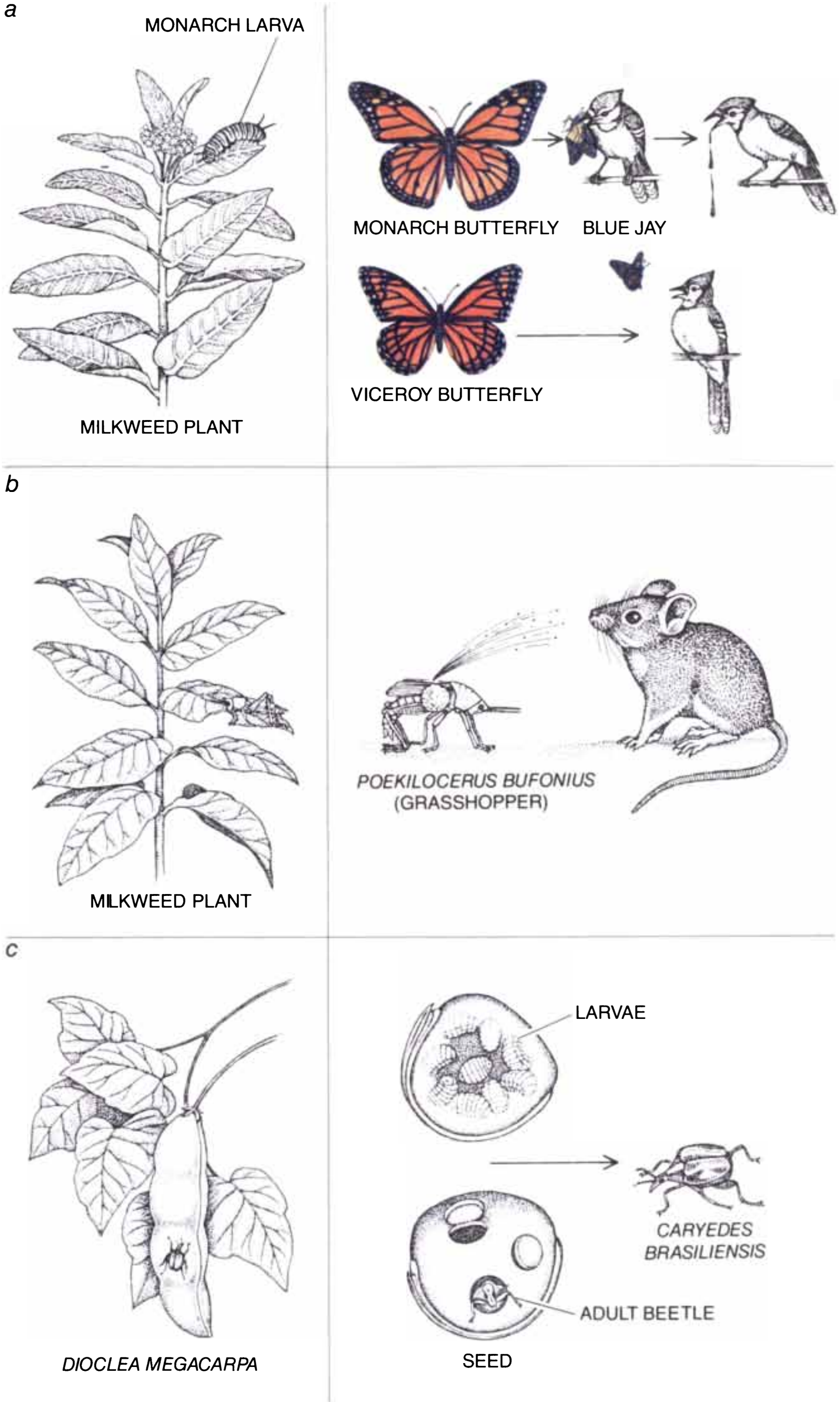
So far I have discussed mainly the allelochemicals that are constitutive to a plant, that is, produced whether or not the plant is under attack by a predator. Many plants rely instead on so-called inducible defenses: protective compounds that are synthesized only in response to an attack. An excellent example comes from the work of Clarence A. Ryan and his colleagues at Washington State University. They have found that when some plants, such as the tomato, are attacked by chewing insects, they release a substance that travels from the site of the wound through the plant and initiates the production of at least two kinds of macromolecule called proteinase inhibitors. Proteinase inhibitors impede the insects' ability to break down proteins they have ingested from the leaf. The leaf thus makes itself a less acceptable source of food.

The study of inducible defense was

recently taken a dramatic step further by the work of David F. Rhoades of the University of Washington, as well as that of Jack C. Schultz and Ian Balwin, then at Dartmouth College. It has been established that when the Sitka willow, *Salix sitchensis,* is attacked by insects, its leaf quality (a measure of its suitability as a food resource for insects) deteriorates. Rhoades noted, however, that the leaf quality of nearby willows—even willows that had not been attacked—also seemed to deteriorate. Perhaps, Rhoades suggested, the attacked tree produces a signal, analogous to an insect's alarm pheromone, that travels through the air to induce defensive responses in neighboring, unattacked trees.

Schultz and Balwin tested Rhoades's hypothesis by planting seedlings of the sugar maple, *Acer saccharum,* in two separate growth chambers. They found that plants that had been damaged intentionally, as well as plants in the same growth chamber as the damaged ones, tended to produce greater quantities of tannins and phenolics, two defensive compounds, than plants grown in a separate chamber. Although these studies have not definitively proved the existence of communicative chemical defense by plants, they have generated great interest in attempts to prove that pheromonal communication between trees may actually take place.

The chemical and biochemical study of the protective allelochemicals of higher plants can make important contributions to future efforts to control depredations on crops by insect pests. A recent report by the National Academy of Sciences emphasized the rapidly growing number of insects that have developed resistance to currently available chemical agents. Biological and chemical studies, the stuff of chemical ecology, may lead to effective pesticides that are less hazardous to the environment and not as easily circumvented. Natural products possess the advantage of structures that have amply proved effectiveness. They also offer excellent opportunities to develop experimental systems for probing the capacity of insects to cope with toxic compounds and may thereby make it possible to undermine that capacity. Natural products may also make it possible to develop novel, nonpesticide approaches to the control of herbivores. It should be possible to exploit natural compounds that deter feeding or the deposition of eggs, or even to grow crops from which the natural substances that attract herbivores or stimulate their feeding behavior have been eliminated.



INSECTS' RESPONSES to plants' chemical defenses often include exploiting for some useful purpose the very chemicals meant to ward them off or to kill them. Larvae of the monarch butterfly (*a*) feed on the milkweed *Asclepias curassavica,* a plant that manufactures toxins called cardenolides. The butterflies store the cardenolides and are therefore unacceptable as prey for blue jays; jays that eat them become violently sick and eventually will not accept monarch butterflies as food. The viceroy butterfly mimics the appearance of the monarch and so, although the viceroy stores no toxins, jays avoid it too. The grasshopper *Poekilocerus bufonius* (*b*) incorporates the cardenolides it ingests when eating milkweeds into the poisonous spray it uses to protect itself from predators. Larvae of the bruchid beetle *Caryedes brasiliensis* (*c*) feed on seeds of the leguminous vine *Dioclea megacarpa.* The seeds contain large amounts of L-canavanine, an amino acid that is not incorporated into proteins and that is often a potent insecticide. The beetle's larvae, which live inside the seed and feed on nothing else throughout their larval life, are able to obtain ammonia from canavanine. They then incorporate the nitrogen of the ammonia (nitrogen that was originally part of the canavanine molecule) into amino acids to be incorporated into proteins.

Radiocarbon Dating by Accelerator Mass Spectrometry

The radioactive carbon 14 is isolated from the other atoms in a sample, making it possible to derive more accurate chronologies from much smaller archaeological or anthropological specimens

by Robert E. M. Hedges and John A. J. Gowlett

Radiocarbon dating is the principal technique by which archaeologists and physical anthropologists construct their calendar of the past 50,000 years. The method relies on the fact that a radioactive carbon isotope, carbon 14, is assimilated into the molecular structure of living organisms. Because a population of radioactive atoms decreases at a regular rate, such atoms can serve as a clock by which to measure how long ago a plant or animal died. Indeed, carbon dating is so widely applied that one seldom sees an archaeological report that lacks a reference to it.

Yet for all its importance the tool, as conventionally applied, has severe limitations. One is that a rather large amount of material is needed for reliable measurement, whereas many artifacts of great archaeological value can be found only as very small remnants today. Other objects whose remnants are found in greater quantities may have a complicated "carbon history": sources of various ages may have contributed to the total carbon content of the object before it became buried. Moreover, the object may have been contaminated, after its burial, by "modern" carbon. Although sophisticated laboratory processes can remove unwanted contaminants, such purification calls for even larger quantities of the raw archaeological material.

The conventional technique is also limited in its range of operation: reliable datings are possible only for articles less than about 40,000 years old; the radioactive emanations from older objects are difficult to separate from background radiation. This limitation is frustrating for archaeologists and anthropologists because the period between 25,000 and 75,000 years ago was a crucial phase in the emergence of modern man and was accompanied by much interesting climatic and faunal change.

A new way to do radiocarbon dating has been developed that does not suffer from these limitations. It employs a particle accelerator in conjunction with a mass spectrometer. The method enables the investigator to count the number of carbon-14 atoms directly instead of measuring their decay rate and therefore requires a far smaller sample than conventional radiocarbon dating. Although its operative range and precision is today comparable to that of conventional dating, as much as a doubling of the range and a halving of the error margin could soon be achieved.

Radiocarbon dating of any kind is possible because about one out of every trillion (10^{12}) carbon atoms in organic matter is a carbon-14 atom. Although in most circumstances it is chemically indistinguishable from the other two stable carbon isotopes (carbon 12 and carbon 13), carbon 14 can be detected by the radioactivity it exhibits as it decays to nitrogen 14.

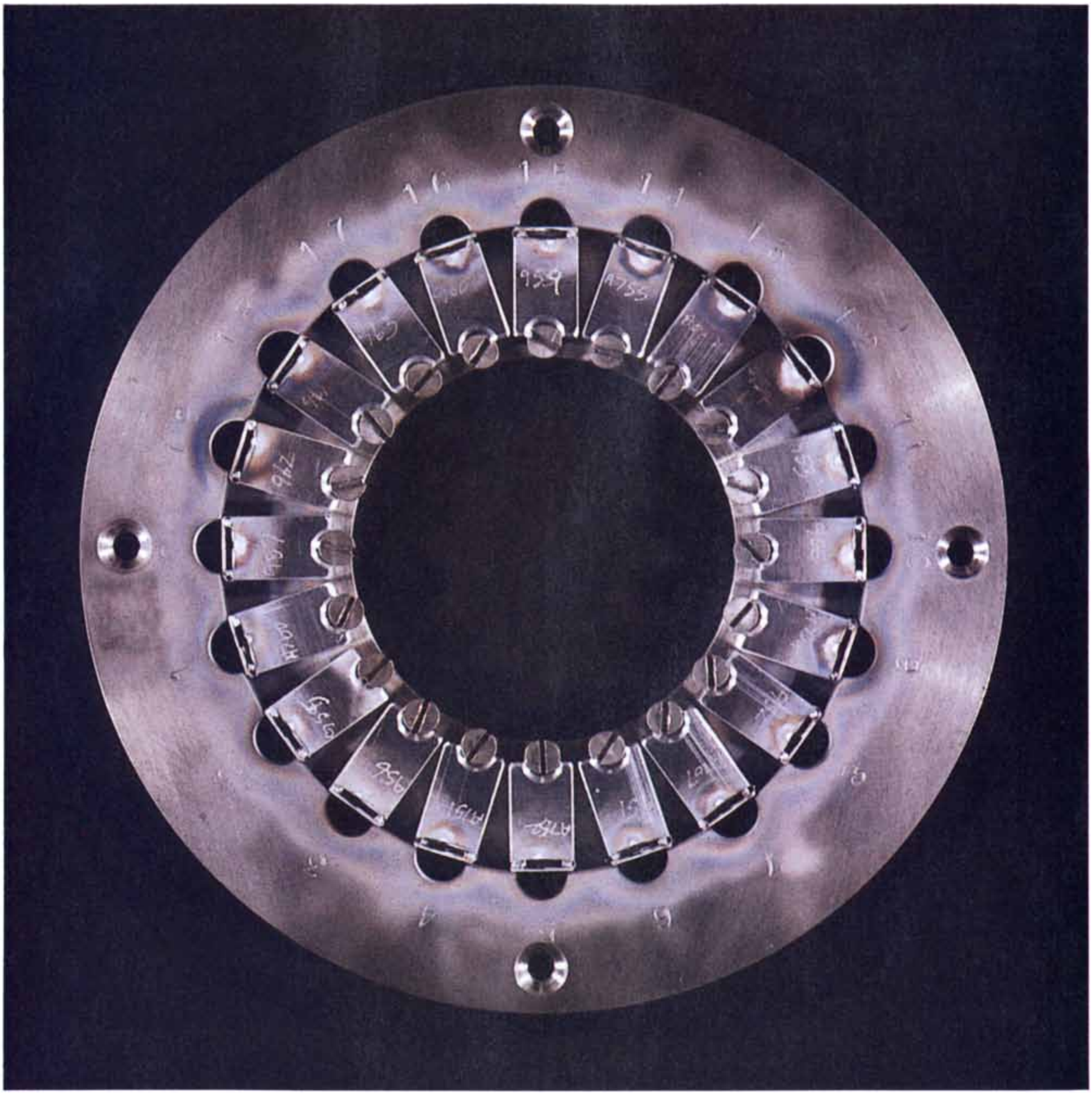
The rate at which a population of radioactive atoms decays is expressed in terms of a half-life. The half-life for C-14 is about 5,700 years. This means that a sample originally consisting of 10,000 C-14 atoms would contain only half this amount, or 5,000 atoms, after 5,700 years (the other 5,000 having turned into nitrogen 14). After another 5,700 years only 2,500 C-14 atoms would remain, and so on. If one knows the original amount of C-14 in a sample, one can easily calculate the age of the sample by determining the number of half-lives that must have elapsed in order to account for the sample's residual level of C-14 radioactivity.

Fortunately the original percentage of C-14 atoms among all the carbon atoms in any organic substance is fairly well known. The reason is that the amount of C-14 in the biosphere has been remarkably constant through the ages. C-14 is continuously produced in the upper atmosphere by the interaction of cosmic rays with nitrogen atoms. The C-14 then combines with oxygen to form radioactive carbon dioxide, which diffuses to the lower atmosphere and enters the biosphere primarily through plant respiration. From the plants the C-14 is passed up through the food chain to higher organisms, including man. The coupled effect of the radioactive loss and stratospheric production has established a constant, although minute, equilibrium concentration of C-14 in the biosphere. Every living organism therefore has (and has had) the same ratio of C-14 to C-12 in its body.

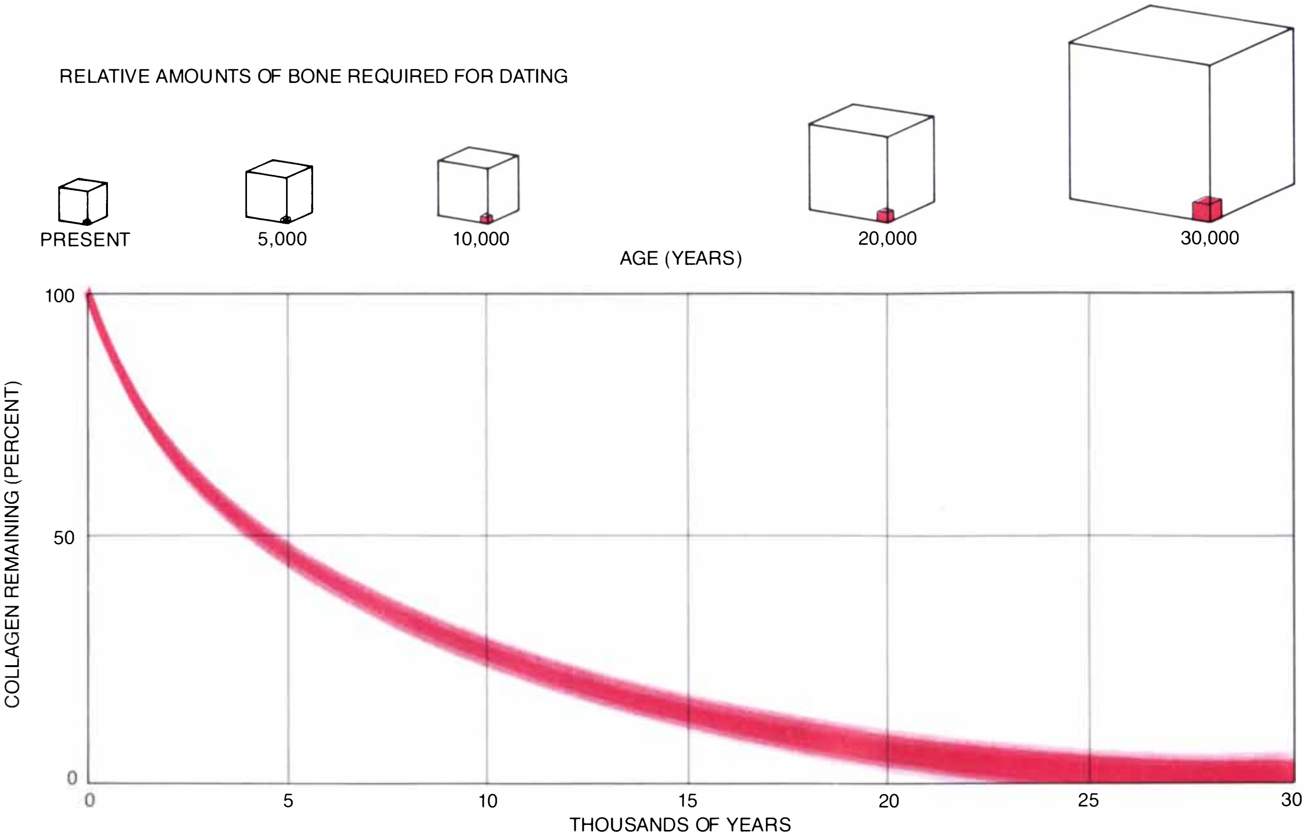
Once an organism dies it can no longer ingest carbon compounds. Consequently the C-14 in its body is not replenished and the amount of C-14 begins to fall as a result of radioactive decay. As long as the organism's remains are not contaminated by modern C-14 compounds (deposited by bacteria, for example), a measurement of the present-day ratio of C-14 to C-12 in its remains is sufficient to determine when the organism ceased to live.

The chronologies established by radiocarbon dating of wood samples have been checked against chronologies obtained by counting annual tree rings in the samples. The results indicate that the C-14/C-12 ratio has not always been strictly constant. The number of cosmic rays that impinge on the atmosphere to produce C-14 is affected by solar magnetic field disturbances and perhaps by fluctuations in the earth's magnetic field as well. In any event, the C-14/C-12 ratio can be adjusted to correct for these deviations. A greater problem is presented by the relatively large amounts of material needed to calculate the C-14 fraction of a sample by means of the conventional technique: measuring the radioactivity of the sample.

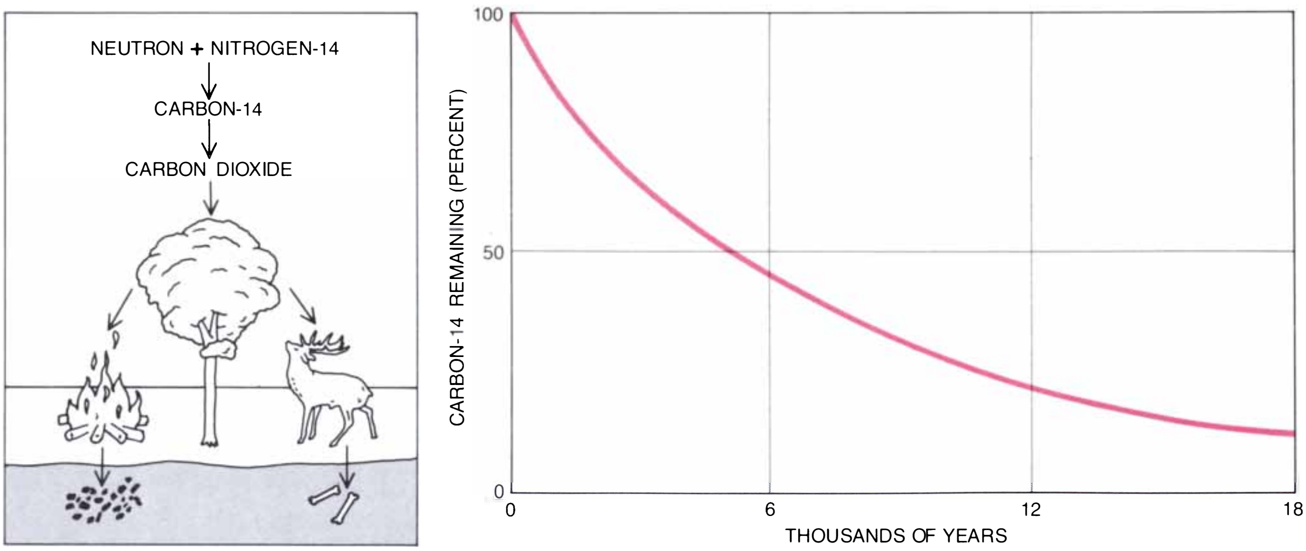
The weak radiation C-14 emits is not easily distinguished from ambient background radiation (mostly cosmic rays that have traveled through the atmosphere). In fact, if the object is more than 37,000 years old, the residual C-14 radioactivity is generally too weak to be discerned against the background radiation level. Hence a large number of C-14 atoms (or, conversely, a long measuring time) is needed before the spontaneous radiation can be adequately gauged by even the most sensitive of radioactivity detectors. Even more C-14 atoms are needed for reasons of measurement precision. (The precision of the calculated age is inversely proportional to the square root of the number of "counts" of radioactivity.) To detect about 10,000 radioactive disintegrations of C-14 in a reasonable amount of time (enough for a dating within 80 years of the correct age) requires between one gram and five grams of pure carbon. Of course, to obtain this much carbon, one might require a sample of organic matter sev-



SAMPLE WHEEL for the radiocarbon accelerator in the authors' laboratory carries 20 metal tabs, arranged like petals, that hold carbon samples from specimens to be dated (and samples of known age or carbon-14 content for calibration). The carbon is deposited as a thin layer of graphite on a short length of tantalum wire, which is placed in the groove at the outer end of each tab. The wheel (shown at about its actual size) is loaded into the accelerator's ion source, where each sample is ionized in turn (*see illustration on page 104*). Carbon-14 atoms are sorted out from among the ions thus produced and are counted; their number is a function of the specimen's age.



QUANTITY OF BONE needed for radiocarbon dating depends on the bone's actual age, its state of preservation and the dating method employed. The relative amount of mass required for dating by accelerator mass spectrometry (*colored cubes*) is 1,000 times less than the mass needed for dating by conventional radiocarbon methods (*larger cubes*), regardless of the bone's age. To avoid measuring the carbon-14 content in "modern" carbon (carbon that is somehow deposited in the bone after it became buried), which would result in underestimation of ages, what is measured is the carbon-14 content of purified amino acids derived specifically from a single component of bone: the protein collagen. Because collagen deteriorates in the course of time (*bottom*), a larger sample is needed for older bones. The curve's fuzziness reflects the fact that the percentage of collagen can vary, depending on the condition of preservation.



CARBON 14, a radioactive isotope of carbon, is produced in trace amounts when cosmic-ray neutrons interact with nitrogen atoms in the atmosphere (*left*). The C-14 then combines with oxygen to form carbon dioxide. Like normal carbon dioxide (containing carbon 12), the radioactive kind enters the biosphere through plant respiration, and it undergoes the same carbon chemistry. By working its way up the food chain C-14 is incorporated into the tissues of higher organisms as well. The continuous infusion of C-14 into the biosphere, balanced by the continuous loss through radioactive decay (*right*), establishes a fairly constant C-14 concentration in all plants and animals. When an organism dies, however, the C-14 in its tissues can no longer be replenished and the radioactive disintegration dominates. Because the C-14 concentration in living organisms is known, the age of organic remnants can be derived by measuring the C-14 content in a sample of the remnants and calculating how much time must have elapsed to account for the observed decline.

eral times larger: from 25 grams to as much as a kilogram, depending on the carbon content of the material.

Clearly, if one could separate C-12 and C-13 atoms from C-14 atoms in a carbon sample of known mass and then count all the C-14 atoms (not just the ones that disintegrate radioactively), much greater efficiencies would be possible. Greater efficiency in measuring the C-14 content of a carbon sample could in turn be translated into an expanded operative time range and improved accuracy in dating. This is what we have in part accomplished at our radiocarbon accelerator facility at the University of Oxford.

First the carbon atoms of a sample are ionized. Then they are accelerated to energies of about 10 million electron volts. The "beam" of high-energy ions thus created travels through a magnetic field that forces the beam to curve. Lighter atoms tend to turn more sharply than heavier ones, and so they move to the inside of the diverging beam. In this way the C-14 ions can be isolated from the other atoms in the sample and counted. (The amount of C-12 and C-13 can then be derived from the difference in mass between the original carbon sample and the total mass of the detected C-14.)

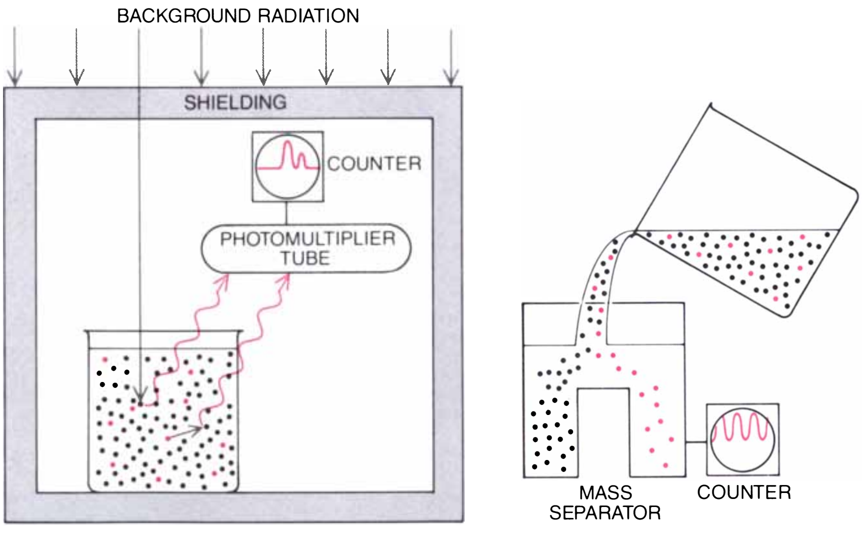
An electrostatic accelerator such as the Oxford Radiocarbon Accelerator makes use of the force between electric charges to accelerate atoms. This is why the atoms of a sample to be analyzed must first be ionized. In the case of the Oxford accelerator a positive charge amounting to some 2.5 million volts first "pulls" negative carbon ions and then, after their electrons have been stripped off, "pushes" them the rest of the way as positive ions [*see illustration on next page*]. The stripper, which removes four electrons from the carbon ions to give them a triple positive charge (as measured in units of electron charge), also causes any stray molecular ions to break up into separate atomic ions. This eliminates the many molecules of mass 14 that might otherwise be indistinguishable from single C-14 atoms.

A mass spectrometer also takes advantage of magnetic forces. Magnets are used to focus and direct the beam of ions as well as to sort the ions by mass. If ions of similar charge and velocity but different mass are subjected to a magnetic field perpendicular to their direction of travel, only ions of like mass would pass through a slit at a given radius of curvature downstream from the magnet. In this way a magnet could "filter out" all charged particles except those of atomic mass 14. Of course, the accelerated ions do not have identical charge and velocity, and so ions that vary in mass, charge and velocity actually make it through the slit behind the magnet. Nevertheless, interfering ions are reduced by a factor of 10 billion by two such magnets in our accelerator mass spectrometer: one acting on the ions before their acceleration, one after.

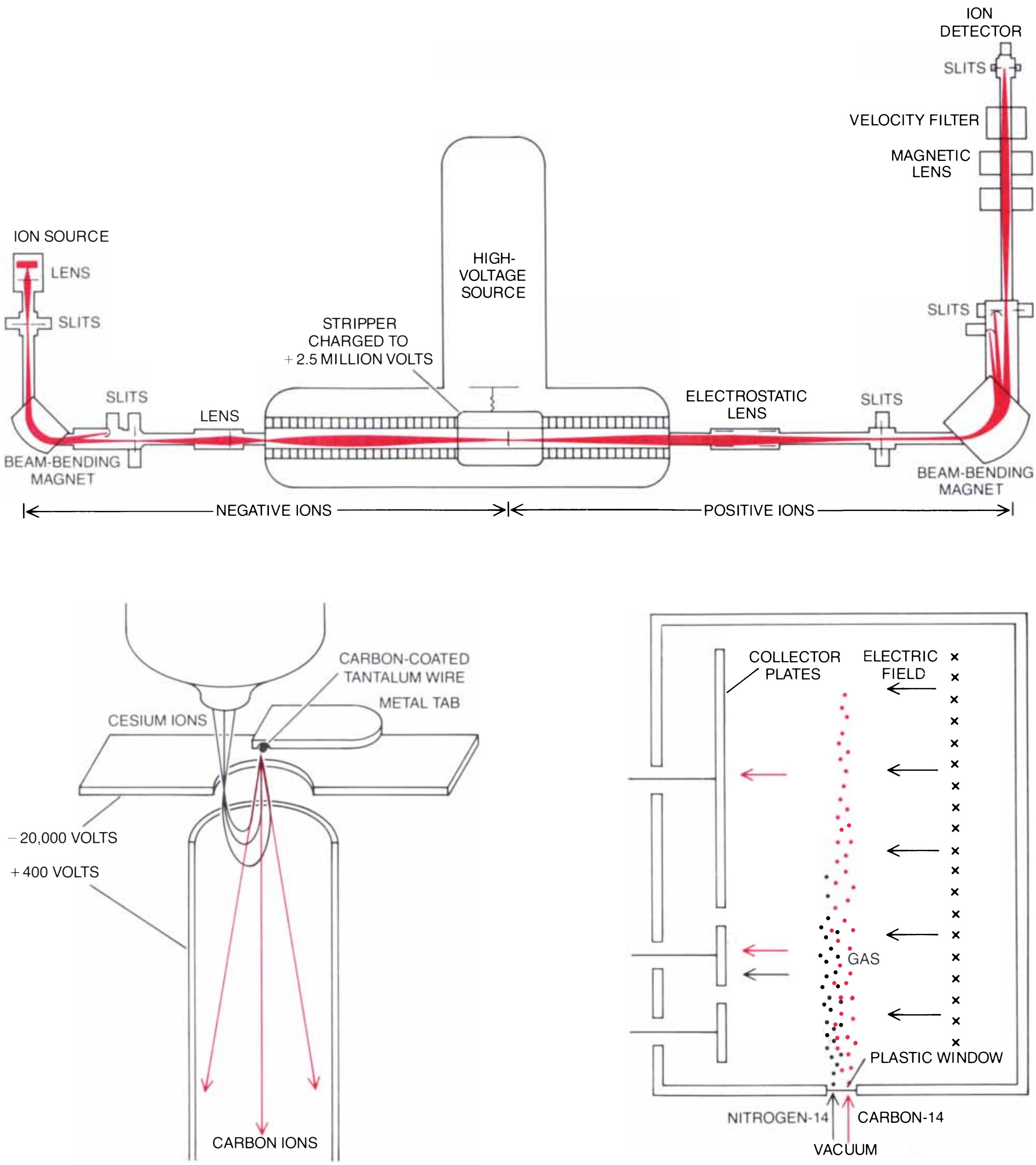
The accelerated particles are further culled by a velocity filter consisting of uniform magnetic and electric fields perpendicular to each other as well as to the ion beam. This reduces the unwanted ions by another factor of about 1,000 before triply charged positive ions of atomic mass 14 (namely C-14 ions) are finally "counted" by an energy detector. The detector can distinguish ions of like mass and velocity but different nuclear charge by the rate at which they lose energy as they collide with molecules of gas. The detector, in fact, is so sensitive that it can identify almost all C-14 ions even when they constitute less than 1 percent of the incoming ion beam, as would be the case for a very old sample (say 60,000 years old).

At present accelerator mass spectrometry allows up to 2 percent of all the C-14 atoms in a carbon sample of "average" age (about 12,000 years old) to be detected within a time period measured in hours. Although this may seem to be a small percentage and a lot of time, it would take 170 years before 2 percent of the C-14 atoms of a sample manifests itself in the guise of radioactivity, regardless of sample size. To detect 10,000 atoms of C-14 in such an average sample by accelerator mass spectrometry, between half a milligram and five milligrams of pure carbon is sufficient, a factor of 1,000 less than is required for conventional radioactivity measurements of C-14.

Before the carbon atoms in the material to be dated can be ionized and accelerated in our machine, the material must be purified (modern carbon removed) and converted into graphite, a form of solid carbon. By treating the specimen with chemicals that react selectively with certain types of carbon compounds, a source of carbon unique to a given archaeological artifact can be extracted for dating. For example, bones can pick up carbonates from groundwater, fall prey to fungal or bacterial attack or absorb organic compounds from the soil. All these processes introduce extraneous carbon into bone specimens that can give mis-



ESSENTIAL DIFFERENCE between conventional radiocarbon-dating methods (*left*) and those making use of a mass spectrometer (*right*) is in the way the C-14 concentration is measured. Conventionally the number of C-14 atoms (*colored dots*) in a sample is estimated from the number of radioactive disintegrations detected within a certain amount of time. Here the radiation (*black arrow*) from a C-14 atom causes a fluorescent dye molecule to emit a flash of light (*colored wave*) that is recorded by a photomultiplier tube. Natural background radiation can also cause the dye molecules to scintillate and can thereby distort the results. Shielding attenuates this radiation to a level lower than that of the radioactivity of samples of "average" age, so that it can normally be subtracted out. Old carbon samples are less radioactive, however; their level of radioactivity is close to that of the ambient background radiation. For objects older than about 37,000 years the background radiation effectively obscures the radioactivity of the sample. In an accelerator mass spectrometer the C-14 atoms are separated from all other atoms and are directly "counted." For this method the limiting factor is the chemical contamination of the sample by modern carbon.



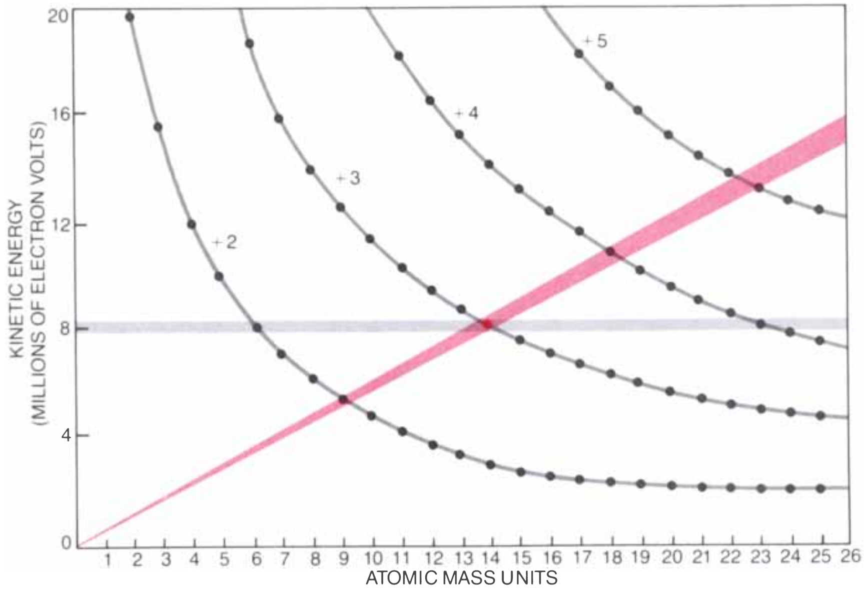
RADIOCARBON ACCELERATOR (*top*) accelerates carbon ions by applying a positive charge of 2.5 million volts to the stripper. The imparted charge first "pulls" the negatively charged ions to the stripper and, after the stripper converts them into positive ions, then "pushes" them the rest of the way. Most of the contaminating ions are filtered out by two beam-bending magnets. These magnets separate the ions by mass: lighter ions tend to turn more sharply than heavier ones. Ions of atomic mass 14 can therefore be singled out by placing a slit at the proper radius of curvature downstream from the magnet. In the ion source (*bottom left*) carbon atoms of the sample to be tested (deposited as a thin layer on a tantalum wire) are bombarded with positive cesium ions to produce negatively charged carbon ions. The electric charges applied to the metal plate and cylinder (*shown here in cross section*) in the device direct the cesium ions to the carbon sample and the carbon ions away from it. The C-14 ions are finally distinguished from the remaining contaminating ions by a detector (*bottom right*). After passing through a plastic window the accelerated ions travel through a gas-filled chamber, where they produce plumes of secondary ions. Because the velocity and atomic charge of the incoming ions are similar, the rate of interaction between the ions and the gas molecules depends primarily on the nuclear charge of the ion. Hence carbon ions (*colored dots*), which have a nuclear charge of +6, interact at a lower rate than nitrogen ions (*black dots*), which have a nuclear charge of +7, and typically travel farther down the length of the chamber before ionizing the gas. An electric field causes the ion plumes to drift over to collector plates, where their average lengths are then measured. From these measurements the C-14 ions are identified.

leading C-14 readings, particularly if the specimen is very old. Bone's chief organic material, however, is the protein collagen. One of the amino acids that make up collagen, hydroxyproline, can for all practical purposes be found only in bone. Hence if a hydroxyproline sample is prepared from a bone specimen, it is almost certain that the carbon contained in the samples was originally incorporated into the bone during its formation and not afterward.

Once the compound of interest has been isolated, it must be converted into a sample of pure graphitic carbon. We do this by burning the compound to produce carbon dioxide gas; the gas is then converted into acetylene (C_2H_2), which deposits a pure layer of carbon on a heated tantalum wire.

Negatively charged carbon ions can then be produced from the carbon layer by bombarding it with positive cesium ions. (Such an ion source is called a sputter source.) By starting with negatively charged carbon ions, we remove most of the problematic nitrogen-14 contamination (N-14 has a mass almost identical with that of C-14) because nitrogen does not exist as a negatively charged ion. Although the sputter source does provide a stable beam of ions that can subsequently be accelerated, only about 10 percent of the carbon deposit is converted into ions. (We are developing a carbon dioxide ion source that would not only be more efficient but also greatly simplify the chemistry involved in preparing the sample for ionization.)

Because an accelerator mass spectrometer allows the detection of almost any isotope down to abundance levels of one part in 1,000 trillion (10^{15}), it might be thought that this apparatus could extend radiocarbon dating to objects at least 80,000 years old. A very real limitation at present, however, is the introduction of minute amounts of modern carbon contamination during the sample processing. This limits the dating to within 40,000 years or so, comparable to the limit for conventional radiocarbon dating. In the case of accelerator mass spectrometry the limit is not fundamental; better laboratory procedures will improve the sensitivity.



SELECTION OF TRIPLY CHARGED C-14 IONS is possible through the application of three devices: a beam-bending magnet, a velocity filter and an energy detector. The radius of the curve traced by an ion subjected to a magnetic field is a function of the ion's mass, charge and velocity or, equivalently, its kinetic energy. The ions that actually pass through a slit downstream of the magnet (*dots*) therefore exhibit a distribution of masses, charges and energies. When they are plotted, they describe a family of hyperbola (*gray curves*), one for each possible ionic charge. The kinetic energy of the ions that pass through a velocity filter is a linear function of atomic mass (*band of pale color*). Hence when the filter is operated in conjunction with the beam-bending magnet, it selects one ion from each of the possible energy-v.-mass hyperbolas. The final selection is made by the energy detector (*gray band*). It is set to detect only ions with a kinetic energy of eight million electron volts, and so a triply charged ion of atomic mass 14—namely a carbon-14 ion (*red dot*)—is singled out.

Currently the cost of radiocarbon dating by means of accelerator mass spectrometry is between 1.5 and 2.5 times greater than the cost of conventional radiocarbon dating. For materials older than about 10,000 years the accuracies of both dating methods are comparable. Although the possibility of achieving better dating accuracies by means of the accelerator technique does exist, at present the method is valued primarily because it can derive dates from samples 1,000 times smaller than those required by traditional techniques. Accelerator radiocarbon dating is therefore most effective where sampling must be minimized to limit destruction of the artifact; where only tiny samples exist to begin with, and where stringent stratigraphic or chemical selectivity of samples is demanded.

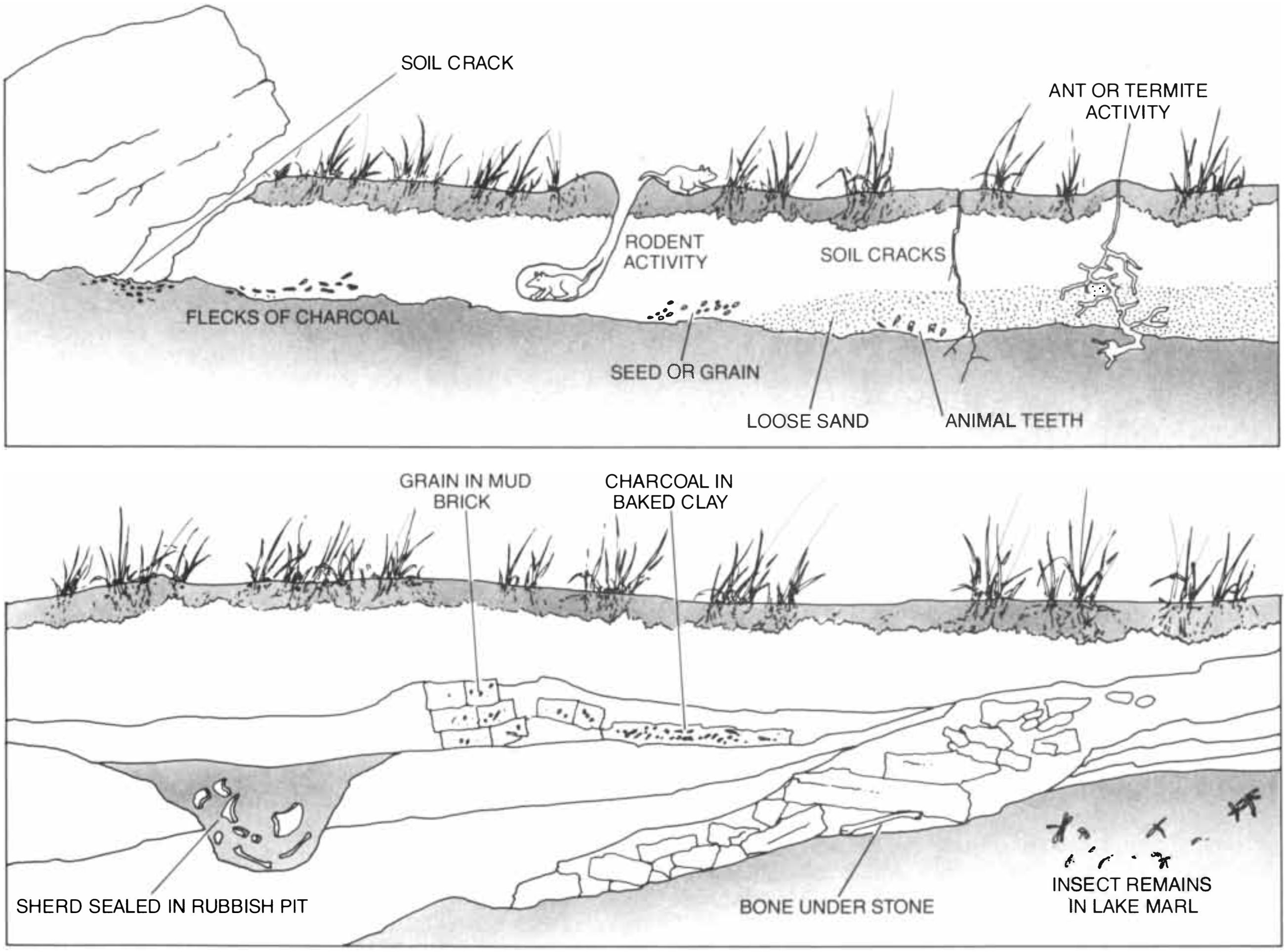
Many ancient manuscripts, for example, are considered too valuable to sacrifice to the "ravages" of conventional radiocarbon dating. Accelerator dating can minimize the amount of material that would have to be sacrificed. A piece of an early medieval *mappa mundi* (world map) belonging to the Duchy of Cornwall and preserved through its reuse as a book binding has already been dated from a small scrap. It was shown to be almost certainly older than the only other extant example in Britain, in Worcester Cathedral. Applied in this way, the radiocarbon accelerator method can verify much paleographic dating (dating based on the study of script styles) and expose forgeries. Early manuscripts of the New Testament or fragments of papyrus would be ideal materials for testing in this manner.

At many archaeological sites materials for dating are too small to serve as individual samples. They are therefore often combined to accumulate a sample that is large enough to date. The problem with this approach, of course, is that components from altogether different time periods can be mixed together and sampled collectively. In many instances accelerator mass spectrometry provides an alternative by enabling workers to date the individual samples rather than their aggregate. This is particularly the case for the fragments of bone or the antler tools made during Paleolithic times, more than 10,000 years ago. A decorated horse mandible uncovered in northern Wales a century ago could have been as old as 25,000 years or as "young" as 5,000 years, based on current knowledge of ancient art. By means of accelerator dating the artifact was found to be about 10,000 years old, and it is therefore an important addition to the catalogue of late Paleolithic artwork. This dating also helps to document the persistence of the Pleistocene horse into more modern times.

Bone or antler harpoon heads of the late Ice Age are even more delicate objects, but they are generally well preserved, so that a date can be obtained from a single small drilling. Stone tools, the most enduring technology of the stone ages, can sometimes be dated from the hafting resin or binding string that connected the wood handle to the worked stone. Bronze spearheads and axheads from later times occasionally can be found with small pieces of their wood handles still attached. A spearhead from Jackscar Cave in northern England has been dated by measuring the C-14 content in the traces of its wood handle. A stone axhammer of the early Bronze Age was similarly dated; it came from a preserved forest bed at Cleethorpes on the east coast of England. It is interesting to note that the age of an adjacent tree stump was found to be 1,000 years less than that of the axhammer. In the past the tree stump could easily have been used to date the axhammer by association.

The ages of bone projectile points and harpoon tips are providing important clues in reconstructing the northward spread of the human species in Europe during the late glacial period; bits of reindeer bone and degraded charcoal that are found at ancient campsites can also yield such information. Sites in the open are far more numerous than cave sites, but they are less likely to offer fragments suitable for conventional carbon dating. Accelerator dating, however, has made it possible to derive a solid radiocarbon chronology from the bones found at the key open sites of Pincevent, Etiolles and Marsangy. Pincevent is now considered to be a little more than 12,000 years old, and this accelerator-determined age agrees reasonably well with some recent conventional datings. The radiocarbon-accelerator measurements also show that Pincevent is about 1,000 years younger than Etiolles but older than Marsangy. Pincevent is thereby seen to be approximately contemporaneous with the archaeological finds in Gough's Cave in southwestern England and the remains of an elk in Lancashire, both of which were also dated by accelerator. The elk skeleton in turn dated a pair of bone projectile points embedded in it, the earliest such points yet dated.

In the past a single plant grain or animal bone could be dated only from the archaeological stratum in which it was found. Such dating by context is important in setting the time frame for the domestication of plants and animals. Yet soil cracks and rodent, worm, ant or human activity can disperse these bits of organic material throughout the soil, not necessarily in the chronological order of the surrounding geological deposits. Accelerator dating is ideally suited to validate or invalidate this form of inferential dating. For instance, measurements made in our laboratory and in a similar facility at the University of Arizona showed that grain kernels and date stones found on a 17,000-year-old site



STRATIGRAPHIC CONTEXT in which ancient organic remains, such as bone, teeth, charcoal or seeds, are embedded must be carefully examined before the results of their carbon dating can be confidently generalized to the strata in which they lie. Contexts of doubtful validity (*top*) include those where one or another agent, such as those shown, is likely to have dispersed the organic material. Sound contexts (*bottom*) include those where the material to be carbon-dated has been sealed in an immobilizing matrix. Because the sealed material is likely to be of small size, accelerator mass spectrometry is the technique best suited to date the strata by context.

in Egypt were in fact only a few hundred years old. On the other hand, grain from other sites in the Middle East has been dated as being many thousands of years old, although it is not of domesticated form. This stratigraphic selectivity is an important advantage of accelerator dating; through it small samples of interest in themselves can be matched to their stratigraphic context.

Another aspect of selectivity that accelerator dating offers is in the sample treatment. Because far less carbon is needed for a measurement, one can afford to be much more discriminating in choosing the carbon compound to be tested. Parchment, like the collagen of bone, can be made to yield an uncontaminated sample of carbon from its amino acid constituents. A piece of ivory can be similarly dated by measuring the C-14 content of the amino acids extracted from its dentine. In this way the extinction date of the mammoth in western Europe could be worked out much more closely from small pieces of tusk preserved at Paleolithic sites. It is also possible to prepare several carbon compounds for testing from a single sample. This can give an indication of the pattern of contamination from other sources of carbon. For example, we were able to determine the various times at which carbon entered lake sediments by analyzing the various lipid fractions they contained.

In general, conventional radiocarbon dating has effectively mapped out the past 30,000 years in areas where organic preservation has been favorable. Yet beyond a threshold of about 20,000 years the number of objects whose age can be determined with certainty diminishes rapidly. Accelerator mass spectrometry can push the threshold further back in time, however.

At a horizon of 40,000 years the amount of carbon 14 in a bone or a piece of charcoal can be truly minute: such a specimen may contain only a few thousand C-14 atoms. Consequently equally small quantities of modern carbon can severely skew the measurements. Contamination of this kind amounting to 1 percent of the carbon in a sample 25,000 years old would make it appear to be about 1,500 years younger than its actual age. Such contamination would, however, reduce the apparent age of a 60,000-year-old object by almost 50 percent. Clearly proper sample-decontamination procedures are of particular importance in the dating of very old artifacts. Unfortunately stringent specimen handling and treatment procedures ultimately result in relatively small sample sizes, which conventional radiocarbon dating is poorly equipped to handle. Moreover, conventional radiocarbon dating would still face the insurmountable problem of discriminating the radioactivity of the sample from ambient background radiation.

Neither sample size nor background radiation present problems to radiocarbon accelerator dating, and so relatively minor improvements in sample chemistry can lead to sharper and more extensive chronologies. For example, accelerator dating of purified amino acids from bones more than 25,000 years old showed that their age had previously been consistently underestimated by 1,000 or more years. The earlier, conventional measurements had been based on whole collagen, which cannot be guaranteed to be free of modern carbon.

One controversy of long standing on which the new dating technique has already had a major impact concerns the first human migrations to the New World. An accurate time scale for the colonization of America is crucial in order to assess how quickly the first Paleo-Indian hunters and gatherers dispersed, settled and developed their ethnic and linguistic diversity. Most observers agree that the earliest human inhabitants of America came from northeastern Asia probably between 25,000 and 12,000 years ago, crossing over a land bridge that then connected Siberia with Alaska. Nevertheless, human skeletons were found in the New World at disparate locations, such as Canada and Peru, that seemed to be considerably older than expected. Accelerator mass spectrometry has refuted these claims: no skeletal remains yet found in America appear to be more than 12,000 years old. If other such finds are made in the future, accelerator mass spectrometry will quite probably be the dating technique called on to determine the skeleton's age directly.

Accelerator dating may also prove to be invaluable in establishing an authoritative chronology of Neanderthal man. The archaeological evidence available indicates a rather abrupt disappearance of this human subspecies at the beginning of the Upper Paleolithic, about 35,000 years ago. Although it is too early to be sure, dating by accelerator mass spectrometry may reveal that the Upper Paleolithic and the period preceding it, the Middle Paleolithic, were unduly compressed by the limitations of conventional dating methods. If this was the case, there would have been much more time for the Neanderthal's disappearance than has commonly been supposed.

Kin Recognition in Tadpoles

Tadpoles of the Cascades frog prefer to associate with siblings, which they distinguish from nonsiblings. The ability to recognize kin is not based on familiarity; it may have a genetic component

by Andrew R. Blaustein and Richard K. O'Hara

The members of many animal species form social groups. There are flocks of birds, schools of fish, troops of apes. The question arises: May some of these groups be based on family ties? In other words, are some animals able to recognize kin? Do they capitalize on this ability in order to associate with kin rather than with unrelated individuals and also to avoid mating with close relatives? Kin recognition has recently been the focus of a number of studies by animal behaviorists. The studies lead eventually to consideration of an evolutionary approach to understanding cooperation and altruism among animals. Before these theoretical implications can be considered, however, one needs to know whether members of a particular species do indeed recognize their kin. If they can recognize kin, how do they do it?

For the past several years we have been investigating kin-recognition behavior in tadpoles of several anuran (frog and toad) species that abound in the lakes and ponds of the Pacific Northwest. We find that these tadpoles do recognize their siblings, and that in one species in particular the ability to distinguish between kin and nonkin is extremely sensitive. Tadpoles of the Cascades frog, *Rana cascadae,* are able to recognize their siblings even if they have never seen them before. They prefer to associate with their siblings rather than with nonsiblings, and this preference persists through their metamorphosis into frogs.

Kin recognition is not a trivial characteristic. It is intricately bound to the cooperative and "altruistic" behavior many social animals preferentially direct toward their relatives. The classic example of this behavior is that of social insects such as bees that live in an ordered, highly specialized society characterized by cooperation and altruism. In such communities sterile individuals (the workers) labor in behalf of a single fertile individual (the queen) that is their genetic sister. Why, however, should animals favor their kin, considering that such behavior may often be costly to the individual?

Much of the current understanding of the evolution of altruism derives from the work of W. D. Hamilton of the University of Oxford. His theories to explain altruistic behavior predict that individuals will aid or cooperate with close relatives rather than with distant relatives or unrelated individuals because to do so gives them an evolutionary advantage. The more closely related individuals are to one another, the more genes they have in common. An altruist will aid relatives, even if such behavior may be risky for the individual, because by doing so the altruist is increasing the probability that genes identical with its own will be passed on to future generations.

Hamilton's term for the reproductive success an organism shares when a genetic relative survives and reproduces is "inclusive fitness." If relatives live nearby (perhaps because dispersal from birthplaces or hatching sites is low), specific recognition of kin may not be necessary for what is in effect kin-directed behavior to operate. If interactions with nonkin are also likely, however, the ability to recognize kin could prevent the misdirection of helping behavior or ensure that the cooperative effort is exerted in behalf of the proper individual, namely a close relative. In general individuals that misdirect their altruistic acts would leave fewer copies of their genes to future generations than animals capable of directing their altruistic actions to relatives would. Natural selection should therefore favor the evolution of kin recognition along with altruism.

There appear to be three basic behavioral mechanisms that enable an individual to recognize its kin. First, if relatives are predictably found together in social circumstances, recognition can come about through the mechanism of social learning, or familiarity. That is, individuals of the same family group learn to recognize one another early in life. A second mechanism, phenotype matching, occurs if an individual learns and remembers a specific characteristic of itself or of its relatives. Such a characteristic might be an odor, a plumage color or a particular marking. The individual can then identify others as kin or nonkin by noting similarities and differences between the characteristic it remembers and that of an unfamiliar individual. Phenotype matching is fundamentally different from recognition based on familiarity because it enables an individual to recognize unfamiliar animals.

The third mechanism, which depends on specific "recognition genes," also enables an individual to recognize unfamiliar animals. This hypothetical mechanism of kin recognition is, however, a purely genetic one; no learning is involved. Recognition genes (or, more specifically, recognition alleles, which are alternative forms of genes) are expressed as a phenotypic character such as odor. Individuals carrying copies of such an allele recognize, and so tend to favor, other individuals carrying the same allele. It is important to remember that any of these mechanisms of kin recognition may operate alone or in conjunction with one another.

We undertook the study of kin-recognition behavior in our laboratory at Oregon State University several years ago largely because we had a likely subject: observations of *Rana cascadae* tadpoles had suggested they would be ideal for such an investigation. Two aspects of the natural history of the species particularly struck us. The first one was that Cascades frog tadpoles are unusually social. Whereas most frog tadpoles are not known to form aggregations or schools, Cascades frog tadpoles form cohesive aggregations of fewer than 100 individuals. Moreover, because one of us (O'Hara) had studied the natural history of these frogs, we knew that Cascades frog tadpoles do not disperse very far from their hatching site. This suggested there must be ample opportunity for kin to interact with one another; it was possible that the aggregations we saw were composed mostly of kin. We speculated that kin recognition might give Cascades frog tadpoles the ability to form and maintain such cohesive groups.

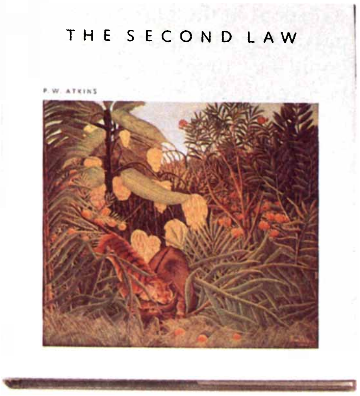
For an inquiry into the nature and the mechanism of kin recognition we needed to design an experimental protocol that would let tadpoles choose between kin and nonkin that were either familiar or unfamiliar to them. We had in mind three basic questions animal behaviorists often ask when approaching any behavioral study for the first time: How does this behavior develop? What is the sensory basis of this behavior? What is its adaptive value?

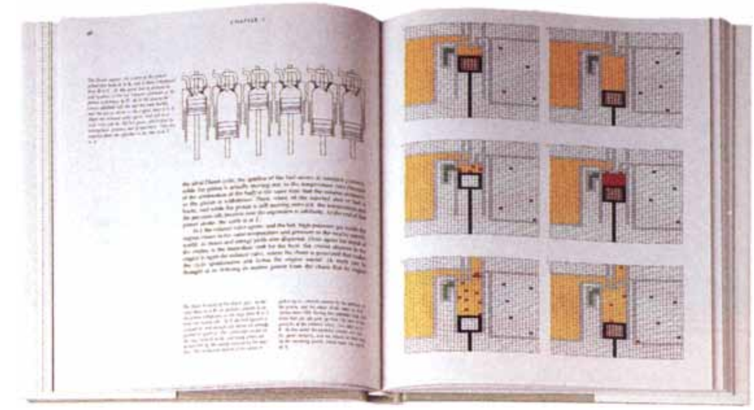
We began by collecting clutches, or egg masses, of Cascades frog eggs laid by different females and fertilized by different males and brought them into the laboratory. (A Cascades frog clutch contains between 400 and 800 eggs; all tadpoles hatched from a single clutch are siblings.) We also collected mating pairs of frogs and allowed egg laying and fertilization to take place in the laboratory. We then subjected developing embryos to one of four different rearing regimes before testing the tadpoles for ability to identify kin.

In the first regime individuals were reared in aquariums with members of their own clutch only, so that they were exposed only to their siblings. In the second regime we reared an equal number of tadpoles from each of two different clutches on opposite sides of an aquarium partitioned by a central mesh screen. Water could flow freely between the two groups, allowing members of both groups to exchange chemical as well as visual cues. We call this a mixed-rearing regime because members of two different kin groups are essentially reared together.

In the third regime we removed newly fertilized eggs from a clutch and

TADPOLES of the frog species *Rana cascadae* form small, cohesive schools in a pond in the Cascade Range of Oregon. Apparently these tadpoles can distinguish siblings from nonsiblings, and they choose to associate with siblings. This preference, which remains with an individual through metamorphosis into a frog, means it is likely that the schools of these tadpoles consist mostly of kin.





YOUR PREMIER VOLUME

THE SECOND LAW

P.W. Atkins

"The interested person can do no better than to peruse this fascinating new volume by Peter Atkins . . . . Even those who feel they already know all about the Second Law can hardly fail to find new angles and fresh insights into what is going on in the world around them. I hope that this remarkable book will be widely read by scientists as well as by the non-scientists for whom it is primarily intended."
NEW SCIENTIST

"A lovely book, beautifully illustrated and presented, and clearly commensurate with its companion volumes in the Scientific American Library. *It should engender affection as a sophisticated latterday Hogben, complete with social allusions implicit in the call to conserve, not energy, but entropy, and to use the virtuous entropy-conserving heat pump rather than crude and profligate combustion...."*
NATURE

In his famous essay on the two cultures, C.P. Snow made the Second Law of thermodynamics the litmus test of scientific literacy: Not to know the Second Law is the same as not having read a work of Shakespeare.

This is the law, of course, that explains why hot objects grow cool whereas cool objects do not spontaneously become hot; why a bouncing ball must come to rest and a resting ball cannot, of its own, begin to bounce. To some people, the Second Law conjures up visions of clanking steam engines, intricate mathematics and incomprehensible physical theory. It is better known to many in its restatement as Murphy's second law: "Things, if left alone, will gradually go from bad to worse!"

That restatement captures, in its way, the universality and power of the law that inspired Arthur Eddington to call it "the arrow of time."

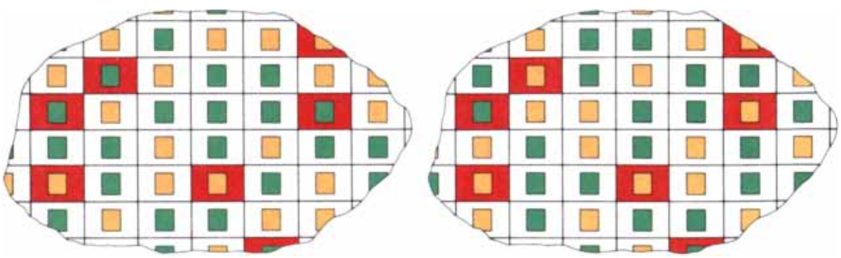
In *The Second Law* P. W. Atkins, of Oxford University, breaks through the mathematical barriers to understanding of the law. By vivid example he shows it at work in heat engines, refrigerators, and heat pumps; the cosmos, in the ecosphere and in the living cell; in the physical, chemical, and life processes that drive the changing world in all its richness and variety.

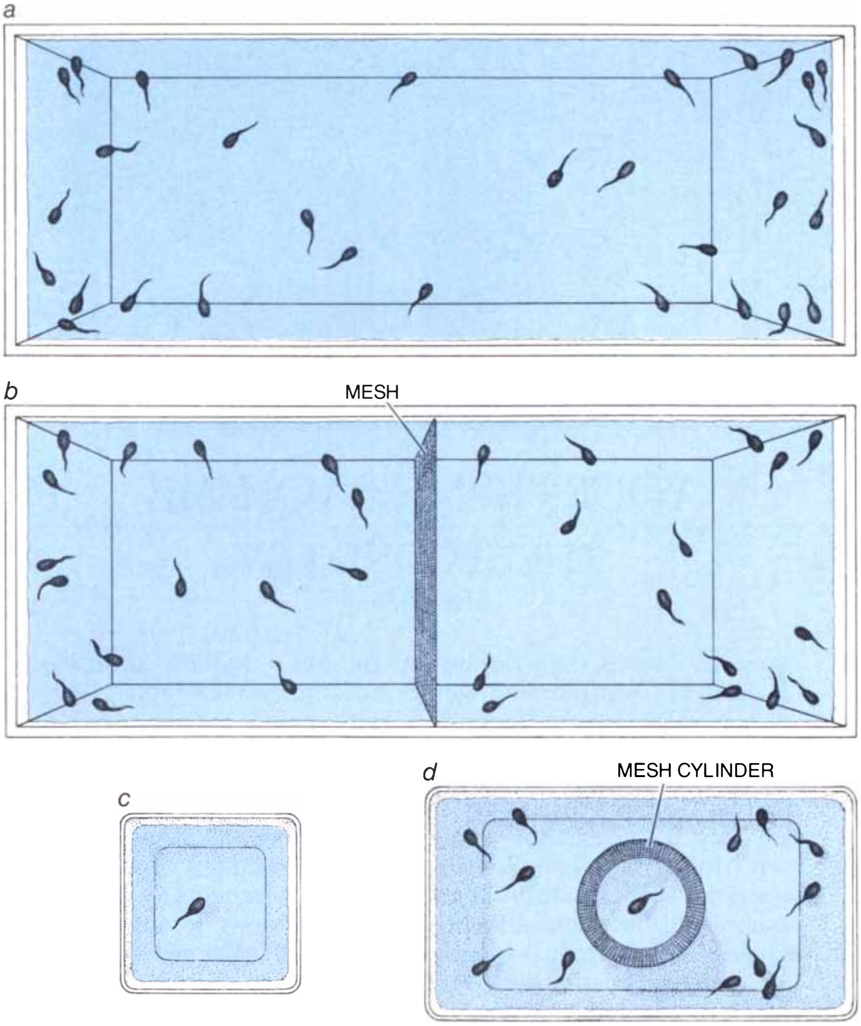
A special appendix equips readers who have computers with a kit of tools to put the Second Law to work.

230 pages, 104 illustrations

P. W. ATKINS is a lecturer in physical chemistry at the University of Oxford and a fellow of Lincoln College. He is the author of numerous distinguished works in his field, including *Physical Chemistry,* Second Edition, *The Creation,* and *Molecular Quantum Mechanics,* Second Edition.

The Mark II universe model of a paramagnetic material at some arbitrary temperature. In the absence of a magnetic field, the opportunity for electrons to be UP (yellow) or DOWN (green) is an additional contribution to the entropy of the sample. Atoms may still be ON (red) and OFF (white), but the spins they carry may be UP or DOWN at random. The two illustrations show the same distributions of ON and OFF, but different distributions of UP and DOWN.





FOUR REARING REGIMES expose developing tadpoles to different environments before they are tested for the ability to recognize kin. In one case embryos taken from a single egg mass are raised together in an aquarium, so that the tadpoles are exposed only to siblings (*a*). In a "mixed rearing" regime (*b*) embryos from two different egg masses are raised on each side of a mesh screen. Members of the two groups can exchange visual and chemical cues and thus are exposed to both siblings and nonsiblings. In a third regime a fertilized egg is put in an opaque dish (*c*), so that the tadpole is raised in isolation. In the fourth regime a single fertilized egg is placed inside a mesh cylinder (*d*), with 12 embryos from a different egg mass outside the cylinder: the enclosed tadpole is exposed only to nonsiblings.

enclosed each egg in an individual opaque container so that the tadpole would be reared in total isolation from other tadpoles. In the fourth rearing regime we enclosed a single newly fertilized egg in a mesh cylinder positioned at the center of a rearing aquarium. We then placed developing embryos from another clutch in the water surrounding the test embryo. Individuals reared in this regime were thus chemically and visually exposed only to unrelated tadpoles prior to being tested. In summary, the four regimes allowed us to work with four different sets of tadpoles: those reared with siblings only, those reared with both siblings and nonsiblings, those reared in total isolation and those reared with nonsiblings only.

To see what effects the various rearing regimes had on tadpole kin recognition we tested the four sets of tadpoles by allowing them to choose to associate with one or the other of two groups of stimulus tadpoles. The stimulus animals varied in their relatedness to and familiarity with the test individuals. Our standard tests were carried out in a rectangular tank divided into two basins by a central watertight, opaque barrier. We could test two tadpoles simultaneously but independently on opposite sides of the barrier. The ends of the tank were partitioned with plastic mesh to create small chambers in which stimulus tadpoles (usually 25 of them) swam. The test tadpoles could see the stimulus animals; exchange of chemical cues was possible, but there could be no tactile contact.

In a series of repeated experiments we recorded the time each test individual spent in the part of the tank near a particular stimulus group. That time could be compared with what would be expected if movement were totally random (50 percent of the total test time). The number of tadpoles that spent more of their time near one end of the test tank than they did near the other was also recorded as an additional measure of preference. Appropriate control tests revealed no biases in our procedures or apparatus.

Our first series of experiments established the basic fact that Cascades frog tadpoles can distinguish between siblings and nonsiblings. Some 80 percent of the tadpoles reared only with full siblings spent more time—a statistically significant amount more—near familiar siblings than they did near unfamiliar nonsiblings. Tadpoles tested both early (from 11 to 19 days after hatching) and late (from 34 to 40 days after hatching) in development displayed this preference.

The next step was to try to determine the influence of familiarity with other tadpoles on kin-recognition behavior. This might yield insights into the development of such behavior in the species and perhaps give us a handle on the mechanism controlling the behavior. What we did was to test three groups of tadpoles: those from mixed-rearing regimes, those reared with nonsiblings only and those reared in total isolation. After being reared in one of these regimes for an average of four weeks, test tadpoles were given a choice of associating with either an unfamiliar full-sibling group or an unfamiliar nonsibling group.

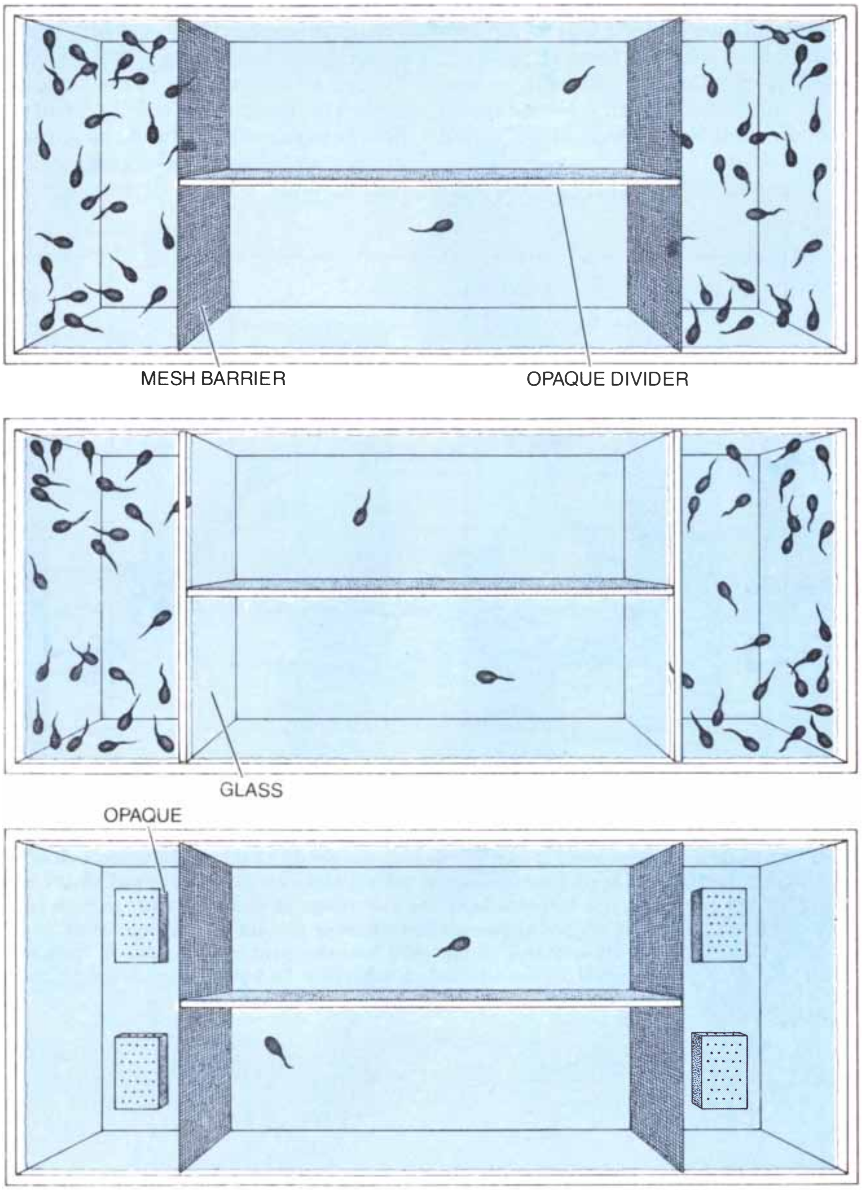
The results from these experiments showed us that, like tadpoles reared with full siblings only, tadpoles reared in a mixed-rearing regime also spent most of their time near their siblings rather than near nonsiblings. What was particularly significant in this series of experiments was that test tadpoles could distinguish between unfamiliar siblings and unfamiliar nonsiblings. Tests of tadpoles reared with only nonsiblings gave the same result: they preferred to associate with unfamiliar siblings rather than with unfamiliar nonsiblings. It appears, then, that exposure to unrelated tadpoles during development does not alter the preference for association with siblings. Moreover, early familiarity with siblings is not a prerequisite for identifying them later in life.

Finally, there were the tadpoles reared in total isolation, which had never associated with, or even seen, other free-swimming tadpoles before we tested them. We found they too preferred to spend most of their time swimming near their siblings. Apparently no experience with any other individual is necessary for kin-recognition behavior to develop. Because tadpoles could identify siblings they had not been in contact with since they were in the egg mass together as embryos, we conclude that familiarity cannot be the prime mechanism of kin recognition in this species. Phenotype matching, recognition genes or both must be implicated.

The degree of sensitivity a kin-recognition system displays may help to show what mechanism is operating. A highly sensitive system may well suggest a genetic component. How sensitive is the *R. cascadae* kin-recognition system? To answer the question we decided to see whether tadpoles could distinguish their full siblings from half siblings and half siblings from nonsiblings. We produced paternal half siblings by fertilizing the eggs of two females with sperm from one male, maternal half siblings by fertilizing one half of a female's clutch with sperm from one male and the other half with sperm from a second male.

In a series of experiments similar to our standard ones, we found that individual tadpoles preferred to associate with full siblings rather than with either maternal or paternal half siblings. Maternal half siblings were preferred over paternal half siblings or nonsiblings and paternal half siblings were preferred over nonsiblings. Apparently maternal or paternal cues alone are sufficient for kin recognition, and Cascades frog tadpoles can distinguish between individuals of varying degrees of kinship; this is indeed a remarkably sensitive kin-recognition system. The fact that test tadpoles preferred maternal to paternal half siblings suggests that maternal cues exert a stronger influence on recognition. Perhaps the reason is that the mother contributes more to the developing embryo (certain cytoplasmic factors in the eggs, for example) than the father, whose contribution is primarily genetic information. Still, the fact that paternal half siblings are distinguished from nonsiblings suggests that at least some part of kin recognition in this species is related to a genetic factor.

What can be the value for Cascades frog tadpoles of any kin-recognition system at all, let alone such a sensitive one? The most obvious possibility is that kin recognition is the basis of these tadpoles' ability to form small, dense schools. Some advantages of living in schools rather than alone are an enhanced ability to find and share a limited food supply, to avoid discovery by predators and to detect and escape from an approaching predator. Recent evidence from our laboratory shows, for example, that tadpoles under attack from their insect predators release chemicals that induce others within the group to flee. We think the prime function of kin recognition in Cascades frog tadpoles is to promote the cohesion of related individuals, possibly in order to facilitate feeding but more probably to avoid predators. To be sure, individuals could gain the group benefits described by being in a group with nonkin. By helping or warning kin, however, they help those



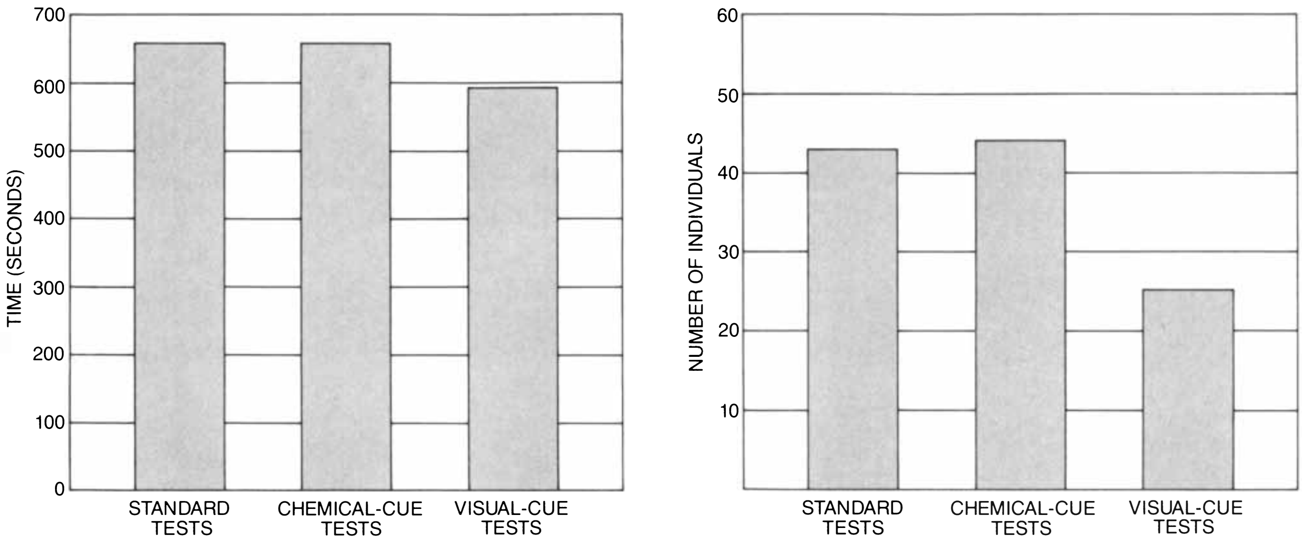
TADPOLE PREFERENCE TESTS are done in an aquarium 1.2 meters long. In the standard test (*top*) the aquarium is divided by an opaque central barrier so that two tadpoles can be tested simultaneously but independently. Plastic-mesh partitions create two end chambers, each of which is occupied by 25 "stimulus" tadpoles: either siblings or nonsiblings of a test animal. The individual being tested can choose to swim near either of the two groups of stimulus tadpoles. The amount of time it spends near each group is recorded. Most of the tadpoles from all four of the rearing regimes chose to spend more time swimming near their siblings than near nonsiblings. The sensory basis of kin recognition is determined by modifying the standard test tank in two ways. Watertight glass panes that replace the mesh end partitions (*middle*) allow the test tadpoles to see the two stimulus groups but prevent the exchange of chemical cues. Under these conditions there is no preference for siblings. When the stimulus groups are enclosed in boxes fitted with an opaque front facing the test animals but with perforated sides and tops (*bottom*), the test tadpoles cannot see the stimulus groups, but they can still sense chemical cues. Under these conditions test tadpoles showed a preference for their siblings, indicating that kin recognition is mediated by chemical cues.

with whom they share genes, thereby enhancing their inclusive fitness.

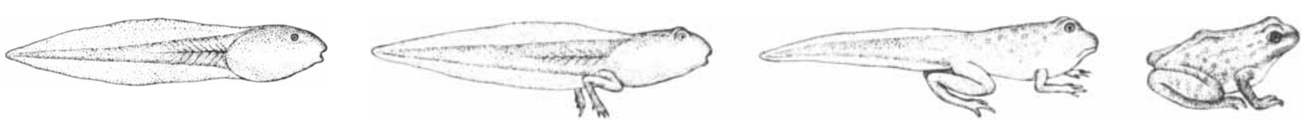
In the tests we have described so far the tadpoles had the full range of their senses. To identify the sensory basis of their kin-recognition behavior we undertook a new series of experiments. First we investigated the possibility of sound communication. We listened to individual tadpoles with a hydrophone and an oscilloscope, monitoring frequencies from 10 hertz (cycles per second), which is known to be at the lower limit of the range of hearing in animals, to 100,000 hertz. The tadpoles did not emit any sounds at the tested frequencies.

By modifying the procedures of our standard tests we then sought to learn whether the tadpoles communicate by means of visual cues or chemical cues. In one test we put stimulus animals behind a watertight glass pane instead of a plastic mesh, so that they and the test animals could see one another but could not communicate chemically. In another test we put stimulus tadpoles in chambers fitted with opaque fronts facing the test animals but with perforated sides and tops. The chambers were placed behind the plastic mesh of the end compartments in the standard tank. This design allowed chemicals to flow between stimulus and test animals but prevented any visual contact. The test animals could not identify kin when they could only see them. On the other hand, kin recognition was effective when only chemical cues could be sensed. Clearly chemoreception mediates kin recognition in these tadpoles. They produce some kind of chemical cue unique to their kin group, which they smell, taste or sense in some other way to distinguish kin from nonkin.

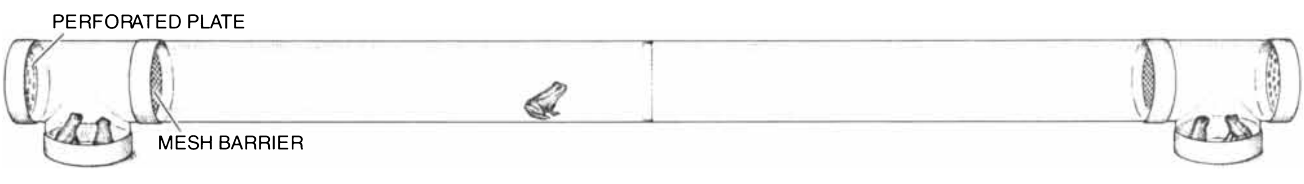
We wondered whether kin-recognition behavior is able to survive the drastic changes in anatomy, physiology, ecology and behavior that accompany metamorphosis, when the aquatic, primarily gill-breathing, omnivorous tadpole changes into an amphibious, lung-breathing, carnivorous



KIN-PREFERENCE BEHAVIOR of *R. cascadae* tadpoles was studied by allowing tadpoles to choose to swim near siblings or near nonsiblings. The average number of seconds tadpoles swam at the end of the test tank nearest their siblings is shown (*left*) for the standard test (in which the tadpoles have the full range of their senses) and for the chemical- and visual-cue tests. Testing was for a total of 1,200 seconds (20 minutes), so that 600 seconds spent on each side of the tank would represent random behavior. In both the standard and the chemical-cue tests the tadpoles spent an average of 660 seconds at the end of the tank close to kin, but when the tadpoles had to rely on visual cues, the time spent near siblings dropped to the random level. A total of 60 individuals were observed in each test; the number that spent more of their time near siblings than near nonsiblings is another measure of preference (*right*). In the standard and the chemical-cue tests some 75 percent of the individuals tested spent more of their time near their siblings.



DEVELOPMENT of tadpoles into frogs proceeds over a span of from six weeks to more than eight weeks. Four representative stages (in a continuum of 26 stages) in this process are shown. (The age at which individual tadpoles reach a particular stage is extremely variable.) *R. cascadae* displays kin-recognition behavior at each of the four developmental stages, including the froglet stage (*right*).



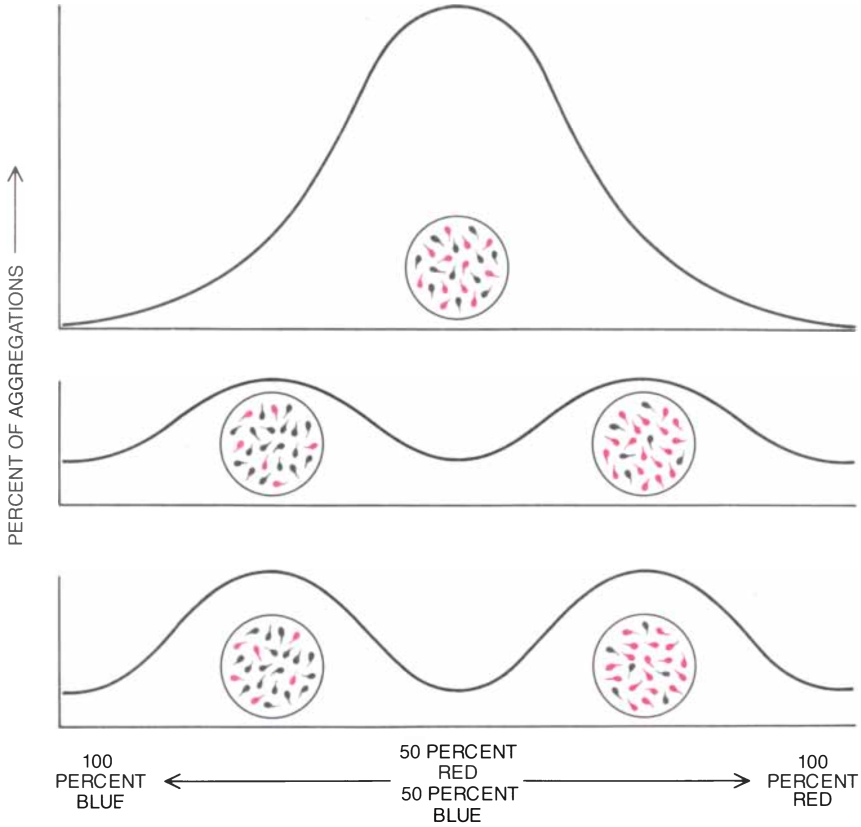
TEST CHAMBER was designed to determine whether kin recognition persists after metamorphosis. The chamber at one end of the clear plastic tube holds siblings of the test frog in the tube; the chamber at the other end holds nonsiblings. Kin preference is determined by the amount of time the test frog spends near the end of the tube housing its siblings compared with the amount of time it spends near its nonsiblings. The test showed that *R. cascadae* continues to exhibit kin-preference behavior even after metamorphosis.

frog. To address this question we undertook the formidable task of caring for hundreds of newly metamorphosed frogs. Then we tested froglets from four to 12 days after metamorphosis and again from 39 to 47 days after metamorphosis for their ability to identify kin.

We did this in a way comparable to our standard preference tests for tadpoles but in an apparatus that would accommodate frogs. We placed individual froglets in a moist, transparent plastic tube that was partitioned by screens near both ends to create two stimulus compartments and a test compartment. One stimulus compartment held siblings of the test froglet and the opposite one held nonsiblings. We recorded the amount of time froglets spent in the part of the tube near their siblings compared with the amount of time spent near nonsiblings. The results suggested that the froglets had apparently retained the ability to identify kin in spite of the drastic changes they had undergone in anatomy and physiology.

What advantage is there in kin recognition for Cascades frogs after metamorphosis, considering that they are not known to form aggregations as adults? Perhaps an ability to recognize kin confers no selective advantage on adult frogs and is retained after metamorphosis simply because there is no particular selective pressure against retaining it. It is also possible that the adults depend on kin-recognition ability either in reproductive behavior (to avoid breeding with close relatives) or in social organization, perhaps to find their home pond after having been displaced.

Laboratory experiments can only approximate and simplify the complex social and ecological conditions animals encounter in nature. For this reason during three consecutive summers we did experiments in natural ponds and lakes inhabited by Cascades frog tadpoles. We got tadpoles for field experiments by collecting egg masses from natural breeding ponds and rearing the individuals in the laboratory either with siblings only or in mixed-rearing regimes. We marked tadpoles from two kin groups with different colors by immersing them in a harmless red or blue dye. Then we intermingled members of the two groups and released them in natural ponds in the Cascade Range. We considered the possibility that tadpoles might form groups based on their new color rather than on kinship. As a control we therefore divided some single kin groups, dyed half of the members of each group red and the other half blue and released the groups in control ponds.



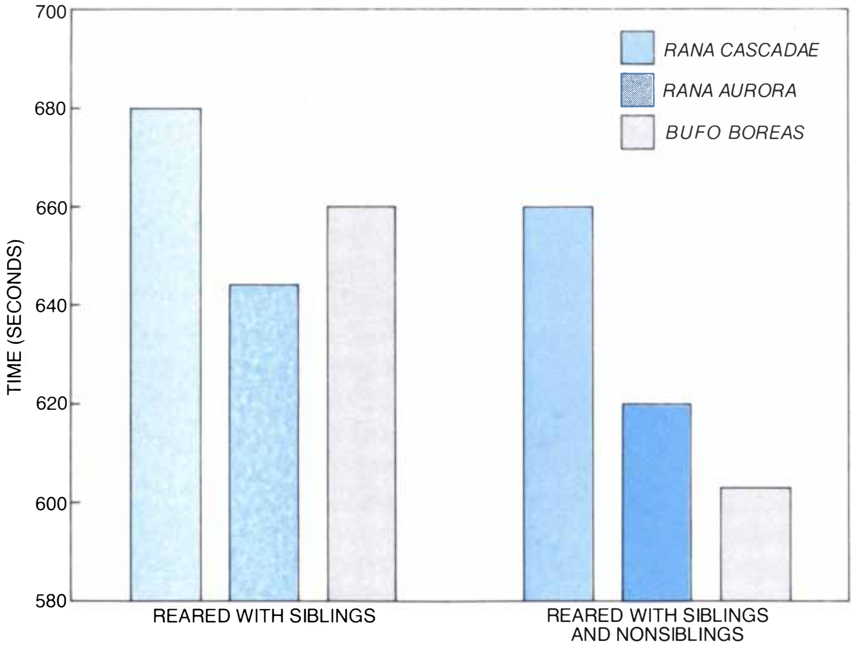
SCHOOLING PREFERENCES of *R. cascadae* tadpoles in a natural setting rather than in the laboratory were studied by marking the animals with a blue or a red dye, releasing them in ponds and determining the color composition of the aggregations they formed. As a control, half of the members of single clutches were dyed blue and the other half were dyed red. The siblings were released in ponds and the number of tadpoles of each color in the resulting schools was counted. When the results of a series of control runs were plotted, they yielded a bell-shaped curve (*top*): most of the schools were roughly half blue and half red, indicating that color in itself does not influence aggregation. For the actual tests, members of one clutch were dyed blue and those of a different clutch were dyed red. When the results of a number of tests were plotted, the color composition of the schools these tadpoles formed was seen to approach a "bimodal" distribution: whether the siblings had been reared apart (*middle*) or together (*bottom*), most of the schools were predominantly blue or predominantly red. In other words, most of these aggregations consisted of kin.

It was not an easy task to approach and collect aggregations of tadpoles. If they detected us moving toward them, they would disperse rapidly. Eventually, however, we did determine that red and blue tadpoles from the same kin group aggregated randomly in control ponds: dye color had no effect in 103 aggregations we sampled. In contrast, the distribution of red and blue tadpoles from different kin groups was not at all random: the 353 aggregations we sampled in the experimental ponds were dominated by one of the two colors. The results of these field experiments lead us to believe that Cascades frog tadpoles can identify their kin, and prefer to associate with them, in nature as they do in the laboratory.

We had originally speculated that certain aspects of the behavior we had noted in Cascades frog tadpoles might be connected with an active kin-recognition system. Having shown that to be true, we looked at other anuran species displaying different behavioral characteristics.

Red-legged frogs (*Rana aurora*) are closely related to the Cascades frog. When we put them through a series of standard preference tests, however, the results were very different. The red-legged tadpoles can identify kin only early in the free-swimming stage, and then only if they have been reared with full siblings; they lose their ability to recognize kin as they develop. Western toad (*Bufo boreas*) tadpoles can recognize kin through the tadpole stage, but only if they have not been exposed to nonkin. If they have been exposed to nonkin, they rapidly lose their ability to identify kin.

Our results with Western toads are similar to those reported by Bruce Waldman of Cornell University, who investigated tadpoles of the closely re-



SPECIES DIFFERENCES in kin-recognition behavior are revealed when tadpoles of two frog species and a toad species are given the standard kin-recognition tests. All three species prefer to associate with kin if they have been raised with siblings only; that is, they spend more than the random amount of 600 seconds at the end of the tank nearest siblings. (This is true for early and late developmental stages of *R. cascadae* and the toad *Bufo boreas* but is true only for the early larval stages of *Rana aurora*.) For tadpoles reared with both siblings and nonsiblings the amount of time spent with siblings drops to the random level in *R. aurora* and *B. boreas;* the preference for kin remains significant in *R. cascadae.*

lated American toad, *Bufo americanus.* In laboratory experiments he found that American toad tadpoles generally cannot distinguish between kin and nonkin if they are reared in mixed-rearing regimes. It is likely that in both toad species, and possibly in the red-legged frog, familiarity influences kin identification.

It was not surprising to learn that red-legged frog and Western toad tadpoles do not have a well-developed ability to recognize kin. The tadpoles of both species disperse rapidly from their hatching sites. Moreover, red-legged frog tadpoles are not known to aggregate, and Western toad tadpoles are usually found in huge schools (up to a million tadpoles) composed of individuals from hundreds of different clutches. Probably tadpoles of these species have not had enough opportunity for interaction with kin during development to have evolved a sensitive kin-recognition system that allows them to form cohesive groups.

When the results of kin-recognition studies in various animal species are surveyed, certain patterns emerge. For example, aquatic forms such as fish and tadpoles depend on water-borne chemical cues. Birds depend on fixed genetic cues such as plumage coloration or variable cues such as vocalizations. Olfactory cues are the basis of identifying kin in most species of insects and mammals. It is now clear that the commonest of the three mechanisms for identifying kin is familiarity, which is efficient when there is a high probability that the individuals a particular animal interacts with are relatives. Kin recognition is influenced by familiarity in the Western toads we examined; familiarity has at least some role in the recognition abilities of most of the birds and small mammals that have been studied. Familiarity is probably an important means of avoiding inbreeding in many species.

Although we are sure familiarity is not an essential part of the kin-recognition mechanism in Cascades frog tadpoles, we cannot be sure whether the mechanism is learned phenotype matching or derives from innate recognition genes. It is often difficult to distinguish between the two, although there are some fairly clear-cut cases of phenotype matching. For example, Gregory R. Buckle and Les Greenberg of the University of Kansas showed that sweat bees (*Lasioglossum zephyrum*) guarding nest entrances distinguish among unfamiliar bees on the basis of phenotypic similarity to individuals with which they were reared. The guards learn what their nestmates smell like and admit only those unfamiliar individuals that smell like the nestmates. Richard H. Porter of Vanderbilt University has demonstrated that spiny mice (*Acomys cahirinus*) that are artificially "odorized" prefer to associate with unfamiliar individuals that have been given the same odor.

Clear cases of a recognition-gene mechanism are harder to establish, but recent studies of mating preferences in the house mouse (*Mus musculus*) have yielded strong evidence that is consistent with a genetic recognition system [see "The Chemosensory Recognition of Genetic Individuality," by Gary K. Beauchamp, Kunio Yamazaki and Edward A. Boyse; SCIENTIFIC AMERICAN, July, 1985]. The results showed that mice can choose between potential mates that differ at a single genetic site in the Major Histocompatibility Complex (*H*-2), which specifies the antigens responsible for distinguishing between self and nonself. Male mice prefer to mate with females whose *H*-2 type is different from their own. The preference is based on the genetically determined scent associated with particular *H*-2 alleles.

The sensitive kin-recognition system of Cascades frog tadpoles may also be a genetic recognition system. One cannot, however, rule out the possibility that a Cascades frog tadpole, even one reared in total isolation, may learn its own genetically determined odor by experiencing itself. It could then match the odors of two unfamiliar groups and compare them with its own odor; the group having the more familiar odor would be taken to be more closely related. It may be impossible to determine experimentally whether phenotype matching or an innate genetic recognition system is operating in Cascades tadpoles; it is conceivable that both are operating simultaneously. In fact, one proposition holds that the differences between the two mechanisms are trivial. Both mechanisms enable an animal to identify an unfamiliar relative, and so the evolutionary consequences are the same.

In addition to the theoretical implications bearing on familial altruism, kin recognition may yield important practical benefits. Knowledge of how members of a particular species recognize one another is critical in animal husbandry and in the propagation of certain endangered species, particularly where "familiarity" with potential mates influences mating behavior. The newly demonstrated connection between a kin-recognition system and histocompatibility genes suggests a very different kind of benefit: the studies of kin-recognition behavior may shed some light on the workings of the immune system.


USA
TRAINING GEAR
ellesse

No other system of keeping up can compare with SCIENTIFIC AMERICAN *Medicine.*

YOUR SYSTEM: a time-consuming, futile struggle to keep up with the information explosion.

The classic texts are convenient references—but the information they contain is obsolete before publication.

Like many physicians, you probably rely on the texts you first used in medical school. But even using the most recent editions, you find material that no longer reflects current clinical thinking—and they lack the latest information on such topics as herpes, oncogenes, AIDS, and photon imaging.

Reading stacks of journals alerts you to recent developments—but can't give you quick answers on patient management.

Struggling through the hundreds of journal pages published each month—even on only the really significant advances in the field—is arduous and memory-taxing. And it's a task that costs physicians valuable time—their most precious resource.

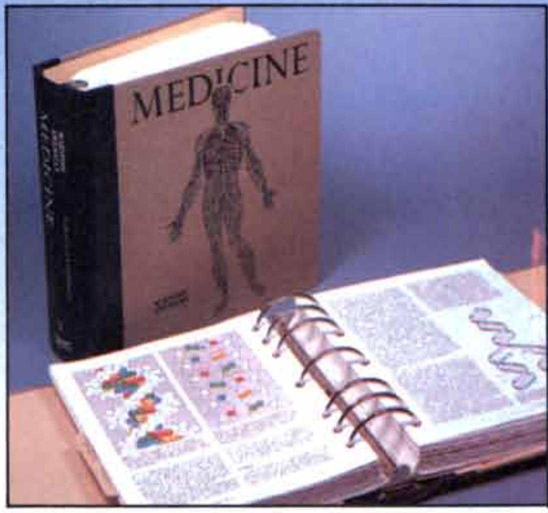
Review courses cover clinical advances—but, months later, do you recall the details of a new procedure or unfamiliar drug?

Seminars can also be costly and make you lose valuable time away from your practice—expenses that may amount to several thousand dollars. And, the speaker's skill often determines how much you learn.

SCIENTIFIC AMERICAN

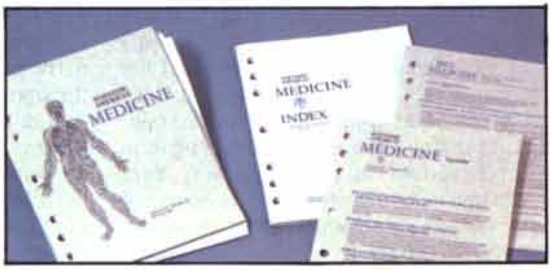
MEDI

OUR SYSTEM: a rewarding, efficient way to keep yourself up-to-date—and save hundreds of hours of your time for patient care.



A comprehensive, 2,300-page text in two loose-leaf volumes, incorporating the latest advances in medical practice as of the month you subscribe.

This superbly designed, heavily illustrated resource, called "the best written of all [the internal medicine] books" by *JAMA* (251:807, 1984), provides a practical, comprehensive description of patient care in 15 subspecialties. And, because the text is updated each month, the clinical recommendations reflect all the current findings. A practice-oriented index and bibliography of recent articles further enhance the efficiency of the text.



Each month, six to nine replacement chapters to update your text *plus* new references, a four- to five-page news bulletin, and a completely new index.

You'd have to read hundreds of journal pages each month—and memorize the contents—to get the same information SCIENTIFIC AMERICAN *Medicine* contains. With our updated text, you read only the information you really need. Our authors, largely from Harvard and Stanford, sort through the literature and monitor developments, incorporating the significant advances into our chapters.

At no additional cost, a 32-credit CME program, to save you valuable patient-care time and the expense of attending review courses.

Earn 32 Category 1 or prescribed credits per year with our convenient self-study patient management problems; each simulates a real-life clinical situation. Choose either the complimentary printed version or, at a modest extra charge, the floppy-disk version for your IBM® PC/PC Jr or Apple® IIe or II+ (or compatible models).

Your subscription includes:
the two-volume, 2,300-page loose-leaf text
and, each month, six to nine replacement chapters,
a newsletter, new references,
and a completely revised index.

call toll-free: 1-800-345-8112

(in Penn. 1-800-662-2444)

CINE

SCIENTIFIC AMERICAN MEDICINE
415 MADISON AVENUE, NEW YORK, N.Y. 10017

- ☐ **Yes,** I'd like to try the SCIENTIFIC AMERICAN *Medicine* system. Please enter my subscription at a first-year price of US$245*.
- ☐ Also enroll me in the CME program. I'd prefer:
  - ☐ the printed version at **no additional charge.**
  - ☐ the floppy-disk version at $US87* additional.
    Computer type or compatible:
    - ☐ IBM® PC/PC Jr. (128K RAM, DOS 2.0/2.1)
    - ☐ Apple® IIe (64K RAM)
    - Apple II+ (64K RAM) 80-col. card by: ☐ Apple ☐ Videx™
- ☐ Enroll me in the floppy-disk CME **only** at US$148.* (Note model above.)

☐ Check enclosed* ☐ Bill me

Signature ____________

☐ VISA ☐ MasterCard Exp. Date ____________

Account No. ____________

Name ____________

Specialty ____________

Address ____________

City ________ State ________ Zip Code ________

*Add sales tax for Calif., Ill., Mich., or N.Y. Allow 8 weeks for delivery. Add US$10 for shipping to Canada. IBM is a registered trademark of International Business Machines Corporation. Apple is a registered trademark of Apple Computers, Inc. Videx is a trademark of Videx, Inc. 7A

THE AMATEUR SCIENTIST

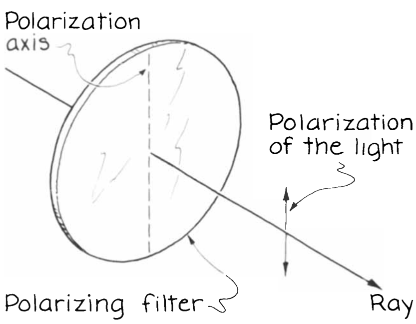
An inexpensive homemade polarimeter can analyze optically active compounds

by Jearl Walker

Solutions containing optically active compounds such as sugar rotate the polarization of light passing through them. The rotation reveals asymmetries in the construction of the compounds. This phenomenon also has practical applications, such as controlling the concentration of sugar in food products and sugar refining. Sam Epstein of Los Angeles has designed an inexpensive polarimeter that measures optical activity.

A classical model of light describes it as a moving wave of oscillating electric fields. Often the light is depicted as a ray in order to indicate its direction of travel. Superposed on the ray are vectors representing the direction and strength of the associated electric field. The vectors are always perpendicular to the ray, and they vary in direction and strength in such a way that the composite resembles a wave. The electric field appears to oscillate as the light moves past a given point.

When the light is not polarized, the field can oscillate in any direction perpendicular to the ray. If the light passes through a polarizing filter, the oscillations are restricted in such a way as to be parallel to a single axis perpendicular to the ray. The light is said to be polarized. Its polarization is represented by a double-arrow vector. An imaginary line (the polarization axis) parallel to the polarization of the light represents the effect of the filter.



A filter's polarizing action

If a second polarizing filter is placed in the path of the beam, the intensity of the light that passes through it depends on the filter's orientation. If its polarization axis is parallel to the oscillations of the incident light, the light is fully transmitted. If the axis is perpendicular to the oscillations, the light is fully blocked. Intermediate orientations of the filter pass intermediate intensities of light.

The filter closest to the light source is called the polarizer; the other filter is the analyzer. What you see in using the instrument is the light emerging from the analyzer as it is rotated about the original ray. When the axes of the two filters are parallel, you see the brightest light. After a 90-degree rotation of the analyzer you see no light. This position of the analyzer is called the endpoint.

Solutions of optically active compounds such as lactic acid, tartaric acid, nicotine, turpentine, amino acids and vitamins rotate the polarization of light passing through them. They are distinguished from other compounds by their three-dimensional structure. An optically active compound has one or more carbon atoms, each of which is attached to one of four different types of atoms or groups of atoms. As light passes a carbon atom and its attachments, the electric field of the light interacts with the atoms in a way that rotates the polarization of the light about the ray.

Suppose a cell that contains a solution of an optically active compound is placed between the filters. When the polarized light passes through the solution, its polarization is rotated about the ray. Hence when it reaches the analyzer, it has an orientation different from the one it had before the cell was introduced. To block the light the analyzer must be rotated to a new endpoint. You study the optical activity of solutions by measuring the angle through which the analyzer must be turned to block the light.

The direction of rotation of the light's polarization is specified from the perspective of the observer. The solution is said to be dextrorotatory if the rotation is clockwise and levorotatory if it is counterclockwise. The extent of the rotation is determined by how many of the optically active molecules the light passes on its way through the cell. Longer cells and higher concentrations of molecules increase the rotation of the polarization. The rotation also varies with the wavelength of the light.

In order to describe how much a certain compound rotates the polarization of light, one speaks of "specific rotation." In Epstein's work this measure is the angle through which the polarization rotates when the light passes through one decimeter (.1 meter) of a solution in which the concentration is 100 grams per 100 cubic centimeters. (Some references define specific rotation in terms of other units.) The wavelength of the light is usually taken to be 589 nanometers, that of the yellow emission line of sodium. The temperature of the solution is usually 20 degrees Celsius.

Epstein's polarimeter operates on light from a 60-watt bulb. The light passes through a color filter and a collimating lens and then into a housing, where it is reflected upward from a mirror. In the housing it travels through a polarizer and a cell holding the solution of interest. Thereafter the light proceeds through a condensing lens and an analyzer, finally reaching an eyepiece through which the endpoint is determined.

The bulb and its socket are mounted on a wood support and covered with an inverted fruit can. The can is mounted about an inch above the support so that air can flow into the can. Holes in the top of the can allow air heated by the bulb to escape. Extending from a hole punched in one side of the can is a length of polyvinyl chloride (PVC) pipe of one-inch internal diameter. The outer end of the PVC tube is covered with a thin plate of ground glass.

Since most data on specific rotation for optically active compounds are listed for the yellow emission line of sodium, Epstein filters the white light emitted by the bulb. He avoided the cost of a professional color filter by making a filter. His rig is made with two microscope slides that serve as windows in a rectangular cell. The top of the cell is made of plastic fitted snugly into place to reduce evaporation. The rest of the cell is glued together with epoxy. The cell contains a water solution of potassium dichromate at a 10 percent concentration (10 grams per 100 milliliters of solution). The cell is placed in a protective covering made of sheet metal or quarter-inch Masonite, and the assembly is mounted on a wood pillar so that the windows are at the proper height to intercept the light.

The potassium dichromate solution acts as a filter because it transmits a narrow range of wavelengths close to the yellow emission line of sodium. Therefore Epstein's combination of a white-light source and a color filter yields approximately the same light as a professional's sodium emission lamp. (The color filter can also be made from a square of orange No. 22 Kodak Wratten gelatin filter.)

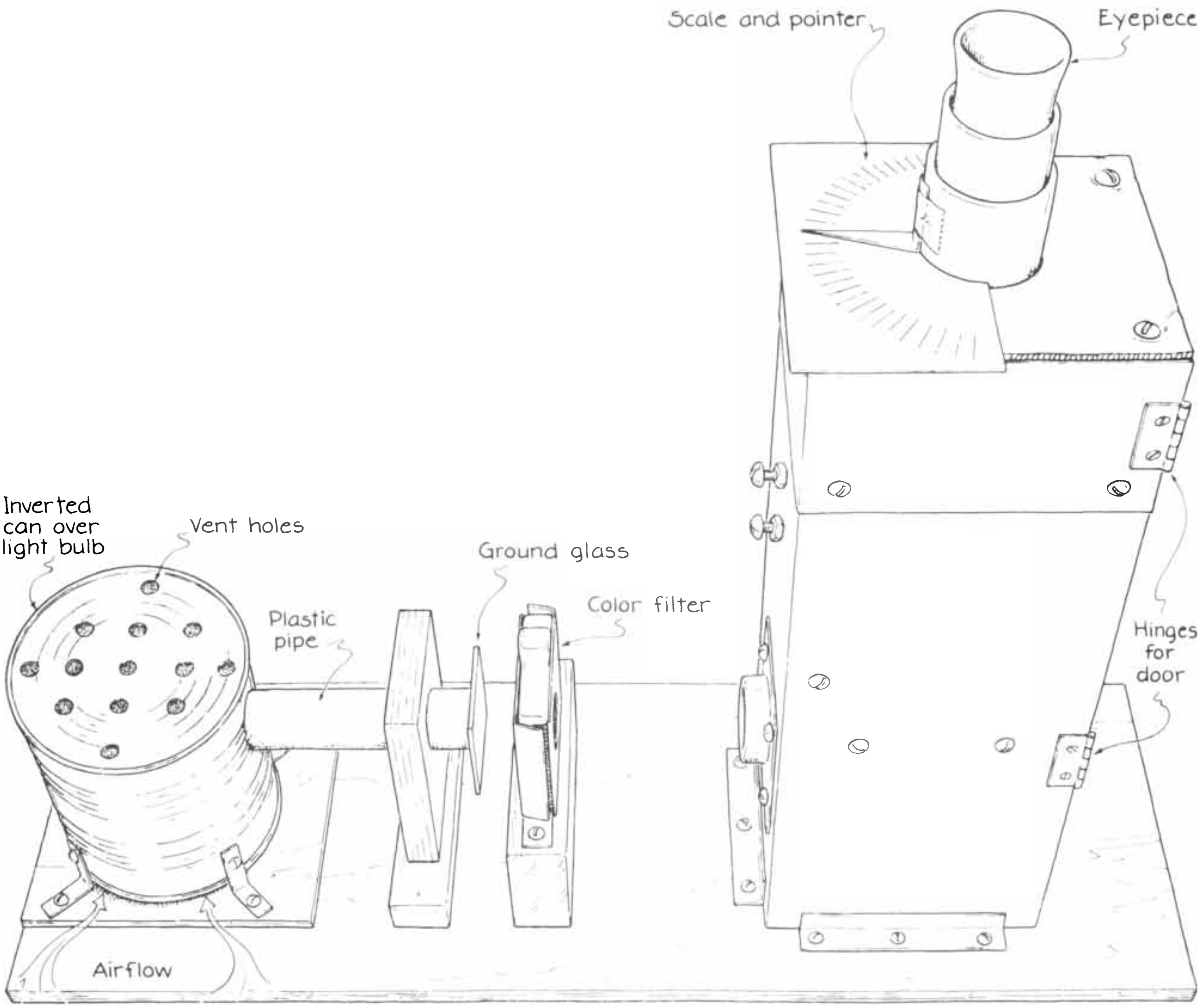
The housing for the rest of the apparatus is made from quarter-inch plywood or Masonite to form a rigid support. The interior is painted flat black to eliminate stray light. A door on one side of the housing provides access to the sample cell.

The distance between the color filter and the collimator lens is equal to the focal length of the lens, so that the light from the filter passes through the rest of the apparatus approximately as a beam. The lens is mounted in a section of PVC pipe that shades the lens from extraneous light. The light shines on a front-surface mirror glued with epoxy to a wood support. The mirror is mounted at an angle of 45 degrees with respect to the horizontal. Epstein cautions that proper alignment of the optical path is essential.

Just above the mirror is a partition made from quarter-inch plywood or from Masonite. The mirror directs light through a hole five-eighths of an inch in diameter cut into the partition. The polarizer is glued to the bottom of the partition.

Epstein made the polarizer by sandwiching a one-inch square of polarizing filter between two microscope slides. The position of the square is maintained by two sections of file-card stock. The edges of the two microscope slides are covered with transparent tape that extends one-eighth of an inch inward to keep the polarizing filter from sliding out of position. A spot of epoxy is put on both sides of each piece of file-card stock to glue the sandwich together.

A microscope slide is glued across the hole in the partition to protect the polarizer sandwich and the mirror against leaks of the solution being tested. On top of the slide is a one-by-two-inch Masonite pad. A hole in the pad matches the hole in the partition. The sample cell containing the test solution



The polarimeter designed by Sam Epstein rests on the pad and is held in the path of the light by a six-inch length of PVC pipe, sawed so that the cross section of the top five inches is a half circle. The pipe is attached to the housing of the apparatus with two 10–32 brass bolts.

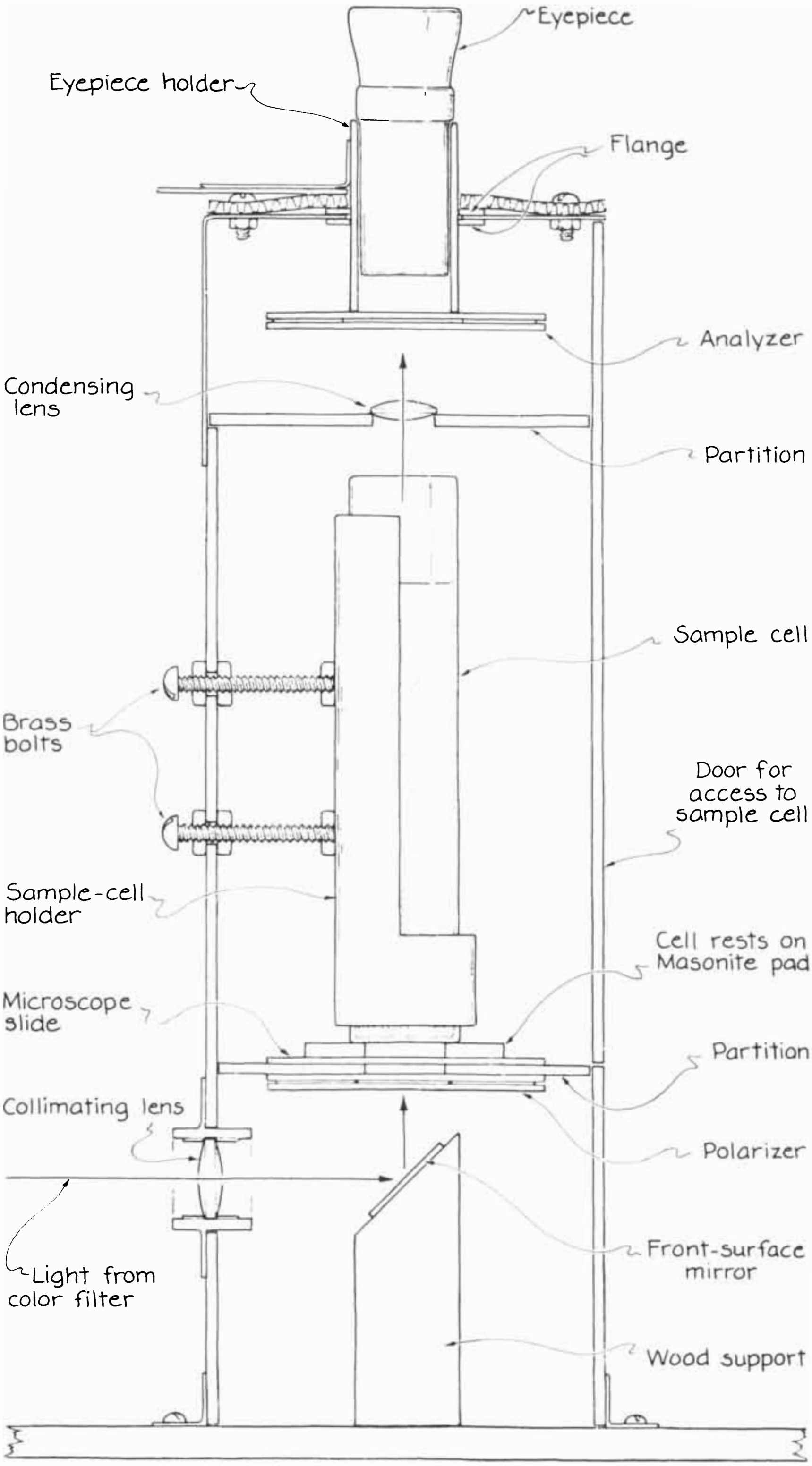
A sample cell must have a flat bottom to avoid distorting the light beam. Epstein uses 50-milliliter color-comparison tubes (Type EXAX, low form) that are available from most laboratory-supply houses. The tubes slide into the PVC holder and rest on the Masonite pad.

Because the tubes cost about $8 each, Epstein suggests that you build your own cell from a transparent plastic tube that is one inch in diameter and has a wall 1/16 inch thick. Cut off a six-inch length of the tube and grind one end flat. Glue a 1½-inch length of microscope slide to the flat end. The slide should be centered on the tube. The glue should be epoxy and must provide a watertight seal.

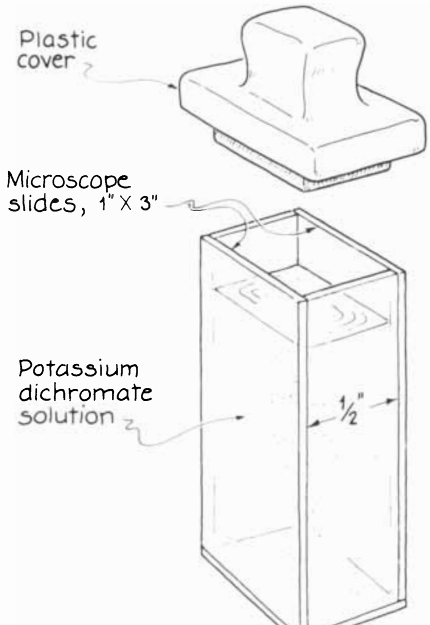
This design has two drawbacks. You will need to modify the PVC holder so that the cell fits into place. A more serious problem is that the cell may be ruined if you do experiments with some types of organic solvents.

Above the sample-cell holder is a condensing lens that directs light from the cell through the analyzer and the eyepiece. The analyzer is a sandwich made in the same way as the polarizer is made with one important difference: the filters are skewed. Begin with a 1¼-inch square of polarizing filter. Cut a five-degree triangular segment out of the center. Slide the remaining parts of the filter together. Trim them so that they form a one-inch square. Sandwich this arrangement between the microscope slides.

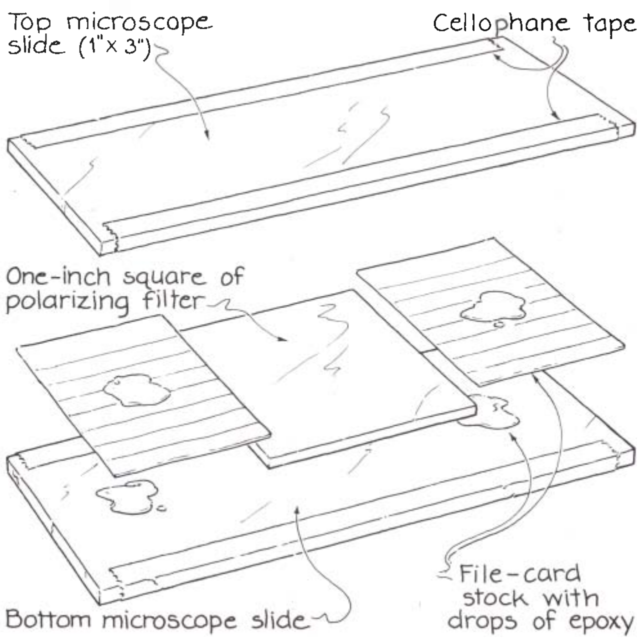
The skewed arrangement of filters in the analyzer makes the determination of the endpoint easier. Otherwise you must guess what position of the analyzer best eliminates the light passed by the test solution. With the skewed filters you merely compare the relative brightness of the light passing through each part of the filter arrangement. The endpoint is achieved when the two parts are equally bright. If you



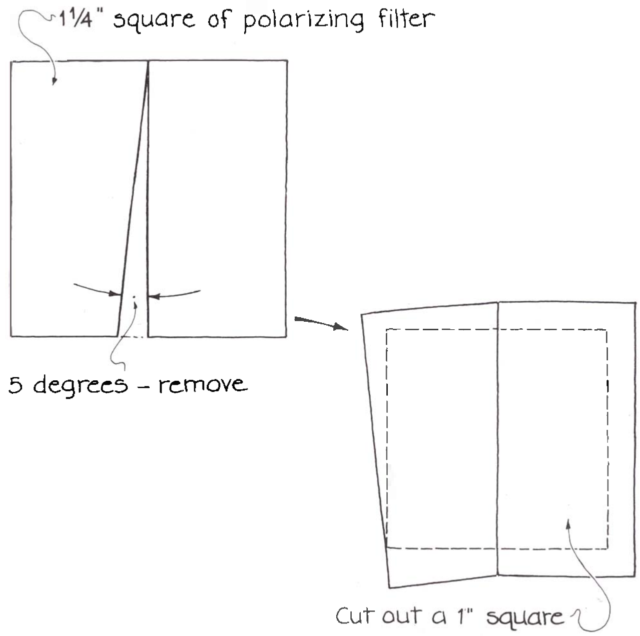
Details of the apparatus holding the sample cell



Epstein's color filter



How the polarizer is assembled



Skewed parts of the analyzer

rotate the analyzer in either direction from the endpoint, one part brightens and the other part darkens. Hence the skewed arrangement of filters enables you to fine-tune the determination of the endpoint.

The analyzer sandwich is glued to the bottom of a holder for the eyepiece. The holder passes through a hole in the housing. It is held in place with two flanges but is still free to turn in the hole. The eyepiece can be a single lens or a low-power compound-lens system. A scale marked in degrees surrounds the holder.

To set a pointer on the scale switch on the light bulb and turn the eyepiece holder (and thus the analyzer), monitoring the brightness of the light as you look through the eyepiece. Find the position where the two parts of the analyzer are equally bright. Then attach a pointer to the eyepiece holder so that it points to zero on the rotation scale.

The bottom of a sample cell often has optical imperfections that alter the polarization of light. To eliminate this possibility fill each cell with 50 milliliters of water and mount it in the apparatus. Rotate the cell in its holder, monitoring the light through the eyepiece, until you get the best overall image and endpoint. Permanently mark the side of the cell holder. Make a corresponding mark on the cell. Whenever you use that cell, place it in the holder with the marks aligned. Repeat this procedure for each cell.

Epstein suggests testing the optical activity of a solution of sucrose (ordinary sugar). Begin with a sample cell containing 50 milliliters of water. Place the cell in the polarimeter with the mark on its side aligned with the mark on the cell holder. Rotate the eyepiece to find the endpoint. Note where the pointer lies on the circular scale of degrees.

Exchange the water for 50 milliliters of sugar solution at a concentration of 20 grams per 100 milliliters. Again check for the endpoint. Epstein finds that it is rotated clockwise by about 17 degrees from its position on the scale when only water is examined.

The rotation can also be ascertained by a calculation. The specific rotation for sugar is 66.5 degrees clockwise. In Epstein's experiment the length of the light path is 1.3 decimeters. To compute the expected rotation multiply the specific rotation, the concentration (grams per milliliter) and the length (decimeters). The expected rotation is clockwise by about 17 degrees. You might like to determine how the rotation depends on the sugar concentration. Begin with the strongest solution and gradually dilute it as you measure the rotation.

Epstein also investigated a well-known reaction in which sucrose is broken down into two simpler sugars, dextrose and levulose. (The process is called inversion.) Dissolve 20 grams of sucrose in 50 milliliters of water in a 100-milliliter flask. Mix it well. Prepare 10 milliliters of a mixture of hydrochloric acid in the ratio of one part acid to three parts water. Add the mixture to the flask and again mix the solution. Put a thermometer into the flask and place the flask in a water bath at 70 degrees C. Monitor the temperature of the solution until it reaches 67 degrees. Keep it in the bath for five more minutes. (Do not let the temperature of the solution exceed 69.5 degrees.)

Transfer the flask to another water bath at 20 degrees C. Remove the flask from the bath when the temperature of the solution falls to 35 degrees. When the temperature of the solution reaches 20 degrees, rinse the thermometer with about 25 milliliters of water, letting the water run into the flask. Add enough water to bring the volume in the flask to 100 milliliters. Again mix the contents.

Place the flask in the 20-degree bath for another 15 minutes. If necessary, again add enough water to the flask to bring the volume to 100 milliliters. Mix the contents and pour 50 milliliters of it into a sample cell. Position the cell in the polarimeter and measure how much the solution rotates the polarization of the light. Epstein measures a rotation of about 2.8 degrees counterclockwise.

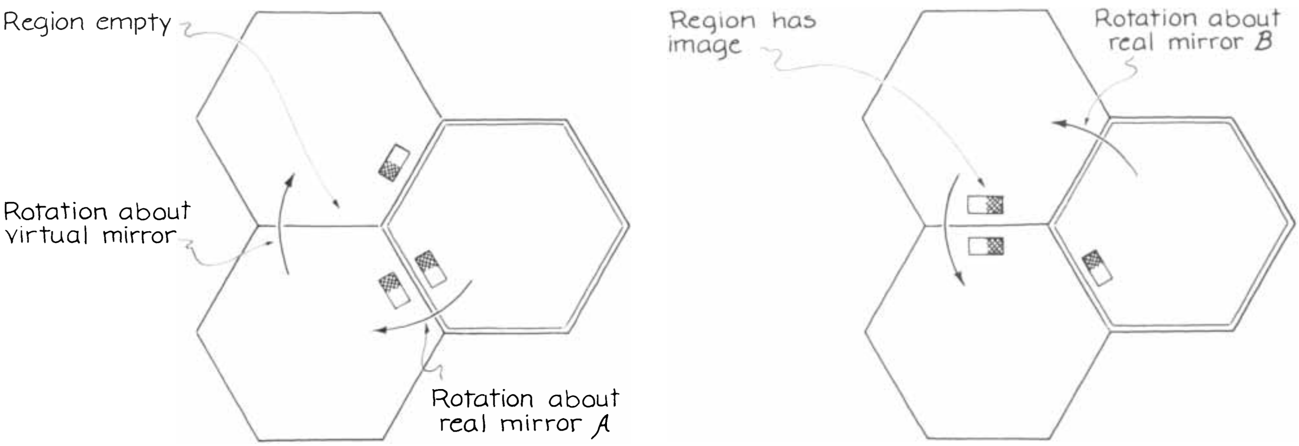
The breakdown of sucrose is catalyzed by the hydrochloric acid, producing 10.5 grams each of dextrose and levulose. The specific rotation of dextrose is 52.5 degrees clockwise, the specific rotation of levulose 93 degrees counterclockwise. Calculate the rotations created by the two products in the sample cell. The dextrose should rotate the polarization of light by about 7.16 degrees clockwise and the levulose should rotate it by about 12.7 degrees counterclockwise. Since the rotations are in opposite directions, the calculation of the average rotation amounts to subtracting the two numbers and dividing by two. The result is about 2.8 degrees counterclockwise, just as Epstein measured.

You might also enjoy studying the optical activity of such substances as corn syrup, maple syrup and pancake syrup. Add 25 milliliters of the syrup to a 100-milliliter volumetric flask, along with about 25 milliliters of rinse water to ensure the complete transfer of the syrup. Add two drops of concentrated ammonium hydroxide to serve as a catalyst. Mix the contents well, add enough water to bring the volume to 100 milliliters and mix again. Pour 50 milliliters into a sample cell and measure the rotation of polarization. The molecules of the sugars in the syrup rearrange themselves in a process called mutarotation until an equilibrium is reached. Before equilibrium the optical activity of the solution varies. Epstein suggests that you take readings on the polarization rotation about every 15 minutes until the variations disappear.

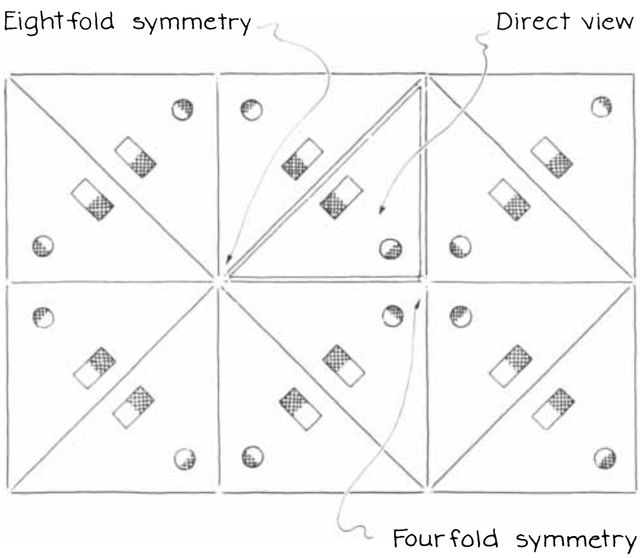
With a similar procedure you might also study the optical activity of different types of honey. The procedure for sugar inversion can also be employed to study the optical activity of gelatin, soft drinks and other liquids containing sugar. If the gelatin is flavored, make sure it is the orange variety in order to obtain a color close to that of the sodium yellow light. You might also compare the polarization rotations of both pure and sweetened fruit juices.

Last month I examined the optics of kaleidoscopes consisting of either two or three mirrors. In each type one can have a direct view of the objects at the far end of the kaleidoscope and see additional images reflected from the mirrors. In a two-mirror system the images lie in pie-shaped sectors clustered around the vertex at the intersection of the mirrors. In a three-mirror system the entire field of view is filled with images. In both types the images appear to lie in a flat plane called the image field that extends through the direct view.

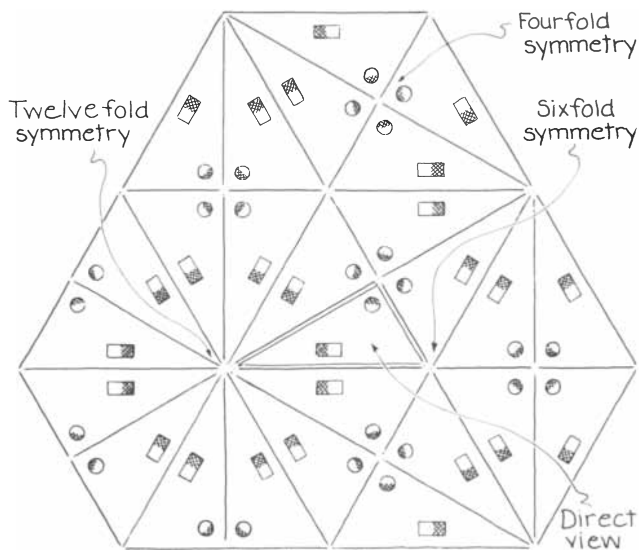
Most arrangements of three mirrors create image fields that are ambiguous in the sense that the content in any area of the field depends on your angle of view into the kaleidoscope. Suppose you see a red bead at a certain place in the field. If you change your perspective, something else may replace the bead. The image fields in a kaleidoscope whose mirrors form an equilateral triangle or a rectangle are unam-



Ambiguity in a hexagonal-mirror system



A kaleidoscope with two types of symmetry



A kaleidoscope with three types of symmetry biguous in the sense that their content is independent of your angle of view.

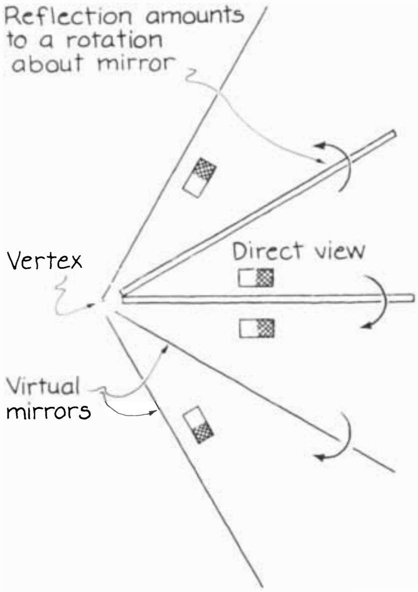
What other mirror systems yield unambiguous image fields? Initially I thought the only requirement was that an optical system should fill the image field with nonoverlapping copies of the direct view, much as one might fill a floor with identical tiles. I discovered that I was wrong by considering the hexagonal arrangement of mirrors shown in the upper illustration on the opposite page. The direct view and two reflected hexagons appear in each part of the illustration. An easy way to derive the reflected hexagons is to rotate the direct view about an edge until it lies again in the image plane. Then rotate the reflected hexagon about one of its edges to form a second reflected hexagon.

The first part of the illustration indicates that I began with the direct view and proceeded clockwise to find the two reflected hexagons. I rotated the direct view about real mirror *A* to form the reflected hexagon at the lower left. Then I rotated that hexagon about its top edge to form the hexagon at the upper left.

An edge of a reflected hexagon is called a virtual mirror because it effectively functions as a mirror even though it is only an image of one. Hence the second reflected hexagon arises from an edge serving as a virtual mirror. My procedure results in two reflected hexagons that would be seen by an observer looking into mirror *A* toward the position of the hexagons in the image field.

The second part of the illustration shows how again I began with the direct view but proceeded counterclockwise to find the reflected hexagons that would be seen by an observer looking into mirror *B*. Note that the two sets of reflected hexagons differ in content. The image field is ambiguous.

After playing with hexagons and other polygons I finally understood what arrangements of mirrors yield unambiguous image fields. For any arrangement the key is to examine each vertex at which two mirrors meet. Looking into the system, you see a sector of the direct view; around the vertex are reflections of that sector. To determine the reflections imagine rotating the sector of the direct view about one of its sides until it is again in the image field. Then you rotate the new sector about its side. Continue the rotations both clockwise and counterclockwise until the sectors begin to overlap. If their contents overlap precisely, the image field around the vertex is necessarily unambiguous. The only vertex angles that yield such a result are even divisors of 360 degrees.



Determining the sectors of a cluster

If a mirror system is to yield unambiguous image fields, every vertex in the system must meet this requirement. As I already knew, an equilateral triangle and a rectangle qualify. Surprisingly, there are only two more polygons, both right triangles, that meet the requirement. One of them has angles of 45 degrees and the other has angles of 60 and 30 degrees. I do not know whether these systems have already been discovered.

The four polygons generating unambiguous image fields differ in the types of symmetry that appear in the fields. An equilateral triangle gives rise to clusters that have only sixfold symmetry. Each cluster consists of six images that are either exact or reflected copies of the direct view. The image field from a rectangle of mirrors consists of clusters that have only fourfold symmetry. The right-triangle systems offer clusters displaying more types of symmetry. In one of them the right-angle vertex produces a fourfold symmetry and the 45-degree angle produces an eightfold symmetry.

The system with 60- and 30-degree angles represents the most beautiful kaleidoscope design because it offers three types of symmetry, the maximum in any unambiguous image field. The right angle creates a fourfold symmetry, the 60-degree angle a sixfold symmetry and the 30-degree angle a twelvefold symmetry. A kaleidoscope of this type is easy to construct because the mirrors form a right triangle with a hypotenuse twice the length of the shorter leg.

BIBLIOGRAPHY

Readers interested in further explanation of the subjects covered by the articles in this issue may find the following lists of publications helpful.

COMPUTER RECREATIONS

AN APL2 GALLERY OF MATHEMATICAL PHYSICS—A COURSE OUTLINE. Gregory J. Chaitin. IBM Research, P.O. Box 218, Yorktown Heights, N.Y. 10598; 1985.

DIRECT N-BODY CALCULATIONS. Sverre J. Aarseth in *Dynamics of Star Clusters,* edited by Jeremy Goodman and Piet Hut. D. Reidel Publishing Company, 1985.

A UNIFIED N-BODY AND STATISTICAL TREATMENT OF STELLAR DYNAMICS. Alan P. Lightman and Stephen L. W. McMillan in *Dynamics of Star Clusters,* edited by Jeremy Goodman and Piet Hut. D. Reidel Publishing Company, 1985.

SPACE SCIENCE, SPACE TECHNOLOGY AND THE SPACE STATION

MISSION TO EARTH: LANDSAT VIEWS THE WORLD. Nicholas M. Short, Paul D. Lowman, Jr., Stanley C. Freden and William A. Finch, Jr. NASA SP-360, U.S. Government Printing Office, 1976.

SPACE SCIENCE COMES OF AGE: PERSPECTIVES IN THE HISTORY OF THE SPACE SCIENCES. Edited by Paul A. Hanle and Von Del Chamberlain. National Air and Space Museum, Smithsonian Institution Press, 1981.

SPACE RESEARCH IN THE ERA OF THE SPACE STATION. Kenneth J. Frost and Frank B. McDonald in *Science,* Vol. 226, No. 4681, pages 1381–1385; December 21, 1984.

ASTROPHYSICS TODAY. Edited by A. G. W. Cameron. American Institute of Physics, 1985.

GROWTH, DIFFERENTIATION AND THE REVERSAL OF MALIGNANCY

ERYTHROLEUKEMIC DIFFERENTIATION. Paul A. Marks and Richard A. Rifkind in *Annual Review of Biochemistry,* Vol. 47, pages 419–448; 1978.

CONTROL OF NORMAL CELL DIFFERENTIATION AND THE PHENOTYPIC REVERSION OF MALIGNANCY IN MYELOID LEUKEMIA. Leo Sachs in *Nature,* Vol. 274, No. 5671, pages 535–539; August 10, 1978.

CONSTITUTIVE UNCOUPLING OF PATHWAYS OF GENE EXPRESSION THAT CONTROL GROWTH AND DIFFERENTIATION IN MYELOID LEUKEMIA: A MODEL FOR THE ORIGIN AND PROGRESSION OF MALIGNANCY. Leo Sachs in *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 77, No. 10, pages 6152–6156; October, 1980.

NORMAL DEVELOPMENTAL PROGRAMMES IN MYELOID LEUKEMIA: REGULATORY PROTEINS IN THE CONTROL OF GROWTH AND DIFFERENTIATION. Leo Sachs in *Cancer Surveys,* Vol. 1, No. 2, pages 321–342; 1982.

THE STRUCTURE OF COMET TAILS

INTERPLANETARY GAS. XXIII. PLASMA TAIL DISCONNECTION EVENTS IN COMETS: EVIDENCE FOR MAGNETIC FIELD LINE RECONNECTION AT INTERPLANETARY SECTOR BOUNDARIES? M. B. Niedner, Jr., and J. C. Brandt in *The Astrophysical Journal,* Vol. 223, No. 2, pages 655–670; July 15, 1978.

STRUCTURES FAR FROM THE HEAD OF COMET KOHOUTEK, II: A DISCUSSION OF THE SWAN CLOUD OF JANUARY 11 AND OF THE GENERAL MORPHOLOGY OF COMETARY PLASMA TAILS. Malcolm B. Niedner, Jr., and John C. Brandt in *Icarus,* Vol. 42, No. 2, pages 257–270; May, 1980.

THE COMET BOOK. Robert D. Chapman and John C. Brandt. Jones and Bartlett Publishers, Inc., 1984.

APPLICATIONS OF OPTICAL PHASE CONJUGATION

NONLINEAR OPTICAL PHASE CONJUGATION. Special issue of *Optical Engineering,* Vol. 21, No. 2; March/April, 1982.

OPTICAL PHASE CONJUGATION. Edited by R. A. Fisher. Academic Press, 1983.

NONLINEAR OPTICAL PHASE CONJUGATION. David M. Pepper in *The Laser Handbook: Vol. 4,* edited by M. L. Stitch and M. Bass. North-Holland Publishing Co., 1985.

PRINCIPLES OF PHASE CONJUGATION. B. Ya. Zel'dovich, N. F. Pilipetsky and V. V. Shkunov. Springer-Verlag, 1985.

MINERAL DEPOSITS FROM SEA-FLOOR HOT SPRINGS

HYDROTHERMAL PROCESSES AT SEAFLOOR SPREADING CENTERS. Edited by Peter A. Rona, Kurt Boström, Lucien Laubier and Kenneth L. Smith, Jr. NATO Conference Series, IV Marine Sciences, Vol. 12. Plenum Press, 1983.

HOT SPRINGS ON THE OCEAN FLOOR. John M. Edmond and Karen Von Damm in *Scientific American,* Vol. 248, No. 4, pages 78–93; April, 1983.

THE CHEMICAL DEFENSES OF HIGHER PLANTS

HERBIVORES: THEIR INTERACTIONS WITH SECONDARY PLANT METABOLITES. Edited by Gerald A. Rosenthal and Daniel H. Janzen. Academic Press, 1979.

CHEMICAL ECOLOGY OF INSECTS. Edited by William J. Bell and Ring R. Cardé. Sinauer Associates, Inc., 1984.

INSECTS ON PLANTS. D. R. Strong, J. H. Lawton and R. Southwood. Harvard University Press, 1984.

RADIOCARBON DATING BY ACCELERATOR MASS SPECTROMETRY

WORKING WITH RADIOCARBON DATES. H. T. Waterbolk in *Proceedings of the Prehistoric Society,* Vol. 37, Part 2, pages 15–33; December, 1971.

RADIOISOTOPE CLOCKS IN ARCHAEOLOGY. R. E. M. Hedges in *Nature,* Vol. 281, No. 5725, pages 19–24; September 6, 1979.

PROCEEDINGS OF THE ELEVENTH INTERNATIONAL RADIOCARBON CONFERENCE—SEATTLE. In *Radiocarbon,* Vol. 25, No. 2; 1983.

KIN RECOGNITION IN TADPOLES

GENETIC CONTROL FOR SIBLING RECOGNITION? Andrew R. Blaustein and Richard K. O'Hara in *Nature,* Vol. 290, No. 5803, pages 246–248; March 19, 1981.

RANA CASCADAE TADPOLES AGGREGATE WITH SIBLINGS: AN EXPERIMENTAL FIELD STUDY. Richard K. O'Hara and Andrew R. Blaustein in *Oecologia,* Vol. 67, No. 1, pages 44–51; 1985.

KIN RECOGNITION IN ANIMALS. Edited by David J. C. Fletcher and Charles D. Michener. John Wiley & Sons, Inc., 1986.

THE AMATEUR SCIENTIST

OPTICAL METHODS OF CHEMICAL ANALYSIS. Thomas R. P. Gibb, Jr. McGraw-Hill Book Company, Inc., 1942.

FUNDAMENTALS OF OPTICS. Francis A. Jenkins and Harvey E. White. McGraw-Hill Book Company, Inc., 1957.

"Station wagon" somehow doesn't say nearly enough. There should really be another term for this remarkable new Mercury Sable model.

It's the result of a five-year, three-billion-dollar effort to bring forth a fresh new kind of automobile. And it looks it, with its flush glass "greenhouse" and striking front-end "laser" lightbar. The Sable wagon is, in fact, the most aerodynamically pure wagon ever produced in its class. It's a car that moves through the air with ease, and over the road with the utmost assurance. Sable truly is a great road car that just happens to be a wagon.

Power is supplied by a new 3.0-liter V-6 that was designed and built from scratch for this car. Its electronically controlled fuel injection is of advanced sequential port design.

Road management is handled very competently by a four-wheel independent suspension with gas-pressurized shocks.

Inside, the driver has no doubts about this being a road car of the first magnitude. The backlighted instruments obviously mean business. Controls are right where your hands and feet would like them to be. Power-adjustable lumbar support is available for both bucket and split-back front seats.


Buckle up—together we can save lives.
1986 Mercury Sable. A great road car that just happens to be a wagon.
Sophisticated new shape. The 1986 Mercury Sable.
On the station wagon side of its personality, Sable has a very practical 60/40 split rear seat, permitting part or all of the seat back to fold flat for cargo.
And there's a wide-coverage rear window wash and wipe system available, as well as a lockable storage area. A unique optional picnic tray is a handy Sable addition.
But perhaps most important of all, Sable is a Mercury. And Mercury's commitment to quality isn't just something written on paper—it's built into the car.
In every sense, the car you're ready for is now ready for you. The 1986 Mercury Sable wagon.
For more information, call 1-800 MERCFAX
LINCOLN-MERCURY DIVISION Ford
MERCURY. The shape you want to be in.

Why buy a video camera that just records? Panasonic presents OmniMovie.™ It plays back your magic moments. And Hollywood's too.

Now bring the joy and magic of your best scenes, or Hollywood's, to the home screen. With OmniMovie. The spectacular new video camera. VHS™ recorder. And playback unit all in one.

Shoot your own movies.

OmniMovie makes it simple. Just touch a button and it zooms in for a closeup. With auto-focus and auto-exposure, it's automatically great. Shoot indoors or out with the light of just a few birthday candles.

OmniMovie even has direct playback through the viewfinder. So you can see what you've shot. Right on location. After all, even the best directors need a second take sometimes.

And OmniMovie, unlike some camcorders, uses standard VHS tape. Now you can shoot for up to two hours* on a single charge of its battery.

Watch Hollywood movies.

And OmniMovie uses the popular VHS format. So it's easy to find thousands of

prerecorded VHS tapes to rent or buy. Everything from Hollywood movies to rock concerts and videos.

Exchange your tapes. Share your big splash, or Hollywood's, with relatives and friends.

Now you don't have to miss another magic moment. Panasonic OmniMovie.

SPLASH is a Touchstone Home Video release. © MCMLXXXIV Buena Vista Distribution Co., Inc. TV Pictures Simulated. *Depending upon use.

THE PANASONIC LAS VEGAS PRO-AM.
APRIL 30-MAY 4, 1986. CALL 1-800-722-GOLF
SEE IT ON NBC

Panasonic.
just slightly ahead of our time.™