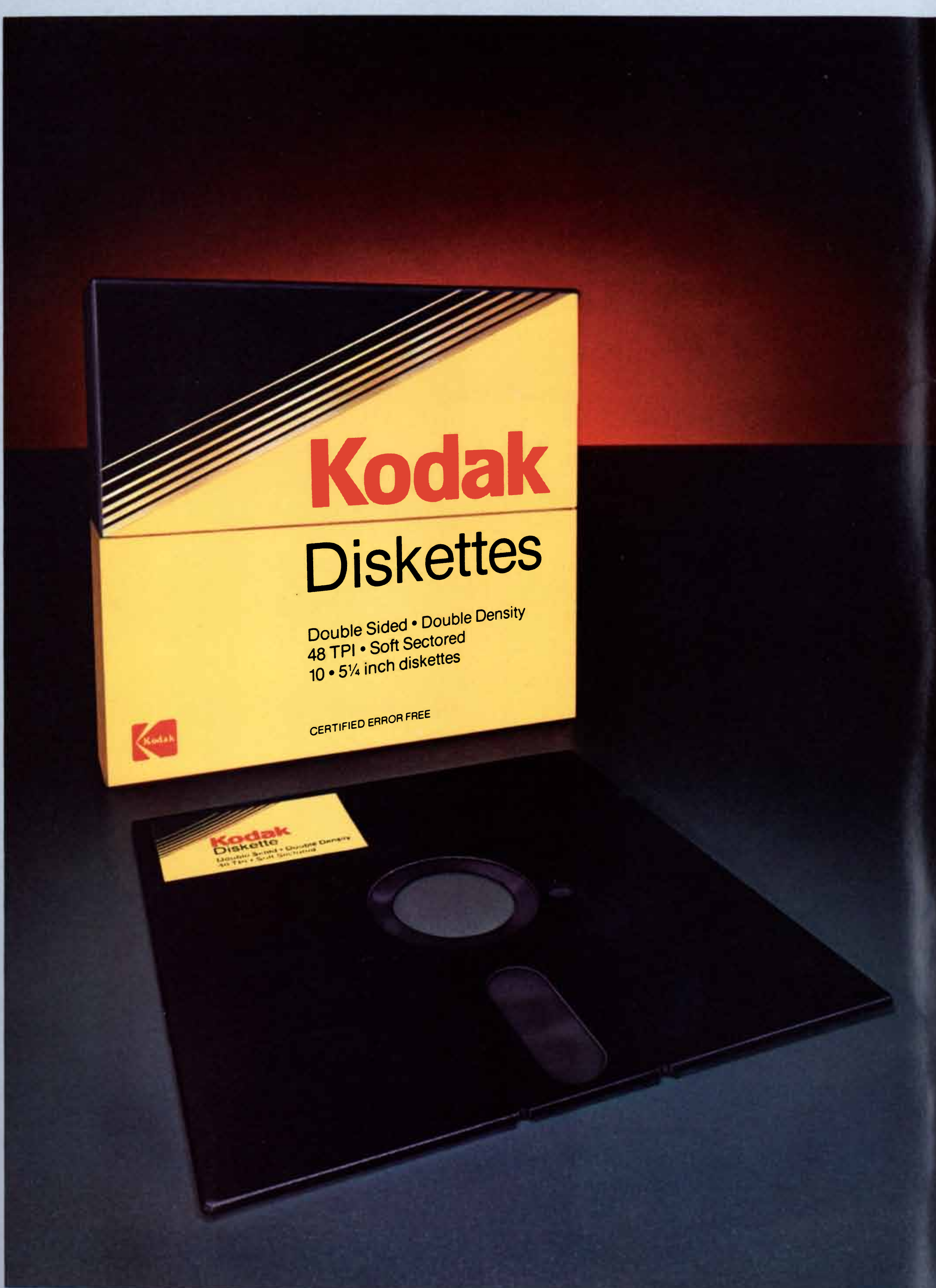
SCIENTIFIC AMERICAN

STAINED GLASS IN PERIL

$2.50

May 1985


Kodak
Diskettes
Double Sided • Double Density
48 TPI • Soft Sectored
10 • 5¼ inch diskettes
CERTIFIED ERROR FREE
Kodak
Kodak
Diskette

A NEW LEGEND BEGINS.

INTRODUCING NEW KODAK DISKETTES.

For as long as anyone can remember, the world has trusted Kodak film to capture its memories. Now the world can trust legendary Kodak quality to capture its computer data.

Introducing Kodak diskettes. And the beginning of a new legend.

We know you expect nothing less than extraordinary performance from a Kodak product. We didn't disappoint you.

These remarkable new diskettes are so thoroughly tested, they're certified error-free.

Every Kodak diskette has a highly burnished head surface for optimum read-write accuracy. And every standard diskette is made to withstand 4½ million passes before significant wear occurs.

With accuracy and durability like that, we can offer this no-questions-asked replacement policy:

This KODAK Diskette will be free from manufacturing defects, or we will replace it.

Kodak diskettes for home and business PC use are available in standard 8- and 5¼-inch formats, high-density 5¼-inch diskettes, and 3½-inch micro diskettes in our HD 600 Series.

New Kodak diskettes. Because the only thing that can follow a legend is another legend.

KODAK.
The name says it all.

TEST DRIVE A SAAB. IT WILL TELL YOU A LOT ABOUT YOURSELF.

Thirty minutes in the driver's seat of a Saab 900 can be more revealing than a session on a psychiatrist's couch.

Will you look at Saab's front-wheel drive as a way to get through winter and the foul weather of the other seasons? You're ruled by your intellect.

Or will you look at front-wheel drive as the means to give a sedan the road-hugging, toe-curling, cornering ability of a sports car? You're ruled by your emotions.

But maybe you see front-wheel drive as both. You're ruled neither by your intellect nor your emotions. You are ruling them.

The ego and the id.

The id, the repository of your instinctual impulses, will want to know, on a test drive, how good a Saab is at, well, burning rubber. Don't repress that feeling.

All the practical considerations for buying a Saab (economy, the reasonableness of its price, the active and passive safety features, the durability in a world where disposability is a perverse virtue) are just as real and just as practical after testing its acceleration as they were before.

Sure, a car's reason for being is to get from Point A to Point B. But a Saab's reason for being is to do that as responsibly as possible without ignoring the romance in the possibilities of Points Q, R, S, T, not to mention X, Y, and Z.

An interpretation of your dreams.

Is the Saab 900 the car of your dreams?

Who knows?

We do know it's the car of our engineers' and designers' dreams. There are, you know, seemingly disparate elements on a Saab.

Rack-and-pinion steering, disc brakes on all four wheels, front-wheel drive, 53 (!) cubic feet of cargo space, aerodynamic body, incredible fuel efficiency* considering its performance.

On a Saab, though, they add up to a whole greater than its parts. Saabs somehow feel better to Saab owners than other cars they've owned.

For the first time, they got the performance car they wanted with the responsible car they needed.

They found out the joy of not following the crowd, but of starting a crowd of their own.

Truth is, a 30-minute test drive may not be able to tell you all this. But we'll bet that three years owning one will.

SAAB 900

The most intelligent car ever built.

**Saab 16-valve Turbo 5-speed: 19 EPA estimated city mpg, 25 estimated highway mpg. Use estimated mpg for comparison only. Mileage varies with speed, trip length and weather. Saabs range in price from $11,850 for the 900 3-door, 5-speed to $18,620 for the 900 4-door, 5-speed, 16-valve Turbo. Manufacturer's suggested retail prices. Not including taxes, license, freight dealer charges or options.*

Established 1845 **SCIENTIFIC AMERICAN** May 1985 Volume 252 Number 5

ARTICLES

DEPARTMENTS

SCIENTIFIC AMERICAN (ISSN 0036-8733), PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. SECOND-CLASS POSTAGE PAID AT NEW YORK, N.Y., AND AT ADDITIONAL MAILING OFFICES. AUTHORIZED AS SECOND-CLASS MAIL BY THE POST OFFICE DEPARTMENT, OTTAWA, CANADA, AND FOR PAYMENT OF POSTAGE IN CASH. SUBSCRIPTION RATE: $24 PER YEAR IN THE U.S. AND ITS POSSESSIONS; $33 PER YEAR IN ALL OTHER COUNTRIES. POSTMASTER: SEND ADDRESS CHANGES TO SCIENTIFIC AMERICAN, 415 MADISON AVENUE, NEW YORK, N.Y. 10017.

SCIENTIFIC AMERICAN

CORRESPONDENCE

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

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 photograph on the cover documents an intermediate stage in the effort to restore some of the oldest stained glass in the world (see "The Restoration of Medieval Stained Glass," by Gottfried Frenzel, page 126). The pane, made in about A.D. 1130, depicts the prophet Hosea. It was installed, with four other images of prophets, in the south clerestory windows of Augsburg Cathedral. The right side of the pane has been cleaned of corrosion; the left side remains severely blackened. Two pieces in the pane (the nose and a part of the background at the lower right) are 19th-century replacements of the original stained glass. Presumably they were matched to the transparency of the original a century ago; thus they suggest the severity of damage inflicted since then by 20th-century air pollutants, notably sulfur dioxide.

THE ILLUSTRATIONS

Cover photograph by Gottfried Frenzel

Page	Source
11	SCIENTIFIC AMERICAN
18–24	Ilil Arbel
26–30	James K. Park
36	*Pre-Columbian Plant Migration.* © 1984 by the President and Fellows of Harvard College
41	*Colours of the Stars.* © 1984, Cambridge University Press
46	Edward Bell
47–51	Ilil Arbel
53–54	Patricia J. Wynne
55	G. Nigel Godson, New York University School of Medicine
56–57	Patricia J. Wynne
58	Edward Bell
59	Patricia J. Wynne
61–68	Ian Worpole
89	Roquefort Association, Inc.
90–92	James Kilkelly
97	Roger Maret, Nestlé, S.A. (*top*); Quesada/Burke (*bottom*)
98–99	Quesada/Burke
102	James N. Cameron, University of Texas at Austin
104–108	Tom Prentiss
110–117	Allen Beechel
118–119	Quesada/Burke
120–124	Carol Donner
127	Sonia Halliday Photographs (*left*); Westfälisches Amt für Denkmalpflege (Westphalian Bureau for the Conservation of Art) (*right*)
128–131	Gottfried Frenzel
132	Sally Black
135	Gottfried Frenzel
138–143	Michael Goodman

LETTERS

To the Editors:

"The Crossbow," by Vernard Foley, George Palmer and Werner Soedel [SCIENTIFIC AMERICAN, January], was excellent as a scientific study. It does, however, leave the false impression that the crossbow was a superior weapon to all other types of bow. History dramatically records otherwise. The English longbow was far superior.

It was first used by the English at the Battle of Falkirk in 1298. Until that time war had been the exclusive domain of the armored knight. Archers, along with pikemen, were considered the lowest among soldiers. With the advent of the longbow the situation changed, at least as far as the English were concerned. Because it took from two to three years to train a longbowman, his value increased, and he was harder to replace.

Sir Winston Churchill devoted an entire chapter to the longbow in his *History of the English-speaking Peoples.* Describing one instance of the longbow's effectiveness, he wrote: "One of the knights had been hit by an arrow which pierced not only the skirts of his mailed shirt, but his mailed breeches, his thigh, and the wood of his saddle, and finally struck deep into the horse's flank. For the first time infantry possessed a weapon which could penetrate the armour of the clanking age, and which in range and rate of fire was superior to any method ever used before, or ever used again until the coming of the modern rifle."

Dr. Isaac Asimov writes of the longbow: "The longbow was lighter than the crossbow and had an even longer range, up to twelve hundred feet maximum. Much more important, the longbow could fire very rapidly. The longbowman, reaching over his shoulder for arrows in the quiver he carried on his back, could fire five or six accurate shots in the time it took the crossbowmen to reload."

Finally, from the *New Columbia Encyclopedia* on the Battle of Crécy:

"English longbowmen, firing from fixed positions, so thoroughly outclassed Genoese crossbowmen fighting for the French that the longbow replaced the crossbow as the dominant European projectile weapon."

It remained dominant for some 300 years until 1595, when it was replaced by the musket as the standard weapon of the British infantry.

R. A. JOSLIN

Victoria, B.C.

To the Editors:

We should like first of all to call the reader's attention to an error in the caption for the table summarizing our wind-tunnel results. What was given as the ratio of mass to drag should have been given as the reciprocal, the ratio of drag to mass. We greatly regret any inconvenience caused by this mistake.

Where the question of longbow versus crossbow is concerned, we certainly yield on the matter of rate of fire. With respect to range, however, we have not been able to find any sources with good credibility that suggest the longbow could exceed 300 meters. Two hundred and forty yards was considered a long shot in the time of Henry VIII. It is also worthy of note that components of medieval fortresses were purposely sited beyond longbowshot. Finally, nothing suggests the longbow could match Payne-Gallway's distance feat with the crossbow (420 meters).

Since killing power is a function of both mass and velocity, we assume that a weapon with a longer range may have a higher terminal velocity, and at less than maximum range its arrow is likely to be going faster. The crossbow bolt is shorter than the longbow arrow but thicker, and their weights are about the same.

With respect to the famous English victories at Crécy, Poitiers and Agincourt, we find that the Genoese crossbowmen were badly used at Crécy. They were tired from a day's march and protested the order to fight. The English were better rested. As the Genoese came into position, they found themselves shooting into the setting sun, which certainly did not help their aim. Finally, it is likely that they were using composite bows, which Genoa had known for more than a century. It was too early for the steel bow.

A weakness of the composite bow is its susceptibility to dampness, and a short but heavy rainstorm swept the battlefield just before the engagement began, wetting everyone to the skin. The chronicler Jean de Venette says the soaking damaged the bowstrings of the Genoese, but in our view the "sinew" backing of their bows was more likely to have been affected.

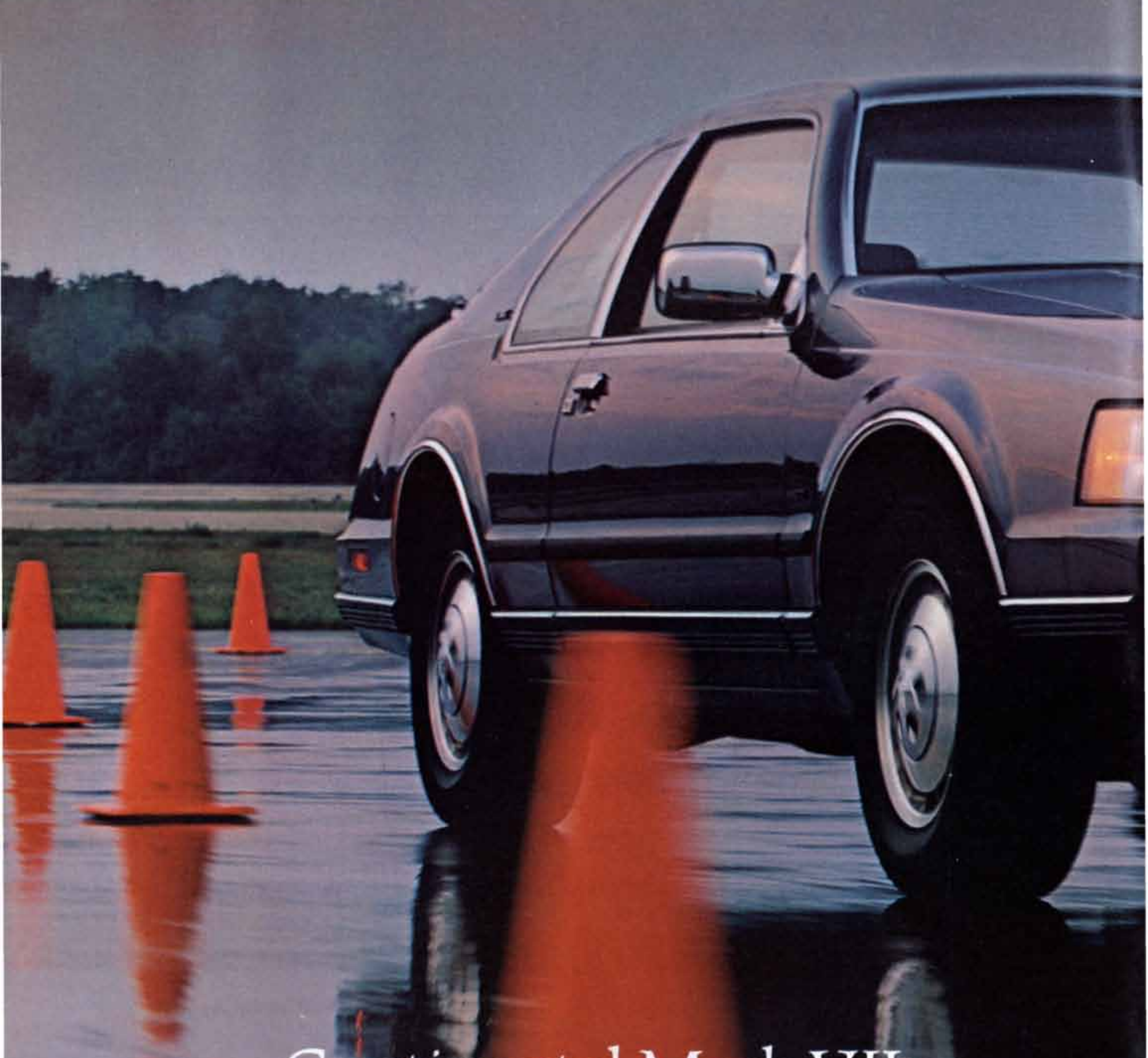
In all three battles French leadership was very bad. Eventually the French neutralized the longbow and went on to win the Hundred Years War.

VERNARD FOLEY

GEORGE PALMER

WERNER SOEDEL

West Lafayette, Ind.



Continental Mark VII. Anti-lock brakes enable it to do what other domestic luxury cars cannot.

Now select models of the Continental Mark VII and Continental are equipped with revolutionary technology called an Anti-Lock Brake System. Or ABS, for short. Incidentally, Lincoln is the only domestic automaker to offer it.

In addition to giving you braking with short stopping distances, our Anti-Lock Brake System was designed to prevent wheel lock-up —even in a panic reaction to an emergency situation. As a result, it helps the driver maintain lateral stability and steering control.

Experienced drivers know the quickest, most efficient way to stop an automobile in an emergency is to pump the brake pedal. The way our ABS system works to assist the driver is by pumping the brakes automatically. It does so precisely, efficiently, and up to ten times per second—far faster than humanly possible.

To achieve this, magnetic sensors in each wheel of the Mark VII LSC

constantly monitor individual wheel speed. The data is then relayed to a central microprocessor. In a panic-stop situation the system is instantly activated by the driver's reaction of applying hard, steady pressure on the brake pedal. At that point, ABS automatically pumps the brakes to maximize brake force and minimize the chances of brake-induced skidding.

In addition, the ABS electronic control system on the Mark VII operates on a redundancy principle with twin parallel microprocessors constantly monitoring each other's performance. This allows the Anti-Lock Brake System to revert to an operational mode exactly like that of a car without ABS in the unlikely event of an interruption in the system.

ABS technology has been used successfully for years on the wheel brakes of commercial aircraft landing gear to maintain control. And it is now available from an American luxury automaker. Couple this with the fact that Lincoln automobiles have been rated the highest quality luxury cars designed and built in the U.S., based on an average of problems in the prior six months reported in a survey of owners of 1981-1983 luxury cars conducted in 1984.

Continental Mark VII and Continental.

As impressive as the way they move is the way they stop.

Get it together–buckle up

LINCOLN

Lincoln-Mercury Division Ford

Best wishes for a speedy recovery

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

MAY, 1935: "The airlines of the United States again set an enviable record for safe flying during 1934 with 4,878,655 miles per fatal accident. It was noticeable, however, that with improvement in aids to bad-weather flying, more bad-weather flying was attempted. In view of certain of the accidents it is difficult not to feel that the eagerness to maintain schedule was manifested at some cost to safe procedure. No one close to the subject of the progress being made in the instrumentation of planes and in the radio aids to flight doubts that blind flying and actual blind landings will soon become a commonplace. The danger lies in anticipating this condition before it has fully arrived."

"In the course of a carefully planned campaign to observe double stars Gerrit Kuiper at the Lick Observatory set out to examine with the 36-inch telescope all stars known to be within 80 light-years of the sun. The results have been amazing. More than 100 new double stars have been found. Half of these are faint red dwarf stars whose companions appear so close that their real distances probably average less than that of Saturn from the sun. A few decades of observation should give good orbits and more than double our knowledge of the masses of these faint red stars."

"A new cable containing only a single wire, centrally located within the sheath, over which it is possible to transmit hundreds of telephone messages at one time, has been developed by the Bell Telephone Laboratories. This cable, which may profoundly affect future developments in long-distance telephony and also provide a television channel giving size and clarity of vision hitherto unknown, is called the coaxial cable."

SCIENTIFIC AMERICAN

MAY, 1885: "What is now before the statesmen of our country for their prime consideration? It is this: How shall the vast increase of our next century (if it be in proportion to the last in per cent.) be supplied with homes and honest employment? The 3,070,000 population of 1780 has grown in one hundred years to 50,156,000 and has extended from the confines of the thirteen Atlantic States, containing an area of 800,000 square miles, to one of 3,110,061 miles embracing thirty-eight states and seven Territories. If the 3,070,000 population of 1780 made an increase to 50,156,000 in a century, the total in the next century will reach at the same ratio 852,000,000."

"Four expeditions have been sent to Alaska within two years, and they have succeeded in giving us a knowledge of the magnitude and possibilities of that once despised possession, which is inspiring lofty dreams of national and private wealth. Its fisheries have returned the government an interest of nearly five per cent a year on the $7,200,000 which Secretary Seward paid Russia for Alaska in 1867, as a delicate acknowledgment of our gratitude for that nation's firm friendship during the rebellion, and now it is found that the possession which we then did not want especially contains vast rivers, mountains, forests and mines of undreamed-of riches."

"On July 1 an important change will be made in the rates of postage. A special stamp of the value of ten cents may be issued, which when attached to a letter, in addition to the lawful postage thereon, will entitle the letter to immediate delivery at any place containing 4,000 population or more according to the Federal census. Messengers for this special delivery are to be paid eighty per cent. of the face value of all the stamps received and recorded in a month, provided that the aggregate compensation paid to any one person for such service shall not exceed $30 per month."

"Railway tonnage has reached its present magnitude in this country by a rapidity of development little dreamed of in the first stages of its growth. Articles are transported every year that were never transported before; and if the cars already in use are not adapted to the new traffic, special cars are soon devised and built. An illustration of this is afforded in the remarkable growth of the transportation of perishable commodities within the past few years by means of refrigerator cars. The shipment of dressed meats from Chicago and other points farther west to the Eastern seaboard has already grown from small beginnings to a heavy traffic. The semi-tropical fruit products of Cuba, Florida, Mexico and southern California are finding their way to Northern markets during the warm season in larger quantities every year in refrigerator cars so well adapted to the purpose as to make the losses from the perishable nature of the freight comparatively light."

"Cotton picking is now almost universally done by hand and is a slow, tedious and expensive operation. A machine which would take the place of hand picking has long been needed. The perfect machine should remove all the fiber from every pod, leave the plants uninjured, require a minimum amount of care and be rapid in operation. The cotton harvester illustrated below will pick about four bales per day, doing the work of sixty hands."

A mechanical cotton picker invented by R. K. Charles

announces hundred year

we've designed a family of mainframes that costs less to run, needs fewer support people, and allows users to increase processing power up to 70 times without reprogramming.

Plus, we've developed software that brings the programming and operational ease of personal computing to mainframes. And equally remarkable software that increases programmer productivity up to an astonishing 1000%.

Of course, Burroughs has the office and data processing products and supplies to go with our systems.

And we offer a lot more.

Like our System Development Corporation—an acknowledged leader in secure networks and systems integration. And our Memorex unit—offering highly advanced storage devices for Burroughs and other systems complemented by a wide range of peripherals.

Finally, to help our people put our plan into practice, we've recently spent $67 million on personnel training and more than $100 million on service and support facilities.

In an industry where many companies don't last a decade, Burroughs is proud to be launching our second century.

The experience and insights we've gained are helping us develop new solutions that will benefit our customers and their profitability for the next hundred years.

Burroughs

THE AUTHORS

EDWARD E. DAVID, JR. ("The Federal Support of Mathematics"), is president of the Exxon Research and Engineering Company. He is a graduate of the Georgia Institute of Technology and the Massachusetts Institute of Technology, where he got a doctorate in electrical engineering in 1950. He then joined the Bell Telephone Laboratories, where he was a pioneer in computer time-sharing and made crucial contributions to the development of cellular mobile radio. In 1970 he entered Government service as Science Adviser to the President and director of the White House Office of Science and Technology. He returned to industry in 1973 as executive vice-president of Gould, Inc., and in 1977 he joined Exxon. David has served as president and chairman of the American Association for the Advancement of Science, and he is a member of the White House Science Council, U.S. representative to the NATO science committee and a member of the governing boards of several universities.

G. NIGEL GODSON ("Molecular Approaches to Malaria Vaccines") is professor and chairman of the department of biochemistry at the New York University School of Medicine. He was educated at the University of London, which granted him a Ph.D. in biochemistry in 1962. He served on the scientific staff of the Chester Beatty Research Institute of the Institute of Cancer Research in London, then did postdoctoral work at the California Institute of Technology and at the Yale University School of Medicine. In 1969 he was appointed to the faculty at Yale. He taught in the department of radiology there until 1980, when he accepted his present position at New York University.

HANS A. BETHE and GERALD BROWN ("How a Supernova Explodes"), longtime associates in the study of supernovas, are respectively professor emeritus of physics at Cornell University and professor of physics at the State University of New York at Stony Brook. Bethe was born and educated in Germany. In 1935, after work at universities there and in Great Britain, he joined the faculty at Cornell. In 1967 Bethe was awarded the Nobel prize in physics for his descriptions of the nuclear reactions through which stars generate energy. Brown earned his doctorate at Yale University in 1950. He then went abroad, teaching first at the University of Birmingham and in 1960 at the Nordic Institute for Theoretical Atomic Physics in Copenhagen, where he was professor. After his return to the U.S., he taught at Princeton University before joining the faculty at Stony Brook in 1968. For help with their article Bethe and Brown are grateful to Stanford Woosley of the University of California at Santa Cruz, who provided much information and constructive criticism.

FRANK V. KOSIKOWSKI ("Cheese") has spent 45 years teaching and doing research as professor of food science at Cornell University. He got his bachelor's degree at the University of Connecticut in 1939 and went on to earn master's and Ph.D. degrees in dairy science from Cornell in 1941 and 1944 respectively. His research and teaching have focused on food-processing fermentations such as the one that is crucial to cheesemaking. He has also been active in various capacities in international food policy. In 1960 he served the government of Puerto Rico as an adviser on food problems, and in 1963 he took part in international food development at the headquarters of the Food and Agriculture Organization in Rome. As a volunteer in the International Executive Service Corporation he spent much of 1970 in Iran as a technical adviser; he also studied cheesemaking among nomads of the Azerbaijan region.

JAMES N. CAMERON ("Molting in the Blue Crab") is professor of marine studies and zoology at the University of Texas at Austin. He earned a bachelor's degree (1966) at the University of Wisconsin at Madison and a Ph.D. (1969) from the University of Texas at Austin. He spent two years as a National Institutes of Health postdoctoral fellow at the University of British Columbia in Vancouver. In 1971 he continued north to join the faculty of the Institute of Arctic Biology at the University of Alaska in Fairbanks, and in 1975 he accepted a position at Texas. He has a particular interest in partially air-breathing creatures and what they reveal about how air-breathing forms evolved from aquatic ancestors. That interest has taken Cameron to the Pacific islands of Eniwetok and Palau for a study of land crabs and to the Amazon for research on lungfish.

ROBERT M. HAZEN and LARRY W. FINGER ("Crystals at High Pressure") have worked together in high-pressure crystallography for a number of years. Both are at the geophysical laboratory of the Carnegie Institution of Washington, where Hazen is an experimental mineralogist and Finger is staff crystallographer. Hazen earned bachelor's and master's degrees in earth sciences from the Massachusetts Institute of Technology; he received a Ph.D. in mineralogy from Harvard University in 1975. In 1976, after a stint as a NATO postdoctoral fellow in the department of mineralogy and petrology of the University of Cambridge, he joined the Carnegie Institution. Hazen is a professional symphonic trumpeter and with his wife, Margaret, has written several books on the history of geology. Finger was educated at the University of Minnesota, which granted him a Ph.D. in 1967. He has been at the Carnegie Institution since then, with an interruption from 1975 to 1976 for work at the State University of New York at Stony Brook as a visiting professor. From 1975 to 1979 he served as secretary of the Mineralogical Society of America.

HANS WALLACH ("Perceiving a Stable Environment") is professor of psychology at Swarthmore College. Working under Wolfgang Köhler, the founder of Gestalt psychology, he earned his doctorate from the Institute of Psychology of the University of Berlin in 1934. In 1936, following Köhler, Wallach moved to Swarthmore. He began his career there as a research assistant but soon was teaching full time. He held a concurrent post at the New School for Social Research from 1947 to 1957 and also spent a year, from 1954 to 1955, as a visiting psychologist at the Institute for Advanced Study in Princeton. He retired from teaching in 1975 to devote himself to his research interests.

GOTTFRIED FRENZEL ("The Restoration of Medieval Stained Glass"), a historian and restorer of art, is head of the Institute for Stained Glass Research and Restoration in Nuremberg. He studied painting and graphics in Dresden; art education, archaeology and psychology occupied him at Darmstadt. While employed in a glassworks, designing stained-glass windows in a modern idiom, he did doctoral research at the University of Nuremberg, where he got his degree in 1954 for a thesis on the late-medieval stained glass of that city. Since then Frenzel has been engaged in the restoration of medieval stained glass, in research on glass technology and in efforts to preserve historical monuments in Germany and throughout Europe.

BUICK

Electra. Not just lavishly appointed, lavishly engineered.

Buick Electra. The name alone speaks volumes.

Per tradition, this Electra is very smooth, comfortable, and roomy (for six adults) and very, very luxurious. With supple seats, rich upholstery and a high level of convenience.

But we know that simply being luxurious isn't enough for a luxury car anymore. So the Electra is engineered with front-wheel drive, four-wheel independent suspension, plus a 3.8-litre V-6 with multi-port fuel injection.

It is an automobile with ample performance as well as ample room, and with road manners as sophisticated as its appointments. And the Electra was tested in the heat of Australia and the hairpin turns of the German Alps to make sure it's ready for the ultimate test: yours.

Buckle up and visit your Buick dealer.

To ask any questions, request a brochure or test drive, call the Buick Product Information Center, 8 a.m. to 8 p.m. Eastern time, weekdays:

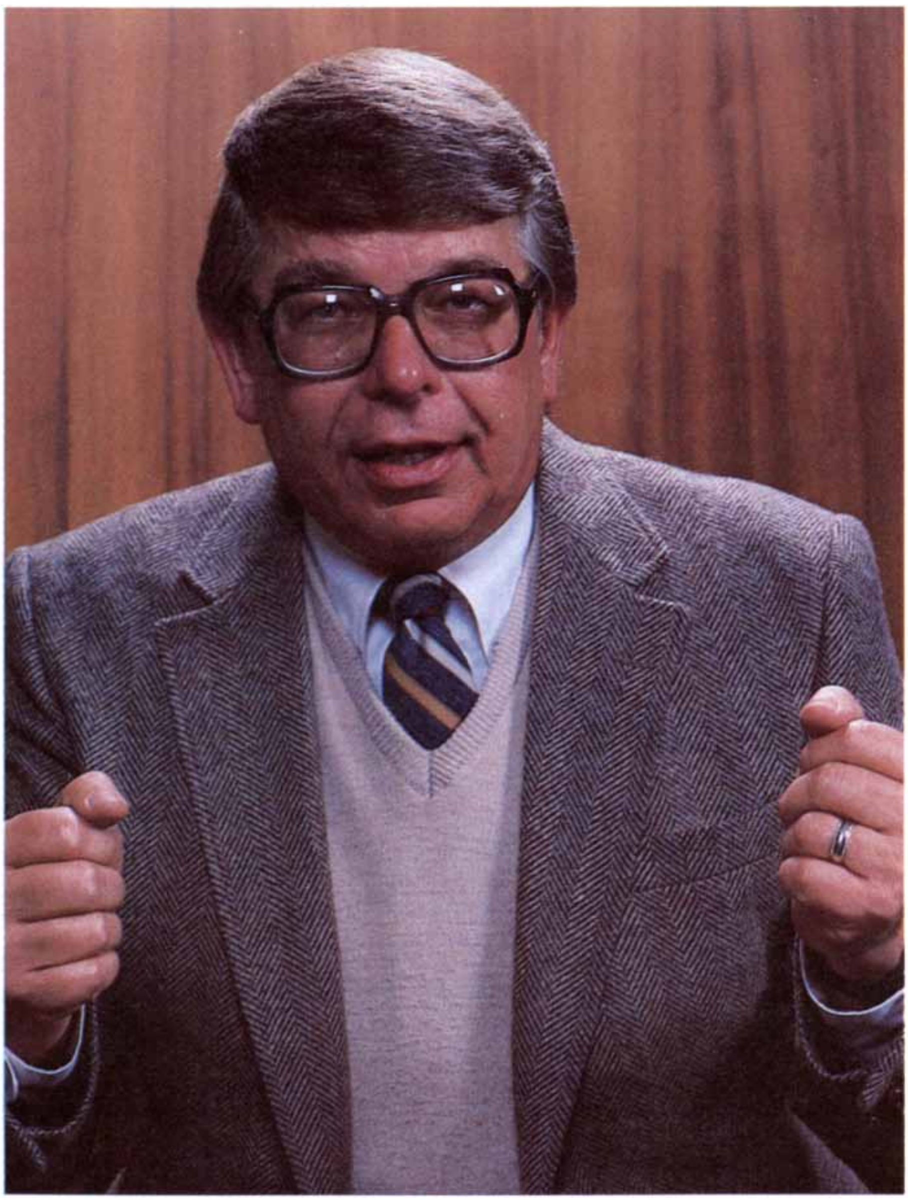
1-800-85-BUICK (1-800-852-8425).

Wouldn't you really rather have a Buick?

The ultimate Electra: Buick Park Avenue.

GM

WHAT SHOULD YOU WORKING "PARTNERSHIP"



"At Ford Aerospace, HP automatic test sets cut test time by 90%, saving $3 million a year.

"We've used HP instruments and computers for 10 years. In that time, we've quadrupled the circuits on our communication satellites while cutting test time by 90%. By automating with HP test sets we save up to $3 million a year."

Ford Aerospace designs and manufactures commercial communication satellites. In Palo Alto, Ca., John Merizon, Manager of Systems Test & Operations says: "HP's engineering staff helped us assemble HP instruments and computers to build 4 universal satellite test systems.

"Because satellite testing is a 24 hours a day, 7 days a week operation, downtime is irretrievable and translates directly into schedule slips which could cost us up to $30,000 a day. HP saves us the time and money of routine service calls by training our staff to calibrate and maintain HP equipment in-house. Because our HP systems deliver 99% uptime, we've cut down on repeat testing and virtually eliminated costly delays.

"HP is available for each launch with spare parts and a support team to back us up. Based on the flexibility and ruggedness of HP instruments and HP's consistent high performance, we plan to expand our test program with additional HP equipment."

EXPECT FROM A
WITH HEWLETT-PACKARD?

"Crown Zellerbach's ECZEL Corporation uses an HP Office Information Network and speeds response to customers, saving $1.5 million in just 10 months.

"A network of 23 HP 3000 computer systems link our 22 distribution centers across the country. HP 3000 Software for Distributors automates data processing for marketing, sales, warehouse functions and management planning."

ECZEL produces and distributes supplies for computer and word processing systems.

In San Francisco, California, Linda Bos, Vice President of Sales, says: "We're growing in excess of 10% a month. Ten months ago we made a major investment in HP. The return on that investment is already apparent. Our HP systems can expand with us as we grow. By networking the systems and using HP's word processing and electronic mail software, we've streamlined communications. HP's reliability has enabled us to improve productivity throughout our company saving $1.5 million this year alone.

"HP responds quickly and effectively to our needs. Based on this response, we are confident in HP as an office system supplier and plan to expand our distribution network with additional HP 3000 computer systems and software."

Isn't Your Business Worth HP?

HEWLETT PACKARD

0002507

COMPUTER RECREATIONS

Building computers in one dimension sheds light on irreducibly complicated phenomena

by A. K. Dewdney

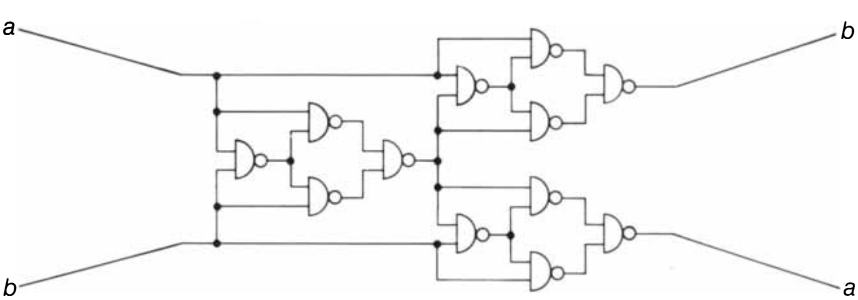
Immersed as we now are in a world of artificial computers it is interesting to consider the possibility that we are also surrounded by natural ones. Computers made of water, wind and wood (to mention just a few possibilities) may be bubbling, sighing or quietly growing without our suspecting that such activities are tantamount to a turmoil of computation whose best description is itself. This is not to say that such natural systems compute conventionally, only that their structure makes computation a possibility that is latent.

An eloquent exponent of this insight is Stephen Wolfram, a theoretical physicist at the Institute for Advanced Study in Princeton. Wolfram notes that a turbulent flow of fluid or the growth of a plant consists of rather simple components whose combined behavior is so complex that it may not be reducible to a mathematical statement—its best description is itself. The irreducibility of a natural system would follow from a demonstration of its ability to store, transmit and manipulate information; that is, to its ability to compute. In the September 1984 SCIENTIFIC AMERICAN Wolfram describes the use of cellular automata to explore this possibility. He proposes to find a cellular automaton that both computes and mimics a natural system. Wolfram's search focuses on the simplest of all possible cellular automata, those that have only a single dimension.

Such automata consist of simple elements that combine to generate complexity. Wolfram suggests that lurking among them are true computers, vast linear arrays of cells blinking from state to state and churning out any calculation a three-dimensional computer is capable of. Wolfram, currently searching through the myriad of one-dimensional cellular automata, is not above enlisting the help of amateurs in this daring and sophisticated enterprise. I shall describe the search and its consequences for natural computers in more detail below.

Before embarking on that adventure readers are invited on a short journey (computationally speaking) from the land of three dimensions down to the unbelievably narrow confines of one dimension. A good jumping-off place is three-dimensional computers, the ones currently inhabiting offices, factories and homes. They consist of fairly simple elements linked in a complex way. I speak here not of input or output devices but of the heart of the machine, a tiny silicon chip housing thousands of logic gates, memory elements, registers and other components. All of these are connected by an elegant pattern of tiny wires. The fact that the circuitry clings to a silicon surface does not mean it is two-dimensional. For one thing, when two connections cross, one must pass under the other. In addition, the silicon substrate of the circuitry mediates the function of every logic component.

Two-dimensional computers are to be found only in spaces with two dimensions, such as the Planiverse [see "Bibliography," page 144]. This realm is inhabited by a race of beings called Ardeans. The Ardeans have apparently succeeded in constructing a two-dimensional computer using just one type of logic element. We call it a NAND gate. Its output is a 1 if at least one of its inputs is a 0. Not only can a computer be constructed entirely from such gates but also the thorny problem of crossed connectors can be solved. The Ardeans create a special plane circuit from 12 NAND gates so that two signals entering the circuit from the left in order *ab* leave on the right in order *ba* [*see illustration on this page*]. Hence two signals may be crossed even if connections cannot. At the same time, the number 12 seems a bit excessive and the Ardeans would be grateful to any readers who can find a simpler NAND circuit that still enables signals to cross.



How to make signals cross in a two-dimensional computer

There is also a discrete two-dimensional space called Life, a game invented by John Horton Conway, the well-known University of Cambridge mathematician. Readers will remember this engaging exercise from "Mathematical Games" columns by Martin Gardner. As recently as October, 1983, Brian Hayes described in this department how the game could be realized in a spreadsheet program. Life is an infinite two-dimensional lattice of square cells whose states are influenced by the states of neighboring cells. Time is also discrete and from one tick of a cosmic clock to the next each cell is either alive or dead depending on a set of very simple rules:

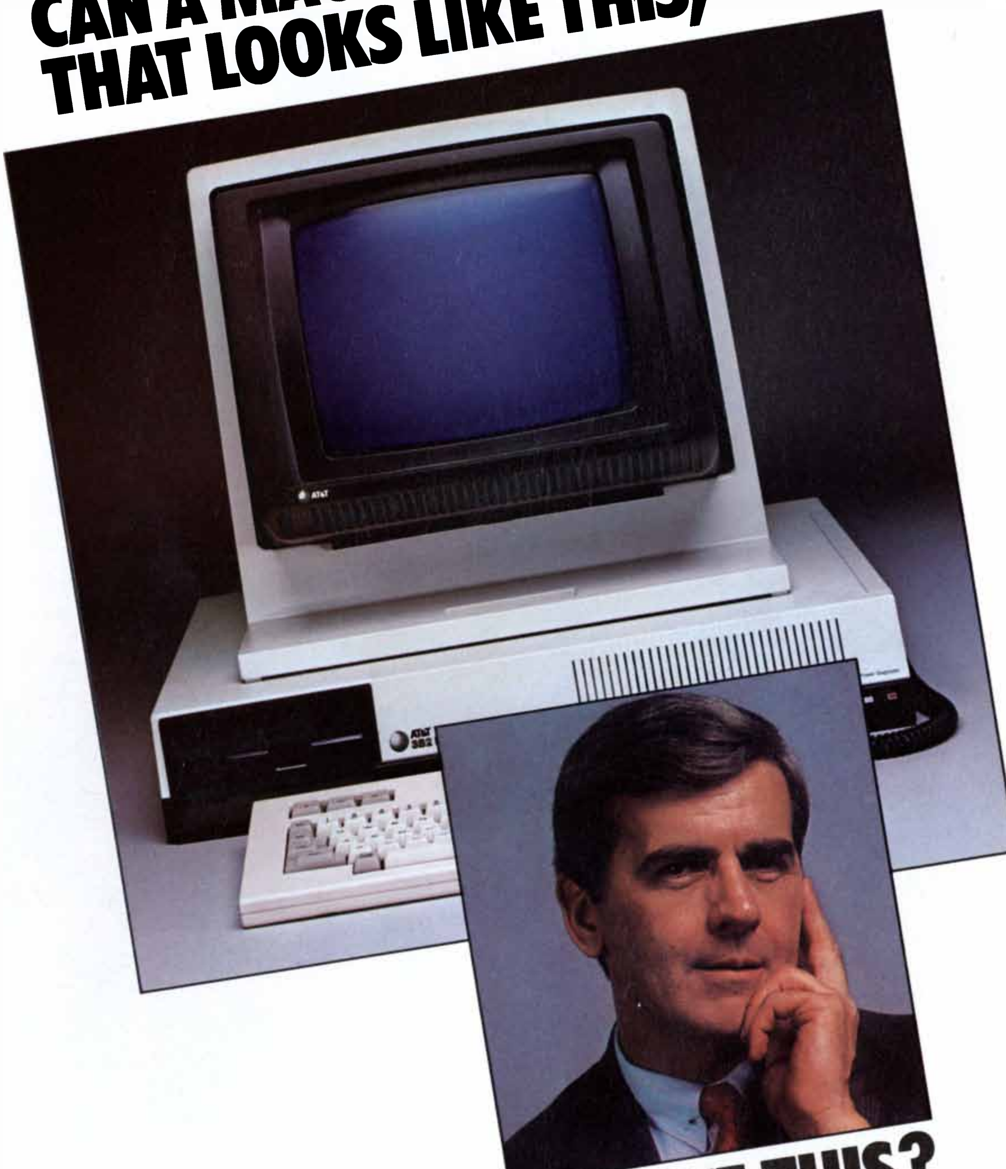
If a cell is dead at time t, it comes alive at time $t + 1$ if, and only if, exactly three of its eight neighbors are alive at time t.

If a cell is alive at time t, it dies at time $t + 1$ if, and only if, fewer than two or more than three neighbors are alive at time t.

With this set of rules everywhere in effect on Life's lattice, an initial configuration of live cells may grow interminably, fall into a cyclic pattern or eventually die off. Through more than a decade of experimentation by Life enthusiasts it has become clear that Life is far more complicated than anyone had thought.

For one thing, it has turned out that computers can be constructed within Life's cellular space. The discovery

Grand Marnier
Liqueur
Triple Orange
750 ml (25.4 fl.oz) · 80° Proof
Paradise found.
For gift delivery anywhere call 800-528-6148 (except where prohibited by law).
Product of France. Made with fine cognac brandy 80 proof. ©1985 Carillon Importers, Ltd., N.Y.
Gallardo



Not yet, but we're working on it.

Computers only look smart. Actually, they can't tell the players, or anything else, without a program.

But, until recently, even the most sophisticated programs only allowed computers to operate in "yes" or "no" terms. Judgment was still an exclusive domain of the human brain.

Now, however, the rapidly developing science of artificial intelligence is moving in on the mind.

At AT&T Bell Laboratories, using the new approaches to computer programming of knowledge engineering and rule-based programming, we have developed software that can cope—like a human—with incomplete and uncertain information, such as incorrect spelling and improper abbreviations.

The Expert with an ACE in the Hole

Called expert systems, this ambiguity-tolerant software is a new approach to artificial intelligence in industry. And, like the subject matter experts it mimics, our first expert system, ACE, works with equivocal data, but produces expert judgments.

ACE, for "Automated Cable Expertise," is a software system that contains the distilled knowledge—in the form of "if-then" rules—of the people who know cables best: telephone company cable maintenance experts. ACE differs from other expert systems in two ways: it manipulates massive amounts of data, and driven by the UNIX™ Operating System on a 3B2 computer, it obtains this information automatically from the data bases of other computers.

Developmental ACE software has been working as an "assistant" for over two years now to the cable maintenance force of the Southwestern Bell Telephone Company. Every night it monitors and analyzes the performance of cable systems serving over half-a-million customers in several metropolitan areas.

But ACE does more than analyze; it makes recommendations.

The Expert at Work

Unlike a conventional computer system, ACE isn't programmed with all logical answers to all possible problems. Instead, it's given a set of about 500 rules to follow.

ACE can run through the cable records of a city the size of Fort Worth overnight, a job that would take a human up to a week. By collecting its information from other computer programs, detecting recurring patterns, requesting additional data, and testing these data against its expert-derived rules, ACE can often isolate problems much earlier than its human counterparts. It provides information on both specific trouble types and locations—such as a break in cable insulation at the corner of 3rd and Elm. And when ACE has a recommendation to make, rather than generating a mound of paper, it communicates via a CRT.

ACE frees humans to work on the causes of problems, not the symptoms.

The Experts Behind the Experts

AT&T's skill in software development rests on a solid base of accomplishments in the computer field. Forty million lines of our software direct the world's largest computer system, the U.S. telephone network. And that experience is helping to make us a leader in new software research and design.

Today, software systems are being developed to help computers write their own software. While others make it possible to type data base requests in plain English. Our Rex expert system, for Regression Expert, aids in statistical analysis. And we've got silicon compilers that speed computer-aided design. There's even a program that allows a computer to derive new rules based on the rules already in its software!

All these systems are just a few examples of software being developed today to meet our customers' needs for tomorrow.

As you cross from industry to industry, tech-talk sometimes turns into bafflegab. We can help. In our new booklet, "Tech Talk," you will find the latest terms and acronyms from the fields of microelectronics, photonics, computers and telecommunications.

To order your copy, send $2.00* for postage and handling to AT&T, Department A, P.O. Box 4, Hawthorne, NJ 07507.

*Please add 6% sales tax for all orders originating in New Jersey.

AT&T

They were revolutionaries from the beginning—created to be the best six passenger cars you could buy. And for 1985, they still are. Aries 2-door, 4-door, and wagon offer the lowest prices, highest standard mileage rating, longest warranty, and best overall value.* All the great qualities that have helped Aries develop a tradition of solid resale.

REVOLUTIONARIES.

But while the practicality under the sheetmetal has remained intact, *in* the sheetmetal, a lot has changed. And the changes are simply beautiful. Aries now adds a sophisticated new look to a proven tradition of over 15 billion front-wheel-drive miles.

Aries LE gives you over 40 standard features. Some 73 improvements and innovations—both standard and optional. And, with offerings like the Premium Equipment Discount Package, savings of $225 when you order an AM/FM radio, rear window defroster, power steering, automatic speed control, tilt steering wheel, and other popular options.**

Ask any of over 450,000 American families who've owned Aries what they think. Then ask your Dodge dealer about buying or leasing* a new Aries. It's traditional K-car value in a beautiful new style that makes Aries 2-door, 4-door, and wagon true revolutionaries again for 1985.

Dodge
DIVISION OF CHRYSLER CORPORATION

AN AMERICAN REVOLUTION

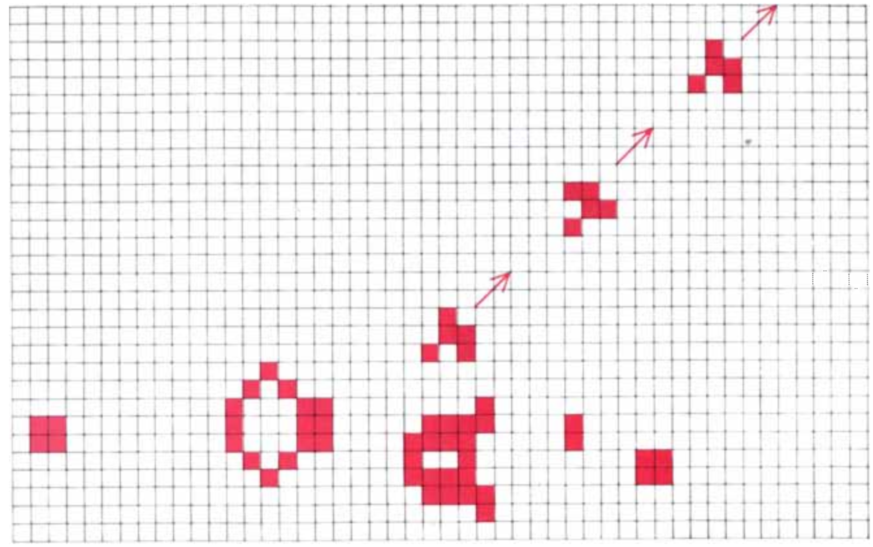
*1985 6-passenger model comparisons: lowest base and comparably equipped sticker prices; highest standard mileage—26 EPA city est. mpg, use to compare, actual mpg will vary with options, driving habits and conditions, and vehicle's condition, CA est. lower; longest warranty—5 years/50,000 miles, whichever comes first. Limited warranty on powertrain and outer body rust-through, excludes leases, deductible applies. Ask for details. **Aries LE savings based on sticker prices of package items if purchased separately. Ask for details.

BUCKLE UP FOR SAFETY.

came incrementally over a period of a few years. In 1969, shortly after he designed the game, Conway discovered a curious little pattern now called a glider. It blinked through four generations only to recover its original form at a location displaced diagonally by the space of one cell. On a monitor displaying the output of a particularly fast Life program a glider resembles a small creature from an exobiologic fantasy, wiggling its tail as it crawls across the screen. It travels at one-fourth the speed of light (in a cellular sense).

Although no one recognized it at the time, here was a medium of communication for use in a two-dimensional Life computer: instead of electronic pulses, use gliders!

The next step in the construction came in 1970 with the discovery by R. William Gosper, Jr., and several colleagues of a glider gun [*see illustration on this page*]. Gosper, then a student at the Massachusetts Institute of Technology, was keen on collecting the $50 prize that Conway had offered in Gardner's column. The prize would go to the first person who demonstrated conclusively that an initial configuration could grow without limit. Gosper's glider gun spewed out a new glider every 30 moves. The gun and the gliders constituted an ever growing population of live cells.



R. William Gosper's glider gun in action

Besides Gosper there were other students at M.I.T. pursuing Life in their spare time. One of these was Michael D. Beeler, who particularly enjoyed the analogy between Life and particle physics. Beeler trained a beam of one or more gliders on various targets, carefully noting the sometimes boring and sometimes spectacular results of the collisions. He even sent two beams of gliders into each other in various ways. His persistence was rewarded by some useful observations.

One such observation was that two gliders could collide and annihilate each other. This made possible the construction of the next component of a computer, logic gates. The simplest of these was a NOT gate. It changes one logic signal into another; 0 input becomes 1 output and vice versa. Life's NOT gate is constructed as follows. Set up a glider gun to send gliders in a specified direction. Binary numbers to be input to the NOT gate are encoded in a second stream of gliders aimed at right angles to the first one. In the input stream a glider may be present (a 1) or absent (a 0). This stream intersects the stream from the glider gun in such a way that when two gliders collide, they annihilate each other. This means that a glider in the input stream punches a hole in the glider-gun stream, converting a glider (1) into a nonglider (0) in the process. The absence of a glider in the input stream allows a glider from the gun to pass through unmolested. In this way a 0 is converted into a 1. Interestingly, the NOT gate has no structure to speak of: apart from its resident glider gun, it is merely a place where gliders meet [*see bottom illustration on next page*]. The construction of other gate types such as an AND gate and an OR gate also involves interacting streams of gliders, but it is too complicated to present here.

Memory and other registers are constructed through the interaction of gliders with four-cell configurations called blocks. A single bit of memory is encoded by the position of a block. The block is moved backward or forward by teams of gliders. Two gliders on appropriately chosen courses suffice to move the block three spaces in one direction after crashing into it. Ten gliders directed with equal care will return it to its original place.

There is much ingenuity and interest in the rest of the construction. It is all available in the delightful book *Winning Ways for Your Mathematical Plays,* by Elwyn R. Berlekamp, John H. Conway and Richard K. Guy. The book is divided into three sections: two-person games, one-person games and no-person games. Life is found in the last section.

In descending the final step down to one dimension there are only cellular spaces to consider; it is hard to imagine how the Ardeans could reduce their computer to a single continuous line. At first glance cellular space seems nearly as restrictive as that line. We are compensated for decreased dimensionality, however, in not being limited to a single set of rules. Instead we are given the opportunity to make our own rules. Additional compensations arise from the very simplicity of such linear spaces and from the fact that hundreds of generations can be viewed at a glance: place an initial pattern on a line and compute successive generations on successive lines down the page or display screen. A space-time diagram will develop.

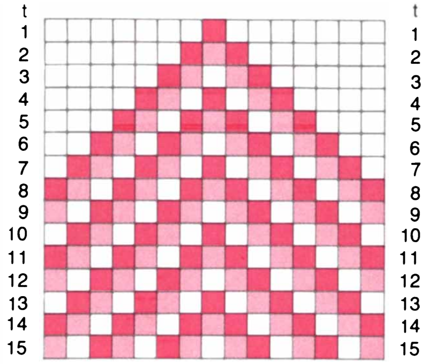
A one-dimensional cellular automaton (hereafter called a line automaton) consists of an infinite strip of cells changing states according to a given set of rules. As in the game of Life a cosmic clock ticks away and at each tick every cell enters a state determined by its previous state and the previous states of cells in its neighborhood. A line automaton is specified by giving two numbers called k and r as well as a set of rules for deriving the next state of a cell. The first number, k, determines how many states are allowed for each cell. In Life there are just two states and so k is equal to 2; among the line automata to be considered, higher values of k are common. The second number, r, refers to the radius of neighborhoods used to compute the next state of a cell. The present state of a cell and the states of its r neighbors on both sides determine the next state of the cell. For example, if r is equal to 2 and k is equal to 3, a certain rule might specify that when a cell's neighborhood looks like

0	2	1	1	0

the next state occupied by the central cell would be

2

The set of rules that defines a given



A glider gun in the line automaton Ripple

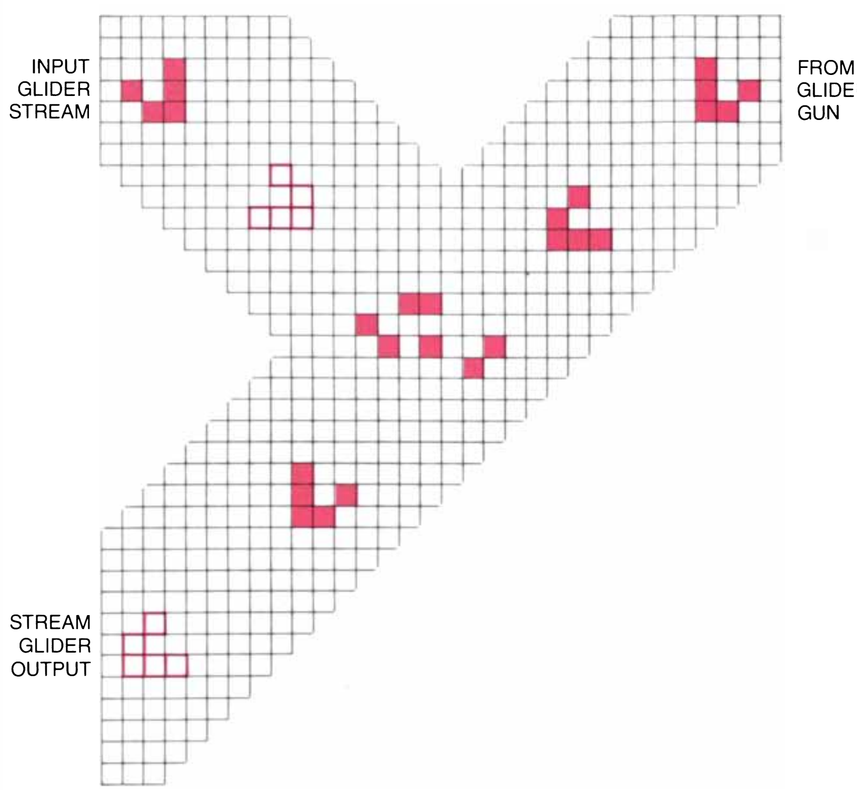
line automaton must decide the fate of a cell for every possible configuration of states inhabiting its neighborhood. Depending on the size of k and r the number of possible rule sets to consider can be enormous. For example, given the modest values of k and r described above, there are more line automata than there are atoms in the known universe.

Clearly each person on this planet can pick and choose his or her own personal line automaton. Indeed, I have already selected one for myself. For reasons that will soon be clear it is called Ripple. Ripple allows three states for each cell; each neighborhood consists of three cells, a central cell and two cells flanking it. The rules of Ripple are reasonably straightforward and easily programmed:

1. If a cell is in state 0, its next state will be 2 if its flanking states add up to 2 or more. Otherwise its next state will be 0.
2. If a cell is in state 1, its next state will be 0.
3. If a cell is in state 2, its next state will be 1 if either of the flanking cells is in state 0. Otherwise its next state will be 2.

I am sure that Ripple is not about to replace Life; Ripple was designed to illustrate some of the more interesting possibilities line automata offer. For example, Ripple has simple gliders and an even simpler glider gun [*see top illustration on this page*].

Assume that Ripple's cellular space is entirely in state 0 except for two adjacent cells. The cell on the left is in state 2 and the other is in state 1. At the next generation this pattern will have shifted by one cell to the left. Undisturbed, the two-cell glider will ripple silently to the left forever. Interchange the states of the two cells and a right-moving glider is created. The glider gun consists of a single cell in state 2. That cell cycles through states 1, 0 and then back to 2, issuing a pair of gliders on each cycle. I wonder if anyone can find a gun in Ripple that emits gliders in a single direction only.



A NOT gate in Life: when gliders are present in the input, they are absent from the output

Simultaneous activation of a pair of such glider guns produces strange effects. In the process of mutual annihilation the resulting explosions send gliders off in both directions. If the guns are an even number of cells apart, this number of gliders go off in both directions. Otherwise a single gun remains in the middle and shoots out gliders in an interminable stream.

Ripple is just one line automaton. What of the others? The number of possible rule sets to consider is greatly reduced by adopting the ones Wolfram calls totalistic. Here the next state of a given cell is determined only by the sum of the states in the cell's neighborhood. The sum includes the state of the given cell. The number of possible sums varies from 0 to m, where m is the largest value of state multiplied by the size of the neighborhood. If one specifies how these sums become the central cell's next state, one has specified the line automaton completely.

For example, there is a very interesting line automaton governed by totalistic rules given in this table:

sum	5	4	3	2	1	0
next state	0	1	0	1	0	0

Here k and r are both equal to 2; the possible values of the sum of states in a five-cell neighborhood vary from 0 to 5. Wolfram calls this set of rules number 20 because the six next-state values in the table represent the number 20 written in binary notation.

There are 64 ways in all that the second row of the table can be filled in. Each one results in a line automaton and Wolfram has examined all 64 of them. Needless to say, a computer is essential in such an investigation. To determine the behavior of a given line automaton, Wolfram forms an initial pattern of some 100 cells in random states and then turns the automaton loose on the pattern. To nullify the effects of the vast arrays of zeros on each side of the pattern, Wolfram makes the left end of the initial pattern contiguous with the right end. This turns the pattern into a circle and its history becomes a cylinder, yet the result is the same as if the initial random pattern were repeated indefinitely in both directions of the original cellular space. In any event the effect of line automata on such random input patterns is surprisingly uniform. Each automaton will fall into one of four broad classes constructed by Wolfram:

Class 1. After a finite number of generations the pattern deteriorates into a single homogeneous state endlessly repeated.

Class 2. The pattern evolves into a

number of unvarying or periodic subpatterns.

Class 3. The pattern never develops any structure. Space-time diagrams look chaotic.

Class 4. The pattern develops complex localized subpatterns, some of them long-lasting.

Of the 64 automata in which k and r are both equal to 2, 25 percent are in class 1, 16 percent are in class 2, 53 percent are in class 3 and only 6 percent are in class 4. Wolfram suspects that computers might be found in the fourth class. For example, the line automaton with code number 20 is in class 4.

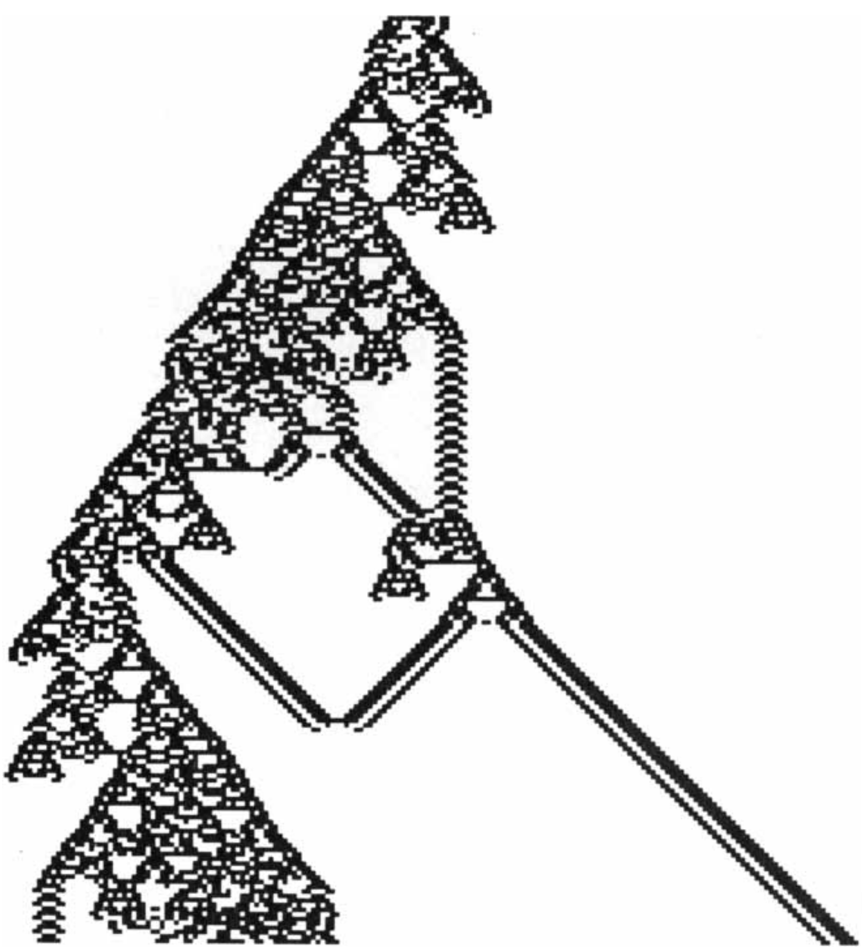
As though encouraging Wolfram in the search, the code-20 automaton has cheerfully disclosed some gliders. They are 10111011 and 1001111011. Both gliders move to the right. Totalistic rules are always symmetrical; to obtain left-moving gliders merely reverse these patterns. Is there a gun for such gliders in the code-20 space? Wolfram thinks there is. But there is another space where he has yet to find even a glider! It is the code-792 space. When we write this number in ternary notation, we obtain a set of rules for a 3-state line automaton:

sum	6	5	4	3	2	1	0
next state	1	0	0	2	1	0	0

The search for gliders and glider guns calls for the advice of a separate document, which I have persuaded Wolfram to write. It describes an algorithm that hunts for what Wolfram calls persistent structures. Adventurous readers may write to "Glider Gun Guidelines," *Scientific American,* 415 Madison Avenue, New York, N.Y. 10017. Please send $3 to cover the cost of mailing.

The search for gliders and glider guns focuses on a number of line automata thought to be computation universal. In other words, these are line automata capable of acting like a computer. Besides the code-20 and code-792 automata already discussed, there is the two-state line automaton ($r = 3$, code number 88) in which a glider gun has already been found.

James K. Park, a student at Princeton University, found the gun in that automaton. Readers who would like to witness Park's glider gun in operation must write a simple program for displaying the generations of a line automaton and for deriving one generation from the next. It is an easy matter to implement line automaton 88 with such a program: when you are ready, input the initial pattern 1111111111011 and watch it expand and contract. The gun spews out a glider in each direction once every 238 generations.

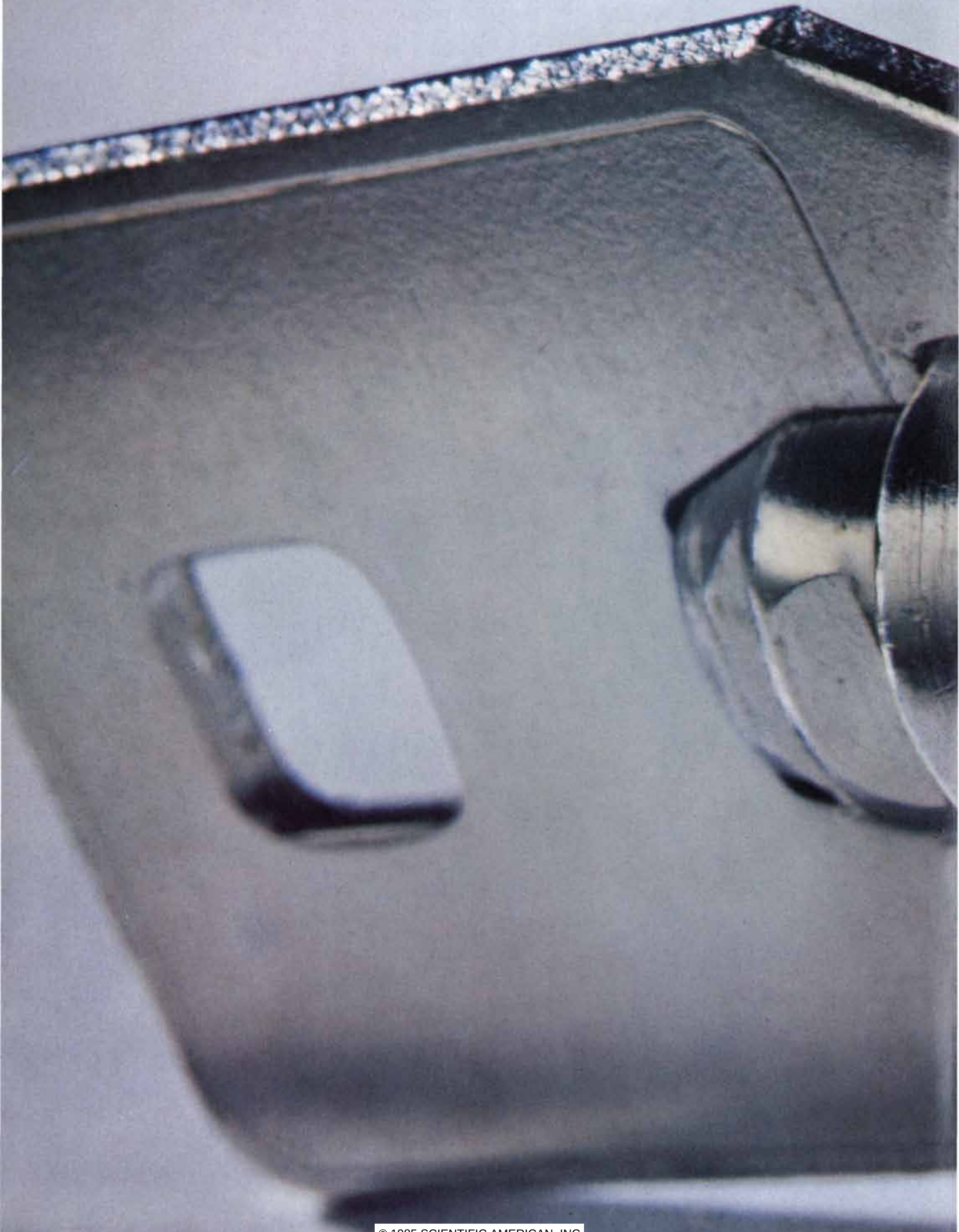


A close-up of James K. Park's one-dimensional glider gun

Rather than implement a specific automaton, readers are advised to write a program that takes arbitrary totalistic rules as input. This is easy to do for fixed values of k and r and almost as easy if these parameters are allowed to vary. Use a special array called *table* for the rules and two linear arrays called *newcells* and *oldcells* to hold the new and the old patterns currently being processed. Make the two arrays as wide as you like but bear in mind that a display screen (depending on the type of display) may show only a limited part of these arrays. Display the largest middle part of *newcells* that fits and cycle through the following steps: Transfer the contents of *newcells* into *oldcells.* Scan *oldcells;* for its ith member add up the value of *oldcells* from $i - r$ to $i + r$. Look up the resulting sum in *table* and transfer the state value so found to the ith member of *newcells.* Next reenter the computational cycle in the display phase and repeat it. Displays can be successive or stationary. In the former case a space-time diagram results; in the latter case one watches a kind of one-dimensional motion picture.

As noted above, a line automaton is called computation-universal if as generation succeeds generation there is an initial pattern capable of acting like a computer. Part of the initial pattern is the input and part of some later pattern is the output. Is it really possible to build such a computer out of glider guns and various other pieces of cellular hardware? Wolfram thinks it is. In one of the line automata currently under investigation gliders have been found that pass through certain stable subpatterns. This gives hope that the transmission of information within a linear computer need not be blocked by components unconcerned with that information. In addition to Wolfram there are other workers currently experimenting with line automata. Kenneth Steiglitz of Princeton is one of them. He has found gliders that have properties like those of solitons. Such gliders can even pass each other!

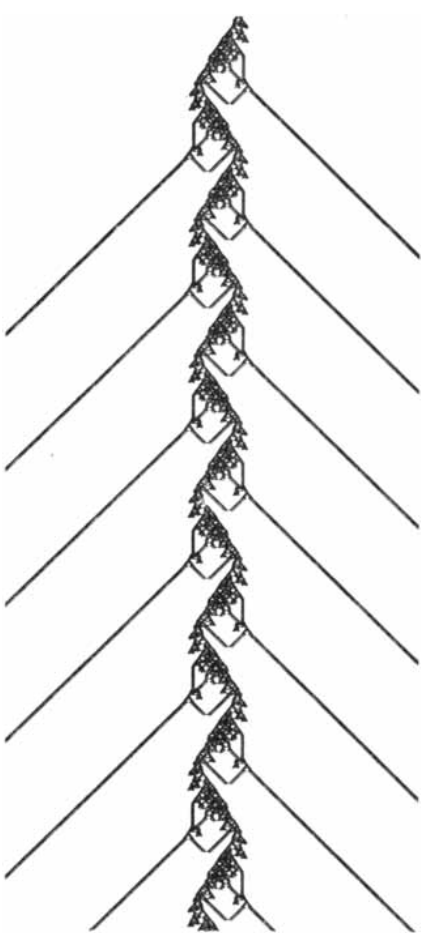
Will we eventually find persistent structures capable of moving, storing and manipulating information in a one-dimensional cellular realm? Perhaps not. It may be that there is some-




CLICK.
How do you give information access to some and deny others?
That's the secret that makes SCOMP click. This data security technology
produced by Honeywell is the first computer system ever awarded
the top A1 security rating by the Department of Defense.
Put a lock on sensitive data as strong as the Pentagon's. Talk to us.
For a white paper on our data security technology, write Security,
c/o Honeywell, Honeywell Plaza, Minneapolis, MN 55408.
Or call 1-800-328-5111, ext. 2760.
Together, we can find the answers.
Honeywell

thing inherently quite different about cellular computation in one dimension, something that requires us to look at computing in an entirely new way.

Wolfram has fantasized freely about how line automata illuminate the behavior of automata in general, and how they in turn provide insight into natural processes. Suppose one were to find a line automaton that closely mimics some natural process, for example a turbulent flow of liquid or gas, the motion of particles under the influence of forces, or even a biological process such as growth. Suppose further the automaton turned out to be computation-universal. In other words, not only can its space contain an explicit structure that computes but also the same space contains the computer implicitly: any attempt to predict the behavior of the automaton would by definition be attempting to predict the actions of a general-purpose computer. This in general cannot be done except by the same kind of device, another general-purpose computer. It follows that no short cut would be available for predicting the behavior of the corresponding natural system. Its processes would be computationally irreducible in the sense that the predicting mechanism (whether formula or computer) must simulate the system in question more or less directly.



Park's gun spewing gliders left and right

This conclusion brings us back to Ripple. An initial population of gliders going in random directions bounce off one another. Sometimes their collisions result in a glider gun, which produces more gliders, and sometimes their collisions produce nothing. It reminds me of a miniature one-dimensional universe filled with elementary particles rippling back and forth. Thus I have reversed the order of Wolfram's search. Starting with a strip automaton that behaves vaguely like a system of particles (odd ones at that), I have a dream that Ripple is computation-universal. Since it is my own personal automaton, I may be dreaming alone.

Last June's column on analog gadgets is producing such a deluge of material from different parts of the world that the subject seems worth revisiting, even though the September 1984 column also carried a reprise.

It is interesting to note that some gadgets of truly practical value have emerged. By a practical analog gadget I mean something like the atmospheric control system described by Homer B. Clay, a consulting engineer in Phoenix, Ariz. Clay was called in to help a printer control temperature and humidity in his plant. Temperature and humidity are critical because they affect the state of the carbon black that coats large copper rollers used in some printing processes. An experienced printer can hold a sheet of carbon black and tell by the amount of droop whether it is in a suitable state. When it is not suitable (because of incorrect moisture content or temperature), the carbon black wrinkles or cracks. Clay devised a more or less standard circuit to sense the temperature and humidity conditions and turn a spray system on or off to control them. The system apparently worked for a short time, until unseen variables threw it out of kilter.

Returning to the printing plant one day to evaluate his control system, Clay was led back to the printing area by a grinning plant engineer. Clay had some trouble believing what he saw: "Our little system was disconnected and in its place was a small strip of carbon black, one end fixed to the shelf, the other end free to move. When the droop was 'just right,' the free end closed a pair of contacts, turning on the sprays." The plant engineer rubbed it in by commenting, "Works like a charm."

Back in 1948 Robert Heppe of Fairfax, Va., was a freshman electrical engineer at the Queens, N.Y., plant of the Sylvania Electric Products Company. Heppe, assigned to assist in the design of vacuum tubes, found the process onerous. The problem was that one had to specify the size, shape and placement of the grids and beam-forming plates on paper. The design was then manufactured in the form of a single tube and tested. This could take several days. His supervisor, Gerald Rich, improved efficiency by suggesting a certain analog gadget.

The gadget consisted of a rubber sheet, a dowel, some plywood and several boxes of toothpicks. The rubber sheet clamped into a large ring represented the tube cross section magnified many times. The cathode was a wood dowel poking up in the center of the sheet. Arrays of toothpicks represented various grid designs. Negative grids tented the sheet up from below; positive grids depressed the sheet from above. Other aspects of tube geometry were captured by plywood shapes also imposed from below or above. Electrons pouring from the cathode were simulated by slowly emptying a can of BB's over the dowel. "It can be shown," writes Heppe, "that the slope of the rubber in such a gadget represents the electric field, and the height represents the voltage in the space between the electrodes.... The BB's rolled down the sheet [as in] a pin-ball game, some collecting at the plate, some at the positive grids. If we didn't like how many arrived at the various electrodes, or which way they went, we could move things around, change sizes, etc., and try it again." Promising configurations were embodied and tested in real tubes.

Many other readers offered examples. There was a hemocrit to measure solid fractions in human blood, invented by Dr. Alan Kwasman of Loma Linda, Calif. Mathias Soop of the European Space Agency in Darmstadt, West Germany, presented a string, nail and paper-clip gadget. The device illustrates an analog solution to the problem of minimizing thruster fuel consumption while controlling satellite attitude.

Readers also recalled analog gadgets of wider utility: the three-arm protractor used in coastal navigation, the range finder employed by U.S. naval forces in World War II and the planimeter, a marvelously simple instrument that serves to measure the area of a plane surface.

The best and biggest of analog gadgets are analog computers. These are alive and well at Electronic Associates, Inc., a manufacturer of analog equipment in West Long Branch, N.J. I was even invited to the plant to witness the superiority of these machines over the humble devices presented last June.

BOOKS

Vaporized diamonds, coca leaves and amaranth, orange stars, Rockaway memoirs, mǎhou'bpabū

by Philip Morrison

DIAMOND, by Gordon Davies. Adam Hilger Ltd., distributed in the U.S. by International Publishers Service, P.O. Box 230, Accord, Mass. 02018 ($28). An optically worked disk of superb gem diamond about the size of a nickel has lain in the open quite unwatched for the past half-dozen years. Covetous collectors may see its photograph here, but they can hardly act; that specimen is on Venus. It serves as a strong window to shield the infrared detector on our lander there from the pressure of the hot atmosphere.

Professor Davies, a King's College, London, physicist long engaged in research on diamond, has written an admirably nontechnical yet full and up-to-date study of the archetypal adamantine substance (from the Greek for unconquerable). He focuses on its marvelous nature and its technical uses rather than on its status as costly jewel. For all that, he does not neglect the natural gemstone. There is a fine chapter on mines and mining, and on the physics that underlies the art of diamond cutting.

Diamonds are enormous organic polymers of pure carbon; each atom links covalently to its four spatial neighbors as though it were in methane. The history of that finding begins with a futile attempt by the workers of the Accadèmia del Cimento to vaporize the unconquerable. They focused the Tuscan sun through their large burning glass on one of their liege's diamonds. Sir Humphry Davy ended the matter by burning diamonds in pure oxygen under careful control; "merely a solution of diamond in oxygen, without any change in the volume of the gas," he wrote. He confirmed the product as carbon dioxide in his usual bold style by smelling and tasting a solution as well as by utilizing less intimate means.

X-ray analysis revealed the carbon lattice to the Braggs in 1913. It was not until 1945, however, that the form of common diamond was established as unique. C. V. Raman had suggested that the electronic bonds might be directional, to allow a number of diamond structures, all compatible with the known atomic positions. His aim was to account for the observation that with respect to infrared transparency high-purity diamonds came in two types. The eminent crystallographer Kathleen Lonsdale quickly noted a minor but unmistakable spot in the diffraction pattern of a sample of the rare infrared-passing variety; the spot established the presence of a symmetry Raman's model would not fit.

The mineralogists have since appropriately named the shock-formed allotrope of diamond lonsdaleite; its hexagonal array is identical with the ordinary form as far as the tetrahedrons of the nearest neighbors go. At the second-nearest neighbors the tetrahedrons make a half turn to form the unusual phase, revealed by a differing spot diagram and another domain of stability. So far this arrangement has been seen in only a few micron-size specimens.

It turns out that the best colorless natural diamonds are quite pure; apart from occasional mineral inclusions and voids, only a few atoms among 1,000 or so are aliens. Genuinely foreign atoms, say iron or calcium, so distort the lattice that their energy is high. They can enter in parts per million only; they are as rare as diamonds themselves in the rich kimberlites. Parsimonious nature favors nitrogen impurities, which fit quite well. The boron atom is also a good fit, but the element is very rare on the earth; blue diamonds, an unusual variety, are colored by boron, often in striking zonal patterns.

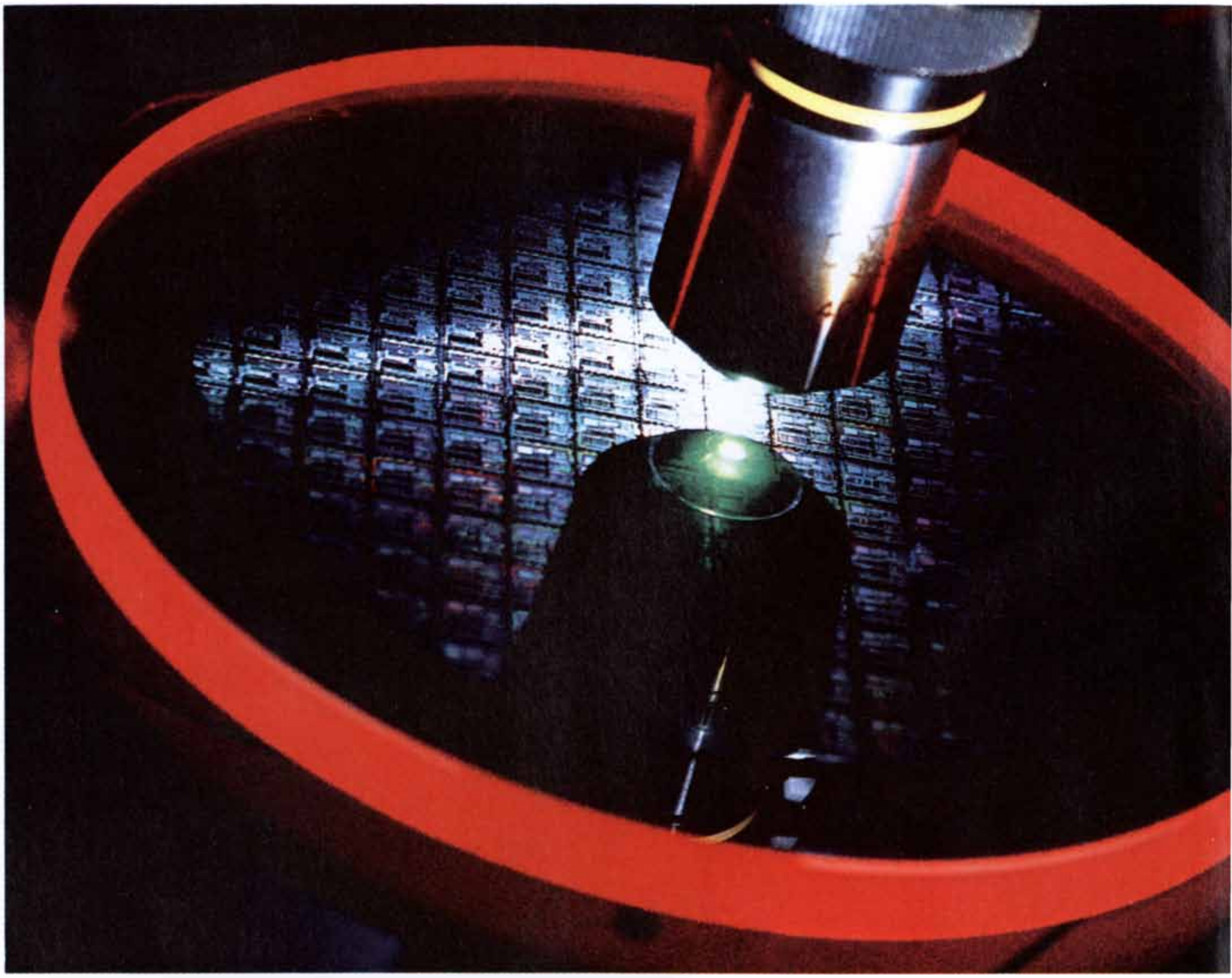
The story of the nitrogen held in diamonds is intriguing. Plenty of nitrogen can make diamonds yellow-brown. Flawless uncolored diamonds show a nitrogen concentration neatly proportional to their infrared absorption (measured at the characteristic peak near the eight-micron wavelength). The long-known difference between types of diamonds that caught Raman is an effect of nitrogen impurity. The rare ones, transparent to infrared, such as the Venus disk are unusually low in nitrogen. Still, they are not as low as indicated by magnetic-resonance techniques that interact with the unpaired electron in each nitrogen atom. These measurements picked up an amount of nitrogen less by several orders of magnitude than the quantity released on destructive heating in vacuum. The nitrogen electrons were magnetically inactive; they were probably canceling out as opposing "adjacent pairs."

That result is now 25 years old. It took nearly two decades to confirm the paired nitrogens. Davies himself probed the electron levels of the "molecule" through its absorption of energetic ultraviolet photons. Isolated nitrogen atoms, N_2-like pairs and even N_3's are now demonstrated. Heat and time shuffle and aggregate the nitrogen atoms; synthetic diamonds that harbor only isolated nitrogen atoms steadily grow pairs at high temperature; the nitrogen atoms apparently diffuse until they can pair within the diamond lattice. A method of dating diamond formation is thus provided. Random patches of atoms, seen as weak diffraction spots and revealed by electron microscopy as well, lie in monolayers along diamond crystal planes; these layers may consist of as many as a million atoms, which may or may not be mainly nitrogen.

The synthesis of diamond, achieved 20 years ago by rational thermodynamic understanding and big presses, is now an industrial process. Fine gem diamonds have been grown, a carat or two requiring a week of press time. Geologic time is cheaper: synthetic diamonds dominate as the hardest abrasives; jewelers deal only in natural stones.

How diamond grit is engineered into shapes that suit its manifold uses is one of the freshest topics here described. Submillimeter single crystals strong enough to work on hard rock are grown in carefully varied conditions. The needle-shaped little diamonds are all mounted edgewise in grinding wheels by exploiting the magnetism of the usual nickel and cobalt catalyst inclusions. Resin is poured on the edge of the wheel. The process is carried out within a radial magnetic field; consequently the cured resin wheel bears an aligned myriad of little sharp flat diamond knives. Surgical scalpels and microtome knives of diamond are individually shaped; almost everywhere tools and dies edged with sizable blocks of polycrystalline sintered diamond mounted in a cobalt matrix draw wire and drill for oil.

The cleaving of diamond by one



A pattern detection device scans millions of microcircuits on a silicon wafer—one of many automatic checking procedures Hitachi uses to ensure absolute quality for every microchip produced.

cunning tap against a steel blade seems almost magical. That blade is not sharp; it acts merely to force open a small guiding scratch, starting the excessive stress. The split travels through the stone as the bonds break successively, the cleavage advancing at speeds of up to seven kilometers per second, limited by the rate of atomic vibration. The plane of cleavage must locally present the fewest bonds per unit area. There are several possible directions, but only the octahedral plane is practical, probably because growth faults often lie there. It is also the hardest face of a diamond. Within about three degrees of that direction no surface can be prepared by polishing; the random orientation of the fine diamonds of polishing grit too infrequently presents a face harder than the surface it must work.

Dr. Davies has a flair for historical analysis. His account of the long struggle to synthesize diamond, a tale full of uncertainty and unwitting fraud, is a pleasure to read. It looks as though the persistent efforts of some very able chemists and engineers were more than once completed by their less devoted staff, who finally sneaked in tiny diamonds to bring to an end their long campaigns of fruitless and even dangerous trials. Davies' book gains special standing from his way of explaining some pretty subtle solid-state physics without resort to mathematics. The market and its history are the burden of most of the fascinating books on diamond written at so readable a level; this is diamond science. Diamonds are mineral messages from an inaccessible depth in the mantle. No book so far has made vivid to the general reader the cold volcanoes of awesome power, the diamond pipes through which those remarkable crystals suddenly entered our low-pressure world.

Post-Columbian print depicts pineapple plant

AMARANTH: MODERN PROSPECTS FOR AN ANCIENT CROP, a panel report of the National Research Council. National Academy Press, distributed by Rodale Press, 33 East Minor Street, Emmaus, Pa. 18049 (paperbound, $6.95). PRE-COLUMBIAN PLANT MIGRATION, edited by Doris Stone. Papers of the Peabody Museum of Archeology and Ethnology, distributed by Harvard University Press (paperbound, $30). The Forty Thieves had a password, "Open sesame!" But the eavesdropper could not get it right; he tried wheat, barley, oats ... The small seeds of sesame are chiefly pressed for their oil, but they are also enjoyed as a grainy treat. That is unusual, all right, for sesame is a broad-leaved annual, not at all a grass like the small-seeded true cereals with which it was confused. Buckwheat and the Andean crop quinoa are other domesticated broad-leaved annuals that are also called pseudocereals because their small nutritious seeds are eaten in pancakes and porridges.

The grain amaranths are pseudocereals too. Their seeds are millimeter-size, often smaller than mustard seeds; leaves, stems and flowers are brightly colored, even to purple and orange, and the seeds too may be pink, yellow or cream. Each seed head, rounded or long, is crammed with thousands of the tiny seeds. Like corn and beans, grain amaranth was a basic food in pre-Columbian America. It is recorded—the relevant page of a Nahuatl codex is reproduced here—that Montezuma was entitled to a yearly tribute that totaled more than 20,000 tons of grain amaranth; this was as much as the supplicants were obliged to offer in maize.

Amaranth had ritual worth among the Aztecs; its ground seeds were mixed with honey (by some accounts, at times with human blood) and shaped into cakes having the forms of beasts and gods. The cakes were eaten on high occasions in temple and home. The extirpation of the old cultures after the conquest seems to have extended to holy amaranth, and so the crop is cultivated today mainly in out-of-the-way places in the mountains, from Mexico south to the Andes. Corn and beans have traveled from the Americas to become familiar food for much of the world, but amaranth has remained largely local and obscure.

That it has value overseas is demonstrated by its widespread cultivation and use in the distant Punjab hills, where the red and yellow crop tinges "with flame the bare mountain slopes." Often amaranth occupies more than half of those upland fields. There they sell light cakes made of the white popped seeds bound with honey; it is just what the hill country markets in southern Mexico sell. (No fancy shapes appear, alas, in either of the up-to-date photographs.) How and when the plant made its way to the foothills of the Himalayas is not known; it has been at home there for a long time.

This brochure outlines the enthusiastic attention now given the amaranths by agronomists and innovative farmers from Argentina to Zambia. A thousand varieties from a handful of species are at present growing in test fields in Pennsylvania, where the adaptive and tolerant crop seems to have reached the threshold of commercial production in the U.S. The genus shares the efficient photosynthesis pathway dubbed C_4 with the tropical grasses corn and sugarcane; that biochemistry promises economy of water use and good response to intense sunlight. This is a crop meant for bright, hot and dry lands.

The nutritionists admire it especially, for amaranth seed protein is closer to human needs in its amino acid mix than any other grain; its mean protein content is above that of the rest as well. Amaranth has too little gluten to make leavened bread, although it is a valuable supplement to flour, particularly enriching that grain with the amino acid lysine.

Taking a new crop into widespread cultivation is not a swift process. Yields, insect resistance, uniformity of height, seed held tightly in the heads to allow machine reaping and many another property of the subtle artifact that is a modern crop are under active improvement by the breeders. Along with the crop, a market for it must grow up as well.

It seems almost too good to be true that the same genus has three or four species yielding broad fresh leaves, to be boiled as potherbs like spinach. But from Greece to Taiwan those amaranths are in wide use; in the humid Tropics of Africa and southern Asia there is perhaps no more widely used vegetable. Some of these species are Old World in origin, some have come more recently from the Americas. During the 19th century a variety of a red-leaved New World grain amaranth spread as a major garden crop into the African interior, outrunning the European explorers. These varieties of amaranth are less in need of technical attention than the conventional grain

SCIENCE/SCOPE

The feasibility of turning sea water into electricity is being studied in fusion energy experiments at Kyoto University in Japan. The studies involve a Hughes Aircraft Company gyrotron, a microwave tube that uses a spiraling stream of electrons to produce extremely high power microwave frequencies. Fusion energy holds tremendous potential because its source of fuel (hydrogen) can be extracted from sea water. It could produce large amounts of power with little or no radioactive waste and no threat of meltdown or explosion. In fusion energy research, the gyrotron's high-power radio waves heat hydrogen particles (plasma) to temperatures of tens of millions of degrees. These particles fuse under pressure, causing a thermonuclear reaction that provides energy for driving steam turbines.

A third communications satellite is being built for Indonesia as a replacement for one rescued from an errant orbit last November by NASA's space shuttle. Palapa B-3, set for launch in 1986, is the third in a follow-on series of spacecraft designed and built by Hughes for Perumtel, Indonesia's government-owned telecommunications agency. The Palapa B model has more than twice the capacity of Palapa A, which in 1972 unified the world's largest archipelago electronically. It can carry 1,000 voice circuits or a color television transmission in each of its 24 transponders.

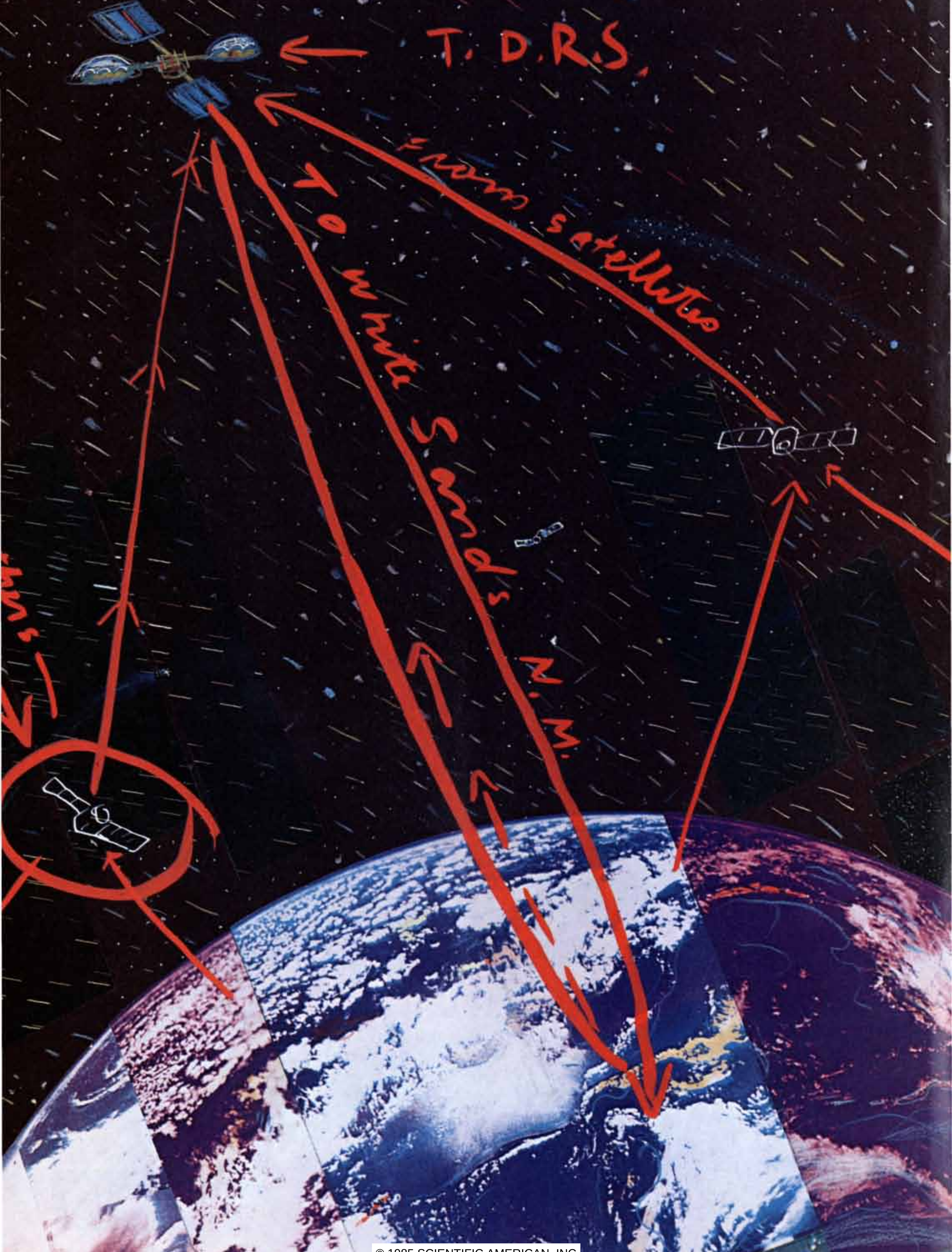
Single-seat military aircraft will be able to fly low-altitude attack missions at night with a system now undergoing evaluation by the U.S. Air Force. The Low Altitude Navigation and Targeting Infrared System for Night (LANTIRN) permits attacks at night and in low-visibility weather while relieving a pilot of many manual targeting functions. Elements include infrared sensors, an automatic multimode tracker, a laser designator/ranger, and a terrain-following navigation system. These components are mounted in two pods installed under the aircraft. Hughes, as subcontractor to Martin Marietta, has supplied five modified Imaging Infrared Maverick air-to-surface missiles along with launchers and a missile boresight correlator (the device which automatically hands off targets from the pod sensor to the missile). LANTIRN is designed for the F-16, F-15, and A-10 aircraft.

Swedish JAS-39 pilots will have better views from their cockpits, thanks to a wide-field-of-view head-up display (HUD) that incorporates diffraction optics technology. The display saves pilots from looking down into the cockpit to read instruments by superimposing data on a clear plate mounted at the pilot's eye level. Compared with conventional displays, the new HUD is clearer and eliminates bulky support structures. Its wide field of view can be used with infrared or low-light-level TV imagery so pilots can fly high-speed low-altitude missions at night. Hughes produces the HUD using a proprietary process involving holographic techniques and lasers. Sweden is the first country to award a production contract for a HUD that uses diffraction optics.

Career growth opportunities exist at Hughes Support Systems for a variety of engineers qualified by degree or extensive work experience. They include systems engineers, radar engineers, and software and hardware design engineers for major simulation and test equipment programs. Also, field engineering posts throughout the U.S. offer travel, autonomy, and responsibility for the life cycle of Hughes electronics systems. Please send your resume to Lowell Anderson, Professional Employment, Dept. S2, Hughes Aircraft Company, P.O. Box 9399, Long Beach, CA 90801-0463. Equal opportunity employer. U.S. citizenship required.

For more information write to: P.O. Box 11205, Dept. 68-10, Marina del Rey, CA 90295

HUGHES
AIRCRAFT COMPANY


T. D. R. S.
From satellites
To White Sands N. M.

species; right now one major form beats spinach in nutrition and yield in warm climates; it can even reseed itself. A recent tasters' test panel in Maryland found it "at least as good as spinach." A century ago peanuts, sunflower and even soybean were important crops only locally; now these harvests help to nourish the continents. Open, amaranth!

The second book assembles a set of nine related papers that together seek to unravel the times before Columbus when important New World crops spread from their place of origin to the rest of the Americas. At the center of attention lie the Amazon basin as source and the shores of the Gulf of Mexico as recipient. All these scholars link botany with archaeology; the two disciplines share the adventurous field style and a high regard for visual evidence. The book, although technical, is both readable for its firsthand qualities and beautifully presented and illustrated.

Bones intimately reveal diet, and diet sheds light on crops. The data reside in the collagen of bone, a protein that resists contaminants. The isotope ratio, carbon 12 to carbon 13, reflects the chain of reactions that fix carbon in the plant food taken during life, whether or not an animal intermediary is involved. The unusual biochemistry of maize, again the C_4 photosynthesis chain, is responsible. A change of staple diet from manioc to maize means that the intake isotope ratio is enriched in the rarer odd isotope by about 1 percent.

In coastal people seafood might confuse the decision, for it too shows higher odd-carbon content. Nitrogen isotope ratios can settle that issue; low nitrogen 15 accompanied by high carbon 13 means high maize consumption. One result from this method is in; five skeletons from the Orinoco show a strong increase in the odd isotope. The archaeology found with the 800 B.C. remains includes graters and griddles; these are the kitchen tools of root eaters. The A.D. 400 bones go with metates and carbonized kernels of maize. Their staple had changed. There was almost no maize in early use; a millennium or so later there was no less than 80 percent dependence on maize. The local population increased 15-fold over the same period: these people had become grain farmers, a change that was perhaps both cause and effect of their increase.

Botanical evidence suggests the wet Amazon forests first gave rise to cacao (chocolate) and pineapple, to the peach palm and starchy manioc. Chilis, coca and maize have diffused in complex paths north and south across the Isthmus or around it, although they did not begin in that same forest. To read here of the two species of coca now chewed in mountain and in forest for its percent or less of the alkaloid cocaine is to admire an ancient stone head from Colombia, its cheeks bulging with the quid of leaf. Now the concentrated alkaloid floods north by airplane.

Why did coca not spread long ago past Central America? It is not reliably reported north of there before Columbus. Its seeds seem unable to survive for long, and no ancient chain of gardeners was present to transfer the living plant, as Columbus carried the pineapple back on shipboard to his monarch.

How is it that cacao was collected in its native rain forests only as another not very succulent wild fruit, whereas the prized stimulant seeds went unused by the people? It seems more than likely that the plant ranged north without help from human hands as far as the wet Costa Rican forest, whence the Mayans long ago learned to use wild seeds for those heady alkaloids. Eventually they took the plant into cultivation so that they could raise the now precious tree itself.

What is certain is that the long symbiosis between plants and humanity holds much more than we yet know. Even today that intricate bond has not lost its endless novelty.

COLOURS OF THE STARS, by David Malin and Paul Murdin. Cambridge University Press ($27.95). Nowadays the astronomers, armed with instruments of unprecedented power (and cost), assemble the images they gather from any wave band into exuberant maps, familiar illustrations in this magazine and in some excellent new books. Meticulous reports of form they are, but the colors are all art and artifice.

Partners Malin and Murdin are respectively a noted photographic specialist at the Anglo-Australian Observatory and the head optical astronomer at the new United Kingdom telescopes in the Canary Islands. They stay much closer to the world of color perceived by the human eye. They work with visual images on photographic plates, not with great strings of bits; their tools are of the darkroom, not of the crammed relay rack. Still, it is no mean arsenal Malin and Murdin draw on: a big telescope is generally at hand, and each photographic emulsion is no less than a compact high-gain analogue amplifier and gigachannel recorder in its own right.

The exposed plate holds within it a physical image, a spatial distribution of tiny grains of silver. Wonders are worked by clever employment of that arrangement in space. As one example, any very faint image is stored closer to the plate surface than are the uniformly distributed grains of background fog. The reason is that the light scatters on its way through the emulsion, and with little light a faint image contains few grains at depth. Contact prints made with a diffuse light source below the plate show up the shallow image sharply but soften the shadows of the deeper-lying fog. The faint features are amplified against the visual noise in such prints. The technique reveals galaxies that have entire onion shells of stars and galaxies surrounded by otherwise-invisible faint halos. A complementary masking technique brings out detail hidden within the overexposed bland regions of brighter objects. There is a striking gallery of such black-and-white victories of unexpected contrast.

But it is the color pictures of stars and clusters, of rich fields of nebulas, and of galaxies that are the main harvest. Photographs using the conventional tripack color film go back to a pioneer, the late William C. Miller, who worked at the Hale Observatory on Mount Palomar before 1960. Miller's color images are found in every textbook.

Commercial films, however, are not tuned to the extreme exposure times of deep-sky astronomy. The scheme here made optimum for astronomical color is in principle that of the first color photography, the colored ribbons projected by Clerk Maxwell and his expert photographer colleague Thomas Sutton before the Royal Society in 1861.

The technique is the direct superposition of three positive separation images made in black and white, both taken and recombined through well-chosen red, green and blue filters. Instead of a final projection the images can be superposed carefully onto a positive-working color-print paper (Malin and Murdin use Cibachrome). No process seems to allow more flexibility for the calibrated corrections needed to regain acceptably faithful color of the background sky. All the black-and-white magic outlined above can be applied to the separation plates.

The gallery of some 60 color pictures is pretty glorious. Blue reflection nebulas, lighted by stars to produce skylight on a light-year scale, jostle star-filled clouds illuminated by the red recombination glow of ionized hydrogen; dark foreground dust modulates the entire painterly scene. "Making color pictures of new areas of sky . . . is always exciting, because one cannot visualize the end result." The nebular regions of Orion, the Trifid

Nebula, the enigmas of the active galaxy Centaurus A and the Magellanic Clouds at scales large and small are a few of the end results to be admired, marvels of the far southern sky.

The question that remains is, of course, how true are such colors? They are clearly not the artist's arbitrary pleasures open to the computer masters, but neither are they what eyes can see. Even if an observer traveled to the spot, these colors would not appear. They are too faint for the cones of the retina to record properly. The colors can nonetheless claim to be simple unbiased extrapolations of the human chromatic sense.

A careful account is given of the difficulties presented by the small-angle, faint, unsaturated objects of the cosmos to our subtle human color vision, adapted to very different scenes. One splendid pair of photographs shows the big dome of the Anglo-Australian telescope, a snapshot by sunlight and a long exposure by full moon. They were shot on the same roll of fast color film but, accounting for both lens opening and shutter speed, the exposure ratio was 10 million to one. The exposures look almost the same; the night sky is almost as clear a blue as the day sky. The star trails are in color, and the Large Cloud of Magellan is pink. Only the swaying gum trees are blurred in the long exposure. Although the filmmakers have their *nuit américaine,* color suppressed, made with a blue filter, no eyewitness can report a blue sky by moonlight. The eye does not accumulate photons. It is storage time that outdoes direct vision.

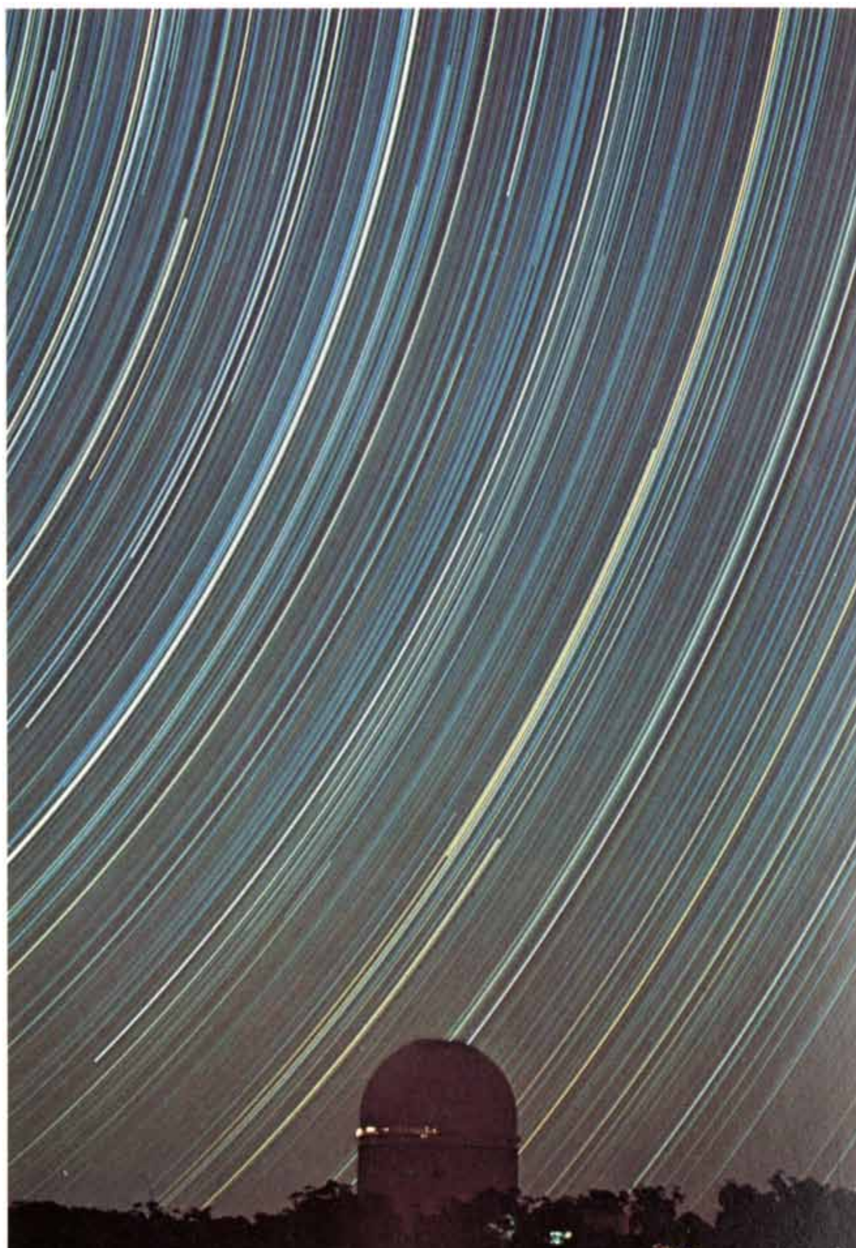
The history of visual star color is a fascinating subplot of this original book. That wonderfully reflective observer, Marcel Minnaert, reported star colors accurately, as he saw everything else. Stars hold every difficulty for color judgment, yet he could pick out white, four shades of yellow, and orange. (He listed Mars and Anatares as orange.) His somewhat bland scale correlates neatly with photoelectric color results.

There have been imaginative reports too, in which stars are described as jacinth and sardonyx. One astronomer spent years noting color changes in a Pointer star. He found a definite period of 35 days, varying from "chrome yellow" to "pale firey red" and back again. The star is a binary, we know now, with a 44-year period. Its color is unvarying. He was looking for the changes in color he imagined might arise from the Doppler shift, not fully understood in the 1870's, when motion was expected to be the origin of star color. "The ingenuity of the theory is extreme but its correctness is more than doubtful," John Tyndall wrote dryly.

Maybe we shall yet see star colors change, Malin and Murdin suggest. We can imagine a marvelous space-made time-lapse film of unique SS-433, its relativistic jets "changing colour from yellow to deepest red, and back, as they precess." It is 100 years since the first deep astrophotographs; what may another century of photography not bring?

"SURELY YOU'RE JOKING, MR. FEYNMAN!": ADVENTURES OF A CURIOUS CHARACTER, by Richard P. Feynman, as told to Ralph Leighton. W. W. Norton & Company ($16.95). Even the official entry for the library card has caught some of the style of the 40-odd pieces that make up this delightful mosaic of an informal autobiography taken from tape recordings: "Science—Anecdotes, facetiae, satire, etc."

The pieces are arranged in order of time. They begin with memories of that bright kid in Far Rockaway who fixed creaky broken radios and proceed through the times of the M.I.T. undergraduate interested only in science. They encompass the experiences of the safecracking young theorist at Los Alamos, who was the only guy who did not just say, "Yes, yes, Dr. Bohr." Ultimately they carry through to that ceremonial evening in Stock-



Setting stars leave time-exposure trails above the dome of the Anglo-Australian Observatory holm with a sample of princesses and the dinner organizer, and they include a serious but far from stuffy commencement address delivered at Caltech a decade ago.

Follow the tip from the Library of Congress. Anecdotes: Perhaps the Las Vegas stories can stand for all. Unlike most of the people in town, the showgirls Feynman met had read *Time* and knew of Pauling and Gell-Mann. One day such a young woman introduced Dick to Nick the Greek, who explained how he could make a living at a game where the odds are 50–50; he battens on side bets he can strike at favorable odds "with people ... who ... will bet ... just to have the chance of telling the story ... of how they beat Nick the Greek."

Facetiae: It is hard to beat the wonderful Italian double-talk. It is the intonation that counts; after all, the words might be quite unfamiliar, in the dialect of some distant region, but that cheerful fellow slapping the back of one hand against the other must be an iTALian! "I call back, 'RONte BALta!', returning the greeting."

Satire: It was the dean's wife at Princeton pouring tea long ago who declared what became the title of this book. The innocent new graduate student, asked whether he would like cream or lemon, had absently answered, "Both." But the imitation Oxford that was Princeton and "what I would call a gold-plated cyclotron" at M.I.T. are both treated with strong but not misleading exaggeration. Pomposity and pretense, from academic gowns to rubber-stamp committees, are forever Feynman's special butts. One concedes that it might not be satire but simple candor to point out how practical gown wearing is at table, for gowns are never cleaned or repaired, and remain correct apparel even in tatters.

Etc.: Feynman is a splendid traveler. He is a graceful and rhythmic dancer and percussionist, an articulate and enthusiastic linguist and mime, witty, unfailingly dexterous with word, hand and mind, and irresistibly led to whatever is lively and a bit raffish, free of prejudices save against empty forms. Perhaps the best story of all is the one he tells of his well-earned part in a winning band one pre-Carnaval in Rio. Marching down Avenida Atlantica among 100 Brazilians from the *favelas* all dressed up in Greek helmets of papier-mâché, Feynman was beating a rhythmic chick-a-chick on a toy frying pan in his own widely admired style. They swing past the waiters in his luxury hotel ... "O PROFESSOR!" It is an ultimate of tourist tales.

Now to science. To his mind the Royal Society was right on: *nullius in verba*. Take nobody's word for it, not even your own. An unhelpful text says that triboluminescence is the light emitted when crystals are crushed. That is not science but a word described in other words. In fact, if you crush a lump of sugar with pliers in the dark, you can see a bluish flash. No one knows why. That is triboluminescence, put so that you can try it at home. "I can't understand anything in general unless I am carrying along in my mind a specific example and watching it go ... I ask a lot of these 'dumb' questions.... But later ... when the equation says it should behave so-and-so and I know that's the wrong way around, I jump up"—and so he signals the error.

A second related theme is the need for recognition of the wider context. Puzzles are best solved not on their own terms alone but in the world with which they interact. The magicians know how powerful is the misdirection of implied or explicit false purpose or gesture. Feynman shared with the professional locksmith at Los Alamos the recognition that the prized safe often still bears the combination set at the factory! One Los Alamos safe in five opened at once by trial of the two factory combinations. Some miracles of solution, with safes or people, take place by sheer chance, others by inference from context and fewest by following the most evident path, one blocked by conscious effort.

Lighthearted, animated by good-natured chutzpah, the book is not a set of jokes. Generally Mr. Feynman is not joking; it is we, the setters of ritual performance, of hypocritical standards, pretenders to care and understanding, who are joking instead. This is the book of a powerful mind honest beyond everything else, a specialist in spade-naming. The reason is clear: "I have to understand the world, you see." These joyful understandings and uncoverings will help and delight others for a long time to come, sharp evocations of a life around and beyond the culture of science in the 20th century.

Ralph Leighton is a Pasadena friend and fellow drummer who compiled the tales for the pleasure of the rest of us. The written lines are not far from the swift, imitative, expressive, colloquial speech in which they were first told by that curious character who drives a sand-colored van painted all over with... Apache pictographs? No, time diagrams of photon-particle interaction.

LANGUAGES OF ASIA AND THE PACIFIC: A TRAVELLERS' PHRASEBOOK, by Charles Hamblin. Angus & Robertson Publishers. Distributed in the U.S. by Merrimack Publishers' Circle, 47 Pelham Road, Salem, N.H. 03079 ($15.95). In our day of easy travel, abundant meetings and international telecommunications the little phrase book that allows swift nodding acquaintanceship with a language is more useful than ever. Such a book also offers a stay-at-home some chance to gain new feeling for a world more diverse than Babel.

In this compact little volume a footloose professor of philosophy from New South Wales has gathered (with help from his knowing friends) no fewer than 25 such phrase books. The reference language of the framework is English, and the list includes Spanish, Portuguese and French, certainly no novelties. The other 22 tongues span the chief speech you might hear among a round two billion people as you move eastward along the shores of the Indian Ocean from the Persian Gulf around the Pacific (except that Russian is omitted).

From Farsi to Hindi and Urdu (taken as one), to Bengali, Singhalese and Tamil, on to five languages of southeastern Asia, to East Asia, and on to islands large and small, from Malay and Indonesian (again treated as one) to Tongan and Tahitian, not forgetting New Guinea and Fiji. Chinese speakers are recognized in three major forms, Mandarin, Cantonese and Hokkien, the last the dialect of Fujian Province, now leading in Singapore and Penang.

For each of the languages there are a couple of pages of dehydrated grammatic tips, then the workhorse words of greeting, counting and so on, a few pages of traveler's sentences, including phonetic renderings, and several hundred words of alphabetized vocabulary, arranged to allow quick consultation. There are seven tonal languages included, with a little first aid for the inadept.

Trace the words for gasoline or tape recorder across two dozen tongues; compare numbers (these are written out in a few important unfamiliar forms, although as a whole the text remains within the Roman alphabet); consent with a casual New Guinea Pidgin *orait* or decline politely in Burmese, *mahou'bpabū*.

There is much encouragement for those readers who seek to extend their horizons: "The beginner will be excused the graded courtesies of Tongan society." Today, however, our polyglot species has little place for linguistic snobbery: "Nearly every language in this book is the vehicle of a substantial literary culture, and the speakers of the others are hard at work building."

Established 1845 SCIENTIFIC AMERICAN May 1985 Volume 252 Number 5

The Federal Support of Mathematics

Between the late 1960's and the start of this decade it declined by about a third, and without new Federal initiatives the future health of mathematics is at risk. A plan for renewal is presented

by Edward E. David, Jr.

November 19, 1969, was a black Wednesday for basic research in mathematics. On that day President Nixon signed the Military Procurement Authorization into law for fiscal year 1970. Section 203 of the document was the so-called Mansfield Amendment, named for Senator Mike Mansfield of Montana, who introduced it. The Mansfield Amendment adopted the seemingly uncontroversial policy of limiting the Department of Defense to "studies and projects that directly and apparently relate to defense needs." Its effect, however, was to cut off a major source of Federal money for basic research. The consequences have been particularly severe and long-lasting for fundamental work in mathematics. Although Federal support for fundamental mathematics had begun to wane as early as 1968, the budget policies articulated by the Mansfield Amendment helped to bring about a 15-year period of neglect of the field.

The Mansfield Amendment was in legal effect for only a year, but its influence has continued to be felt throughout the Department of Defense. Moreover, it has drastically reduced the willingness of many other Federal agencies to fund basic scientific work that cannot be clearly related to their current missions. Between 1968 and 1973 the Federal support of fundamental mathematics was reduced by an estimated 33 percent, and the cuts in the defense budget accounted for a significant part of the reduction.

The intent of the Mansfield Amendment was to transfer responsibility for long-term, basic research to the civilian Federal agencies. In Washington, however, such drastic changes rarely go smoothly and often thwart the original intent. To compensate for the defense cutbacks in all the basic sciences, Congress appropriated $10 million for the National Science Foundation (NSF) in fiscal year 1971 and $40 million in fiscal year 1972. For basic sciences such as physics and chemistry this support proved to be a reasonably effective stopgap. It maintained the general level of productive work until other sources of income were found. For mathematics, however, the only major source of Federal support outside the defense establishment had traditionally been the NSF. Virtually none of the new appropriations for basic science reached the Mathematical Sciences Section of that agency.

On the contrary, from 1967 to 1971 and in several subsequent periods NSF funds for mathematics, measured in constant dollars, actually fell. Between 1969 and 1974 the number of predoctoral fellowships and traineeships supported by the NSF in mathematics declined by more than 90 percent. It is heartening to note that in recent years the NSF, with strong backing from the White House Office of Science and Technology Policy, has substantially increased its funding for mathematics. Nevertheless, the previous cuts were so deep that NSF funding for mathematics regained its 1968 level only in 1984, by which time the number of mathematicians needed to meet the demands of teaching in the universities had more than doubled.

In the universities, where 90 percent of all fundamental work in mathematics is carried out, the initial response to the cutbacks was to absorb them through the redeployment of internal resources. Unfortunately in the mid-1970's many universities began to face serious financial difficulties of their own, largely because Federal funding cuts affected many disciplines. Mathematics was hit particularly hard, since more than 90 percent of its outside support continued to come from the Department of Defense and the NSF; funds from industry for fundamental work were negligible at the time because of a reluctance to finance fundamental work that is many steps removed from application.

The consequences of this series of events are not hard to guess. By various errors of inadvertence and omission the U.S. has progressively undernourished the university mathematics community to the point where, without new Federal initiatives, it will not be able to carry out its essential task. Between 1973 and 1982 Federal support for mathematics grew hardly at all in constant dollars, and so it remained at roughly two-thirds of its level in 1968. The typical investigator in 1982 could expect only a third of the level of support that was granted in 1968. The number of American citizens awarded a doctorate in mathematics fell by more than half in the same period. So few students are now studying for a doctorate in mathematics that, if present trends continue, the university mathematics community will not be able to replace itself with qualified people.

In the spring of 1981 this deteriorating state of affairs prompted action. William Browder of Princeton University, who was then chairman of the Office of Mathematical Sciences at the National Research Council (NRC), approached other scientists at the NRC with an urgent appeal for a study of the prevailing health of mathematics. A committee, made up of many nonmathematicians as well as mathematicians, was appointed to examine the problem closely and make recommendations; I was asked to serve as its chairman. Our detailed findings show that Browder's initial concerns were more than justified. What is now at stake is not only the future of mathematics but also the quality of the undergraduate instruction in mathematics for the students on whom the future of science, engineering, the economy and society in general will depend.

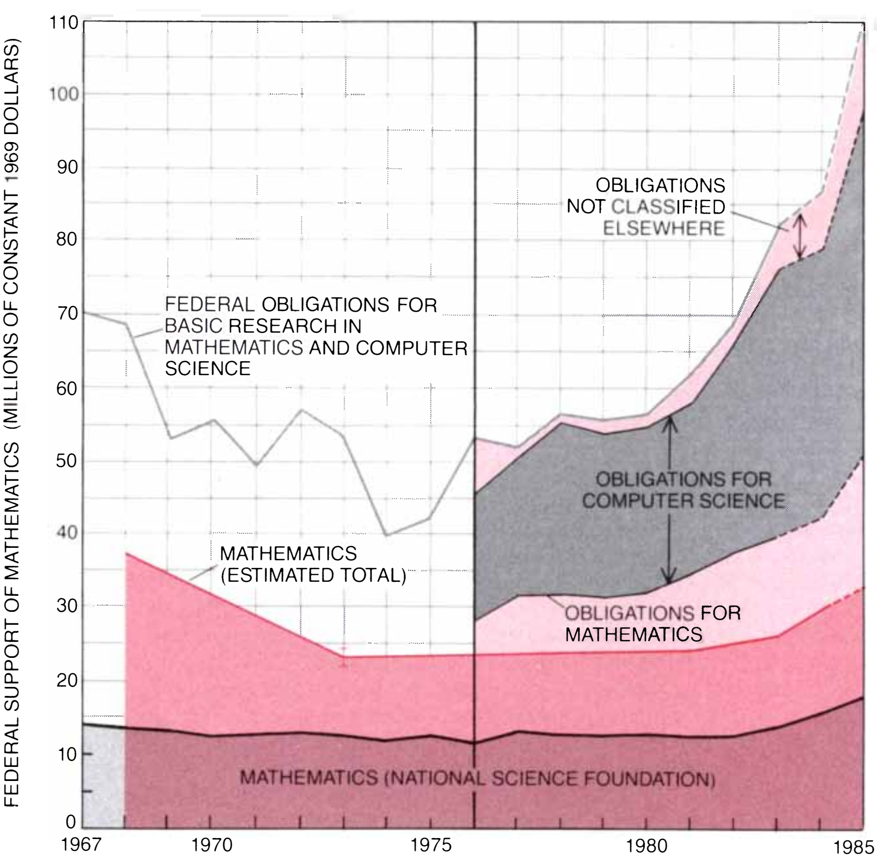
How did such a drastic change in the allocation of society's resources to mathematics go unnoticed for so long? For one thing, the productivity of mathematicians in this country is remarkably high. In the past several decades they have compiled a dazzling record of achievement. The Fields Medal, the equivalent for the mathematician of a Nobel prize, has been awarded to mathematicians in the U.S. 11 out of 27 times. Four Nobel prizes awarded to U.S. scientists in the past six years have recognized work in astrophysics, economics, medicine and physics that was largely mathematical in nature. In spite of serious inroads that have been made on the time the academic mathematician can allot to original work, several of the oldest and hardest problems in mathematics have recently been solved.

The second reason the neglect of mathematics went unnoticed is that exact figures for the support of basic research are hard to assemble. The support of mathematics within the Federal Government is not monolithic. There is no single item in the Federal budget that encompasses all Federally supported work in mathematics. Although, as I have mentioned, the NSF and the Department of Defense administer most of the appropriated funds, the funds are further split up within the defense establishment among agencies of each of the three major services: the Air Force Office of Scientific Research, the Army Research Office and the Office of Naval Research. Each agency has its own mission, accounting methods and research definitions. Moreover, most Federally supported research is budgeted by project, and so the support of mathematics must be estimated as a share of each project. Such estimates are particularly difficult to make for budgets that do not specify the funds allocated to mathematical work.

Our study found that several Federal reporting practices have been inadvertently masking the extent of the decline in the support of mathematics. Until 1976, for example, mathematics and computer science appeared together as a single line item in the statistical summary of Federal obligations for research and development collected each year by the NSF. Since then the NSF has collected funding information about the two fields separately, thereby recognizing that they are intellectually independent.

The first half of the 1970's, however, was a period of extremely rapid growth in computer science; consequently the line item that included the support of mathematics appeared to be growing nicely. The lack of growth in the budget for mathematics during this critical period was not visible even to mathematicians in policymaking positions at the time, and the stagnation cannot be fully documented in retrospect because many program records are no longer available.

The term "mathematical research"



FEDERAL FUNDING in mathematics can only be estimated from the available statistical data. For example, one measure of the support, which is published by the National Science Foundation (NSF), is the total Federal funding obligations for basic research in mathematics and computer science (*solid gray line*). Until 1976, however, no distinction was made in the reporting of this statistic between the support of mathematics and the support of computer science. The more detailed breakdown reported since that year is plotted in the right part of the graph, but even this information is almost certainly misleading. The definition of "mathematical research" has varied according to the person who reports the data from an agency. A careful analysis of the Federal support of fundamental mathematics for selected years (*solid colored line*) suggests it declined by about a third in the late 1960's and early 1970's, in large part because of reductions in defense spending. Moreover, there was an almost complete stagnation in Federal funding throughout the subsequent decade. Although support from the NSF was supposed to offset the losses in defense funding, additional Congressional appropriations for the NSF never found their way into the Mathematical Sciences Section. Hence NSF funding for mathematics also declined in almost every year from 1967 through 1976 and did not begin to rise significantly until 1983 (*solid black line*). Estimated increases in Federal funding for 1984 and 1985 are indicated by the broken lines.

in the NSF statistical summary can also be misleading. Indeed, the summary itself warns its readers that the definition of mathematical research may vary widely among agencies. Typically the information on which the summary is based is supplied anonymously by people with no mathematical or scientific background, who are merely filling out the blanks in a questionnaire as best they can.

It is now clear that a large fraction of the money classified as funding "mathematical research" does not support fundamental work in mathematics. Instead such income supports the application of known mathematical methods in diverse fields. New applications of known mathematics have increased explosively in recent years, and so once again the line item in the NSF summary has falsely suggested that the support of fundamental mathematics is in a healthy state.

In 1979 the NSF took a major step toward improving the reliability, comprehensiveness and consistency of the Federal data on the funding of mathematics. The Interagency Committee for Extramural Mathematics Programs (ICEMAP) was reactivated. The members of ICEMAP are knowledgeable representatives from each of the major Federal agencies that support mathematics; excellent data on funding in the years since 1981 are now available.

Although inadequate funding information accounts for the neglect of fundamental mathematics by the makers of science policy, the neglect of the problem by mathematicians themselves is more difficult to understand. It appears that most mathematicians were generally unaware of what was happening not only for the reasons already mentioned but also because of the role of the universities in masking the initial decline. The universities assumed just enough of the additional burden to obscure the magnitude of the Federal shifts in policy. The support available in universities today is sufficient only to retard the decline in the support of mathematics, not to arrest it and certainly not to reverse it.

Because of the inadequacy of the centralized records, one must resort to surveys in order to gauge the relative levels of Federal and university support. In 1980 a survey by the National Science Board showed that an overwhelming proportion of the work in mathematics was being supported by the universities rather than by the Federal Government. The finding was in sharp contrast to the patterns of Federal support in related disciplines. For example, the number of mathematicians teaching at universities in 1980 who were actively engaged in research and publishing was somewhat greater than the number of university physicists or chemists. Only about two-thirds as many mathematicians were getting Federal support.

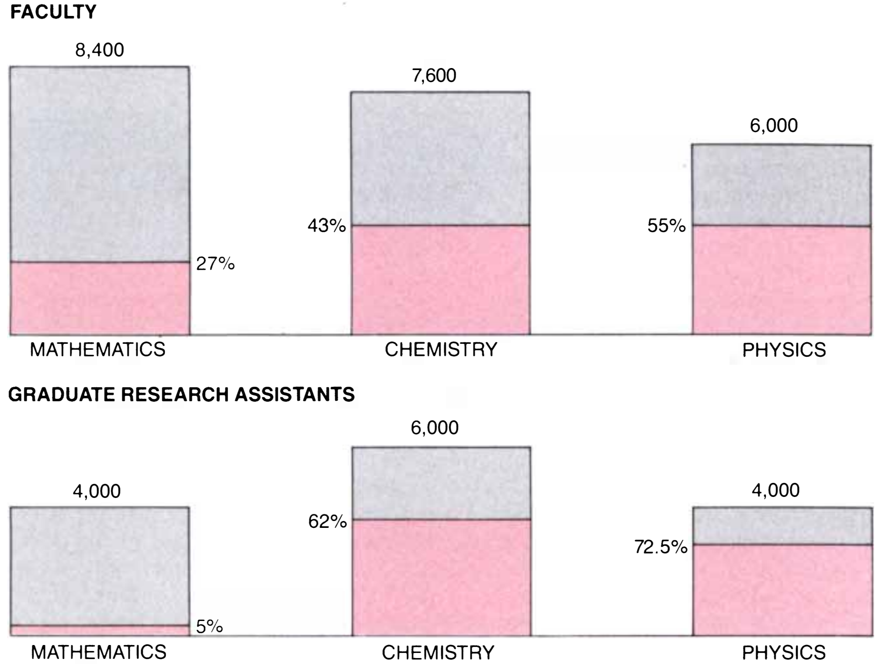
The situation was even worse among graduate students. Although universities conferred roughly as many doctorates in mathematics as they did in physics, the Government supported only 200 graduate research assistants in mathematics in 1980, compared with 2,900 in physics. At the postdoctoral level only 50 students in mathematics were being supported, compared with 1,200 in physics.

Why has mathematics fared so much worse than the other basic sciences? One reason is the time lag between the invention of new mathematical tools and their application. Although the lag has become progressively shorter in recent years, it can still be 20 years or longer. Moreover, applications of fundamental work in mathematics are often made in ways that were unforeseen when the work was initially undertaken. The practice of allocating funds to scientific work on the basis of its foreseeable, short-term applications is unsatisfactory in any basic science, but it is particularly inappropriate in mathematics.

A second reason for the imbalance in Federal funding between mathematics and other basic sciences is that fundamental work in mathematics is much more highly concentrated in the universities than fundamental work in other sciences. Mathematics is therefore greatly weakened by any general cutbacks in the support for academic research.

To compound this effect, mathematicians themselves were concerned at the beginning of the 1970's about an apparent surplus of people with a doctorate in mathematics. Accordingly the mathematical community did not press the NSF for the support of university graduate students and postdoctoral mathematicians, whereas in other fields additional funds were distributed to the universities by the NSF as research grants.

I have noted above that private industry is not likely to take up much of the slack. In recent years, it is true, many companies have increased their support of fundamental mathematics, both in their own laboratories and in the universities. Industry is now the fastest-growing source of funding for research and development in the universities. Nevertheless, in spite of the increasing importance of industrial funding, it is unrealistic to expect industry to supplant government as the primary supporter of fundamental mathematics. The NSF exists in the recognition that only government can



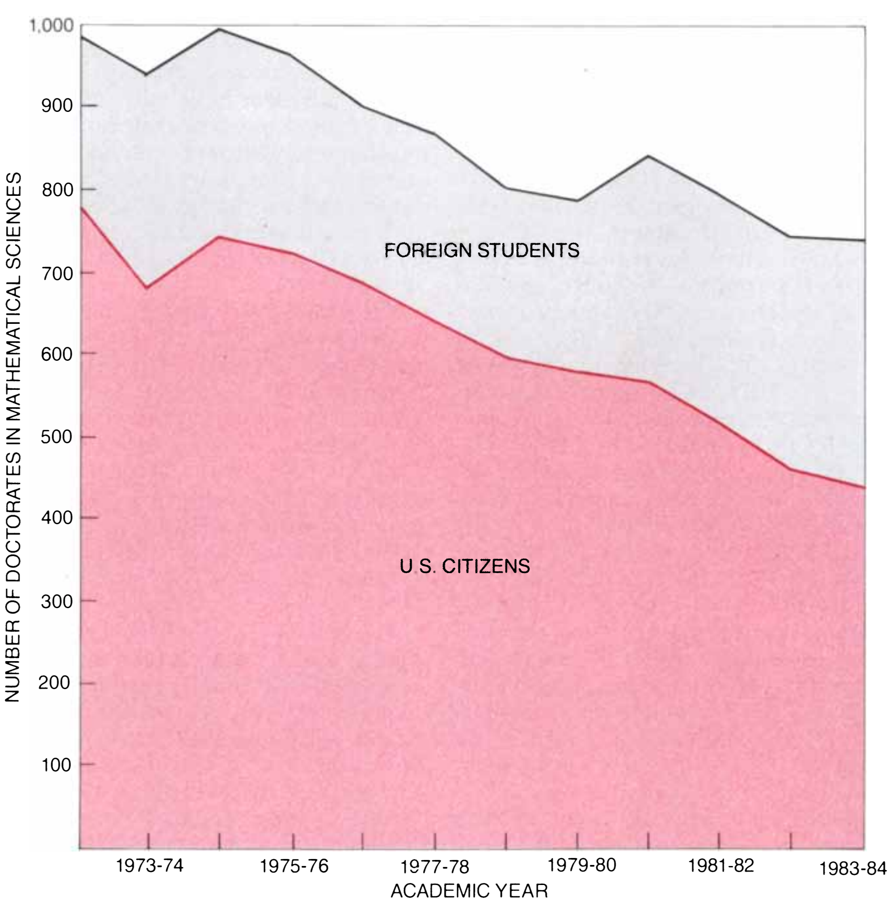
FEDERAL SUPPORT OF MATHEMATICIANS both in senior faculty positions and in graduate schools (*colored parts of bar graphs*) is sharply out of balance with the Federal support of scientists in related disciplines. Compared with the proportion of chemists and physicists being supported, a much smaller fraction of all mathematicians receive Federal money. The data are from a study done by the National Science Board in 1980; it is likely that the proportion of Federally funded mathematicians is now only marginally higher.

take on the major portion of outside investments in basic science and mathematics at the universities.

Indeed, there is now the threat that private industry will effectively weaken the national resource in fundamental mathematics by luring mathematically talented people into relatively high-paid jobs. Such people are flexible and widely employable. In the face of the neglect of mathematics many of them have already opted for other fields, as the decline in the annual output of doctorates makes clear. Students who would have studied mathematics a decade ago now enter computer science or some other more marketable field, and it is not uncommon for them to earn more in their first year on the job than their former professors earn after many years of experience. According to the American Mathematical Society, university mathematics departments are being forced to fill more than half of their teaching positions normally requiring a Ph.D. in mathematics with candidates who lack the degree. If nothing is done, this trend will accelerate dramatically in the 1990's. As the mathematicians trained in the 1950's and 1960's begin to retire, the number of competent university mathematicians may well become critically low.

There is a further and somewhat paradoxical reason for the deterioration of conditions in university mathematics departments. It is the rapid increase in undergraduate enrollment. According to the Conference Board of Mathematical Sciences, the number of students enrolled in mathematics and statistics courses increased by more than 70 percent between 1970 and 1983. Fundamentally this effect is salutary for the mathematics profession, because it shows the strong and growing student interest in mathematics. There is a recognition among students of the increased importance of mathematics in such fields as business and management as well as in the traditional scientific and engineering disciplines.

Yet the large enrollments impose a heavy burden on mathematicians in the universities: from 1970 to 1983 the teaching load for each full-time faculty member in mathematics increased by nearly 30 percent. Even if the present conditions do not cause an exodus of mathematicians from the universities, fundamental mathematics in this country is living on borrowed time. The U.S. is still reaping the harvest of the investment made in mathematics in the 1960's, and that investment is not being renewed.



NUMBER OF DOCTORATES in mathematics conferred each year on American citizens (*colored curve*) has now declined to the point where, without increased Federal support, the field will not be able to renew itself as older mathematicians retire. The number of doctorates in mathematics awarded by U.S. universities to foreign students has been roughly constant for the past 15 years (*gray*). The data are from the American Mathematical Society.

If such stark trends in the mathematics community are to be reversed, the responsible Federal agencies must take part in a sustained program of renewal. But why support mathematics in the first place? Why not allow market forces to hurry the transfer of talent from pure to applied mathematics and computer science?

An important answer is that fundamental mathematics is a cornerstone of our culture. It has long been regarded as one of the noblest activities of the intellect, and certainly as one of the most rigorous and challenging ones. Beyond that, it has contributed profoundly to what Arthur M. Jaffe of Harvard University calls the "ordering of the universe." The abstractions of mathematics have been routinely adopted by scientists to understand the patterns of nature.

One must confess, however, that society has come to expect additional returns from the millions of tax dollars it provides for esoteric mathematical work. In the happy phrase of Eugene P. Wigner of Princeton, it is the "unreasonable effectiveness of mathematics" that captures the popular imagination. New mathematical ideas, often developed with no thought of application, have repeatedly turned out to have immense scientific, technological and economic benefits.

Jaffe notes a striking instance of the unexpected application of mathematical ideas. In 1940 the British mathematician Godfrey Harold Hardy rather haughtily cited his own substantial contributions to number theory as the very antithesis of the utilitarian. Hardy wrote: "I have never done anything 'useful.' No discovery of mine has made, or is likely to make, directly or indirectly, for good or ill, the least difference to the amenity of the world." Forty-five years later, as Jaffe points out, number theory is at the heart of issues in national security: it has become the basis of several new schemes for constructing secret codes.

The practical justification for the support of fundamental mathematics is straightforward: mathematics is extraordinarily cost-effective. In spite of the increased application of computers in mathematics, most mathematical work is still done with pencil and paper or with chalk and blackboard. This simple fact may well contribute to the difficulty mathematicians have had in making their case to the Federal funding agencies. Since their needs in equipment, facilities and technical staff are relatively small, it has often been tempting to postpone them.

Yet consider the scientific concepts, the instruments and the technologies that would not exist without the prior abstractions of mathematics. The general-purpose, programmable computer, for example, was developed by the mathematician John von Neumann and his colleagues, chiefly for the purpose of solving mathematical problems. They were in turn indebted to the mathematical logicians Alan M. Turing, Alonzo Church, Emil Post, Kurt Gödel and others for the framework in which they carried out their investigations. The algorithms, or routine procedures, that are automated on the computer have similar mathematical roots; they can be traced to systematic work in the art of computation and numerical analysis begun by mathematicians such as Newton, Leonhard Euler and Carl Friedrich Gauss.

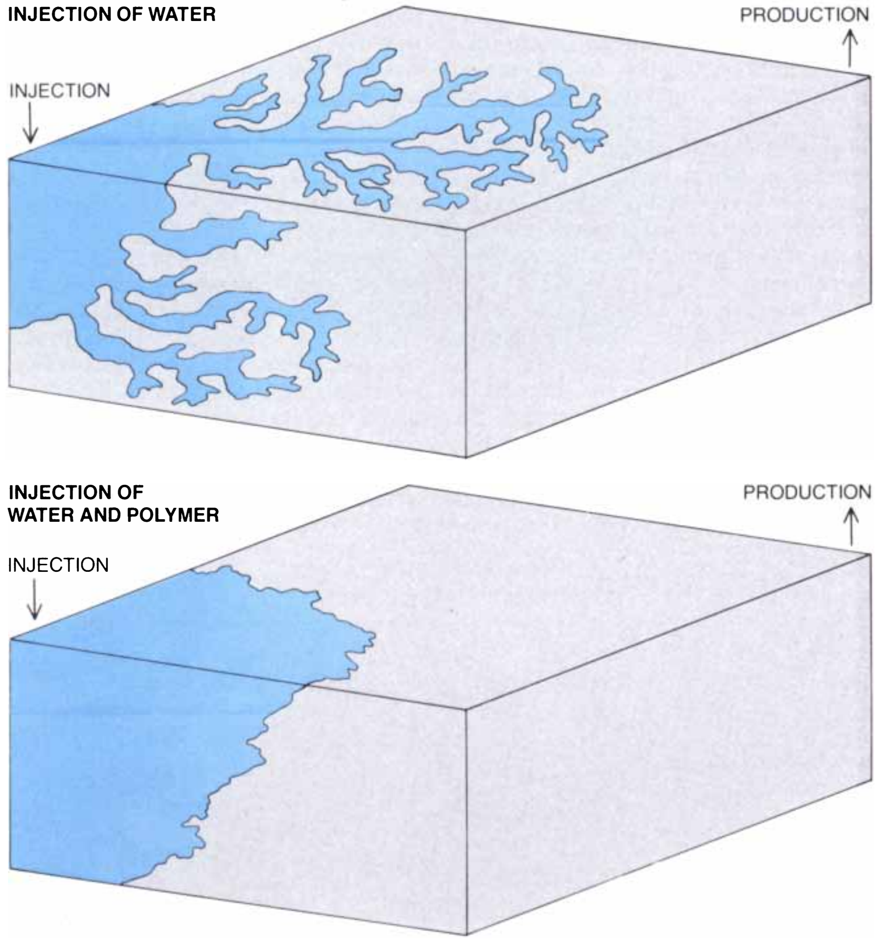
Mathematical analysis continues to be essential to the development of algorithms that take optimum advantage of computational resources. Last year, for example, Narendra Karmarkar, a mathematician at AT&T Bell Laboratories, devised a new algorithm for solving problems in linear programming. In this case the payoff of a mathematical result may be almost immediate. Problems in linear programming, such as the scheduling of machine time and the allocation of resources, arise continually in government and industry. They are currently solved by a fast algorithm called the simplex method. If Karmarkar's new algorithm outperforms the simplex, it could save many millions of dollars in its applications.

One of the most promising current directions in fundamental mathematics is the continued development of the mathematical model. Mathematical modeling of natural phenomena is hardly new. Nevertheless, advances in numerical analysis and the development of the computer have made it possible to simulate processes in ways that are much more complex and more realistic than ever before. Mathematical modeling in partnership with the computer is rapidly becoming a third element of the scientific method, coequal with the more traditional elements of theory and experiment.

It is now possible, for example, to simulate much of the aerodynamics of an airfoil before the airfoil is subjected to expensive testing in a wind tunnel. Mathematical models are employed in the design of a variety of products, including nuclear reactors, automobiles, storage disks for magnetically encoded data, semiconductor chips and pilot plants for the production of fuels and chemicals. Computer-generated models of the atmosphere continue to play a major role in the public debate on such issues as the atmospheric transport of pollutants and the recently recognized threat of a "nuclear winter." Finally, a mathematical simulation of the hypothetical properties of any complex system can guide the investigator in delimiting the range of values for the variables in a theory. The dynamics of stellar interiors, for example, are being explored in this way.

The fit between a mathematical model and the phenomenon it simulates depends, of course, on the degree of distortion introduced by the simplifying assumptions that must be made in the model. There are many complex industrial processes whose control has long been more art than science because realistic assumptions have led to intractable mathematics and mathematically tractable assumptions have led to irrelevant oversimplifications. Recent work in fundamental mathematics is bringing many such processes within the scope of satisfactory models for the first time.

The investigation of nonlinear systems, for example, has enormous potential applications in mathematical modeling. In a linear system a combination of inputs gives rise to a response that is a simple sum of the responses to the individual inputs; in a nonlinear system the response depends on the inputs in much more complicated ways. A detailed physical analysis of a nonlinear system leads in many cases to



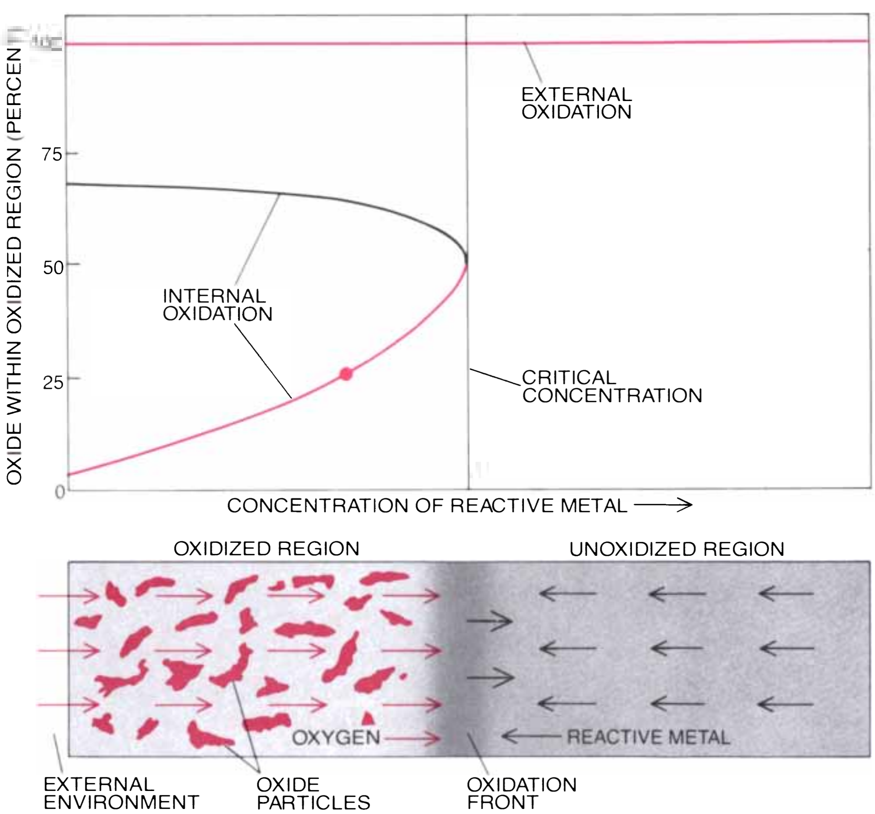
FUNDAMENTAL INVESTIGATIONS now under way into the mathematics of so-called nonlinear systems could lead to better methods for extracting oil from porous rock deep in the earth. Flooding the rock with water injected under pressure is one standard method of recovery; the water pushes some of the oil out of the rock and toward an extraction well. Because the water is less viscous than the oil, the advancing front of water splits up into long fingers that leave behind large volumes of oil (*upper illustration*). It is known that adding a polymer to the water increases the viscosity of the injected fluid. The tendency of the fluid to split into fingers is thereby reduced and a greater fraction of the oil can be recovered (*lower illustration*). A mathematical description that takes account of the interactions of such complex phenomena in a real petroleum reservoir is under development.

nonlinear partial differential equations or nonlinear partial integral equations. The solutions to such equations are mathematical functions that specify the evolution of a particular quantity in time or space, given certain initial conditions. The structure of these solutions is so complex, however, that until recently many nonlinear phenomena were treated mathematically as if they were linear.

Let me give two examples of nonlinear systems that are found in my own industry, the petroleum industry. The first example arises in the attempt to recover the oil trapped in the pores of a rock formation that may lie at a depth of a mile or more. In most such reservoirs only a small fraction of the oil readily flows to the surface. One common method for increasing the fraction of recovered oil has been to pump water at high pressure through the reservoir. The water pushed through the pores of the rock displaces some of the oil and forces it to the surface. Even this method, however, leaves roughly two barrels of oil in the ground for every barrel recovered.

One problem with waterflooding as well as with more sophisticated recovery techniques is a phenomenon called fingering. The boundary between the two fluids becomes unstable because water is less viscous than oil. The front of advancing water splits up into fingers in the rock that bypass volumes of the formation rich in oil. Once the water breaks through the formation to the well for extracting the oil, the recovery of the oil remaining in the rock proceeds much more slowly and eventually becomes uneconomical.

It is known that by adding a polymer to the water the tendency of the injected fluid to split into fingers is substantially reduced. The addition of the polymer makes the resulting fluid more viscous than water. There is no simple relation, however, between the viscosity of the fluid and the fingering pattern. One must also consider factors such as inhomogeneities in the rock and the geometric relations between the injection wells and the production wells. A reasonably detailed mathematical description of all these factors leads to highly nonlinear systems of equations.



MATHEMATICAL MODEL that describes the corrosive oxidation of an alloy gives rise to nonlinear partial differential equations, which have been objects of intense scrutiny by mathematicians. A bifurcation, or branching, diagram (*upper illustration*) summarizes the general properties of the solutions to the equations. In the model it is assumed the alloy is made up of two metals; one metal reacts with oxygen and the other metal does not. For each initial concentration of the reactive metal the diagram shows there is one solution to the equations that predicts an oxidized region in the alloy made up entirely of oxide (*colored horizontal line*). This prediction corresponds to a phenomenon called external oxidation: a coating of oxide is formed on the surface of the alloy, which protects the interior from further oxidation. For each concentration of the reactive metal less than a critical concentration there are two other solutions to the equations. One solution is mathematically unstable and cannot be physically observed (*black curve*); the other solution corresponds to a phenomenon called internal oxidation (*colored curve*). The colored circle on the bifurcation diagram corresponds to the areas of internal oxidation shown in the lower illustration. Oxygen enters the alloy from the external environment and reacts with the reactive metal to form oxide particles (*colored regions*) inside the alloy. The bifurcation diagram shows that alloys whose initial concentration of reactive metal exceeds the critical value must oxidize externally. If the initial concentration is less than the critical concentration, the diagram shows that oxidation can take place in one of two ways: either externally or internally.

My second example arises in the study of corrosion. According to one estimate, corrosion costs the U.S. petroleum industry more than $200 million per year in its refineries alone. Such losses are multiplied many times throughout the rest of the economy because corrosion attacks almost all structures. The need to control corrosion at high temperatures has led to the development of complex metal alloys now used in much refinery equipment. Such an alloy can withstand corrosive environments and high temperatures because a continuous protective layer of metal oxide is formed at the surface of the metal. Corrosive agents such as carbon and oxygen cannot rapidly diffuse through an unbroken layer of metal oxide.

In order to design alloys more resistant to corrosion a key question is how to give a mathematical description of two competing tendencies in the alloy. One is the tendency of the material to form an oxide coating, and the other is the tendency for the corrosive agent to diffuse through the coating and damage the interior. Even for an alloy made up of only two metals, one that oxidizes and one that does not, the mathematical description leads to a set of nonlinear equations that have only recently begun to yield to new mathematical methods.

The analysis of nonlinear systems is one of the most important current topics in fundamental mathematics. For example, there is much interest in solutions to nonlinear equations that oscillate periodically or, for certain initial conditions, become chaotic. Mathematicians are now seeking to classify the kinds of solutions to which these equations can give rise [*see illustration at left*]. The interest arose partly from the effort to understand nonlinear physical systems and partly from questions that have developed within fundamental mathematics.

As nonlinear systems become better understood mathematically, the opportunities for applying the new mathematical ideas will be abundant. Nature typically exhibits nonlinear behavior, and so do the complex physical and chemical systems in industry. I have already mentioned two kinds of nonlinear system that arise in the petroleum industry, and there are many others. Indeed, to realize mathematical descriptions that take account of all the relevant interactions in, say, a real petroleum reservoir, advances in topology, network theory and random processes, as well as in nonlinear differential equations, will be required.

How much support is enough for mathematics? There is no way to derive the "right" levels of support from a set of self-evident axioms, but there are a few principles that seem easy to accept. One principle is that funding should be commensurate with the apparent opportunities for productive work in the field. Such opportunities are in turn dependent on the state of knowledge within the field and the caliber of the people attracted to it. The level of support should enable the research community to renew itself by attracting talented young people. The funding structure should rely on the initiative of the individual investigator and it should stress long-range goals.

The level of support should also reflect the mutual effect that every discipline has on every other discipline. Mathematics and the sciences tend to advance together. A breakthrough in one field can have a tremendous impact on the others. Hence one of the major goals of funding policy should be to estimate the proper balance of support among several fields.

In formulating our recommendations for the support of mathematics we have recognized that in one respect mathematics is a "small" science. Although some branches of mathematics depend on relatively expensive computing time, much of mathematics, unlike physics, chemistry and other "big" sciences, does not require costly instrumentation. Accordingly funding for mathematics need not match these fields dollar for dollar. On the other hand, the needs of people in mathematics are just as great as the needs of people in the other sciences. Mathematicians too require research time, graduate students, postdoctoral assistance, secretarial help and support for travel, conferences, professional journals and reports.

The report of the National Research Council calls on the Federal Government to more than double its current support of fundamental mathematics.

RECOMMENDED FEDERAL SUPPORT IN 1989

		MILLIONS OF CONSTANT 1984 DOLLARS
FACULTY RESEARCH GRANTS	2,600 @ $31,500	81.9
GRANTS FOR YOUNG INVESTIGATORS (3-5 YEARS AFTER Ph.D.)	400 @ $25,000	10.0
POSTDOCTORAL FELLOWSHIPS (TWO-YEAR GRANTS)	200 @ $90,000	18.0
VISITING SCHOLARS	130 @ $90,000	11.7
GRADUATE STUDENT STIPENDS AND TUITIONS	1,000 @ $30,000	30.0
CONFERENCES AND RESEARCH INSTITUTES		11.0
COMPUTATION INITIATIVE AND EQUIPMENT		17.5
	TOTAL	180.1

RECOMMENDATIONS FOR RENEWED SUPPORT of mathematics call for a real increase in Federal funding of 18 percent per year in constant dollars over a period of five years. If the recommendations are adopted by the Government, the cost of fundamental mathematics in 1989 will be slightly more than $180 million in constant 1984 dollars. The money is to be spent according to the breakdown shown in the table. The recommendations were presented last year to the National Research Council by the Ad Hoc Committee on Resources for the Mathematical Sciences, a group that included many nonmathematicians.

When the report was first drafted in 1984, the total Federal support of mathematics was estimated to have been $78.2 million. We recommended that the level be increased by 18 percent per year in constant dollars for five years, which would bring the annual budget to about $180 million in 1984 dollars for fiscal year 1989.

These recommendations are actually quite modest. They are based on a careful analysis of the needs of mathematics in eight general categories. Thus our report urges that by 1989 funding be provided for 1,000 graduate students, 400 postdoctoral students and 2,600 established investigators, as well as for a new initiative in the mathematics of computation. Such numbers are still smaller than their counterparts in physics and chemistry: in physics there were 1,200 postdoctoral students who had Federal support in 1980; in chemistry there were 2,500. Moreover, the goal of funding 2,600 established investigators represents the support of only 40 percent of the field. In physics the corresponding percentage is now 70 percent; in chemistry it is 50 percent.

If our program is put into practice, it will initially arrest the long decline in the number of doctorates awarded in mathematics to American citizens. In the longer term it will also increase their number to a level that will balance the needs of higher education and the national research enterprise in the 1990's and beyond.

Compared with the $7 billion the Federal Government spent on basic research in fiscal year 1984, let alone the more than $50 billion it spent on all research and development, fundamental mathematics is a bargain at the $180 million funding level we have recommended. I have already noted that the NSF has recently taken the lead in increasing the Federal support. The agency's budget for mathematical sciences increased by more than 15 percent per year in current dollars in fiscal years 1984 and 1985, and a similar increase is proposed for fiscal year 1986. The National Science Board, which is the governing body of the NSF, adopted a formal resolution in December of last year urging that all Federal agencies make a concerted effort to bring their support of mathematics back to the proper levels.

Nevertheless, Washington's current preoccupation with the Federal deficit forces new programs of any kind to wage an uphill battle. Bureaucratic inertia in the relevant Federal agencies also hinders effective action. The budget for fiscal year 1985, although increased, was nowhere near our proposed goal of $92 million in 1984 dollars; the proposed budget for fiscal year 1986, although increased again, continues the pattern of undernourishment in mathematics.

What is to be done? Mathematicians, as well as people in science, industry and government whose work depends on mathematics, must continue their efforts to present the case for fundamental mathematics to those who control the budget-making process. Restoring the health of mathematics is essential to the nation's future.

Molecular Approaches to Malaria Vaccines

Study of genes encoding the molecules of the malaria parasite's outer coat reveals a class of proteins forming repeated antigenic sites. They may serve as decoys deflecting the immune response

by G. Nigel Godson

For a brief period in the early 1960's it appeared that malaria, an ancient scourge, might soon be brought under control. Extensive spraying with DDT was reducing the *Anopheles* mosquito population and novel drugs such as chloroquine were available for treating infected patients.

Twenty years later malaria is resurgent. Its causative agent, the protozoan parasite *Plasmodium,* is developing resistance to the drugs, and the parasite's vector, the female *Anopheles* mosquito, is becoming resistant to DDT and other insecticides. Today the disease afflicts some 200 to 400 million people in a broad tropical band around the world. In Africa it kills some 10 percent of its victims directly and debilitates the rest; it is a major cause of early-childhood mortality rates ranging as high as 50 percent. Clearly there is urgent need for the new attack on malaria that is currently under way. The major effort is to exploit the tools of molecular biology to develop antimalaria vaccines and other ways to combat the parasite.

An effective vaccine must stimulate the immune system to make antibodies that can attack and neutralize the parasite. To develop such a vaccine will not be easy. In regions of Africa where the disease is endemic a large proportion of the population is either chronically infected with one or another species of the parasite or is continually reinfected by the ever present mosquito. These people develop antibodies aplenty to the parasite, but few of them develop protective immunity. The reason is that *Plasmodium,* even during the brief periods when it is not hidden from the immune system of its human host (in liver cells or red blood cells), is adept at evading the immune response. Studies of the molecular biology of *Plasmodium* have begun to reveal how it does so and to suggest new approaches whereby the parasite's escape mechanisms may be circumvented. One can hope that in the not too distant future genetically engineered vaccines or other molecular weapons may be developed and that eventually malaria will be eradicated.

Two species of *Plasmodium* are important agents of human malaria: *P. falciparum* (the most prevalent and most lethal) and *P. vivax.* In the course of its life cycle in its mosquito and human hosts the unicellular parasite undergoes an astounding series of developmental and morphological changes. The stage that infects man, the lance-shaped sporozoite, resides in the mosquito's salivary gland and is delivered into the victim's bloodstream when the insect takes a blood meal. Within an hour each sporozoite finds its way to a liver cell. There it undergoes a complex series of transformations. Eventually a giant multinucleate stage, the schizont, fissions into small, roughly spherical merozoites. The result is an enormous amplification of parasites: a liver cell infected by one sporozoite releases into the bloodstream from 5,000 to 10,000 merozoites.

Each merozoite invades a red blood cell, where it multiplies asexually until the cell bursts and releases from 10 to 20 new merozoites that go on to invade more red cells. It is the periodic lysis of the blood cells, with concomitant release of merozoites and toxic waste products, that causes the regular fevers and chills of malaria.

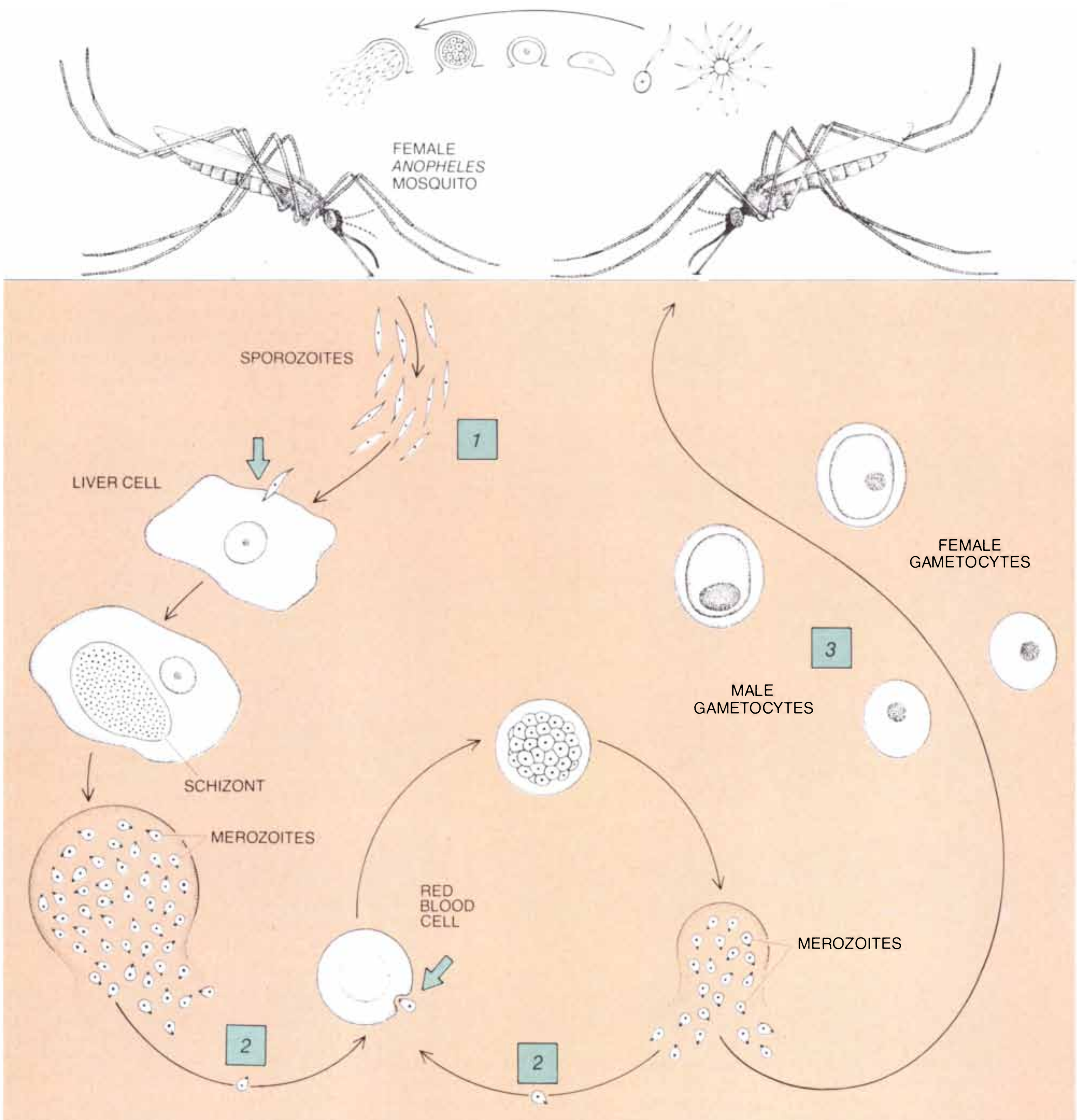
Some merozoites develop into male and female gametocytes (germ-cell precursors), thus initiating the parasite's sexual cycle. The gametocytes are sucked up with red cells by a mosquito, mature in the mosquito gut and fuse to form a zygote. The zygote undergoes yet another series of divisions, transformations and migrations; eventually a mature sporozoite appears in the salivary gland, ready to initiate a new infective cycle.

Each developmental stage of *Plasmodium* has its characteristic shape and distinctive set of functions; it inhabits a particular microenvironment and interacts with a specific target tissue. To the molecular biologist this means that although all the stages have the same genome, or complement of genes, in each stage a different part of the genome is being expressed: different genes are turned on and off in a programmed sequence.

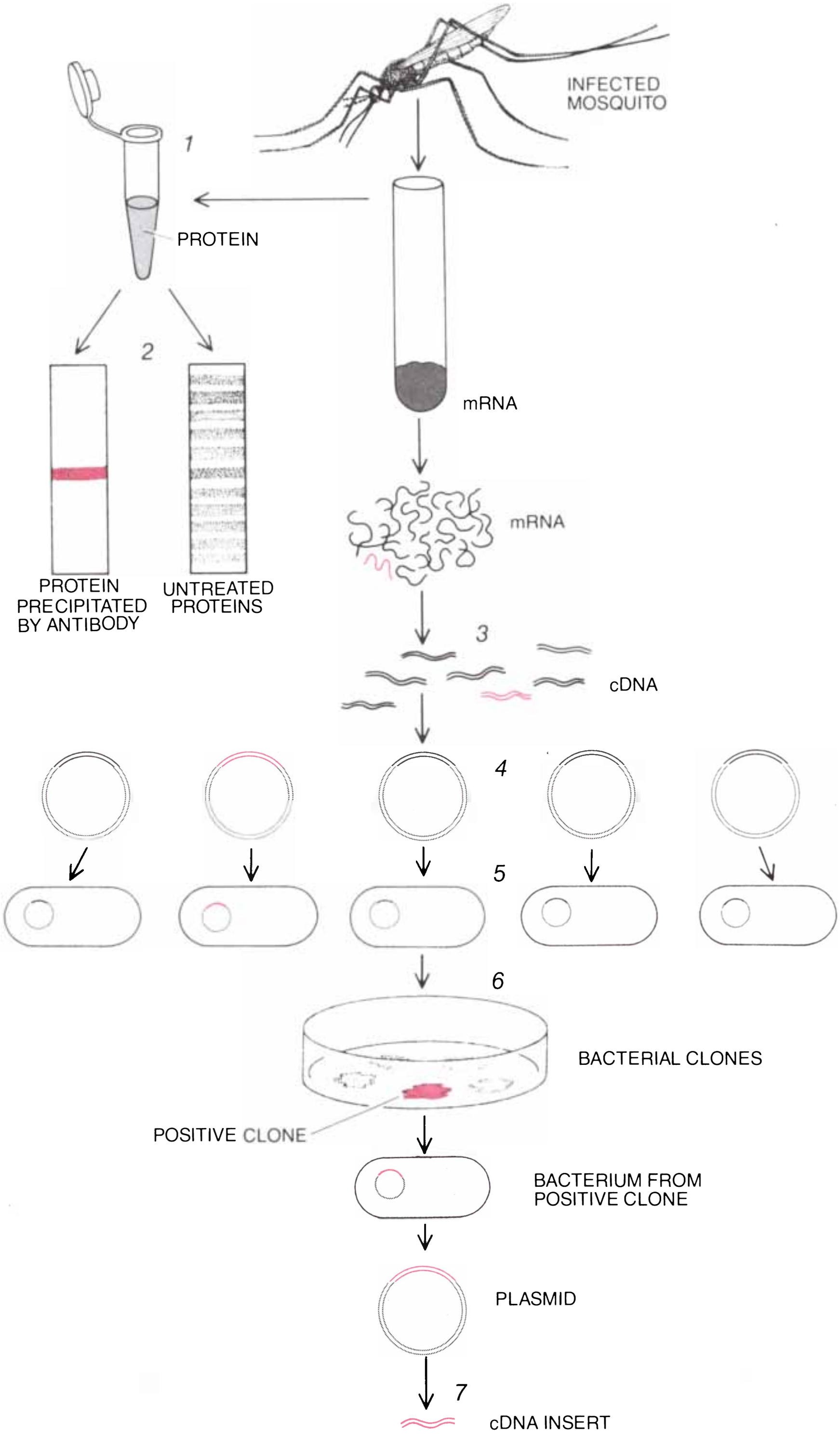
A gene is composed of DNA, a double helix whose two complementary strands are made up of subunits called nucleotides. Each nucleotide is characterized by one of four bases: adenine (*A*), guanine (*G*), thymine (*T*) and cytosine (*C*). Genetic information is encoded in the sequence of the bases. A gene is expressed when one strand of its DNA is transcribed into a complementary strand of messenger RNA (mRNA), which is then translated into a sequence of amino acids, the subunits of proteins.

One way to understand a developing organism at the molecular level is to isolate the genes being expressed at a particular stage of development and study their structure and that of the proteins they encode. In the case of *Plasmodium* such studies have been focused on the parasite's surface. One reason is that the proteins of the cell's outer coat are highly stage-specific: each is expressed in only a single developmental stage. Their genes must therefore be subject to stringent regulation, whose mechanisms are of considerable fundamental interest. The other reason is that these proteins are surface antigens and as such are likely to be implicated in triggering (or in evading) the host's immune response. Studying their genes is therefore important not only for understanding the mechanism of stage-specific gene expression but also for developing stage-specific malaria vaccines.

Some years ago my colleagues and I at the New York University Medical Center set out to isolate and study the gene encoding the major surface antigen of a sporozoite, the so-called circumsporozoite (*CS*) protein. The protein had been studied for many years by Ruth S. Nussenzweig of N.Y.U. and had been shown to be stage-specific: synthesized only in spo-



MALARIA PARASITE *Plasmodium* goes through a number of stages during its life cycle in its vector, a female *Anopheles* mosquito, and in a mammalian host. A sporozoite injected by the mosquito soon invades a liver cell, where it is transformed into a giant, multinucleate schizont. The schizont fissions and the liver cell releases many thousands of merozoites. Each merozoite invades a red blood cell and multiplies; the cell bursts, releasing from 10 to 20 merozoites that invade more blood cells. Some merozoites become male and female gametocytes (germ-cell precursors), which are taken up with a blood meal by a mosquito. After a series of further transformations mature sporozoites appear in the mosquito's salivary gland, where they are available to repeat the infective cycle. A vaccine might be designed to elicit antibodies that attack either sporozoites, merozoites or gametocytes when they are free in the bloodstream (*1, 2, 3*). Agents might also be designed to block the invasion of liver or blood cells (*colored arrows*) or to kill parasites within a cell.



SEARCH FOR GENE encoding the circumsporozoite (*CS*) protein of the monkey-malaria parasite *Plasmodium knowlesi* began with extraction of messenger RNA (mRNA) from parasite-infected mosquitoes. To detect *CS*-protein messenger some of this mixed mRNA was first translated into radioactively labeled protein (*1*). Half of the protein was subjected to precipitation with a monoclonal antibody to the *CS* protein and half was left untreated. The proteins were separated by gel electrophoresis (*2*). The presence of a band of protein that had been precipitated by the antibody (*color*) gave assurance that the *CS*-protein mRNA was present. The total mRNA was reverse-transcribed (*3*) into a DNA copy (cDNA), which was inserted into plasmids (*4*). The recombinant plasmids were introduced into bacterial cells (*5*) and the bacteria were grown (*6*). The resulting clones were tested (*see illustration on page 56*) with the antibody to the *CS* protein. Some clones were found to have expressed the protein (*color*). Plasmids in the positive clones were isolated and examined. Three of them were found to carry parasite-cDNA inserts coding for part of the sporozoite protein (*7*).

rozoites. It is the major protein synthesized by sporozoites in the salivary gland and it covers the entire surface of the cell. We chose to work with *P. knowlesi,* the agent of monkey malaria, largely because the *Anopheles* species that carries it generates some 10 times as many sporozoites as a mosquito infected with one of the human parasites. Infected mosquitoes were supplied by Robert W. Gwadz and Louis H. Miller of the National Institute of Allergy and Infectious Diseases (NIAID), who provide many investigators worldwide with malaria-parasite material.

To isolate an active stage-specific gene one ordinarily begins with the total mRNA of the stage under study, and so we tested several laborious methods for separating sporozoite material from infected mosquitoes. Eventually we found that instead of having to purify sporozoites we could begin with the total mRNA of infected mosquitoes (or of their thoraxes). In the mixture of sporozoite and mosquito mRNA we could detect the specific mRNA encoding the *CS* protein, and so we could clone the parasite gene directly from the total mRNA.

The total mRNA was converted with the enzyme reverse transcriptase into a DNA copy (cDNA). The cDNA fragments were inserted into plasmids (small circles of bacterial DNA) in the middle of a gene coding for a plasmid protein. A recombinant plasmid incorporating the parasite cDNA should therefore express a fusion product, part plasmid protein and part parasite protein.

Recombinant plasmids were introduced into the bacterium *Escherichia coli.* The bacteria were grown and the resulting clones (colonies descended from a single cell) were screened with the monoclonal antibody to the *CS* protein by means of a two-site immunological assay developed by Fidel P. Zvala of N.Y.U. Joan Ellis, a student in my laboratory, found three clones to which antibody bound, showing that these bacteria had synthesized an active fusion protein.

When the plasmids in the positive *E. coli* clones were analyzed, we found that the fragment of sporozoite cDNA inserted in one of them was extremely short: only 340 base pairs, or long enough to encode only about 110 amino acids (since each amino acid is specified by a codon of three nucleotides). This was a serendipitous finding with remarkable ramifications. It meant that this small fragment of sporozoite cDNA must include the region of the gene coding for the immunoreactive part of the *CS* protein: the epitope, or antibody-combining site.

To locate the epitope-encoding region of the small cDNA insert more precisely we turned to transposon mutagenesis. This mapping technique depends on bacterial transposons: bits of DNA, often encoding a gene for antibiotic resistance, that can jump from one plasmid to another almost at random. A transposon inactivates gene function beyond the point at which it is inserted, so that by mapping the insertion sites that result in deactivation one can delimit functional regions of genes. By this means James R. Lupski, another student in the laboratory, was able to show that the antigen-combining site is encoded within a segment some 110 base pairs long at the extreme left-hand end (what is called the 5′ end) of the 340-base-pair insert.

Pamela Svec then determined the nucleotide sequence of the 340-base-pair insert. To our astonishment the entire insert turned out to consist of tandem repetitions of a single sequence 36 base pairs long; there were seven complete repeat units, with incomplete units at each end. When the other two clones to which the *CS*-protein antibody had bound were examined (the ones in which the sporozoite cDNA inserts were larger), we found those inserts also incorporated multiples of the 36-base-pair unit. Since the repeat unit was common to all three positive clones, it seemed clear that it must code for the *CS* protein's epitope. The epitope itself must be a chain of 12 amino acids repeated in tandem—an entirely new and remarkable structure for a surface antigen.

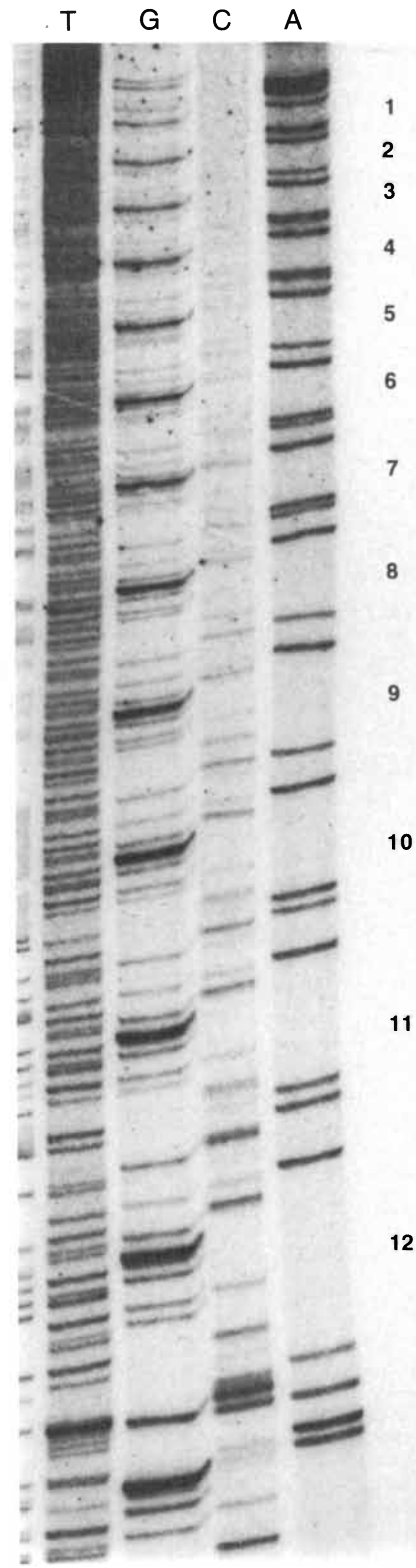
We still did not know the amino acid sequence of the epitope. We knew the sequence of nucleotides but could not translate it into a sequence of amino acids because we did not know the reading frame: the way in which the string of nucleotides should be divided into codons specifying amino acids. Because a codon is a triplet of nucleotides, there are three potential reading frames in each strand of DNA, and so there are six possible reading frames in the double helix.

We established which was the coding strand and deduced the reading frame by finding the junction between the *Plasmodium* DNA and the known nucleotide sequence of the plasmid-protein gene. Knowing the reading frame, we could translate the nucleotide sequence to derive the 12-amino-acid sequence of the epitope. The logical next step was to assemble a synthetic peptide (a short protein chain) corresponding to the derived sequence and see if it could mimic the immune properties of the natural sporozoite surface protein.

David H. Schlesinger of N.Y.U. assembled amino acids to make both the 12-amino-acid epitope and a double-unit peptide 24 amino acids long. The immunological assay showed that the double-unit synthetic peptide did bind to the monoclonal antibody against the *CS* protein; in a competitive assay the single-unit synthetic peptide not only bound to the antibody but also in doing so blocked the antibody's normal binding to the sporozoite surface [*see illustration on page 56*]. These results established conclusively that the 12-amino-acid peptide did either constitute the epitope of the *P. knowlesi* surface antigen or include it. The latter proved to be the case: in further experiments done with Victor Nussenzweig and Schlesinger we showed that synthetic peptides consisting of only eight of the 12 contiguous amino acids carry all the information needed for a complete antigen-antibody interaction.

At this point several advances had been made. The sequence of the epitope had been determined; the reading frame of the entire *CS*-protein gene had been established, and synthetic peptides had been constructed that mimicked the protein's immunoreactive region. When the peptides were injected into mice and rabbits, they were highly immunogenic, that is, they stimulated the formation of antibodies. Whether the antibodies would establish protective immunity (in which case the peptides could be the basis of an antimalaria vaccine) remained to be determined by testing in animals and by further studies of the molecular biology of the surface antigen.

The next objectives were therefore to establish the structure of the complete *CS*-protein gene (not a cDNA construct but the actual gene found in the parasite chromosome) and to deduce the complete amino acid sequence of the protein and then its structure. Luiz S. Ozaki fragmented the parasite's genomic DNA, separated the fragments according to size by electrophoresis and transferred them to a nitrocellulose filter. He probed the filter with one of our plasmid cDNA's. The cDNA probe hybridized with



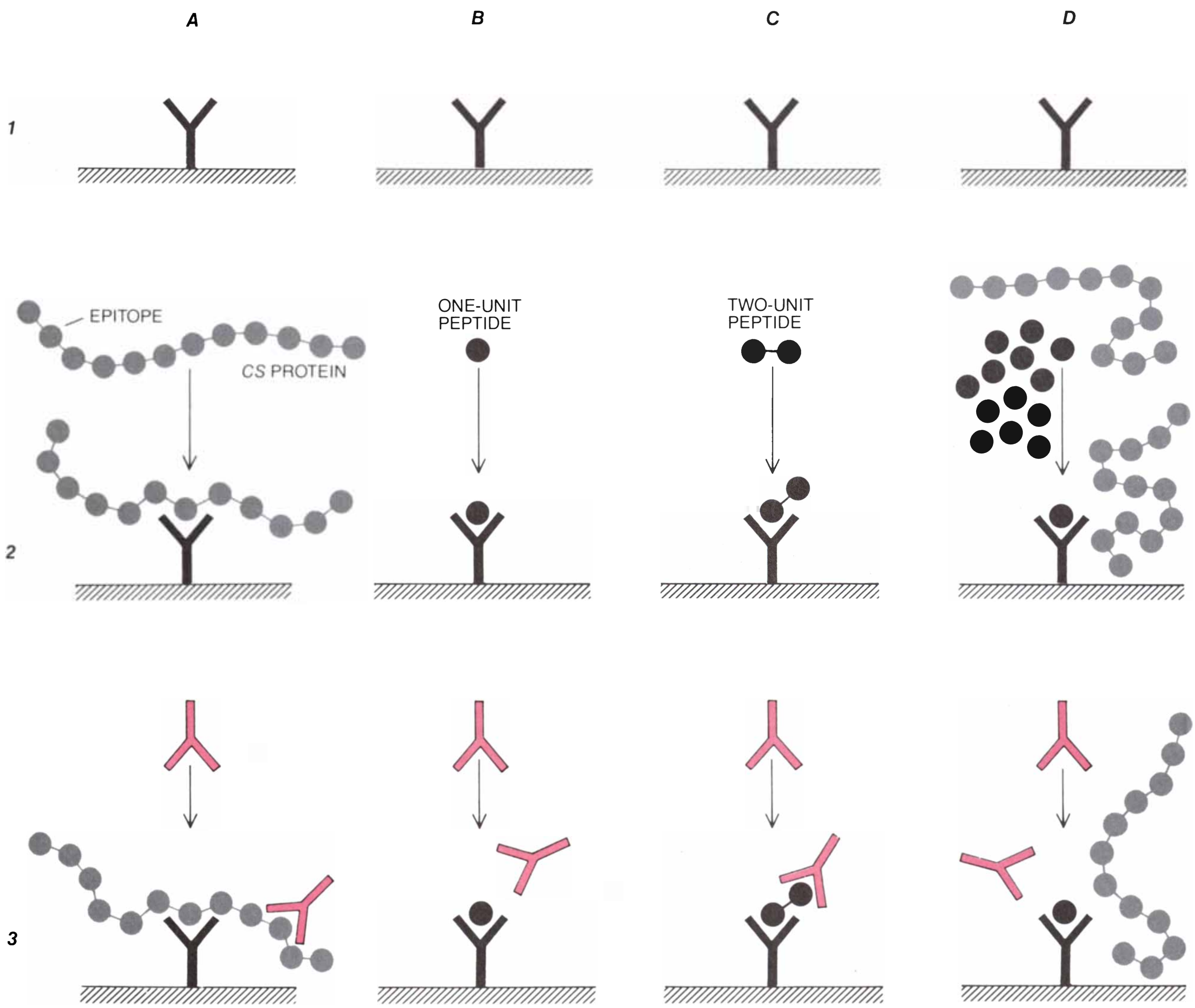
REPEATED PEPTIDES were noted when the nucleotide sequence of a cDNA insert encoding the *CS* protein was being determined. The four columns of the autoradiograph represent occurrences of the four subunits (*T, G, C, A*) of which DNA is composed. Even before the sequence was read off it was evident from the reiterated pattern that this part of the gene encodes 12 tandem repetitions of a short peptide. The peptide is an epitope, or antigenic determinant: a specific site on the protein that binds to an antibody. The region of the protein encoded here is therefore a multiple epitope.

(bound to) a complementary stretch of DNA contained within one segment, 11,000 nucleotides long, of the parasite genome. The gene for the *CS* protein was thereby shown to be within that segment.

After isolating the 11-kilobase fragment (by cloning) and establishing its overall structure (by restriction-enzyme mapping), we determined the full nucleotide sequence of the region containing the surface-antigen gene. The mapping and sequence data led us to two conclusions. One was that, in spite of the vast amplification of *CS*-protein expression known to take place during the sporozoite stage, the parasite genome contains only one gene for the protein. The other conclusion was that the coding region of the gene, unlike most genes in eukaryotes (organisms higher than bacteria), is not interrupted by introns: noncoding intervening sequences that are removed only when the mRNA is processed. The lack of introns suggests that *Plasmodium* may be a very primitive eukaryote.

With the reading frame established and the full nucleotide sequence of the *CS*-protein gene in hand we could deduce the amino acid structure of the complete protein. The central 45 percent or so of the protein turned out to consist of 12 repeats of the 12-amino-acid epitope unit. In some respects other parts of the protein were what might be anticipated in a surface protein. The extreme left end (the NH_2 terminus) has the strongly hydrophobic (water-repellent) signal sequence characteristic of proteins exported through the outer membrane of a cell. The other end (the COOH terminus) is a hydrophobic tail (which typically anchors a surface protein in the cell membrane); it is preceded by four amino acids (cysteines) between which disulfide bonds can form to link segments of the protein into a globular structure or to link adjacent molecules of the protein.

The protein's molecular weight as calculated from the amino acid sequence is, however, substantially less than the molecular weight measured



MONOCLONAL ANTIBODY to the *CS* protein serves to detect clones expressing the protein (*A*) and to test synthetic peptides and so verify the deduced amino acid sequence of the epitope (*B–D*). Unlabeled antibody is adsorbed on a solid surface (*1*). A positive clone is detected (*A*) when an epitope on the *CS* protein it expresses binds to the antibody (*2*); binding of the protein is demonstrated by the binding of a labeled copy of the same antibody to a second epitope on the protein (*3*). When a single synthetic peptide (12 amino acids) binds to the adsorbed antibody (*B*), it offers no second site for attachment of the labeled antibody. The double peptide (24 amino acids) does provide a second site, however (*C*), and so the labeled antibody can bind to it. In a competitive assay (*D*) the single peptide at high concentration saturates the combining sites of adsorbed antibodies and thus blocks the binding of the true *CS* protein. This proves that the synthetic peptide and the parasite protein compete for the same antibody: they must have identical antibody-combining sites.

by the protein's rate of migration in a gel. This finding and others suggested that the *CS* protein must have peculiar physical properties. Some of these properties could be deduced from the amino acid sequence of the repeat unit. The peptide has three glycines, three alanines, three glutamines and an aspartic acid, a proline and an asparagine. They are arrayed in two dimensions in such a way that small polar (hydrophilic) amino acids alternate with large hydrophobic ones. Such alternation is also characteristic in a six-amino-acid repeating subunit of fibroin, the major protein of silk. The fibroin chain doubles back and forth on itself, with successive antiparallel segments linked by hydrogen bonds. This configuration, a beta pleated sheet, gives silk its fibrous, flexible nature.

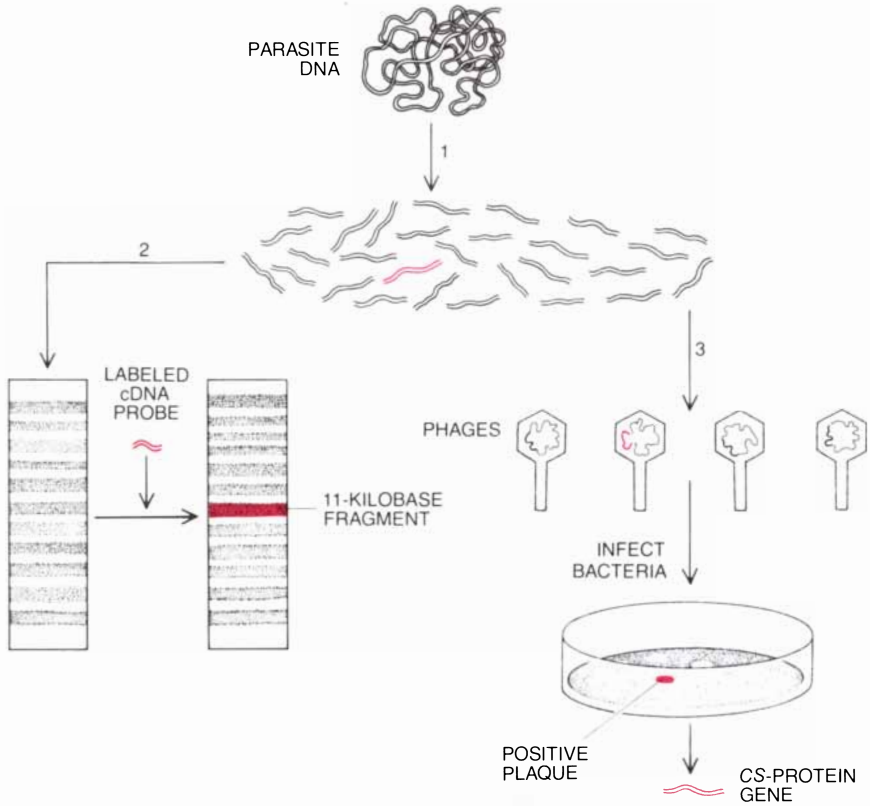
By building molecular models we showed that the *CS* protein's repeating peptides could form a similar pleated sheet. The peptide chain would bend at each proline (where protein chains often turn sharply), so that succeeding units of the repeat zigzag in opposite directions [*see top illustration on page 59*]. Hydrogen bonds form naturally; the sizes of the amino acids' side chains alternate in such a way that abutting side chains do not interfere with each other. Such a beta sheet should be very stable. In collaboration with Schlesinger and Walter A. Gibbons of the University of London, we have recently confirmed some of the predictions of this proposed structure experimentally. When peptides corresponding to two, three and four tandem repeats (24, 36 and 48 amino acids) are synthesized, they do show considerable beta structure, strongly suggesting that this is the case in the natural protein.

If the repeats of a single *CS*-protein molecule can zigzag to form a sheet, adjacent molecules should be able to interact in a similar way (as they do in fibroin). This would make for a surface structure in which molecules fold on themselves and interact with one another to form a network: an ideal protective surface for the parasite. It is more than a physical barrier. The *CS* protein appears to promote the sporozoite's evasion of the host's defenses primarily by "focusing" the immune response, concentrating it on a single target to the detriment of its ability to find other targets. Several lines of evidence point to such a strategy.

When an experimental animal is injected with sporozoites, most of the antibodies it raises are directed against the *CS* protein, and specifically against the repeating epitope; there is little indication that the immune system recognizes any other part of the surface protein. This suggests that the protein chain is folded so that only the 45 percent of the chain carrying the repeating epitope is exposed on the surface. Each molecule of the protein therefore presents to the immune system only one vulnerable site—but that site is repeated 12 times. The immune system sees multiple identical targets on the same molecule; any other potential targets are relatively inaccessible.

The repeating epitope, then, is essentially a multiple decoy. It is also a renewable decoy. Many years ago Jerome P. Vanderberg of N.Y.U. noted that sporozoites exposed to antisporozoite antibodies appeared to slough off a discrete surface coat. The coat is visible in electron micrographs as a thick, fuzzy layer surrounding the cell; presumably it is the network of *CS*-protein molecules. There is evidence that the sporozoite's coat is sloughed off continually and is continually restored by newly synthesized protein secreted to the surface. It is notable that the *CS* protein is manufactured in large amounts, accounting for from 10 to 30 percent of the total protein synthesized by the sporozoite in the mosquito salivary gland. The immune system, then, is presumably forced to mount a larger than normal attack both because there are multiple epitopes and because the surface coat is continually replaced. Moreover, the sloughed-off molecules, particularly if they are present as a network, must act as a further decoy that tricks the immune system into recognizing them as live parasites and thus mops up still more antibodies.

All of this may give sporozoites injected by a mosquito time to reach shelter in liver cells even if circulating antisporozoite antibodies are already present in the host because of previous infection. Such an escape mechanism would be particularly appropriate for a parasite stage such as the sporozoite, which is exposed to the immune system only during a short interval. On the other hand, a parasite exposed to antibodies for a long time, such as the trypanosome, may need to keep changing its surface antigens so that the immune system cannot catch up



ENTIRE *CS*-PROTEIN GENE was isolated. The sporozoite genome (total DNA) was digested with an enzyme (*1*). The fragments were separated by size on a gel (*2*) and a labeled *CS*-protein cDNA insert was used as a probe to show that the gene was within a particular 11-kilobase fragment (*color*). That fragment was isolated by cloning (*3*). The DNA fragments were introduced into a bacterial virus, phage lambda. The phages infected bacteria cultured in a petri dish. Each phage multiplied, leaving a discrete plaque in the lawn of bacteria. With the labeled cDNA insert again serving as a probe, plaques containing the 11-kilobase fragment were identified. The *CS*-protein gene was isolated from a positive plaque.

with it [see "How the Trypanosome Changes Its Coat," by John E. Donelson and Mervyn J. Turner; SCIENTIFIC AMERICAN, February]. Such antigenic variation takes time, however. An immune-decoy protein, on the other hand, can protect a parasite from the moment it appears in the host.

Soon after we described the surface antigen of the *P. knowlesi* sporozoite a remarkably similar one was isolated from a different stage (the blood-stage merozoite) of the human malaria parasite *P. falciparum*. Working with mRNA of the blood-form parasites and the serum of people repeatedly exposed to the parasite, David J. Kemp, Robin F. Anders and Graham F. Mitchell and their colleagues at the Walter and Eliza Hall Institute of Medical Research in Australia were able to clone and express a number of merozoite surface-antigen genes.

One of them was the gene for a merozoite *S* antigen, a surface protein that forms a fuzzy layer covering the merozoite as it emerges from a lysed red blood cell. As in our *CS*-protein gene, there is a multiple repeat. In the *S*-antigen gene the repetitive sequence is 33 base pairs long (so that it codes for a repeating peptide with 11 amino acids rather than 12) and the sequence is repeated in tandem more than 100 times (rather than 12 times). Like the *CS* protein, the *S* antigen seems to be an immune-decoy protein that is continually secreted and shed and presents a repeated epitope to the immune system.

Since the *P. knowlesi CS* protein and the *P. falciparum S* antigen were characterized, five more *Plasmodium* surface proteins have been isolated. The work has been done by Kemp and his colleagues in Australia, by Miller's group at NIAID in collaboration with a group at the Walter Reed Army Institute of Research, by Jeffrey V. Ravetch of the Memorial-Sloan Kettering Cancer Center and Gunter Blobel of Rockefeller University and by Luis Pereira da Silva of the Pasteur Institute and Benno Müller-Hill of the University of Cologne. Every one of these proteins has a tandemly repeated peptide unit [*see bottom illustration on opposite page*]. Most of the repeating peptides include a proline, raising the possibility that zigzagging at a proline may give rise to a beta pleated sheet in all plasmodial surface antigens.

It seems clear that many of the major surface proteins of both sporozoites and merozoites are immune decoys and that the antibodies they induce are not protective ones: they do not incapacitate the parasite. That probably explains why natural infection rarely results in protective immunity. It also leads one to conclude that vaccines designed to stimulate antibodies against these major surface antigens are probably not the best candidates for inducing lasting immunity.

There may nonetheless be effective ways to attack the parasite. Judging by the multiplicity of antiplasmodial antibodies found in the blood of people with malaria, many more stage-specific surface proteins remain to be isolated and characterized. Most of them are probably directed against merozoites or against the surface of infected red blood cells (to which an infecting merozoite somehow exports some of its antigens), but some may be minor constituents of the sporozoite or gametocyte surface; unlike the major antigens, they are presumably not immune decoys. Many laboratories are engaged in an effort to isolate the genes encoding these minor surface antigens by exploiting cloning procedures such

243
...CTCCACATACTTATATACAAGAACAAGATGAAGAACTTCATTCTCTTGGCCGTCTCC
Met Lys Asn Phe Ile Leu Leu Ala Val Ser

300
TCCATCCTGCTGGTGGACTTGCTCCCCACACACTTCGAACATAATGTAGATCTCTCCAGG
Ser Ile Leu Leu Val Asp Leu Leu Pro Thr His Phe Glu His Asn Val Asp Leu Ser Arg

360
GCCATAAATGTAAATGGAGTAAGCTTCAATAATGTAGACACCAGTTCACTTGGCGCACAG
Ala Ile Asn Val Asn Gly Val Ser Phe Asn Asn Val Asp Thr Ser Ser Leu Gly Ala Gln

420
CAGGTGAGACAAAGTGCTAGCCGAGGCAGAGGACTTGGTGAGAAGCCAAAAGAAGGAGCT
Gln Val Arg Gln Ser Ala Ser Arg Gly Arg Gly Leu Gly Glu Lys Pro Lys Glu Gly Ala

480
GATAAAGAAAAGAAAAAAGAAAAAGGAAAAGAAAAAGAAGAAGAACCAAAGAAGCCAAAT
Asp Lys Glu Lys Lys Lys Glu Lys Gly Lys Glu Lys Glu Glu Glu Pro Lys Lys Pro Asn

540 1
GAAAATAAGCTGAAACAACCGAATGAAGGACAACCACAAGCACAGGGTGATGGAGCAAAT
Glu Asn Lys Leu Lys Gln Pro Asn Glu Gly Gln Pro Gln Ala Gln Gly Asp Gly Ala Asn

600 2 3
GCAGGACAACCACAAGCACAAGGAGATGGAGCAAATGCAGGACAACCACAAGCACAGGGT
Ala Gly Gln Pro Gln Ala Gln Gly Asp Gly Ala Asn Ala Gly Gln Pro Gln Ala Gln Gly

660 4
GATGGAGCAAATGCAGGACAACCACAAGCACAGGGTGATGGAGCAAATGCAGGACAACCA
Asp Gly Ala Asn Ala Gly Gln Pro Gln Ala Gln Gly Asp Gly Ala Asn Ala Gly Gln Pro

720 5 6
CAAGCACAAGGAGATGGAGCAAATGCAGGACAACCACAAGCACAGGGTGATGGAGCAAAT
Gln Ala Gln Gly Asp Gly Ala Asn Ala Gly Gln Pro Gln Ala Gln Gly Asp Gly Ala Asn

780 7 8
GCAGGGCAACCACAAGCACAGGGTGATGGAGCAAATGCAGGACAACCACAAGCACAAGGA
Ala Gly Gln Pro Gln Ala Gln Gly Asp Gly Ala Asn Ala Gly Gln Pro Gln Ala Gln Gly

840 9
GATGGAGCAAATGCAGGACAACCACAAGCACAAGGAGATGGAGCAAATGCAGGACAACCA
Asp Gly Ala Asn Ala Gly Gln Pro Gln Ala Gln Gly Asp Gly Ala Asn Ala Gly Gln Pro

900 10 11
CAAGCACAGGGTGATGGAGCAAATGCAGGACAACCACAAGCACAGGGTGATGGAGCCAAT
Gln Ala Gln Gly Asp Gly Ala Asn Ala Gly Gln Pro Gln Ala Gln Gly Asp Gly Ala Asn

960 12
GCAGGACAACCACAAGCACAAGGAGATGGGGCAAATGTACCACGACAAGGAAGAAACGGG
Ala Gly Gln Pro Gln Ala Gln Gly Asp Gly Ala Asn Val Pro Arg Gln Gly Arg Asn Gly

1020
GGAGGTGCACCAGCAGGAGGAAATGAGGGGAATAAACAAGCAGGAAAAGGACAGGGACAA
Gly Gly Ala Pro Ala Gly Gly Asn Glu Gly Asn Lys Gln Ala Gly Lys Gly Gln Gly Gln

1080
AACAATCAGGGTGCGAATGCCCCAAATGAAAAAGTTGTGAATGATTACCTACACAAAATT
Asn Asn Gln Gly Ala Asn Ala Pro Asn Glu Lys Val Val Asn Asp Tyr Leu His Lys Ile

1140
AGATCTAGCGTTACCACCGAGTGGACTCCATGCAGTGTAACCTGTGGAAATGGTGTAAGA
Arg Ser Ser Val Thr Thr Glu Trp Thr Pro Cys Ser Val Thr Cys Gly Asn Gly Val Arg

1200
ATTAGAAGAAAAGCTCATGCAGGTAATAAAAAGGCAGAGGACCTTACTATGGATGACCTT
Ile Arg Arg Lys Ala His Ala Gly Asn Lys Lys Ala Glu Asp Leu Thr Met Asp Asp Leu

1260
GAGGTGGAAGCTTGTGTAATGGATAAGTGCGCTGGCATATTTAACGTTGTGAGTAATTCA
Glu Val Glu Ala Cys Val Met Asp Lys Cys Ala Gly Ile Phe Asn Val Val Ser Asn Ser

1320
TTAGGCTTAGTCATATTGTTAGTCCTAGCATTATTCAATTAA...
Leu Gly Leu Val Ile Leu Leu Val Leu Ala Leu Phe Asn STOP

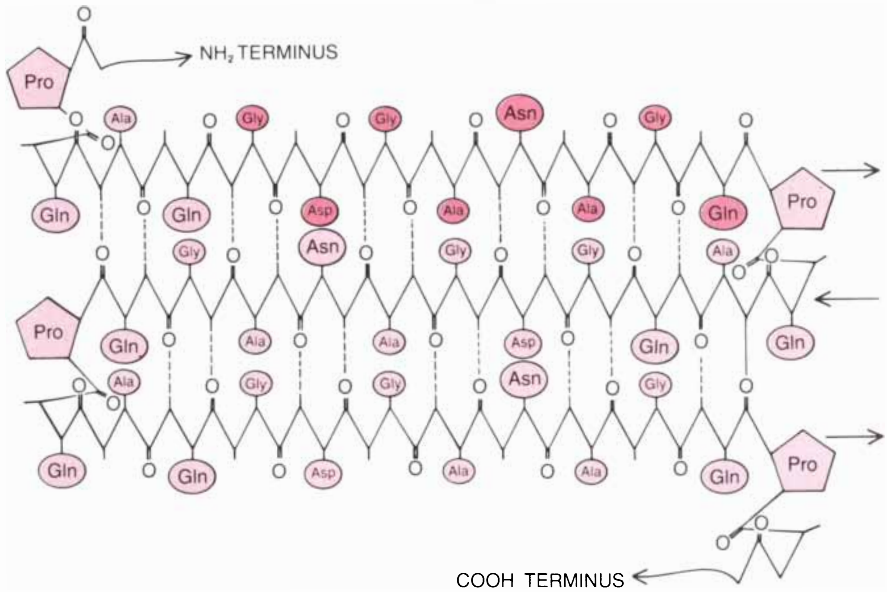
NUCLEOTIDE SEQUENCE of the entire *CS*-protein gene was determined; the translated portion of the nucleotide sequence is shown, and with it the deduced amino acid sequence. The protein has a hydrophobic signal region (*yellow*), which promotes its passage through the sporozoite's cell membrane, and a region of basic amino acids (*green*). Then comes the antibody-binding region (*red*): 12 tandem repeats of a 12-amino-acid epitope. The sequences of the repeated epitopes are almost identical. Most of the few nucleotide differences, such as those in the triplets encoding the central glutamine (*Gln*) and glycine (*Gly*), do not result in amino acid changes; only the final valine (*Val*) is a substitution. Two pairs of cysteines, which can form disulfide bonds, precede the hydrophobic anchor region (*blue*).

as those described above. The problem will then be to learn whether any of them induce protective antibodies.

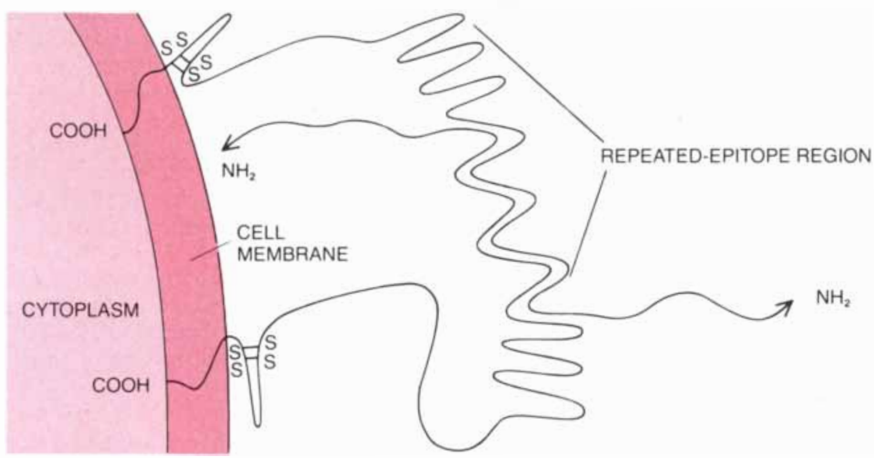
Clearly the parasite is most vulnerable when it is free in the bloodstream, seeking a target liver cell or red blood cell. There are three such stages at which a vaccine can be effective. An antisporozoite vaccine would be ideal: it would break the link between mosquito and man by stimulating the manufacture of antibodies able to attack the sporozoite at the initiation of infection, before it can reach the liver cells. An antimerozoite vaccine would stimulate an attack on the parasite in mid-infection; in conjunction with an antisporozoite vaccine it would establish a second line of defense. An antigametocyte vaccine would break the link between man and mosquito.

There may also be ways to attack the parasite when it is not free in the bloodstream. The most likely opportunity may present itself during the invasion of a liver cell or a red blood cell. To recognize and enter these cells the parasite must detect and exploit some specific receptor on the target cell's surface. The sporozoite or merozoite may be equipped with a recognition molecule or, as Vanderberg has suggested, it may pick up from the host's serum a glycoprotein that serves this purpose. Once the molecular details of the invasion process are understood it may be possible to develop an analogue of the recognition molecule or an antibody to the cell's receptor molecule. Either one might bind to the receptor and thus block invasion. When the precise route of invasion is known, it may even be possible to develop an agent that will kill the parasite while it is inside a liver or blood cell.

The molecular study of *Plasmodium* is still in its infancy, but it has already yielded insights leading to new ways of thinking about the parasite and suggesting new ways to combat it. On the one hand, a molecular approach at the level of DNA seeks to identify mechanisms peculiar to the malaria parasite, one or more of which may prove to be an Achilles' heel. On the other hand, the power of molecular biology lies in the recognition that fundamental mechanisms of gene expression and cellular structure are common to all forms of life, and that what is learned in one biological system is applicable to all systems. Important advances in medicine apply new understanding of general mechanisms to solve specific medical problems, but at the same time they develop information that contributes to the expansion of fundamental knowledge.



CONFIGURATION of the repeated-epitope region (the immune decoy) can be predicted from its amino acid sequence. It is likely to be an antiparallel beta pleated sheet, in which a polypeptide chain folds back and forth on itself and adjacent segments of the chain are linked into a sheetlike structure by hydrogen bonds (*broken lines*). The chain would fold by bending sharply at each proline (*Pro*). Both the size and the charge of abutting amino acids suggest they would not interfere with one another. Experiments have shown that eight contiguous amino acids (*dark color*) suffice to interact with the antibody to the *CS* protein.



BETA SHEET providing an ideal protective surface for an invading parasite could be formed by both intramolecular and intermolecular interactions. The polypeptide chain of a *CS*-protein molecule could fold and be hydrogen-bonded to itself or to the folded chain of an adjacent molecule. The drawing suggests how molecules embedded in the parasite cell membrane might interact to form part of a network constituting the parasite's outer coat.

SPOROZOITES		
P. KNOWLESI	*CS* PROTEIN	Gly Gln Pro Gln Ala Gln Gly Asp Gly Ala Asn Ala (12)
P. FALCIPARUM	*CS* PROTEIN	Asn Ala Asn Pro (41)
BLOOD STAGES		
P. FALCIPARUM	*S* ANTIGEN	Pro Ala Lys Ala Ser Gln Gly Gly Leu Glu Asp (100-110)
	RESA ANTIGEN	Glu Glu Asn Val Glu His Asp Ala
	FIRA ANTIGEN	Val Thr Thr Gln Glu Pro
	PF-11 ANTIGEN	Glu Glu Val Val Glu Glu Val Val Pro
P. LOPHURAE	His-RICH PROTEIN	Ala Pro His_8 Asp Ala His_8

SEVEN SURFACE PROTEINS of various *Plasmodium* species and stages have been isolated and characterized. All of them have repeated epitopes, whose sequences are given here. The number of repeats is also indicated for the three proteins in which it is known.

How a Supernova Explodes

When a large star runs out of nuclear fuel, the core collapses in milliseconds. The subsequent "bounce" of the core generates a shock wave so intense that it blows off most of the star's mass

by Hans A. Bethe and Gerald Brown

The death of a large star is a sudden and violent event. The star evolves peacefully for millions of years, passing through various stages of development, but when it runs out of nuclear fuel, it collapses under its own weight in less than a second. The most important events in the collapse are over in milliseconds. What follows is a supernova, a prodigious explosion more powerful than any since the big bang with which the universe began.

A single exploding star can shine brighter than an entire galaxy of several billion stars. In the course of a few months it can give off as much light as the sun emits in a billion years. Furthermore, light and other forms of electromagnetic radiation represent only a small fraction of the total energy of the supernova. The kinetic energy of the exploding matter is 10 times greater. Still more energy—perhaps 100 times more than the electromagnetic emission—is carried away by the massless particles called neutrinos, most of which are emitted in a flash that lasts for about a second. When the explosion is over, most of the star's mass has been scattered into space, and all that remains at the center is a dense, dark cinder. In some cases even that may disappear into a black hole.

Such an outline description of a supernova could have been given almost 30 years ago, and yet the detailed sequence of events within the dying star is still not known with any certainty. The basic question is this: A supernova begins as a collapse, or implosion; how does it come about, then, that a major part of the star's mass is expelled? At some point the inward movement of stellar material must be stopped and then reversed; an *im*plosion must be transformed into an *ex*plosion.

Through a combination of computer simulation and theoretical analysis a coherent view of the supernova mechanism is beginning to emerge. It appears the crucial event in the turnaround is the formation of a shock wave that travels outward at 30,000 kilometers per second or more.

Supernovas are rare events. In our own galaxy just three have been recorded in the past 1,000 years; the brightest of these, noted by Chinese observers in 1054, gave rise to the expanding shell of gas now known as the Crab Nebula. If only such nearby events could be observed, little would be known about supernovas. Because they are so luminous, however, they can be detected even in distant galaxies, and 10 or more per year are now sighted by astronomers.

The first systematic observations of distant supernovas were made in the 1930's by Fritz Zwicky of the California Institute of Technology. About half of the supernovas Zwicky studied fitted a quite consistent pattern: the luminosity increased steadily for about three weeks and then declined gradually over a period of six months or more. He designated the explosions in this group Type I. The remaining supernovas were more varied, and Zwicky divided them into four groups; today, however, they are all grouped together as Type II. In Type I and Type II supernovas the events leading up to the explosion are thought to be quite different. Here we shall be concerned primarily with Type II supernovas.

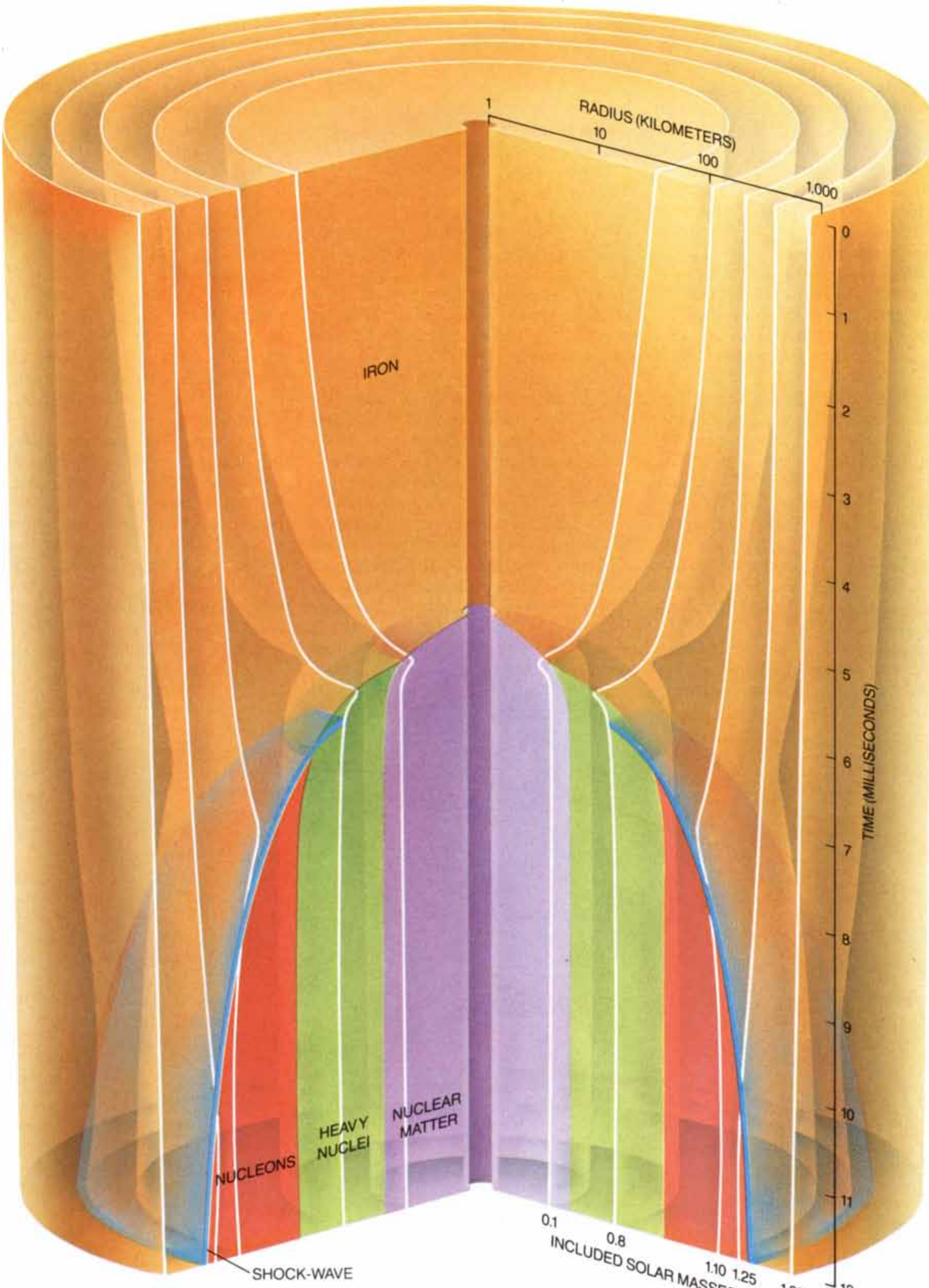
The basis for the theory of supernova explosions was the work of Fred Hoyle of the University of Cambridge. The theory was then developed in a fundamental paper published in 1957 by E. Margaret Burbidge, Geoffrey R. Burbidge and William A. Fowler, all of Caltech, and Hoyle. They proposed that when a massive star reaches the end of its life, the stellar core collapses under the force of its own gravitation. The energy set free by the collapse expels most of the star's mass, distributing the chemical elements formed in the course of its evolution throughout interstellar space. The collapsed core leaves behind a dense remnant, in many cases a neutron star.

A supernova is an unusual and spectacular outcome of the sequence of nuclear fusion reactions that is the life history of a star. The heat given off by the fusion creates pressure, which counteracts the gravitational attraction that would otherwise make the star collapse. The first series of fusion reactions have the net effect of welding four atoms of hydrogen into a single atom of helium. The process is energetically favorable: the mass of the helium atom is slightly less than the combined masses of the four hydrogen atoms, and the energy equivalent of the excess mass is released as heat.

The process continues in the core of the star until the hydrogen there is used up. The core then contracts, since gravitation is no longer opposed by energy production, and as a result both the core and the surrounding material are heated. Hydrogen fusion then begins in the surrounding layers. Meanwhile the core becomes hot enough to ignite other fusion reactions, burning helium to form carbon, then burning the carbon to form neon, oxygen and finally silicon. Again each of these reactions leads to the release of energy. One last cycle of fusion combines silicon nuclei to form iron, specifically the common iron isotope ^{56}Fe, made up of 26 protons and 30 neutrons. Iron is the end of the line for spontaneous fusion. The ^{56}Fe nucleus is the most strongly bound of all nuclei, and further fusion would absorb energy rather than releasing it.

At this stage in the star's existence it has an onionlike structure. A core of iron and related elements is surrounded by a shell of silicon and sulfur, and beyond this are shells of oxygen, carbon and helium. The outer envelope is mostly hydrogen.

Only the largest stars proceed all the



COLLAPSE AND REBOUND are the initiating events in a supernova explosion. Here the core of a massive star is shown as it passes through the moment of "maximum scrunch," when the center reaches its highest density. Each contour represents a shell of matter whose radial position is followed through a period of 12 milliseconds. The included mass, or total mass inside the contour, does not change as the shells contract and expand. Initially the core is iron, but the extreme compression of the collapse converts the innermost few kilometers into nuclear matter, the stuff of the atomic nucleus. Surrounding this region is a shell made up of various heavy nuclei, including iron. At maximum scrunch the contraction stops with a jolt, creating a shock wave (*blue line*) that travels outward at 30,000 kilometers per second or more. In the wake of the shock nuclei are broken up into individual nucleons (protons and neutrons).

way to the final, iron-core stage of the evolutionary sequence. A star the size of the sun gets no further than helium burning, and the smallest stars stop with hydrogen fusion. A larger star also consumes its stock of fuel much sooner, even though there is more of it to begin with; because the internal pressure and temperature are higher in a large star, the fuel burns faster. Whereas the sun should have a lifetime of 10 billion years, a star 10 times as massive can complete its evolution 1,000 times faster. Regardless of how long it takes, all the usable fuel in the core will eventually be exhausted. At that point heat production in the core ends and the star must contract.

When fusion ends in a small star, the star slowly shrinks, becoming a white dwarf: a burned-out star that emits only a faint glow of radiation. In isolation the white dwarf can remain in this state indefinitely, cooling gradually but otherwise changing little. What stops the star from contracting further? The answer was given more than 50 years ago by Subrahmanyan Chandrasekhar of the University of Chicago.

Loosely speaking, when ordinary matter is compressed, higher density is achieved by squeezing out the empty space between atoms. In the core of a white dwarf this process has reached its limit: the atomic electrons are pressed tightly together. Under these conditions the electrons offer powerful resistance to further compression.

Chandrasekhar showed there is a limit to how much pressure can be resisted by the electrons' mutual repulsion. As the star contracts, the gravitational energy increases, but so does the energy of the electrons, raising their pressure. If the contraction goes very far, both the gravitational energy and the electron energy are inversely proportional to the star's radius. Whether or not there is some radius at which the two opposing forces are in balance, however, depends on the mass of the star. Equilibrium is possible only if the mass is less than a critical value, now called the Chandrasekhar mass. If the mass is greater than the Chandrasekhar limit, the star must collapse.

The value of the Chandrasekhar mass depends on the relative numbers of electrons and nucleons (protons and neutrons considered collectively): the higher the proportion of electrons, the larger the electron pressure and so the larger the Chandrasekhar mass. In small stars where the chain of fusion reactions stops at carbon the ratio is approximately 1/2 and the Chandrasekhar mass is 1.44 solar masses. This is the maximum stable mass for a white dwarf.

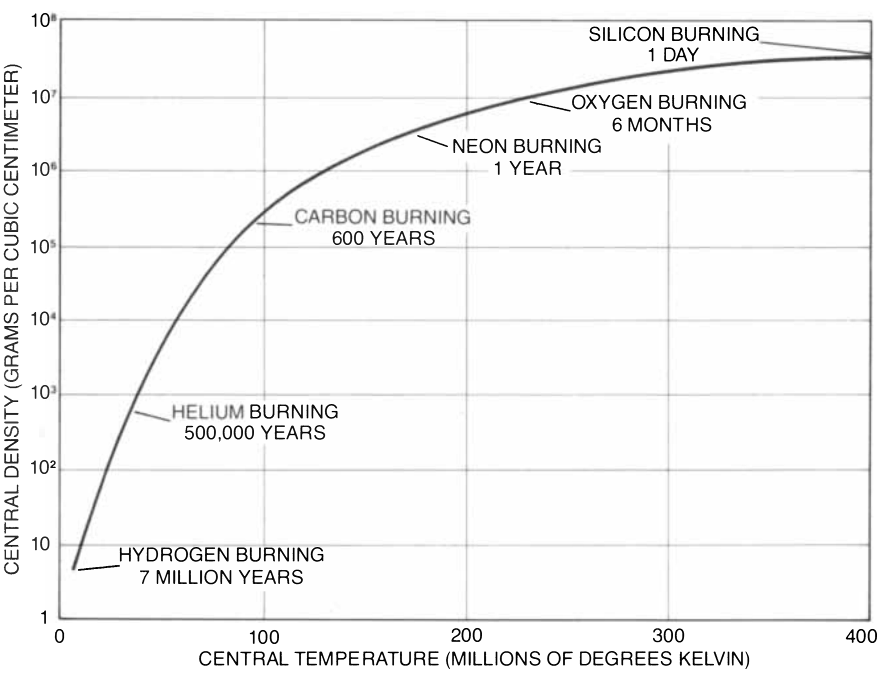
A white dwarf with a mass under the Chandrasekhar limit can remain stable indefinitely; nevertheless, it is just such stars that are thought to give rise to Type I supernovas. How can this be? The key to the explanation is that white dwarfs that explode in supernovas are not solitary stars but rather are members of binary star systems. According to one hypothesis, matter from the binary companion is attracted by the intense gravitational field of the dwarf star and gradually falls onto its surface, increasing the mass of the carbon-and-oxygen core. Eventually the carbon ignites at the center and burns in a wave that travels outward, destroying the star.

The idea that explosive carbon burning triggers Type I supernovas was proposed in 1960 by Hoyle and Fowler. More detailed models have since been devised by many astrophysicists, most notably Icko Iben, Jr., and his colleagues at the University of Illinois at Urbana-Champaign. Recent calculations done by Ken'ichi Nomoto and his colleagues at the University of Tokyo suggest that the burning is actually not explosive. The wave of fusion reactions propagates like the burning of a fuse rather than like the explosion of a firecracker; it is a deflagration rather than a detonation.

Even though the burning is less violent than a detonation, the white dwarf is completely disrupted. The initial binding energy that holds the star together is approximately 10^{50} ergs; the energy released by the burning is 20 times greater (2×10^{51} ergs), enough to account for the 10,000-kilometer-per-second velocity of supernova remnants. In the course of the deflagration nuclear reactions create about one solar mass of the unstable nickel isotope ^{56}Ni, which decays into ^{56}Co and then ^{56}Fe over a period of months. The rate of energy release from the radioactive decay is just right to account for the gradually declining light emission from Type I supernovas.

The Type II supernovas that are our main concern here arise from much more massive stars. The lower limit is now thought to be about eight solar masses.

In tracing the history of a Type II supernova it is best to begin at the moment when the fusion of silicon nuclei to form iron first becomes possible at the center of the star. At this point the star has already passed through stages of burning hydrogen, helium, neon, carbon and oxygen, and it has the onionlike structure described above.



EVOLUTION OF A MASSIVE STAR is a steadily accelerating progress toward higher temperature and density in the core. For most of the star's lifetime the primary energy source is the fusion of hydrogen nuclei to form helium. When the hydrogen in the core is exhausted, the core contracts, which heats it enough to ignite the fusion of helium into carbon. This cycle then repeats, at a steadily increasing pace, through the burning of carbon, neon, oxygen and silicon. The final stage of silicon fusion yields a core of iron, from which no further energy can be extracted by nuclear reactions. Hence the iron core cannot resist gravitational collapse, leading to a supernova explosion. The sequence shown is for a star of 25 solar masses. Data in this illustration and the one on the opposite page are based on calculations done by Thomas A. Weaver of the Lawrence Livermore National Laboratory.

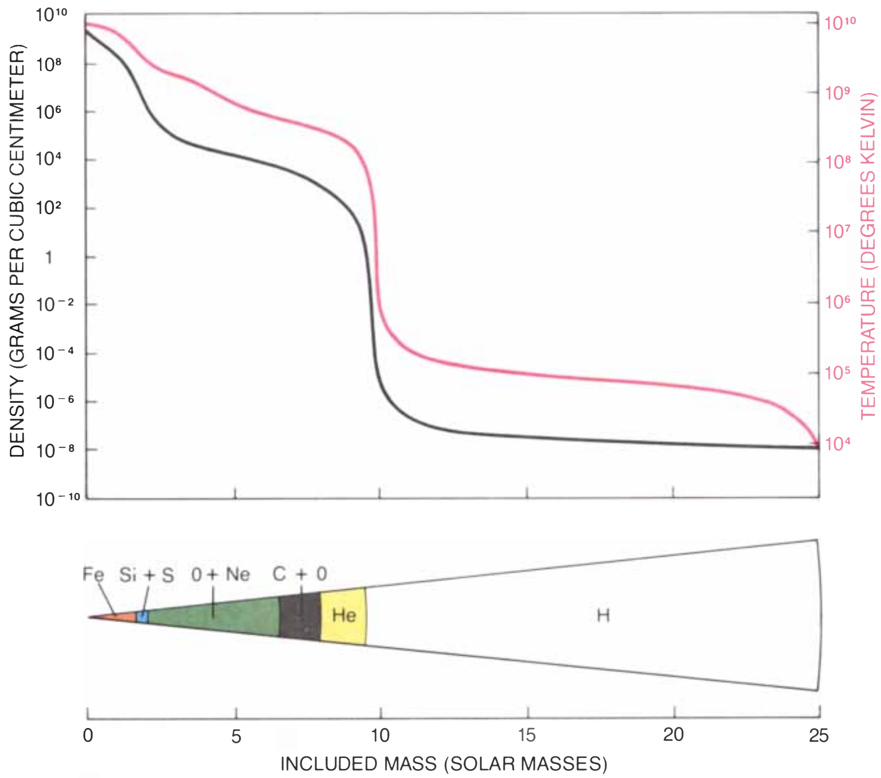
The star has taken several million years to reach this state. Subsequent events are much faster.

When the final fusion reaction begins, a core made up of iron and a few related elements begins to form at the center of the star, within a shell of silicon. Fusion continues at the boundary between the iron core and the silicon shell, steadily adding mass to the core. Within the core, however, there is no longer any production of energy by nuclear reactions; the core is an inert sphere under great pressure. It is thus in the same predicament as a white dwarf: it can resist contraction only by electron pressure, which is subject to the Chandrasekhar limit.

Once the fusion of silicon nuclei begins, it proceeds at an extremely high rate, and the mass of the core reaches the Chandrasekhar limit in about a day. We noted above that for a white dwarf the Chandrasekhar mass is equal to 1.44 solar masses; for the iron core of a large star the value may be somewhat different, but it is probably in the range between 1.2 and 1.5 solar masses.

When the Chandrasekhar mass has been attained, the pace speeds up still more. The core that was built in a day collapses in less than a second. The task of analysis also becomes harder at this point, so that theory relies on the assistance of computer simulation. Computer programs that trace the evolution of a star have been developed by a number of workers, including W. David Arnett of the University of Chicago and a group at the Lawrence Livermore National Laboratory led by Thomas A. Weaver of that laboratory and Stanford Woosley of the University of California at Santa Cruz. They are the "burners" of stars; we and our colleagues in theoretical physics are "users" of their calculations.

The simulations furnish us with a profile of the presupernova core, giving composition, density and temperature as a function of radius. The subsequent analysis relies on applying familiar laws of thermodynamics, the same laws that describe such ordinary terrestrial phenomena as the working of a heat engine or the circulation of the atmosphere.



ONIONLIKE STRUCTURE is characteristic of a massive star at the end of its evolution, just before the gravitational collapse. The iron core is embedded in a mantle of silicon, sulfur, oxygen, neon, carbon and helium, surrounded by an attenuated envelope of hydrogen. Temperature and density fall off steadily in the mantle, then drop precipitously at the envelope. Fusion has stopped in the core but continues at boundaries between layers.

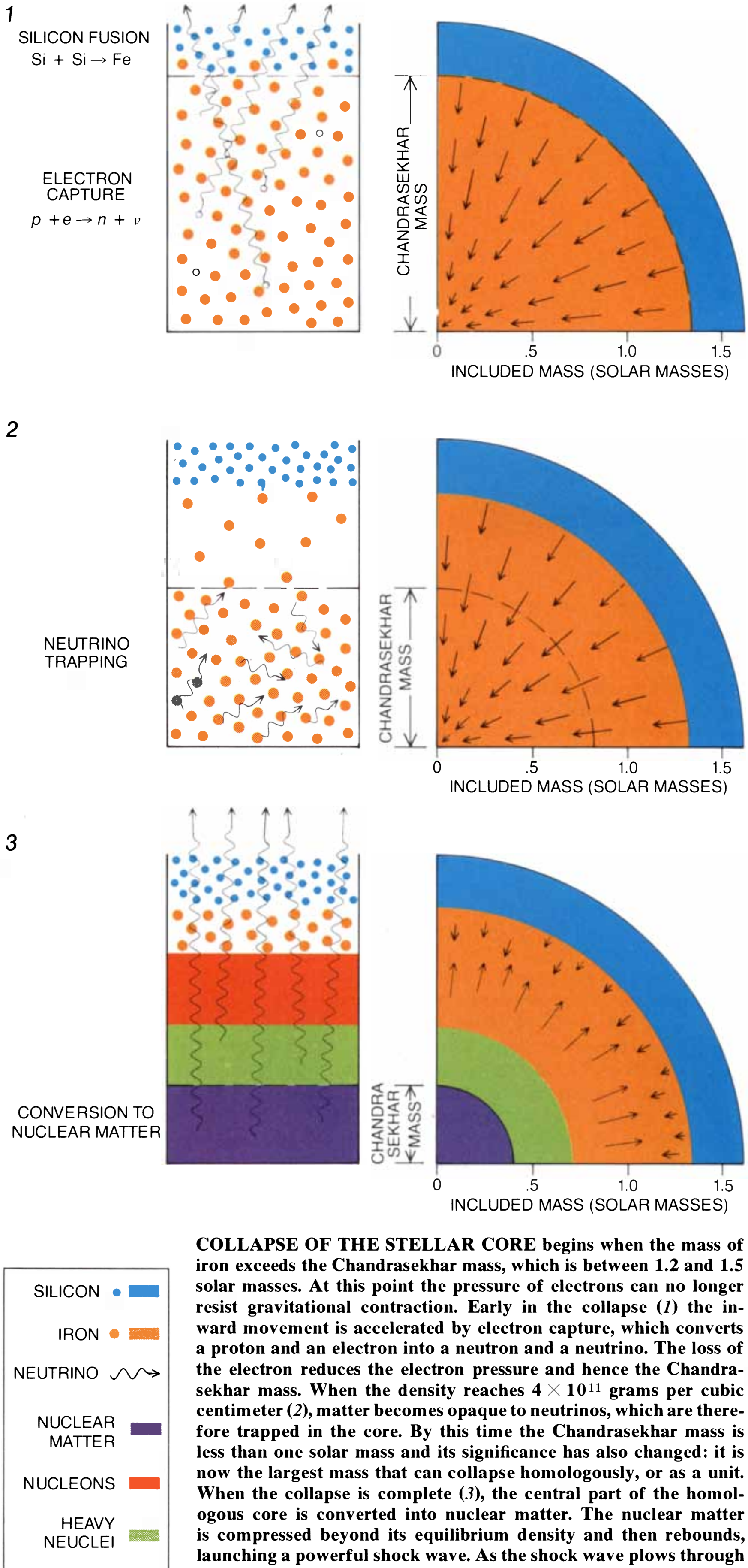
It is worthwhile tracing in some detail the initial stages in the implosion of the core. One of the first points of note is that compression raises the temperature of the core, which might be expected to raise the pressure and slow the collapse. Actually the heating has just the opposite effect.

Pressure is determined by two factors: the number of particles in a system and their average energy. In the core both nuclei and electrons contribute to the pressure, but the electron component is much larger. When the core is heated, a small fraction of the iron nuclei are broken up into smaller nuclei, increasing the number of nuclear particles and raising the nuclear component of the pressure. At the same time, however, the dissociation of the nuclei absorbs energy; since energy is released when an iron nucleus is formed, the same quantity of energy must be supplied in order to break the nucleus apart. The energy comes from the electrons and decreases their pressure. The loss in electron pressure is more important than the gain in nuclear pressure. The net result is that the collapse accelerates.

It might seem that the implosion of a star would be a chaotic process, but in fact it is quite orderly. Indeed, the entire evolution of the star is toward a condition of greater order, or lower entropy. It is easy to see why. In a hydrogen star each nucleon can move willy-nilly along its own trajectory, but in an iron core groups of 56 nucleons are bound together and must move in lockstep. Initially the entropy per nucleon, expressed in units of Boltzmann's constant, is about 15; in the presupernova core it is less than 1. The difference in entropy has been carried off during the evolution of the star by electromagnetic radiation and toward the end also by neutrinos.

The low entropy of the core is maintained throughout the collapse. Nuclear reactions continually change the species of nuclei present, which one might think could lead to an increase in entropy; the reactions are so fast, however, that equilibrium is always maintained. The collapse takes only milliseconds, but the time scale of the nuclear reactions is typically from 10^{-15} to 10^{-23} second, so that any departure from equilibrium is immediately corrected.

Another effect was once thought to increase the entropy, but it now seems likely that it actually reduces it somewhat. The high density in the collapsing core favors the reaction known as electron capture. In this process a proton and an electron come together to yield a neutron and a neutrino. The neutrino escapes from the star, carrying off both energy and entropy and cooling the system just as the evaporation of moisture cools the body. There are several complications to this process, so that its effect on the entropy is uncertain. In any case, the loss of the electron diminishes the electron pres-



COLLAPSE OF THE STELLAR CORE begins when the mass of iron exceeds the Chandrasekhar mass, which is between 1.2 and 1.5 solar masses. At this point the pressure of electrons can no longer resist gravitational contraction. Early in the collapse (*1*) the inward movement is accelerated by electron capture, which converts a proton and an electron into a neutron and a neutrino. The loss of the electron reduces the electron pressure and hence the Chandrasekhar mass. When the density reaches 4×10^{11} grams per cubic centimeter (*2*), matter becomes opaque to neutrinos, which are therefore trapped in the core. By this time the Chandrasekhar mass is less than one solar mass and its significance has also changed: it is now the largest mass that can collapse homologously, or as a unit. When the collapse is complete (*3*), the central part of the homologous core is converted into nuclear matter. The nuclear matter is compressed beyond its equilibrium density and then rebounds, launching a powerful shock wave. As the shock wave plows through the outer core, iron nuclei "evaporate" to form a gas of nucleons.

sure and so allows the implosion to accelerate further.

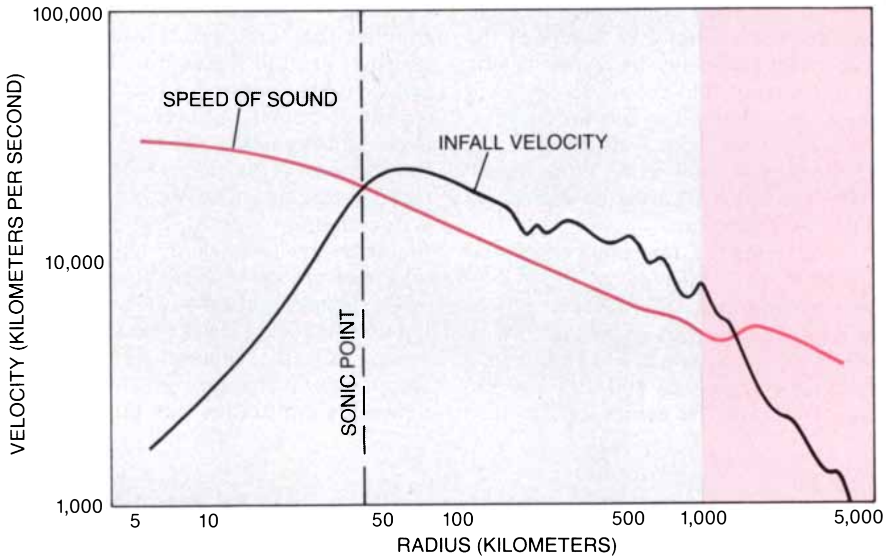
The first stage in the collapse of a supernova comes to an end when the density of the stellar core reaches a value of about 4×10^{11} grams per cubic centimeter. This is by no means the maximum density, since the core continues to contract, but it marks a crucial change in physical properties: at this density matter becomes opaque to neutrinos. The importance of this development was first pointed out by T. J. Mazurek of the Mission Research Laboratory in Santa Barbara, Calif., and by Katsushiko Sato of the University of Tokyo.

The neutrino is an aloof particle that seldom interacts with other forms of matter. Most of the neutrinos that strike the earth, for example, pass all the way through it without once colliding with another particle. When the density exceeds 400 billion grams per cubic centimeter, however, the particles of matter are packed so tightly that even a neutrino is likely to run into one. As a result neutrinos emitted in the collapsing core are effectively trapped there. The trapping is not permanent; after a neutrino has been scattered, absorbed and reemitted many times, it must eventually escape, but the process takes longer than the remaining stages of the collapse. The effective trapping of neutrinos means that no energy can get out of the core.

The process of electron capture in the early part of the collapse reduces not only the electron pressure but also the ratio of electrons to nucleons, the quantity that figures in the calculation of the Chandrasekhar mass. In a typical presupernova core the ratio is between .42 and .46; by the time of neutrino trapping it has fallen to .39. This lower ratio yields a Chandrasekhar mass of .88 solar mass, appreciably less than the original value of between 1.2 and 1.5.

At this point the role of the Chandrasekhar mass in the analysis of the supernova also changes. At the outset it was the largest mass that could be supported by electron pressure; it now becomes the largest mass that can collapse as a unit. Areas within this part of the core can communicate with one another by means of sound waves and pressure waves, so that any variations in density are immediately evened out. As a result the inner part of the core collapses homologously, or all in one piece, preserving its shape.

The theory of homologous collapse was worked out by Peter Goldreich and Steven Weber of Caltech and was further developed by Amos Yahil and James M. Lattimer of the State Uni-



SONIC POINT marks the boundary of the homologous core. It is the radius at which the speed of sound is equal to the velocity of the infalling material. A sound wave at the sonic point moves outward at the speed of sound in relation to the material it is passing through, but since that material is falling inward at the same speed, the wave stands still in relation to the center of the star. As a result a disturbance inside the core cannot reach the outside. The graph is based on calculations done by W. David Arnett of the University of Chicago.

versity of New York at Stony Brook. The shock wave that blows off the outer layers of the star forms at the edge of the homologous core. Before we can give an account of that process, however, we must continue to trace the sequence of events within the core itself.

Chandrasekhar's work showed that electron pressure cannot save the core of a large star from collapse. The only other hope for stopping the contraction is the resistance of nucleons to compression. In the presupernova core nucleon pressure is a negligible fraction of electron pressure. Even at a density of 4×10^{11} grams per cubic centimeter, where neutrino trapping begins, nucleon pressure is insignificant. The reason is the low entropy of the system. At any given temperature, pressure is proportional to the number of particles per unit volume, regardless of the size of the individual particles. An iron nucleus, with 56 nucleons, makes the same contribution to the pressure as an isolated proton does. If the nuclei in the core were broken up, their pressure might be enough to stop the contraction. The fissioning of the nuclei is not possible, however, because the entropy of the core is too low. A supernova core made up of independently moving protons and neutrons would have an entropy per nucleon of between 5 and 8, whereas the actual entropy is less than 1.

The situation does not change, and the collapse is not impeded, until the density in the central part of the core reaches about 2.7×10^{14} grams per cubic centimeter. This is the density of matter inside a large atomic nucleus, and in effect the nucleons in the core merge to form a single gigantic nucleus. A teaspoonful of such matter has about the same mass as all the buildings in Manhattan combined.

Nuclear matter is highly incompressible. Hence once the central part of the core reaches nuclear density there is powerful resistance to further compression. That resistance is the primary source of the shock waves that turn a stellar collapse into a spectacular explosion.

Within the homologously collapsing part of the core, the velocity of the infalling material is directly proportional to distance from the center. (It is just this property that makes the collapse homologous.) Density, on the other hand, decreases with distance from the center, and as a result so does the speed of sound. The radius at which the speed of sound is equal to the infall velocity is called the sonic point, and it marks the boundary of the homologous core. A disturbance inside the core can have no influence beyond this radius. At the sonic point sound waves move outward at the speed of sound, as measured in the coordinate system of the infalling matter. This matter is moving inward at the same speed, however, and so the waves are at a standstill in relation to the center of the star.

When the center of the core reaches nuclear density, it is brought to rest with a jolt. This gives rise to sound waves that propagate back through the medium of the core, rather like the vibrations in the handle of a hammer when it strikes an anvil. The waves slow as they move out through the homologous core, both because the local speed of sound declines and because they are moving upstream against a flow that gets steadily faster. At the sonic point they stop entirely. Meanwhile additional material is falling onto the hard sphere of nuclear matter in the center, generating more waves. For a fraction of a millisecond the waves collect at the sonic point, building up pressure there. The bump in pressure slows the material falling through the sonic point, creating a discontinuity in velocity. Such a discontinuous change in velocity constitutes a shock wave.

At the surface of the hard sphere in the heart of the star infalling material stops suddenly but not instantaneously. The compressibility of nuclear matter is low but not zero, and so momentum carries the collapse beyond the point of equilibrium, compressing the central core to a density even higher than that of an atomic nucleus. We call this point the instant of "maximum scrunch." Most computer simulations suggest the highest density attained is some 50 percent greater than the equilibrium density of a nucleus. After the maximum scrunch the sphere of nuclear matter bounces back, like a rubber ball that has been compressed. The bounce sets off still more sound waves, which join the growing shock wave at the sonic point.

A shock wave differs from a sound wave in two respects. First, a sound wave causes no permanent change in its medium; when the wave has passed, the material is restored to its former state. The passage of a shock wave can induce large changes in density, pressure and entropy. Second, a sound wave—by definition—moves at the speed of sound. A shock wave moves faster, at a speed determined by the energy of the wave. Hence once the pressure discontinuity at the sonic point has built up into a shock wave, it is no longer pinned in place by the infalling matter. The wave can continue outward, into the overlying strata of the star. According to computer simulations, it does so with great speed, between 30,000 and 50,000 kilometers per second.

Up to this point in the progress of the supernova essentially all calculations are in agreement. What happens next, however, is not yet firmly established. In the simplest scenario, which we have favored, the shock wave rushes outward, reaching the surface of the iron core in a fraction of a second and then continuing through the successive onionlike layers of the star. After some days it works its way to the surface and erupts as a violent explosion. Beyond a certain radius—the bifurcation point—all the material of the star is blown off. What is left inside the bifurcation radius condenses into a neutron star.

Alas! Using presupernova cores simulated in 1974 by Weaver and Woosley, calculations of the fate of the shock wave are not so accommodating. The shock travels outward to a distance of between 100 and 200 kilometers from the center of the star, but then it becomes stalled, staying at roughly the same position as matter continues to fall through it. The main reason for the stalling is that the shock breaks up nuclei into individual nucleons. Although this process increases the number of particles, which might be expected to raise the pressure, it also consumes a great deal of energy; the net result is that both temperature and pressure are sharply reduced.

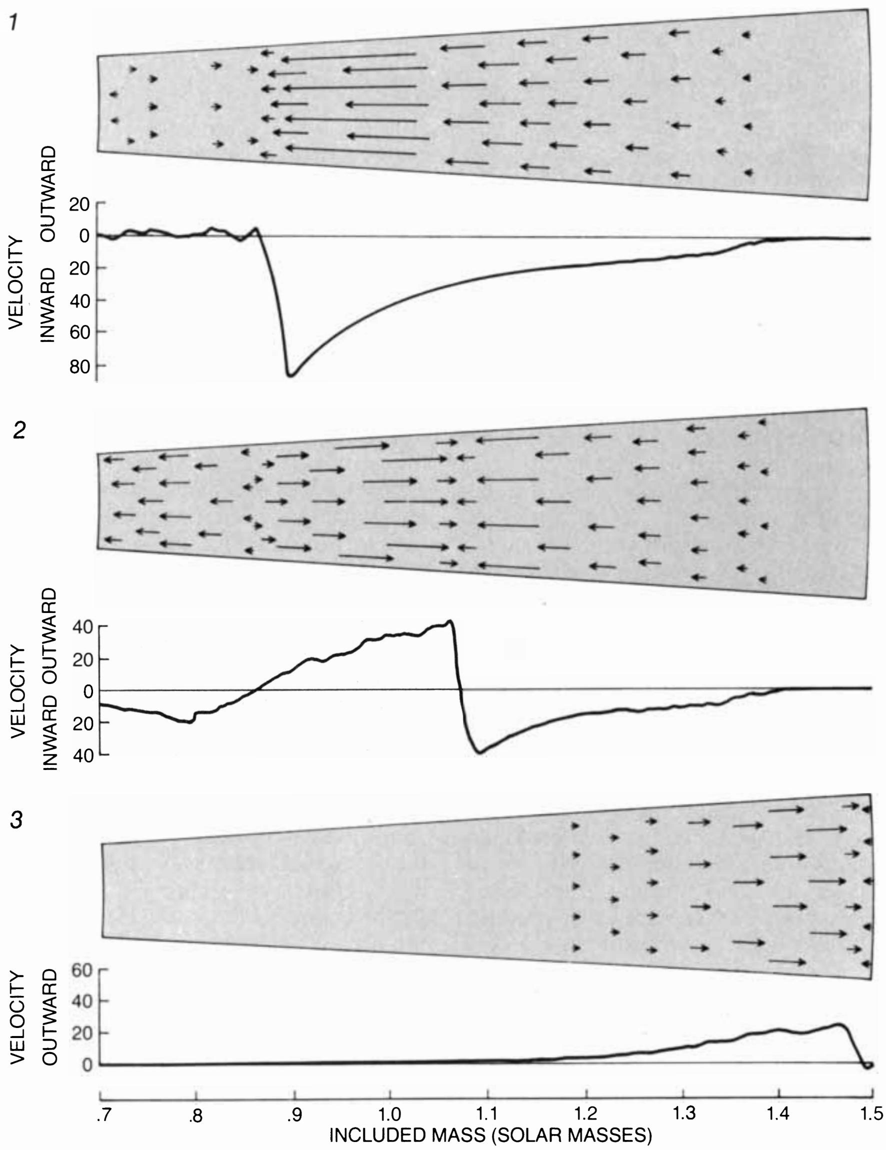
The fragmentation of the nuclei contributes to energy dissipation in another way as well: it releases free protons, which readily capture electrons. The neutrinos emitted in this process can escape, removing their energy from the star. The escape is possible because the shock has entered material whose density is below the critical value for neutrino trapping. The neutrinos that had been trapped behind the shock also stream out, carrying away still more energy. Because of the many hazards to the shock wave in the region between 100 and 200 kilometers, we have named this region of the star the "minefield."

It would be satisfying to report that we have found a single mechanism capable of explaining for all Type II supernovas how the shock wave makes its way through the minefield. We cannot do so. What we have to offer instead is a set of possible explanations, each of which seems to work for stars in a particular range of masses.

The place to begin is with stars of between 12 and about 18 solar masses. Weaver and Woosley's most recent models of presupernova cores for such stars differ somewhat from those they calculated a decade ago; the most important difference is that the iron core is smaller than earlier estimates indicated—about 1.35 solar masses. The homologous core, at whose surface the shock wave forms, includes .8 solar mass of this material, leaving .55 solar mass of iron outside the sonic point. Since the breaking up of iron nuclei has the highest energy cost, reducing the quantity of iron makes it easier for the shock to break out of the core.

Jerry Cooperstein and Edward A. Baron of Stony Brook have been able to simulate successful supernova explosions in computer calculations that begin with Weaver and Woosley's model cores. The main requirement, first surmised by Sidney H. Kahana of the Brookhaven National Laboratory, is that the homologous core be very strongly compressed, so that it can rebound vigorously and create an intense shock. Two factors cooperate to achieve this result in the simulations. The first factor is the use of general relativity rather than the force field of Newtonian gravitation. The second is the assumption that nuclear matter is much more compressible than had been thought.

Baron's first result showed that a star of 12 solar masses would explode if the compressibility of nuclear matter is 1.5 times the standard value. This seemed rather arbitrary, but then one of us (Brown) examined the problem with a sophisticated method of nuclear-matter theory. It turned out that the most consistent interpretation of the experimental findings yields a compressibility of 2.5 times the standard value! We then found that in 1982 An-



SHOCK WAVE can move faster than sound and so it can carry the energy and momentum of the rebound past the sonic point. Just before the bounce (*1*) the inner core has reached the density of nuclear matter and has stopped contracting, but overlying matter is about to fall onto the core at speeds of up to 90,000 kilometers per second. Two milliseconds later (*2*) the core is being driven further inward, but at the same time much of the infalling matter has rebounded to form a shock wave. After 20 milliseconds (*3*) the shock has reached the edge of the core. This mechanism of supernova explosion, in which the shock succeeds directly in bursting through the core, seems to work for stars of between 12 and 18 solar masses. Velocity profiles shown were calculated by Jerry Cooperstein of the State University of New York at Stony Brook. Velocities are given in thousands of kilometers per second.

drew D. Jackson, E. Krotscheck, D. E. Meltzer and R. A. Smith had reached the same conclusion by another method, but no one had recognized the relevance of their work to the supernova problem. We consider the higher estimate of nuclear compressibility quite reliable.

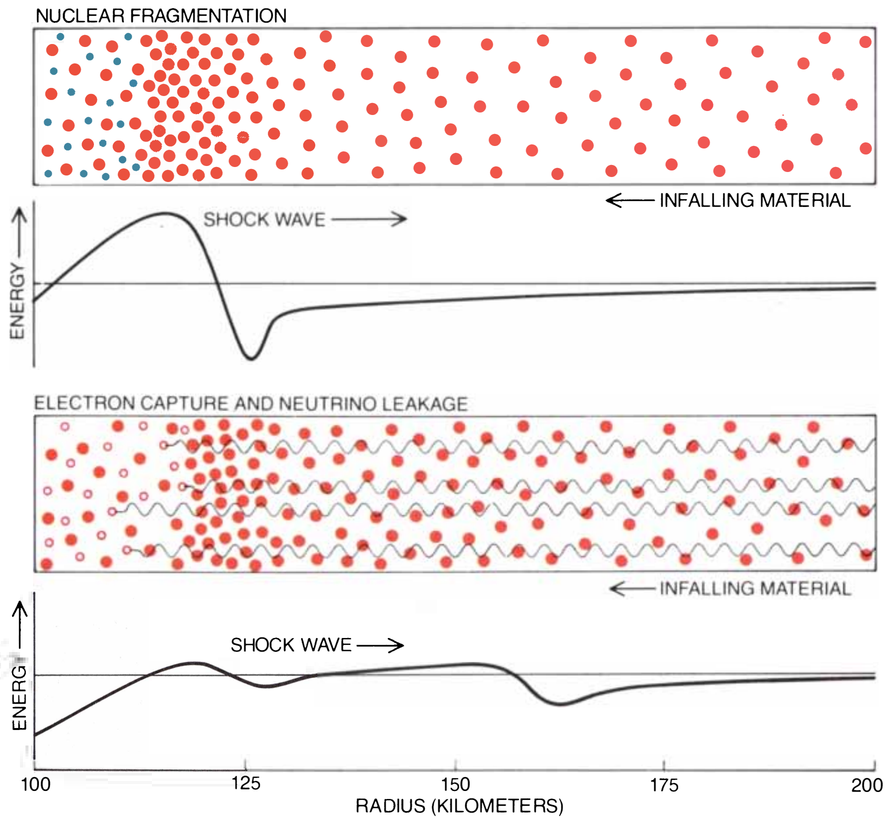
The mechanism described by Baron, Cooperstein and Kahana seems to work for stars of up to about 18 solar masses. With still larger stars, however, even the powerful shock wave created in their simulations becomes stalled in the minefield. A star of 25 solar masses has about two solar masses of iron in its core, and so the shock wave must penetrate 1.2 solar masses of iron rather than .55 solar mass. The shock does not have enough energy to dissociate this much iron.

A plausible explanation of what might happen in these massive stars has recently emerged from the work of James R. Wilson of Lawrence Livermore, who has done extensive numerical simulations of supernova explosions. For some time it had seemed that when the shock wave failed, all the mass of the star might fall back into the core, which would evolve into a black hole. That fate is still a possible one, but Wilson noted a new phenomenon when he continued some of his simulations for a longer period.

In the collapsing stellar core it takes only 10 milliseconds or so for the shock wave to reach the minefield and stall. A simulation of the same events, even with the fastest computers, takes at least an hour. Wilson allowed his calculations to run roughly 100 times longer, to simulate a full second of time in the supernova. In almost all cases he found that the shock wave eventually revived.

The revival is due to heating by neutrinos. The inner core is a copious emitter of neutrinos because of continuing electron capture as the matter is compressed to nuclear density. Adam S. Burrows and Lattimer of Stony Brook and Mazurek have shown that half of the electrons in the homologous core are captured within about half a second, and the emitted neutrinos carry off about half of the gravitational energy set free by the collapse, some 10^{53} ergs. Deep within the core the neutrinos make frequent collisions with other particles; indeed, we noted above that they are trapped, in the sense that they cannot escape within the time needed for the homologous collapse. Eventually, though, the neutrinos do percolate upward and reach strata of lower density, where they can move freely.

At the radius where the shock wave stalls only one neutrino out of every 1,000 is likely to collide with a particle of matter, but these collisions nonetheless impart a significant amount of energy. Most of the energy goes into the dissociation of nuclei into nucleons, the very process that caused the shock to stall in the first place. Now, however, the neutrino energy heats the material and therefore raises the pressure sharply. We have named this period, when the shock wave stalls but is then revived by neutrino heating, "the pause that refreshes."



SHOCK WAVE SEEMS TO STALL in stars whose mass is greater than about 18 solar masses. Several processes sap the wave's energy. The most important is nuclear fragmentation: the energy of the shock is dissipated in breaking up iron nuclei, lowering the temperature and pressure behind the wave. Protons released by the fragmentation provide opportunities for electron capture, which further reduces the pressure. Once the wave enters a region of density less than 10^{11} grams per cubic centimeter, leakage of neutrinos carries off more energy. As a result of these effects the shock wave may slow to the speed of the material falling through it and make no further progress. Because of the various hazards to the shock, the authors call the region between 100 and 200 kilometers the "minefield."

Neutrino heating is most effective at a radius of about 150 kilometers, where the probability of neutrino absorption is not too low and yet the temperature is not so high that the matter there is itself a significant emitter of neutrinos. The pressure increase at this radius is great enough, after about half a second, to stop the fall of the overlying matter and begin pushing it outward. Hence 150 kilometers becomes the bifurcation radius. All the matter within this boundary ultimately falls into the core; the matter outside, 20 solar masses or more, is expelled.

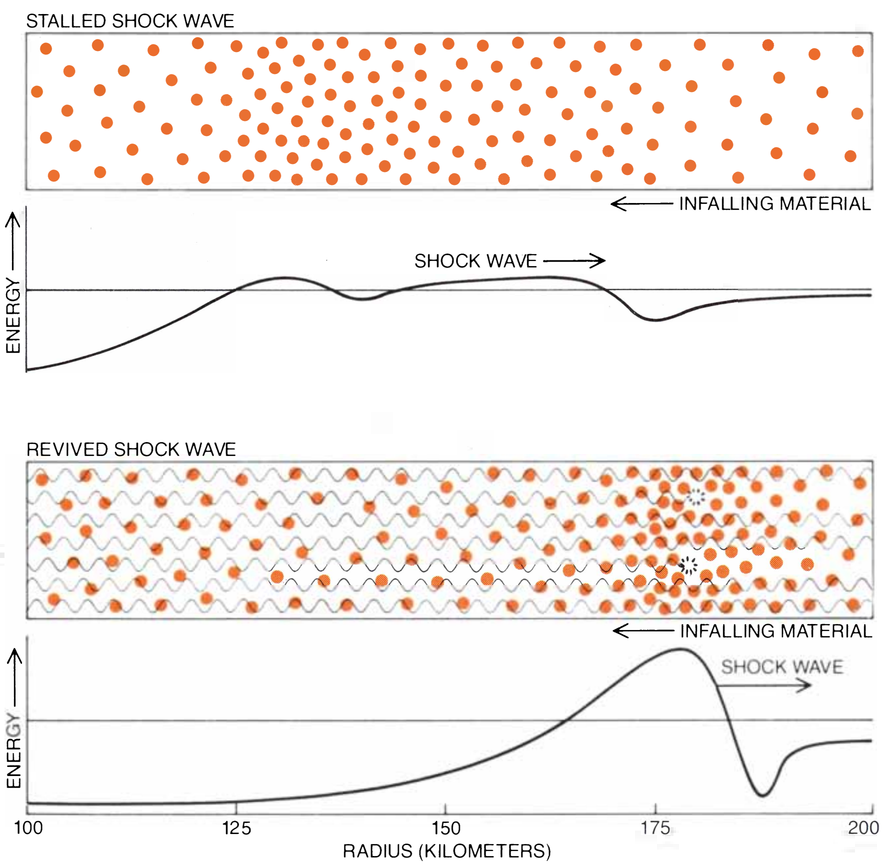
The one group of stars left to be considered are those of from eight to 11 solar masses, the smallest stars capable of supporting a Type II supernova explosion. In 1980 Weaver and Woosley suggested that the stars in this group might form a separate class, in which the supernova mechanism is quite different from the mechanism in heavier stars.

According to calculations done by Nomoto and by Weaver and Woosley, in the presupernova evolution of these lighter stars the core does not reach the temperature needed to form iron; instead fusion ends with a mixture of elements between oxygen and silicon. Energy production then stops, and since the mass of the core is greater than the Chandrasekhar limit, the core collapses. The shock wave generated by the collapse may be helped to propagate by two circumstances. First, breaking up oxygen or silicon nuclei robs the shock of less energy than the dissociation of iron nuclei would. Second, farther out in the star the density falls off abruptly (by a factor of roughly 10 billion) at the boundary between the carbon and the helium shells. The shock wave has a much easier time pushing through the lower-density material.

For a star of nine solar masses Nomoto finds that the presupernova core consists of oxygen, neon and magnesium and has a mass of 1.35 solar masses. Nomoto and Wolfgang Hillebrandt of the Max-Planck Institute for Physics and Astrophysics in Munich have gone on to investigate the further development of this core. They find that the explosion proceeds easily through the core, aided by the burning of oxygen nuclei, and that a rather large amount of energy is released.

Two recent attempts to reproduce the Nomoto-Hillebrandt results have been unsuccessful, and so the status of their model remains unclear. We think the greater compressibility of nuclear matter assumed in the Baron-Cooperstein-Kahana program should be helpful here. Of course it is possible that stars this small do not give rise to supernovas; on the other hand, there are suggestive arguments (based on measurements of the abundance of various nuclear species) that the Crab Nebula was created by the explosion of a star of about nine solar masses.

After the outer layers of a star have been blown off, the fate of the core remains to be decided. Just as gravitation overwhelms electron pressure if the mass exceeds the Chandrasekhar limit, so even nuclear matter cannot resist compression if the gravitational field is strong enough. For a cold neutron star—one that has no source of supporting pressure other than the repulsion of nucleons—the limiting mass is thought to be about 1.8 solar masses. The compact remnant formed by the explosion of lighter stars is well below this limit, and so those supernovas presumably leave behind a stable neutron star. For the larger stars the question is in doubt. In Wilson's calculations any star of more than about 20 solar masses leaves a compact remnant of more than two solar masses. It would appear that the remnant will become a black hole, a region of space where matter has been crushed to infinite density.



REVIVAL OF THE STALLED SHOCK WAVE in heavy stars may be due to heating by neutrinos. Their source is the collapsed core, which radiates the energy equivalent of 10 percent of its mass in the form of neutrinos. Only a small fraction of them are absorbed, but the flux is so intense that many iron nuclei are dissociated. Earlier in the evolution of the supernova the breakup of iron nuclei took energy from the shock wave, but since the process is now powered by external neutrinos, the dissociation no longer decreases shock energy.

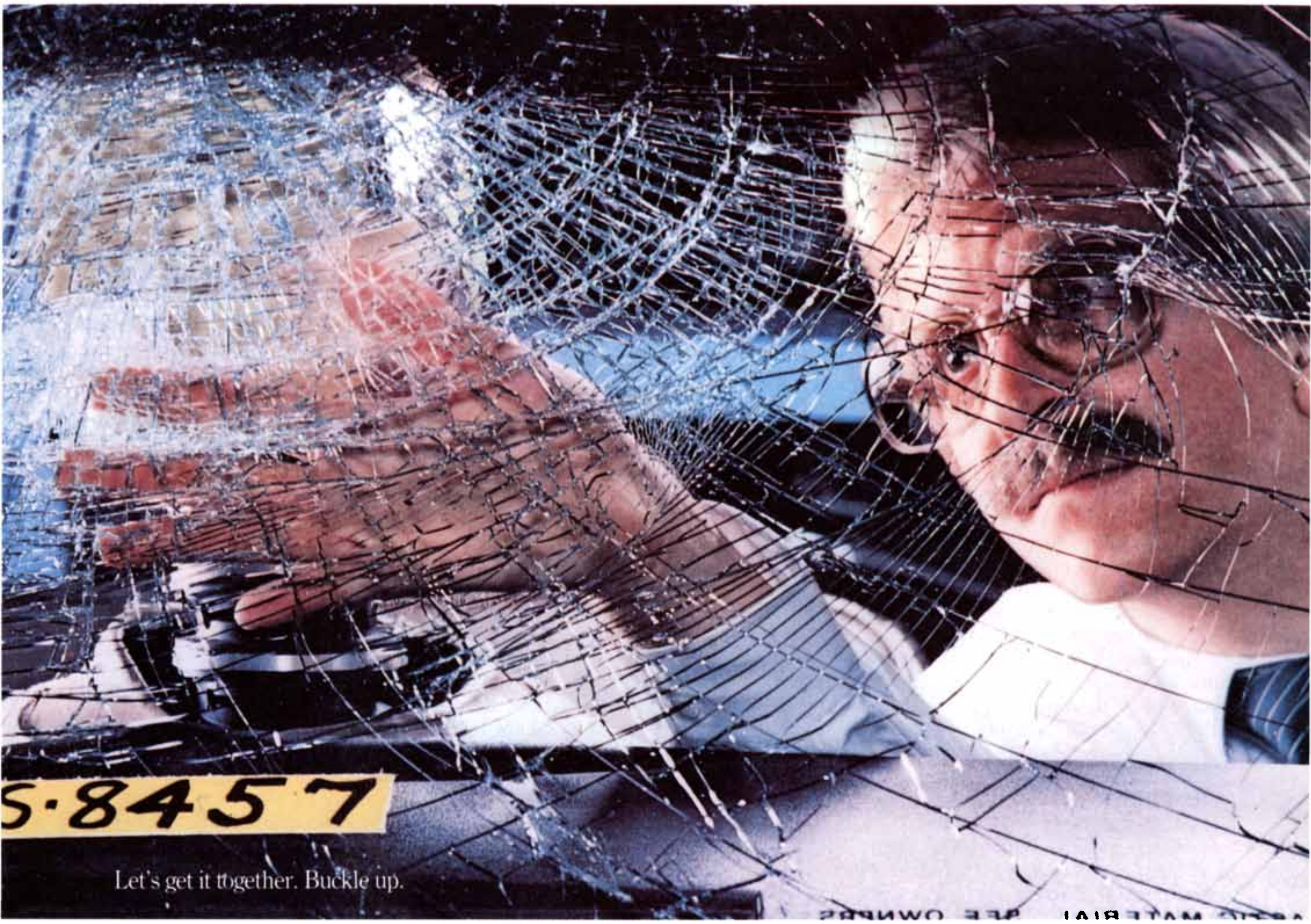
Even if the compact remnant ultimately degenerates into a black hole, it begins as a hot neutron star. The central temperature immediately after the explosion is roughly 100 billion degrees Kelvin, which generates enough thermal pressure to support the star even if it is larger than 1.8 solar masses. The hot nuclear matter cools by the emission of neutrinos. The energy they carry off is more than 100 times the energy emitted in the explosion itself: some 3×10^{53} ergs. It is the energy equivalent of 10 percent of the mass of the neutron star.

The detection of neutrinos from a supernova explosion and from the subsequent cooling of the neutron star is one possible way we might get a better grasp of what goes on in these spectacular events. The neutrinos originate in the core of the star and pass almost unhindered through the outer layers, and so they carry evidence of conditions deep inside. Electromagnetic radiation, on the other hand, diffuses slowly through the shells of matter and reveals only what is happening at the surface. Neutrino detectors have recently been set up in mines and tunnels, where they are screened from the background of cosmic rays.

Another observational check on the validity of supernova models is the relative abundances of the chemical elements in the universe. Supernovas are probably the main source of all the elements heavier than carbon, and so the spectrum of elements released in simulated explosions ought to match the observed abundance ratios. Many attempts to reproduce the abundance ratios have failed, but earlier this year Weaver and Woosley completed calculations whose agreement with observation is surprisingly good. They began with Wilson's model for the explosion of a star of 25 solar masses. For almost all the elements and isotopes between carbon and iron their abundance ratios closely match the measured ones.

In recent years the study of supernovas has benefited from a close interaction between analytic theory and computer simulation. The first speculations about supernova mechanisms were put forward decades ago, but they could not be worked out in detail until the computers needed for numerical simulation became available. The results of the computations, on the other hand, cannot be understood except in the context of an analytic model. By continuing this collaboration we should be able to progress from a general grasp of principles and mechanisms to the detailed prediction of astronomical observations.

Shattering old ideas about safety.



The energy-absorbing steering column, multi-beam headlamps, side-guard door beams and windshield safety glass were GM-pioneered.

Today, we're working to develop new ways to build-in safety.

Our latest effort: the GM "Inner Shield" windshield. It's in limited production now. It's a flexible, two-part plastic layer that's applied to the windshield's inner surface to help protect you against facial cuts and lacerations from broken glass.

Exclusively GM, it's just one more way we are helping to protect the most important part of our cars: You and your passengers.

We believe in taking the extra time, giving the extra effort and paying attention to every detail. That's what it takes to provide the quality that leads more people to buy GM cars and trucks than any other kind.

That's the GM commitment to excellence.

GM

Nobody sweats the details like GM.

Chevrolet • Pontiac • Oldsmobile • Buick • Cadillac • GMC Truck

SCIENCE AND THE CITIZEN

Powering Down

Between 1973 and 1982 energy consumption by U.S. industry, which had increased rapidly since World War II, peaked and then declined. By 1982 (the most recent year for which definitive data are available) the total industrial use of energy was less than it had been in 1973. Since energy conservation was promoted by many public and private organizations in the 1970's, it might be assumed that the saving of energy through increased efficiency was responsible for the decline in the rate of growth. According to a study by Robert C. Marlay of the U.S. Department of Energy, however, only about a third of the deceleration was due to increases in efficiency. The slowing of economic growth in the late 1970's and early 1980's and changes in the composition of industry produced the other two-thirds.

Industry, defined as mining and manufacturing, accounts for about 40 percent of all energy utilized in the U.S. From 1947 to 1973 the amount of energy used by industry grew at a rate varying from 2.7 to 3.6 percent per year. If that rate had persisted, industrial consumption of energy by 1982 would have been more than four billion joules. Actually U.S. industry consumed less than three billion joules in 1982. In Marlay's study, which was published in *Science,* he concludes that of the decline in the rate of growth, 1.4 percent per year was due to the slowing of economic growth, 1.2 percent to increased efficiency in industrial processes and 1.0 percent to changes in the mixture of industrial products.

A change in the distribution of industries has a potent effect on energy demand because some industries consume much more energy per unit of output than others. During the 1970's the declining industries included a disproportionate share of energy-intensive types such as steel production. The industries that grew fastest were those where the greatest investment is in materials rather than in energy, equipment or labor.

Thus the deceleration in the industrial demand for energy reflects in part a fundamental change. The shift away from primary heavy industry toward more technologically sophisticated secondary industry brings with it a reduction in the growth of demand for energy. Since the economy is now growing faster than it was in the early 1980's, the demand for energy may be increasing again. The demand will probably never increase as fast as it did before 1973, however, since the changes in industrial distribution are likely to prove permanent. Marlay concludes: "These observations and the data on which they are based suggest that U.S. industrial production, in both its composition and use of energy, underwent and is perhaps still undergoing a transformation of unprecedented proportions, with implications for both long-term energy demand and a range of other socioeconomic issues."

Death at an Early Age

Adolescent suicide in the U.S. has increased some 300 percent in the past 25 years. During that time there has also been a significant decrease in infant mortality, which means many more infants are surviving adverse conditions in utero and at birth. Are the two trends related? A report in *The Lancet* suggests they may be. It appears to be the case that "individuals whose early life experience included adverse perinatal conditions are more vulnerable to suicide during adolescence."

Lee Salk of the Cornell University Medical College and Lewis P. Lipsitt, William Q. Sturner, Bernice M. Reilly and Robin H. Levat of Brown University studied the prenatal and birth records of 52 adolescents (whose ages ranged from 12 through 19) recorded as having committed suicide in Rhode Island between 1975 and 1983. (There were 43 males and nine females, about the same sex ratio as prevails for suicides in the same age group in the general population.) They also studied two control groups: the 52 individuals, matched for hospital of birth, sex and race, whose birth most closely preceded that of each suicide and the 52 whose birth most closely followed that of each suicide. The suicide group did not differ from either control group in family size, age of parents or father's occupation. For each subject and control the investigators itemized the presence of 36 risk factors affecting the pregnancy and birth and 10 other factors reflecting the infant's condition at birth. The overall incidence of risk factors was found to be significantly higher in the suicide group than it was in either control group.

Salk and his colleagues then isolated the most distinctive risk factors: those in which the incidence was appreciable and at least twice as high in one group as in either of the other groups. There were 10 such "distilled" factors, and the incidence of nine of them was highest in the suicide group. The investigators went on to identify the three risk factors most disproportionately associated with the suicide group. They were chronic disease in the mother (21.2 percent in the suicide group, 0 and 5.8 percent in the two control groups); lack of prenatal care before 20 weeks of pregnancy (30.8 percent as opposed to 3.8 and 11.5 percent), and respiratory distress for more than an hour at birth (19.2 percent as opposed to 7.7 and 3.8 percent). Sixty percent of the suicide victims had experienced at least one of these three risk factors.

A great many infants survive adverse conditions before and soon after birth; few of them commit suicide. Salk and his colleagues are therefore careful not to propose a direct relation between perinatal adversity and eventual suicide. They do, however, conclude that experience before and soon after birth influences the risk of suicide later in life.

Charon's Wake

Astronomers at three observatories have witnessed, for the first time, eclipses of Pluto by its satellite, Charon. The eclipses will continue for several years. Further observations promise to yield better answers to fundamental questions about the outermost planet: How big is it, and what is it made of?

Pluto is on the average 40 times farther from the sun than the earth is, and its faint light is difficult to study even with the most sensitive detectors. Charon, just a bump on the planet's near-infrared image, was not discovered until 1978. Calculations of the satellite's orbit soon showed that the timing of the discovery was extraordinarily fortunate: the orbital geometry is such that eclipses can only be seen from the earth during two brief phases in Pluto's 248-year cycle around the sun, or once every 124 years.

Charon is locked in a synchronous orbit (that is, it is stationary with respect to Pluto) whose period is 6.4 days long; an eclipse of Pluto by its satellite or an occultation of Charon by the planet thus takes place about every three days. The onset of the current series of eclipses was detected in January at the Kitt Peak National Observatory by Edward F. Tedesco of the Jet Propulsion Laboratory and verified in February at the McDonald Observatory by Richard P. Binzel of the University of Texas at Austin and at Mauna

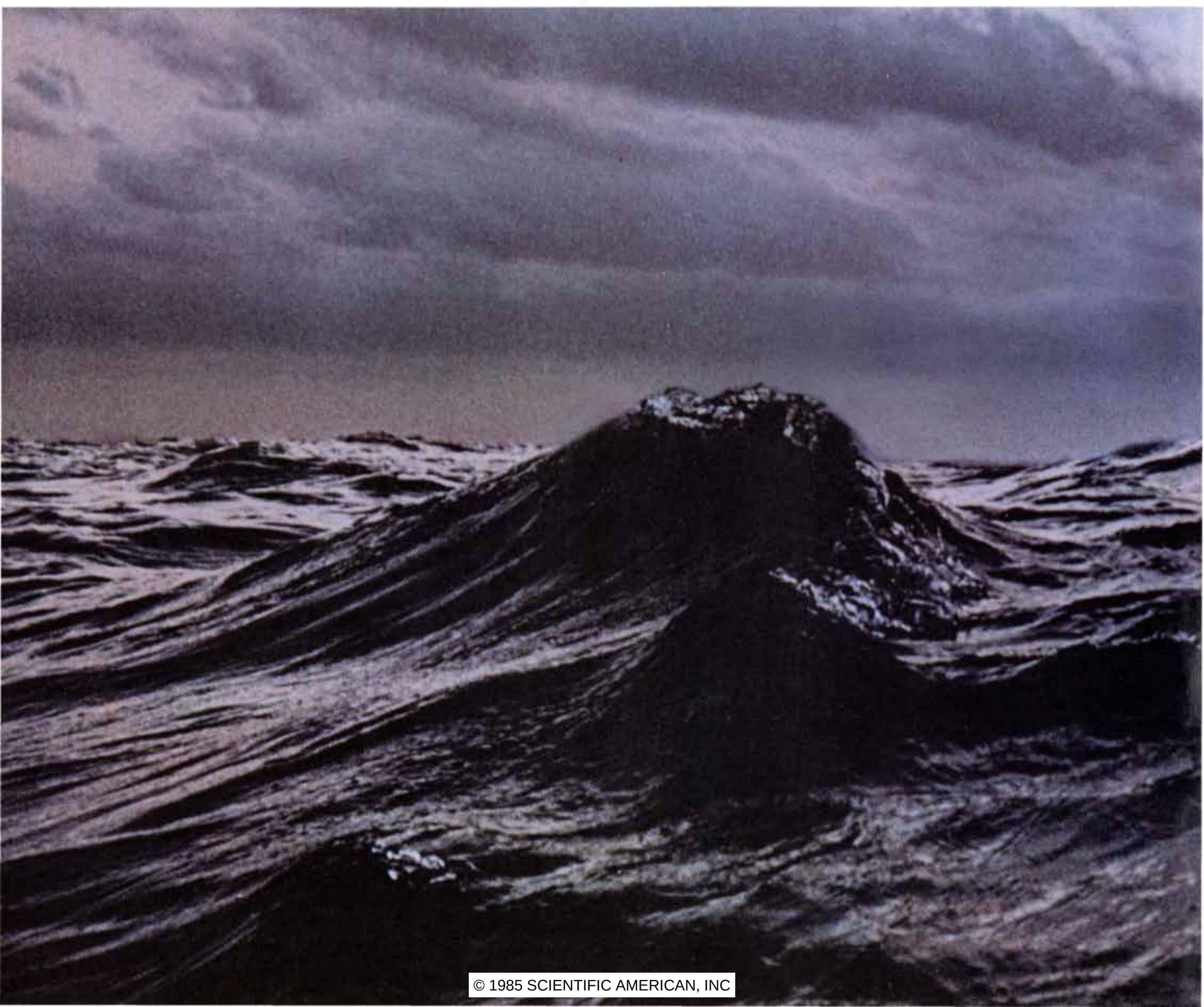

If high technology gives you high anxiety, we cut it down to size.
Stepping into the jungle of confusion surrounding telecommunications is no picnic.
Conflicting advice...demands that you make a decision...it all leads to anxiety. At our NYNEX companies we remove it.
Worried about what equipment to trust? Or who?
Relax. We don't make any of it. So we're free to select the best from different sources.
Can't keep up with new technologies?
We can and do.
By the time you've heard about something new, NYNEX technical experts have already investigated its capabilities, located its weaknesses, and determined its effectiveness.
Worrying about obsolescence?
Don't. As new technologies are approved, we automatically integrate them into our network. So you avoid the downtime needed for an update later on.
No matter how high your anxiety, NYNEX has ways to cut it down. After all, in today's high-tech jungle nobody is more at home than we are.
NYNEX is the parent company of New York Telephone and New England Telephone, plus other companies that offer business communications equipment, cellular services and directory publishing.
For more information, call (800) 441-1919.
In Utah, call (800) 453-9000.
It pays to know us.
© NYNEX Corporation 1985
NYNEX

THE DEFENSE THAT NEVER RESTS.

The Greek word for shield is "Aegis." And that's exactly what RCA gave the cruiser in this picture. A shield that can protect an entire fleet.

In fact, the AEGIS Combat System which RCA designed for the U.S. Navy can simultaneously track hundreds of aircraft and missiles while it maintains continuous surveillance of anything that moves on or above the water.

This makes AEGIS the most advanced missile and air defense system on the high seas. And it makes RCA the first company, ever, to be given total responsibility for the development of a complete shipborne combat system.



with AEGIS. But soon, they'll have a lot of company. RCA is already at work on 11 additional ships' systems.

Which should come as no surprise. RCA is constantly involved in vital electronic and communications work for the armed forces.

RCA

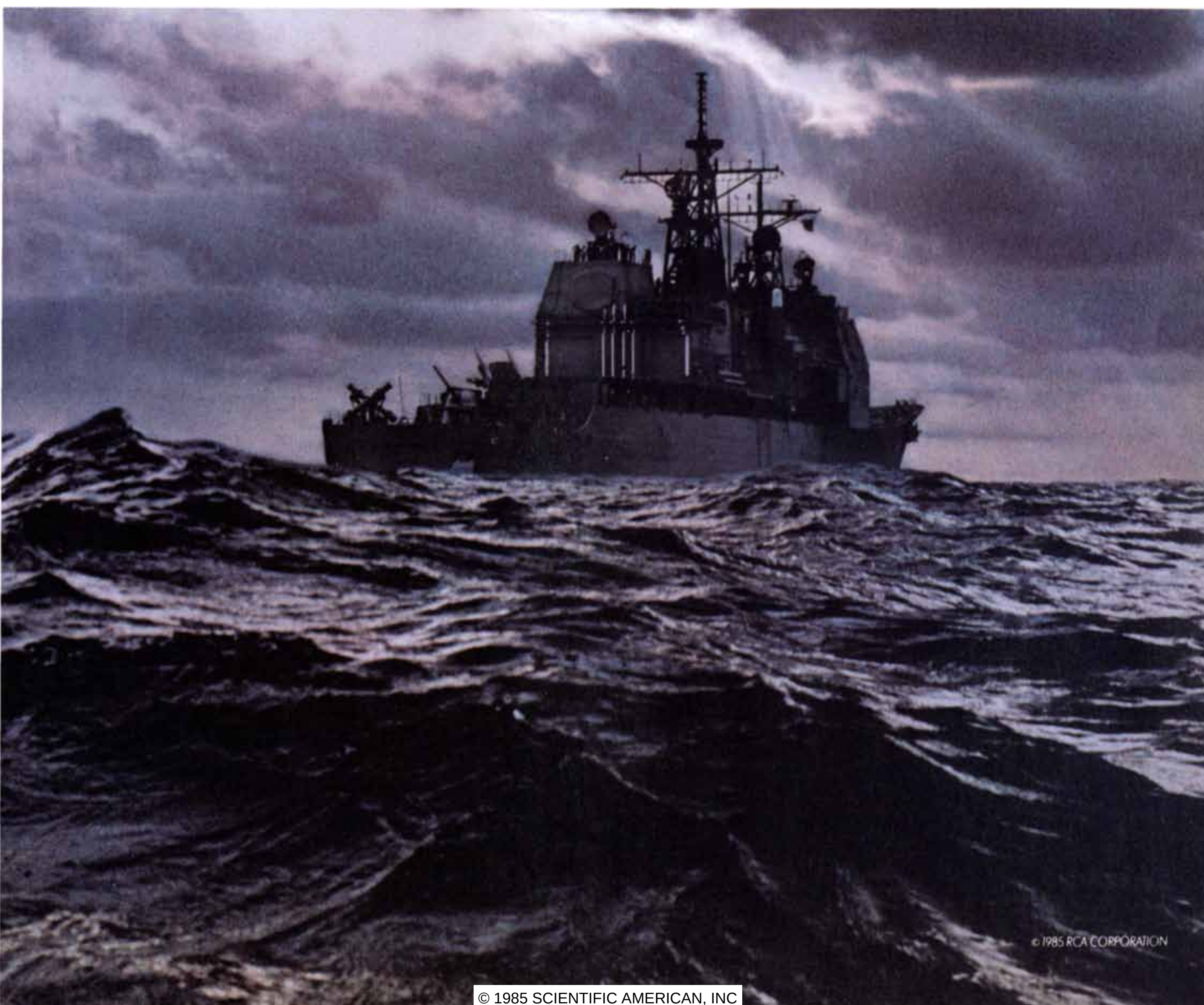
ONE OF A KIND

We're designing and building test equipment for Army tanks and vehicles so the Army can maintain a higher state of readiness.

RCA is also under contract with the Air Force for the Ground Wave Emergency Network (GWEN), which will provide survivable communications throughout the United States in case of a national emergency.

So not only does the defense never rest, neither do we. If you'd like to find out more about what we're working on, just write for "This Is RCA," P.O. Box 91404, Indianapolis, Ind. 46291.

After you read it, you'll know why in electronics, communications and entertainment, RCA is one of a kind.



Kea by David Tholen of the University of Hawaii at Manoa. During the events observed so far the total amount of light emanating from the two bodies has decreased by no more than 4 percent, a small change compared with the planet's normal, cyclical variation in brightness, which is as much as 30 percent. The change is noticeable, however, because Tholen and other workers have over the past few years meticulously plotted Pluto's normal light curve.

Observations of the eclipses will enable astronomers to measure Charon's orbital period and particularly its mean distance from Pluto more precisely. From these two quantities they can calculate (by applying Kepler's third law) the masses of both planet and satellite. The diameters of the two bodies, about which there has been considerable uncertainty, can be determined from the duration of the eclipses and occultations. Mass and diameter together yield mean density, a key indicator of a planet's chemical composition.

Preliminary evidence supports the notion that Pluto is less than 3,000 kilometers in diameter, or significantly smaller than the earth's moon. Spectrographic measurements have identified methane ice as the planet's primary constituent, although there are also less reflective surface regions whose composition is not known. As Charon's shadow sweeps across different regions of Pluto over the next few years, it should be possible to map these dark regions.

Charon itself is thought to be between one-third and one-half the size of Pluto. Its composition is presumably similar, but because its image is difficult to separate from that of Pluto, no direct measurements of its reflectance or of its spectrum have been made. According to Dale P. Cruikshank of the University of Hawaii, workers can now attempt to derive such information indirectly, by "subtracting" measurements made when Charon is hidden by Pluto from those made when both bodies are visible.

Calling Retreat

Five years ago Mark F. Meier and his colleagues at the U.S. Geological Survey predicted that the Columbia glacier, which enters Prince William Sound near Valdez, Alaska, the southern terminus of the Trans-Alaska oil pipeline, would soon begin a drastic retreat. The glacier had been stable throughout this century, but in 1978 its tongue had retreated slightly.

Now the prediction has been confirmed. The forward flow of the glacier is increasing rapidly—but so is the rate of iceberg production: the so-called calving of icebergs from the terminus of the glacier. The production has quadrupled. (On an average day last summer the glacier discharged some 14 million cubic meters of ice.) As a result the glacier is receding at an accelerating pace. In 1984 the retreat amounted to 1.1 kilometers, or nearly twice the total for 1983. Since 1978 the retreat has amounted in places to more than 2.4 kilometers.

Meier made his prediction in response to a request from the Geological Survey. It was important to know whether the Columbia glacier would retreat, advance or remain stable because of the effect these events might have on shipping at Valdez. The investigator based his prediction on the results of a study initiated by a Geological Survey colleague, Austin S. Post. In the early 1970's Post undertook a reconnaissance of all the major glaciers jutting into the sea from Alaska. They seemed not to follow any pattern. One glacier might be advancing; a nearby glacier, exposed to the same local climate, might be stable or even in retreat. As part of the Survey's effort a radar was devised with which the bed of glaciers could be examined through the ice.

Eventually a pattern did emerge: the termini of unstable glaciers lay in deep water. Typically the glaciers were not afloat; still, the presence of seawater and its buoyancy seemed to promote disintegration.

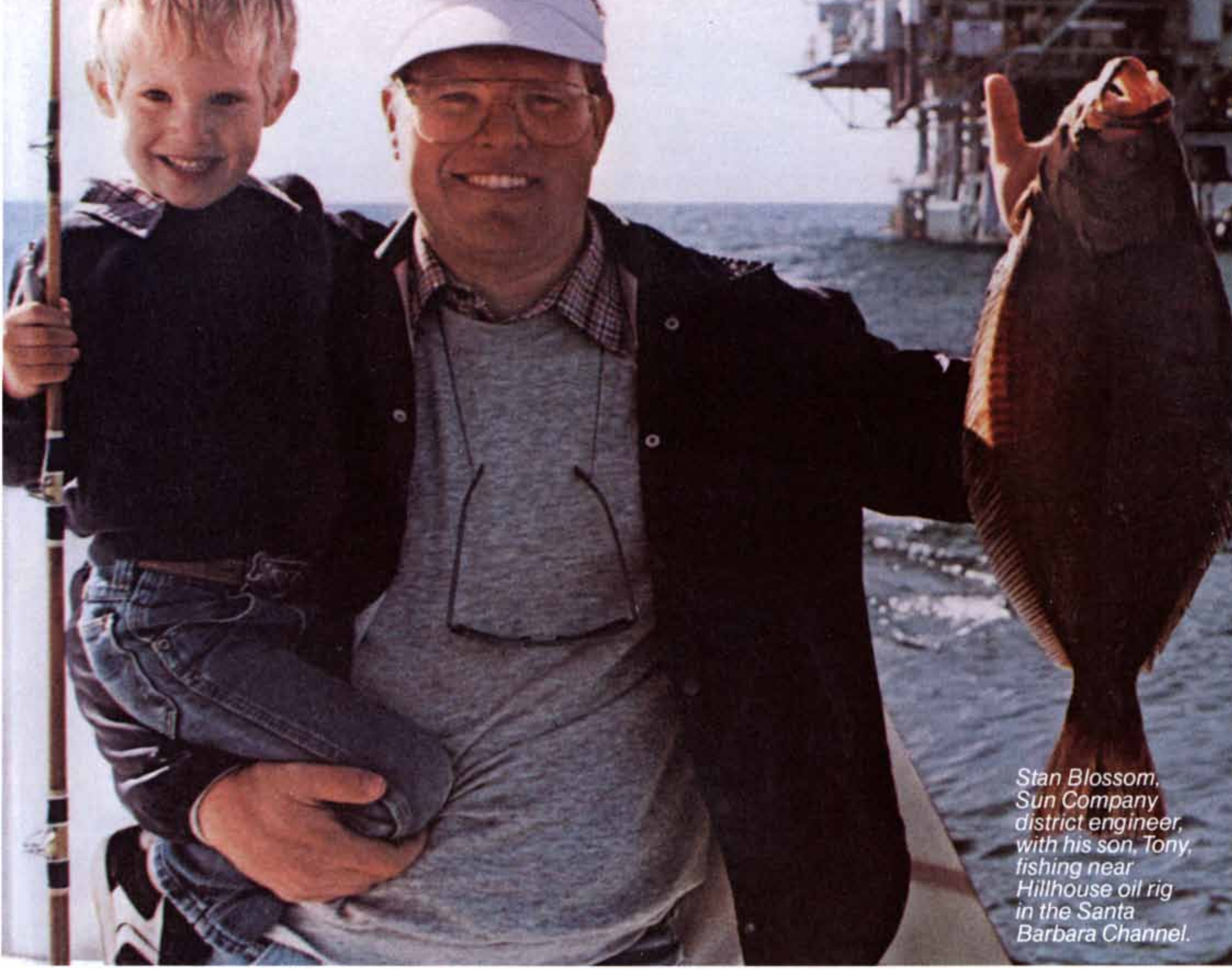
The Columbia glacier proved to be a special case: over thousands of years it had pushed a shoal ahead of its terminal moraine (the mound of debris that precedes an advancing glacier); thus even a slight retreat would put it in deeper water, making it unstable. In 1978 that retreat happened.

Meier expects a further increase in the rate of the glacier's disintegration, perhaps modulated briefly by a short-term winter advance due to the decreased calving of icebergs in winter. The icebergs themselves may not block shipping lanes; a submerged ridge in Prince William Sound keeps large icebergs from floating out to sea. Meier predicts that during the next few decades the Columbia glacier will retreat 20 to 25 miles, or half of its current length. The retreat will thereby expose a fjord that has been covered by ice for centuries. The repopulation of the fjord will be of interest to ecologists. Meanwhile glaciologists should gain insight into how glaciers move.

An iceberg calving from the Columbia glacier's 250-foot cliff

Copping Z's

The heavy, short-lived particles of the weak force, labeled W^+, W^- and Z^0, are as difficult to study in detail as they are important in contemporary efforts to understand the nature of matter. These particles mediate the radioactive decay of nuclear particles, and they relate this and other so-called weak interactions to electromagnetic interactions. Yet the W^+, W^- and Z^0 particles have been detected only a few dozen times, far too infrequently to explore their properties in detail. The scarcity of Z particles is particularly acute: they appear only



Stan Blossom, Sun Company district engineer, with his son, Tony, fishing near Hillhouse oil rig in the Santa Barbara Channel.

SUN, OIL AND THE FISHERMAN'S SON. People get a lot of fish from the Santa Barbara Channel. And this is where Sun Company gets a lot of oil—more than three million barrels a year! Around here oil is our business. But Sun people like Stan Blossom can tell you it's not our only concern.

"I work for Sun, but I also fish in these waters. And I want my son's generation to fish here too. That's why precautions like our containment unit are so important. In case of a spill, it would quickly surround the oil and help prevent damage.

"This is just one of the ways we're helping to make sure there'll be many more years of fishing around here."

At Sun, we think putting energy back into the environment is as important as getting it out.

WHERE THERE'S SUN THERE'S ENERGY.


HEWLETT PACKARD
Integral Personal Computer

one-tenth as often as W particles do.

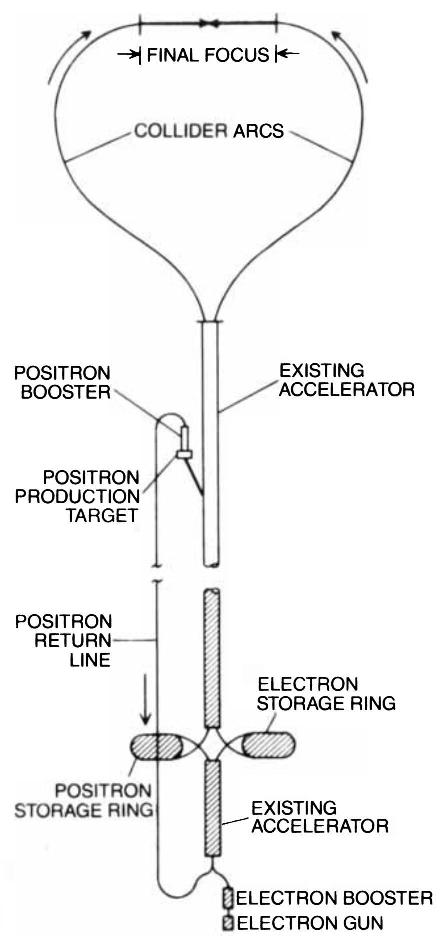
Two projects are under way in an attempt to remedy the problem. At CERN (the European laboratory for particle physics, where the particles were detected) the LEP (Large Electron-Positron collider ring) is being constructed; its completion is scheduled for as early as 1988. It is a ring-shaped electron-positron accelerator, 27 kilometers in circumference, straddling the border between France and Switzerland. The LEP will be the largest ring-shaped accelerator ever built for electrons and positrons. Indeed, it may prove to be the largest such ring of all time: particles constrained to move in a circle lose energy by radiation, so that increasing circles yield decreasing benefits.

At Stanford University the two-mile-long linear accelerator is being modified so that it too can function as a colliding-beam machine. Stanford investigators hope to have the device generating Z's by October, 1986.

Each accelerator will be able to produce thousands of Z^0 particles per day because each should be able to accelerate electrons and positrons so that they collide and release an amount of energy precisely equivalent to the mass of an intermediate vector boson. In the case of the Z^0 this so-called resonance lies at 94 GeV (billion electron volts).

At CERN the energy is achieved by accelerating counterrotating beams of positrons and electrons and then causing them to collide in a detection chamber. The Stanford machine takes a more complicated approach. At the start of each cycle two small storage rings will have been stocked, one ring with two "bunches" of 50 billion electrons, the other with two bunches of positrons. A positron bunch will be extracted, compressed to a length of about a millimeter and injected into the linear accelerator. Behind it will come the two electron bunches, similarly compressed, at a spacing in the accelerator of some 17.6 meters.

The three bunches will be accelerated; then, two-thirds of the way down the accelerator, the trailing electron bunch will be steered into a metal target, where it will create a burst of positrons. (Positrons must be produced; they do not exist in nature.) The remaining two pulses will reach the end of the accelerator. There the positron bunch will be directed magnetically into the north branch of a pair of accelerator tunnels that meet like pincers; the electron bunch will be directed into the south branch. At the meeting of the pincers the two beams will collide. Meanwhile the newly produced bunch of positrons will have been accelerated, sent to the start of the accelerator



The Stanford Z^0 factory

and placed in the positron storage ring, to be employed in a subsequent cycle. The entire cycle will take less than .006 second.

The Stanford investigators hope that the production of "thousands of Z^0's per day" will be yielding insights into the fabric of the universe by early in 1987, thereby stealing a cosmic march on their colleagues across the Atlantic.

Winging It

A thin ring, 13 inches in diameter, has captured the world's record for thrown flight. It covered 349 yards (319 meters) in an open area adjacent to the Rose Bowl. The device, called the Aerobie, was invented by Alan Adler, a lecturer in engineering at Stanford University. Scott Zimmerman, a college student who had practiced assiduously for the event, made the record throw.

Even a less accomplished athlete, according to Adler, could expect to throw an Aerobie two or three times as far as he or she could throw a Frisbee. A liplike structure on the upper, outermost edge of the Aerobie accounts for its superior performance. This "spoiler" helps to reduce the lift generated

The Aerobie: in flight, at rest and in cross section

at the leading half of the ring and so brings it into balance with the trailing half (which flies in the downwash produced by the leading half). Without the spoiler the ring would drift to the left (if it was spinning clockwise as seen from above); the spoiler helps it to fly straight.

Adler set out to design "a flying disk or ring that had the lowest possible aerodynamic drag consistent with stable flight." He arrived at the design with the help of a personal computer. The resulting "wing" consists of an inner core of plastic and a surrounding rubber cushion. The Aerobie's inner diameter is 10 inches; it is .15 inch thick and weighs 112 grams, or slightly less than four ounces.

The precise length of Zimmerman's record throw was 1,046 feet 11 inches, vouched for by the archivist of the U.S. Disc Sports Association. Zimmerman also holds the record for the Frisbee: 456 feet (152 yards). An average person with no special training could throw a Frisbee from 35 to 40 yards; with an Aerobie he or she could expect to achieve between 120 and 150 yards.

Neural Constancy

During gestation the developing human brain generates neurons at an average rate estimated at more than 250,000 per minute. The estimate is based on the assumption that virtually no neurons are produced after birth, even though they are constantly dying. This assumption has been dogma in neurobiology circles, but as Pasko Rakic of Yale University notes in a recent issue of *Science,* it is dogma that "has never been tested in the adult of any primate species," much less in human beings. Moreover, laboratory studies in the past few years have indicated that neurogenesis does occur in other adult vertebrates, including certain fishes, amphibians and birds.

Rakic has now reported the first direct evidence that, as far as primates are concerned, the dogma is accurate. He injected 12 rhesus monkeys, ranging in age from six months to 11 years, with radioactive, tritium-labeled thymidine. Thymidine is a precursor of one of the four nucleotide bases found in DNA and is incorporated only into DNA, not into other parts of a cell. The presence of radioactive thymidine labels a cell's genetic material, and thus the cell itself, as having formed after the thymidine was injected.

Rakic estimates he screened about 100 million cells from all major structures of the brain of each of the 12 monkeys. He found not a single labeled cell resembling a neuron. In contrast, radioactive thymidine was readily detectable in cells known to proliferate throughout adulthood, including the glial cells that fill the space between neurons in the brain. Rakic therefore concludes that the brain of an adult monkey is probably incapable of forming new neurons. Furthermore, what is true for rhesus monkeys is likely to be true for human beings, whose brain is very similar in organization and is thought to follow the same sequence of development.

The notion that neurogenesis ends no later than a few months after birth has profound implications for understanding how the primate brain works. Primates are distinguished from many other animal species not only by their longevity but also by a prolonged learning period preceding maturity. The physical basis of learning is widely assumed to be the setting up of patterns of connections between neurons in the brain. To retain what it has learned, Rakic argues, an animal needs a stable set of neurons.

The evidence for a strict limit on neurogenesis also suggests that the prospects for treating brain and spinal-cord injuries, which at present are generally irreversible, will remain slim. On the other hand, the unique property of neurons—they are the only cells in the adult human body that do not proliferate—makes them immune from cancer. (Brain tumors in adults involve cells other than neurons, such as the glia.) According to Rakic, some workers think there must be a mechanism inhibiting the replication of DNA in neurons. If such a mechanism could be identified, it might lead to a treatment for the malignancies that plague other parts of the body.

Grim Reckoning

When a smoker dies of emphysema, the cause is entered in the official statistics under chronic obstructive pulmonary diseases rather than under smoking. Deaths from other diseases associated directly or indirectly with smoking, alcohol or drugs are also usually attributed to the disease rather than to the underlying

Give people the tools they need, and there is no limit to what they achieve.

Finally, one word processor that has it all.

State-of-the-art and simplicity.

At last, a word processing program that lets you do everything you need a word processor to do.

Yet is so easy, anyone in the office can use it.

New WordStar 2000 is loaded.

With Windows. Undo. Format sheets. Five-function math. A built-in spelling corrector.

WordStar 2000 Plus also lets you manage mailing lists. Even use electronic mail.

Easy-to-learn. Easy-to-use.

With new WordStar 2000, the keys you press are the keys you'd expect to press ("c" for copy, "p" for print, etc.).

Results? They're easier to learn and easier to remember.

Now you have two choices.

New WordStar 2000 becomes the second standard for word processing. Choose WordStar (or WordStar for PCjr) if the original standard suits your needs or budget better.

See your local MicroPro dealer or call (800) 227-6703 [in California (800) 632-7979] for the dealer nearest you.

MicroPro.

WORDSTAR® 2000

Now there are no limits.™

Now available for IBM PC/AT*/XT* and compatibles with 256K RAM. Attention WordStar owners: Special WordStar 2000 offer just for you. Call your dealer or MicroPro Customer Update (800) 227-5609.*

cause. On the basis of a recent survey R. T. Ravenholt of World Health Surveys, Inc., asserts that if underlying cause is taken into account, the abuse of tobacco, alcohol and drugs must be reckoned as the leading cause of death in the U.S.

Ravenholt writes in *Population and Development Review* that "the addictive use of many psychoactive substances—tobacco, alcohol, heroin, cocaine, marihuana, stimulants, hypnotics and hallucinogens"—accounted for about 630,000 deaths in 1980, or nearly a third of all deaths from all causes. He attributed "25 percent to the smoking of tobacco, 5 percent to the drinking of alcohol and 1–2 percent to the abuse of other addictive substances."

Ravenholt arrived at his estimates by examining a number of statistical studies of death and disease rates among smokers compared with nonsmokers. For example, if the relative risk of cancer among nonsmokers is set at 1, the risk among smokers is 2.72. Since 53 percent of men are nonsmokers and 47 percent are smokers, it follows that the number of excess deaths from cancer among men in 1980 was 101,000, or 44.8 percent of all male cancer deaths (225,949). The comparable figure for deaths among females was 34.6 percent.

By similar calculations Ravenholt estimated that smoking accounted for 24 percent of deaths from cardiovascular disease and contributed to death from respiratory diseases other than lung cancer, from digestive diseases, from external causes of injury (as in fires) and from "miscellaneous and ill-defined diseases." He also attributes some infant mortality to smoking by the mother during pregnancy.

Ravenholt summarized his estimates on tobacco: "The total number of excess deaths in the U.S. in 1980 attributable to smoking cigarettes was approximately 485,000. Addition . . . of deaths caused by the smoking of pipes and cigars, the passive inhalation of environmental tobacco smoke and the chewing and snuffing of tobacco probably raises the total . . . to more than a half million—more than one-fourth of all deaths from all causes (1,989,841)." By similar methods he ascribed 100,000 deaths to the effect of heavy consumption of alcohol and 30,000 to the abuse of other drugs.

Asian Roots

Were Adam and Eve Chinese? After analyzing mitochondrial DNA of contemporary ethnic groups, Douglas C. Wallace of the Emory University School of Medicine and

IT'S GETTING TOUGHER FOR NANCY RAY TO BE A MOTHER.

Nancy Ray has multiple sclerosis.

Every day it becomes harder for her to see. Harder for her to walk. Harder to do the things she wants to do for her family.

Four years ago, Nancy Ray was young and healthy. She was a mother with a four-year-old son, a teacher, a writer for educational TV.

Then, like about 200 young adults every week, she learned she had MS.

Multiple sclerosis destroys the myelin covering that surrounds nerve fibers in the brain and spinal cord. Messages to and from the brain get lost as they travel through the damaged areas. So your eyes, arms, legs don't do what you want them to do.

MS can be mild. Or it can be severe. Once you've got it, you've got it for life. Most days, Nancy Ray can get around with a walker. With a lot of time and effort, she can even prepare meals for her family. Some people with MS aren't that fortunate.

The National Multiple Sclerosis Society is helping people with MS live with MS. And funding research for a cure, all over the world. We can do it, but only with your help.

Please give generously. It's tough enough to be a mother without MS.

NATIONAL MULTIPLE SCLEROSIS SOCIETY
205 East 42nd Street
New York, New York 10017

JUST WHEN YOU'RE STARTING TO LIVE, MS CAN STRIKE.

his co-workers have hypothesized that all the groups have common ancestors who lived in Asia between 50,000 and 100,000 years ago.

Mitochondria are organelles in the cytoplasm of the cell that convert energy from nutrients into a form useful in many cellular reactions. They have their own DNA: a double-strand loop that includes about 40 genes. Mitochondrial DNA is rapidly becoming a fundamental tool of inquiry into molecular evolution because it provides clearer information about human lineages than the chromosomal DNA found in the nucleus.

Mitochondrial DNA is simpler to deal with mainly because it is maternally inherited: all the DNA in the mitochondria comes from the mother by way of the cytoplasm of the egg. In contrast, chromosomal DNA comes from both parents in equal proportions. Moreover, in the course of sexual reproduction chromosomal DNA is subject to shuffling processes that complicate the evolutionary record.

Wallace and his co-workers took advantage of the relative simplicity offered by mitochondrial DNA to reconstruct the genetic relations among Bantus and Bushmen in Africa, a sample of Asians (consisting mostly of mainland Chinese), a sample of people of western European descent and Pima and Papago Indians from the southwestern U.S.

Mitochondrial DNA from blood cells was cleaved by restriction endonucleases. These enzymes cut a DNA chain wherever a specific sequence of from four to six nucleotides (the subunits of DNA) appears. The sequence, which is specific to a particular enzyme, is known as a restriction site. A site can be destroyed or created by a mutation changing a single nucleotide in the DNA chain. Hence observing the pattern of restriction sites can yield information about mutations.

The restriction sites themselves are found by observing the pattern of fragments produced by the restriction enzyme. For example, suppose the DNA of two ethnic groups is known to contain many of the same nucleotide sequences. The application of a particular restriction enzyme might cleave the DNA of one group into four fragments and the DNA of the other group into only three. It could then be assumed that the two DNA's differ in one restriction site as the result of a single mutation in the course of evolution.

Such comparisons can indicate the degree of relation among the various mitochondrial DNA's. That information in turn makes it possible to construct a genealogical tree for the groups. In *Journal of Molecular Evolution* and *Endocytobiology* Wallace has published trees for the ethnic groups in his sample. The most probable root, or origin, of the trees is mitochondrial DNA's found predominantly in Asia. Because the mutation rate for mitochondrial DNA is not known precisely such trees provide only very rough estimates of when the various branches diverged. Wallace is currently trying to define a more precise chronology.

Missing Mass

There is more to the universe than meets the eye, however powerfully it is aided by astronomical instruments. On both observational and theoretical grounds astronomers and astrophysicists believe the universe possesses far more mass than is apparent.

To account for the motion of stars within galaxies and of galaxies within clusters of galaxies it is necessary to suppose these objects move under the gravitational influence of more mass than can be detected directly. In fact, dynamical calculations based on the motions indicate that as much as 10 times more "dark matter" than luminous matter may be associated with galaxies and clusters.

What is the nature of the dark matter? New and exotic theories are yielding fresh and sometimes controversial answers to this question. Edward Witten of the Institute for Advanced Study in Princeton has proposed a possibility called strange matter. It would differ from protons and neutrons in consisting not of up and down quarks bound in threes but of nearly equal proportions of up, down and strange quarks bound in aggregations of unlimited size. Strange matter would be extraordinarily dense; scattered lumps of it concentrated around galaxies could solve the missing-mass problem.

Hypothetical particles that are the result of efforts to advance the theoretical understanding of matter have also been put forward as the embodiments of missing mass. One is the axion, an exotic particle required in certain grand unified theories. The axion would interact only rarely with ordinary matter; although light, axions would be abundant enough to account for the missing mass. Another theory, known as supersymmetry, which relates every known particle to a partner that has a different spin and a much greater mass, supplies as a candidate the photino, the supersymmetrical counterpart to the photon. A third possibility is suggested by superstring theory, which ascribes as many as 11 dimensions to spacetime and postulates particles more than 10^{19} times as massive as the proton.

A more mundane proposal is that galaxies are surrounded by large quantities of ordinary matter in forms that are difficult to detect, such as brown dwarfs, red dwarfs or ionized gas. This notion runs afoul of the observed abundance of deuterium (an isotope of hydrogen). If the universe contained enough ordinary matter to account for the observed galactic and stellar motions, the density of the early universe would have been great enough to cause most deuterium nuclei to fuse into heavier elements. Victor E. Viola, Jr., of Indiana University recently confirmed that finding with an analysis of the cosmological abundance of the isotope lithium 7. In a universe containing the necessary amount of ordinary matter far more lithium 7 would have been formed than is observed.

Another venerable hypothesis is still being tested. It holds that neutrinos, which are enormously abundant in galaxies, have enough mass to explain the observed motions. Two years ago a Soviet group concluded from a study of the energy spectrum of the electrons emitted during the decay of tritium (an isotope of hydrogen) that the electron neutrino has a mass greater than 20 electron volts. That result has met with some skepticism; at least 12 different groups in this country and elsewhere are trying to duplicate it.

Mass appears to be missing on theoretical as well as on observational grounds. Specifically, recent models of the early evolution of the universe known as inflationary theories require the current mass density to be very near the critical value for a closed universe: the density at which gravity will ultimately halt the expansion of the universe and cause it to collapse again. The necessary amount of mass is some 20 times more than is visible. Dynamical constraints allow at most half of that mass to be associated with galaxies and clusters; the remaining 50 percent must be distributed uniformly throughout space.

An even cosmological background of neutrinos that move at nearly the speed of light and result from the decay of more massive neutrinos has been proposed by Duane A. Dicus of the University of Texas at Austin, Edward W. Kolb, Jr., of Fermilab and Vigdor L. Teplitz of the Arms Control and Disarmament Agency. Although nearly massless at rest, the neutrinos could have an energy equivalent to enough mass to close the universe. In a scenario put forward by Nicholas Kaiser of the University of Cambridge much of the mass in the universe takes the form of dark galaxies, which are composed of matter too tenuous to shine.

Cheese

Each of the 2,000 varieties is made from mammalian milk in a process consisting of some nine steps. The measures taken at particular steps determine the variety produced

by Frank V. Kosikowski

One could go around the world counting varieties of cheese and come up with a total approximating 2,000. Yet they represent different characteristics emanating from only about 20 basic varieties. Moreover, the 20 basic varieties rest on a single fundamental process. All cheese begins with mammalian milk, usually but by no means always from a cow. Acid or rennet turns the milk into a curd, and the free whey is removed. What happens after that determines whether the product that comes to the table is cottage cheese, a Cheddar, an Emmentaler (Swiss) or something else on the spectrum of varieties.

Cheese is thought to have originated in southwestern Asia some 8,000 years ago. The Romans encouraged technological improvements and stimulated the development of new varieties during their invasions of Europe between about 60 B.C. and A.D. 300. Their influence is reflected in etymology: *caseus,* the Latin for cheese, is the root of the modern name.

Cheese can be broadly and simply classified into two groups: fresh and ripened. Fresh cheese is made from milk coagulated by acid or high heat and must be eaten quite soon after it is made or it will spoil. Cottage cheese is the most familiar example, but the category also includes cream cheese, Neufchatel, ricotta and mozzarella.

Ripened cheese is made from milk fermented by lactic acid bacteria and coagulated by an enzyme preparation. The curd is pressed to remove the whey and is salted; then it is held for an extended period in a controlled environment. During that time various physical and chemical changes take place to give the material the characteristic flavor and texture that cause it to be classified as a distinct variety of cheese. The great majority of cheeses are ripened. Much ripened cheese is sold without further treatment, but considerable quantities are instead ground, heated and emulsified with sodium phosphates and other salts to make processed cheese.

Varieties of Milk

Although cows provide most of the milk for cheesemaking, many parts of the world rely more on other animals. In southwestern Asia and along the Mediterranean sheep and goats constitute the major source. France has more than a million milking goats and large numbers of sheep, whose bluish white milk goes mostly into the manufacture of Roquefort cheese. Among the animals providing milk for cheese in other parts of the world are water buffalo, camels, yaks, reindeer and llamas.

Indeed, the milk of almost any mammal could be made into acceptable and perhaps unique cheese, but the amounts collectible or accessible limit the opportunities. How does one milk a guinea pig or a 100-ton whale careering through heavy seas? (Barbour L. Herrington of Cornell University, who was studying the composition of milk from small mammals, actually designed a successful milking machine for guinea pigs some years ago, but the achievement did not lead to a guinea-pig cheese because of the great number of animals that would have been needed to yield enough milk for even one small wheel.)

Just as the variety of grape influences the flavor and bouquet of wine, so the mammalian origin of the milk influences the flavor and aroma of a natural ripened cheese. Goat's milk gives cheese a spicier and more piquant flavor than cow's milk primarily because its fat has more caproic, caprylic and capric acid. These are fatty acids, distinguished from one another by the length of their carbon side chains: six carbons for caproic, eight for caprylic and 10 for capric acid. Each kind of side chain contributes a different pungency to the milk. Goat's milk contains twice as much caproic acid as cow's milk, three times as much caprylic acid and five times as much capric acid.

Sheep's milk yields cheese of a distinctive flavor because the milk contains six times the caprylic acid of cow's milk and twice that of goat's milk. Furthermore, it has only half the capric acid content of goat's milk. Little of this is noticeable in the taste of the fresh milk. The characteristic flavors appear only when a young cheese made from one of these milks is ripened and fatty acids are released from the fat molecules by lipase enzymes.

The milk also influences the color of natural cheese. The milk of sheep, water buffalo and some goats has little or no beta carotene, a yellow pigment, consequently the cheeses made from them are basically white. Cow's milk has variable amounts of beta carotene depending on the season, the breed of cow and what the cow eats; the natural color of the resulting cheese ranges from straw to buttercup yellow.

Role of Microorganisms

The ripening of cheese is the work of a large number of microorganisms, one of the largest concentrations to be found in any of the basic foods. On the first day of a cheesemaking process the number in the starting material ranges from one to two billion per gram. Thereafter the population declines because of insufficient oxygen, high acidity and the presence of inhibitory compounds that are produced as the cheese ripens. Fortunately the ripening organisms are safe and perhaps beneficial. It is largely the action of their cellular enzymes on lactose, fat and protein that creates the ripened-cheese flavor.

Originally the bacteria and fungi that start the fermentation leading to cheese arrived in the milk on their own, having migrated through the air from their natural habitats in plants

ROQUEFORT CHEESE ripens in one of the caves of the village of Roquefort in southern France. The semisoft, blue-mold cheese is made from sheep's milk and formed into wheels weighing about 2.5 kilograms (5.5 pounds) each. In order to be called Roquefort it has to be manufactured in this way and ripened for several months in one of the village's limestone caves. In France similar cheeses made from other types of milk or ripened elsewhere are called Bleu. The ripening organism is the fungus *Penicillium roqueforti*.

and soils. Between 1890 and 1920 several workers in Europe and the U.S. isolated pure cultures of these microorganisms. For example, the late microbiologist J. M. Sherman of Cornell isolated and cultured a gas-forming microorganism, originally named *Propionibacterium shermanii,* that is essential for giving Swiss cheese its eyes, or holes, and flavor.

It soon became evident that the addition of pure cultures to raw milk of poor quality suppressed spoilage organisms and improved the quality of the resulting cheese. Later, when it became standard practice to heat or pasteurize cheese milk, eliminating most microorganisms, pure cultures were essential to provide the necessary types and numbers of bacteria.

Improvements in the technology of

CHEDDAR CHEESE is made at the plant of Great Lakes Cheese of New York, Inc., in Adams, N.Y. Once the cow's milk has been prepared by the addition of lactic acid bacteria that promote fermentation and of a coloring material, rennet (an enzyme preparation) is added to form a curd (*1*). After about 30 minutes the curd is cut by wire knives (*2*) to increase the surface area. The curd cubes are cooked (*3*) for about an hour and become drier through the loss of whey (*4*). Then they are raked, compacted and turned repeatedly

culturing bacteria have made possible frozen, concentrated starter cultures. Such a culture contains about 400 billion lactic acid bacteria per gram. They begin growing immediately in warm milk, and so they can be added directly to cheese vats without prior culturing. Moreover, since cultures are selected beforehand for their insensitivity to bacteriophage (viruses that can invade and destroy bacterial cells, bringing fermentation to a halt), they make cheesemaking easier and more predictable.

In many varieties microorganisms play a further role as the cheese ripens. Added to the surface of the curd or injected into it, they act to produce the flavor and the texture that identify a specific variety of cheese.

Whatever variety the cheesemaker

(*5–9*) in the crucial cheddaring step, which helps to create the characteristic texture of this cheese. The resulting slabs are milled (*10*). The cheese is salted (*11*), wrapped in cloth (*12–14*) and pressed in forms called hoops (*15*) to expel the remaining whey. When it is removed from the forms and packed in boxes, the immature cheese is cured for from two to 12 months at a temperature between 36 and 50 degrees Fahrenheit, being sampled often (*16*). About two-thirds of the factory-made ripened cheese produced in the U.S. is Cheddar.

sets out to produce, the process usually entails nine steps: (1) Preparing the milk, (2) forming a curd, (3) cutting it, (4) cooking it, (5) separating the whey, (6) salting the residue, (7) applying special microorganisms, (8) pressing the curd and (9) ripening the young cheese. In general it is the amount of emphasis put on some of these steps that leads to a desired cheese.

As a rule the milk for making a ripened cheese is raw or underpasteurized. Fully pasteurized milk also serves, but the underpasteurized kind is the commoner choice. The enzymes from the microorganisms that survive the lower temperature give rise to a better flavor in the cheese. In the U.S. a ripened cheese made from raw or underpasteurized milk must be held for at least 60 days. During that time the salt, the acidity, the metabolic compounds of ripening and the absence of oxygen usually destroy any food-poisoning organisms.

One of the first steps in preparing the milk may be to add a coloring material. The coloring agents include beta carotene and extracts of seeds and plants. Annatto, a yellowish red dyestuff made from the pulp of the fruit of the tropical tree *Bixa orellana,* is one such extract, paprika another.

Next the starter culture is added. The cultures for most ripened cheeses contain bacteria that produce only lactic acid, which is derived from lactose, a sugar of milk, and performs many essential functions. Different cultures produce widely different amounts of lactic acid. The amount of lactic acid strongly influences the flavor and texture of the cheese as well as the extent to which eyes are formed in the cheese.

For many natural ripened cheeses the starter culture consists of bacteria that grow well at moderate temperatures (from 20 to 37 degrees Celsius, or 68 to 98.6 degrees Fahrenheit), such as *Streptococcus cremoris* and *S. lactis.* For cheeses such as Emmentaler whose curds are cooked at relatively higher temperatures the bacteria of the starter culture must be able to grow well at 37 degrees C. or higher; the choice then is *S. thermophilus, Lactobacillus bulgaricus* or *L. helveticus.*

Forming the Curd

Once it is prepared, milk is transformed into a smooth, solid curd by a coagulating enzyme. The enzyme is chymosin, more commonly known as rennin. It comes from rennet, which used to be obtainable only as an extract from the fourth stomach of a calf. Rennet extracted from fungi such as *Mucor miehei, M. pusillus* and *Endothia parasiticus* is also available now. Because rennets of this origin cost less than rennet from calves, they have captured about half of the market.

Rennet, when it works as desired, forms a smooth curd in 30 minutes at a temperature of 32 degrees C. The reaction proceeds in two stages. The chymosin attacks casein, one of the proteins of milk, but not the soluble proteins lactalbumin and lactoglobulin. In the presence of calcium ions the residual fractions of casein coagulate. The resulting gel has a fibrous structure. Its filaments cross-link and bond in a lattice that under quiet conditions becomes smooth and firm, forming a bed of curd. The protein in this chymosin curd is called paracasein, but because it is bound largely with calcium, it appears initially as dicalcium paracasein.

In the third step wire knives or cutting bars are used to reduce the large bed of curd in the cheese vat to cubes about 1.5 centimeters on a side. This step increases the surface area.

During the cooking period that follows these small cubes contract and expel whey. At this point the cheesemaker can influence to a degree the final moisture of the cheese by varying the temperature of cooking and the speed at which the curds and whey are stirred. For Cheddar and related cheeses the optimal cooking temperature is 37 degrees C. Emmentaler and Gruyère curds are cooked at about 54 degrees. Cooking continues for a period ranging from an hour to an hour and a half. Then the whey is removed, leaving a warm or hot curd mass that is a recognizable but still immature cheese structure.

The cheesemaker can improve this structure by turning the curd over repeatedly in the vat or by letting it lie dormant under pressure in large cheese hoops or some other form. During this time lactic acid develops from the activity of the starter culture, and the chemistry of the curd changes.

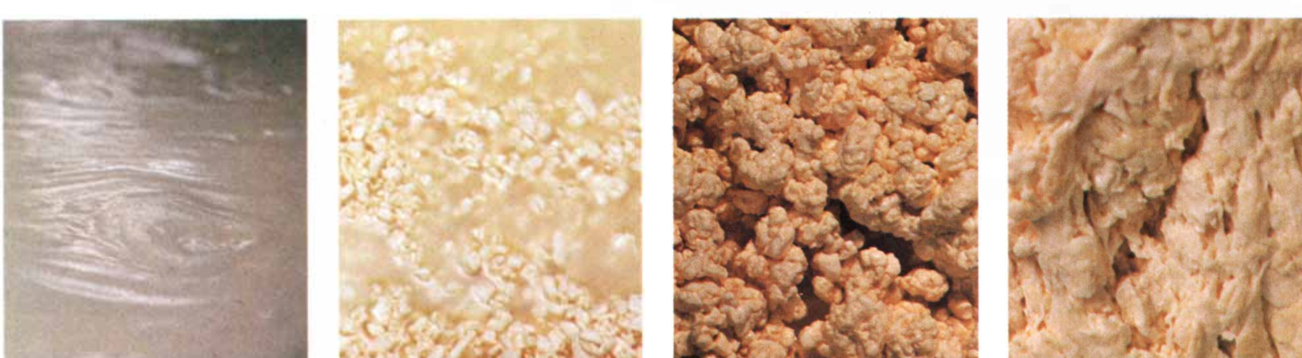
Next comes the salt. In some cases the cheesemaker applies dry salt to the curd. If he has pressed the curd into blocks or wheels, however, he may substitute brine. The practice is to immerse the immature cheese in saturated brine for from two to 72 hours, depending on the size of the cheese.

If the cheese being made requires the addition of special microorganisms to aid ripening, they may be introduced into the brine. They can also be put in the milk during the preparation stage or sprayed on the surface of the immature cheese.

In the pressing stage the wet, warm curd is confined in a wood, plastic or metal form or a cloth bag, sometimes under external pressure. Pressure gives the cheese a compact texture and characteristic shape; it also extrudes any free whey that remains and completes the knitting of the curd.

Pressing is the end of what might be called the preparatory phase of making a ripened cheese. Next the young cheese is put in a controlled environment, where the crucial ripening process takes place.

The pivotal event in ripening is the death of the millions of lactic acid bacteria that are present at the start. The process continues throughout the ripening phase. On death a bacterial cell disintegrates, releasing many intracellular enzymes. These enzymes (togeth-



CHANGING TEXTURE of Cheddar cheese is shown at successive stages of the manufacturing process. The material is the consistency of milk when the rennet is added, becomes puddinglike as a curd and then turns quite firm as increasing amounts of whey are removed.

BY THE YEAR 2000, THE WORLD MAY CATCH UP WITH THE WAY COMPUSERVE'S ELECTRONIC MALL™ LETS YOU SHOP TODAY.

Presenting the computer shopping service that delivers discount prices, name-brand merchandise, and in-depth product information.

To make your computer even more useful, join CompuServe and shop in our Electronic Mall. Easy enough for beginners, it's open 24 hours a day, 7 days a week. And it offers a wide range of goods and services from nationally known stores and businesses including Bloomingdale's, Waldenbooks, American Express and Commodore.

CompuServe's Electronic Mall™ lets you shop at your convenience in all these departments:

The Auto Shop, Book Bazaar, Financial Mart, Leisure Center, Merchandise Mart, Newsstand, On-line Connection, Personal Computer Store, Record Emporium, Specialty Boutique and Travel Agency.

Take the CompuServe Electronic Mall 15-Minute Comparison Test.

What you can do in 15 minutes shopping the Electronic Mall way.

- Access descriptions of the latest in computer printers, for instance.
- Pick one and enter the order command.
- Check complete descriptions of places to stay on your next vacation.
- Pick several and request travel brochures.
- Access a department store catalog and pick out a wine rack, tools, toys…anything!
- Place your order.

What you can do in 15 minutes shopping the old way.

- Round up the family and get in the car.

The Electronic Mall—A Valuable Addition to the Vast World of CompuServe.

CompuServe Information Services bring you information, entertainment, personal communications and more.

You can access CompuServe with almost any computer and modem, terminal or communicating word processor.

To buy a CompuServe Subscription Kit, see your nearest computer dealer. To receive our informative brochure, or to order direct, call or write:

CompuServe

Information Services, P.O. Box 20212,
5000 Arlington Centre Blvd., Columbus, OH 43220
800-848-8199
In Ohio call 614-457-0802

The Electronic Mall™ is a service of CompuServe Inc. and L. M. Berry & Company.

An H & R Block Company

Toronto

Vintage Vantage

And well worth a closer look. Gleaming New World shapes vie side by side with the hallmarks of a bygone era. Come catch a glimpse of Toronto! Lend an ear to our concerts and cabarets or savour the flavours, the creature comforts, the urban delights and the cosmopolitan pleasures. It's a new way of looking at the world – and yours to discover! Ask about the 7% sales tax rebate, 5% room tax rebate and the advantageous exchange on American currency, up to 35%.* For more information, write: Ontario Travel, Queen's Park, Toronto M7A 2E5 Ontario/Canada, or call TOLL FREE 1-800-268-3735.

**Subject to change*

ONTARIO yours to discover!®

er with the residual chymosin and enzymes in the milk) transform the protein, fat and carbohydrate. Consequently the cheese begins to develop its typical flavor and a smooth texture.

In the early stages of ripening the cheese is stiff because of the presence of dicalcium paracasein; the cheese would not melt or become stringy if it were heated. As the amount of lactic acid in the curd increases, more bound calcium dissolves and a new compound emerges: monocalcium paracasein. It is soluble in warm salty water, stretches easily and melts uniformly when it is heated.

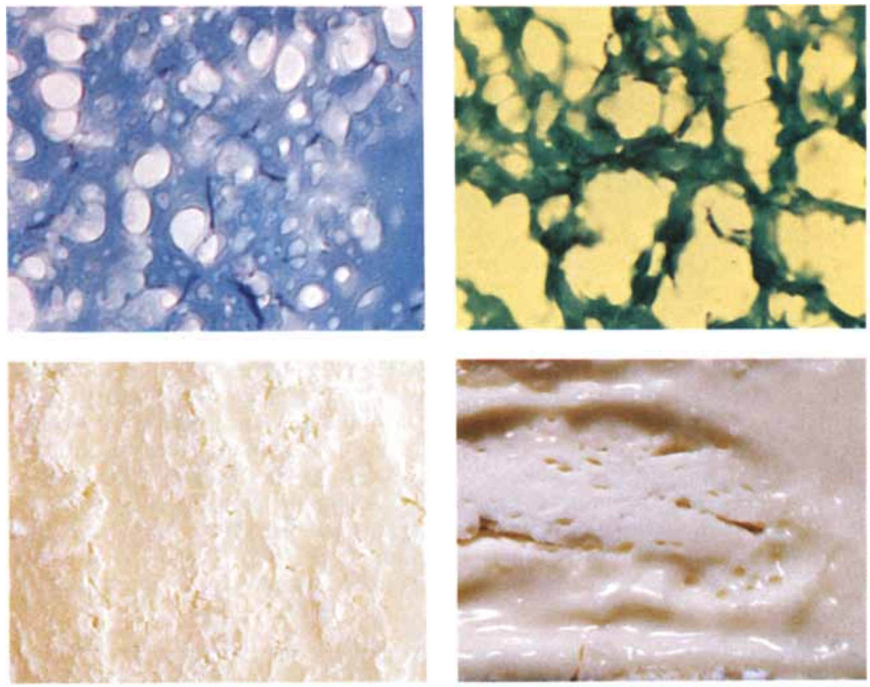
Within 48 hours of pressing much of the dicalcium paracasein has been converted into monocalcium paracasein. In the continued presence of lactic acid more calcium dissolves and some of the monocalcium paracasein becomes simply paracasein. This compound serves as a substrate for enzymes: proteases, which break down proteins into peptones and peptides, and peptidases, which in turn fragment peptides into their constituent amino acids. From these sources in a ripening cheese soluble peptides, free amino acids and amines accumulate, contributing to the typical flavor of the finished cheese. The concentrated action of the enzymes also further breaks down parts of the rigid initial structure of the curd, making the material softer.

The typical flavor of a ripened cheese depends largely on the relative balance of the flavoring components derived from the degradation of protein, fat and carbohydrate. An excess of a derived compound that upsets the balance may lead to undesirable flavors: bitter, rancid, fruity or sulfuric. The art of cheesemaking lies in achieving the desired balance.

If ripening is properly controlled, the fat of a ripening cheese is partially hydrolyzed (fragmented in a chemical reaction involving water) by enzymes from microorganisms and by the lipase in milk. The product of the reaction is free fatty acids. Some of them, such as caproic, caprylic and capric acids, give rise to piquant flavors. Various ketones that provide flavors particularly characteristic of blue cheese come from fatty acids. It is also important that hydrolysis soon slows, otherwise ripened cheese would invariably have a strong, unpleasant taste and a bad smell.

The lactose in a ripening cheese also contributes to flavor. It is converted into lactic acid and lactates. They in turn are changed into other organic compounds (such as diacetyl).

Cheese evolves gas steadily as it ripens. In Cheddar and Emmentaler the gas is carbon dioxide; in Camembert and Brie carbon dioxide may give way to ammonia, decreasing the quality of the cheese. A consistent source of carbon dioxide during normal ripening is the attack on free amino acids by enzymes associated with specific bacterial groups such as the enterococci. Cheese may also generate hydrogen, hydrogen sulfide and an excessive amount of carbon dioxide. Such events indicate an abnormal fermentation and an inferior cheese.



CHEESE STRUCTURE appears in micrographs (*top*) of Parmesan (*left*) and Camembert (*right*). The enlargement is about 2,500 diameters. Below each micrograph is a photograph of the cheese as it appears to the consumer. Parmesan is a hard cheese, Camembert a soft cheese. The micrographs were made by Roger Maret of the Nestlé Co. in Switzerland.

Eyes develop in ripened cheeses that have hard rinds (Emmentaler) or are enclosed tightly in a rubberized-plastic film that is highly impermeable to gas (block Swiss). The phenomenon is particularly conspicuous if *Propionibacterium* organisms are added to the milk and the cheese is kept warm for several weeks. Each eye grows from a bubble of carbon dioxide.

Varieties

The general process I have described differs depending on the specific variety of cheese being made. A few examples from the broad classification of natural ripened cheeses as hard, semisoft and soft illustrate the point.

The hard cheeses include Cheddar, Emmentaler and Provolone. Cheddar, which is named for an English village where its manufacture began in the 17th century, usually ripens in from five to 12 months at a temperature between two and 10 degrees C. In general 100 kilograms of milk yields 9.5 kilograms of Cheddar; the yield varies according to the levels of fat and protein in the milk and the amount of moisture in the finished cheese. The most characteristic feature of the manufacturing process is the cheddaring step. This entails turning for several hours the warm slabs of curd on the floor of the vat.

Emmentaler cheeses acquire their gold-hued rind by daily washing of the surface. The cheese is noted for its eyes, which are larger in the U.S. than they are in Europe. The lactic acid fermentation of Emmentaler is carried out by thermophilic (heat-loving) bacteria. Most of the fermentation caused by the bacteria takes place during the pressing stage.

Provolone is made chiefly in Italy, Argentina and the U.S. In the early stages the steps are much the same as they are for low-moisture mozzarella, the standard cheese for pizza. Provolone is formally described as a *pasta filata,* or pulled-curd cheese. The immature cheese is molded into the shape of a sausage, a pear or a ball and then encased in twine. It is usually smoked and then ripened to develop its distinctive flavor.

Semisoft cheese includes Roquefort and Blue. Limburger, Camembert and Brie are soft cheeses. The two groups have little in common except that they all require air for their special ripening organisms: the fungus *Penicillium ro-*

queforti for Roquefort and Blue; the reddish bacterium *Bacterium linens* for Limburger, and the fungus *P. caseicolum* (known in the cheesemaking industry as *P. candidum*) for Camembert and Brie. Each of these organisms is cultivated in a special liquid medium and transferred to the cheese under sterile conditions.

Roquefort, a blue-mold cheese derived from sheep's milk, is made in a region south of a line between Bordeaux and Grenoble; Corsica is also a source of Roquefort. (French cheeses of the Roquefort type made from other kinds of milk are called Bleu, and similar blue-veined cheeses made in the U.S. and other countries are called Blue.) The standard form for Roquefort is a wheel weighing about 2.5 kilograms (5.5 pounds). To be certified as Roquefort all the wheels must be brought within eight days of manufacture to one of the natural caves of the village of Roquefort and ripened in that cave for from three to four months.

The blue-mold fungus *P. roqueforti* that ripens this cheese requires less air than the white-mold fungus *P. caseicolum;* it is also hardier. Although many strains of *P. roqueforti* exist, only five or six are used in ripening Roquefort. All of them have been isolated from the atmosphere of the caves by a process of natural selection that has gone on for centuries.

The blue-mold spores are applied to the cheese milk or to the curds in the form of a black powder. In the pressed cheese they remain dormant until the carbon dioxide in natural crevices of the cheese or in channels made by inserting steel needles is replaced by air. Sometimes gas-forming *Leuconostoc* bacteria are introduced along with the standard lactic acid culture to create the desired cavities. In a blue-mold cheese held at 10 degrees C. and at high relative humidity the mold spores germinate in approximately 30 days as branched, greenish blue mycelia (filaments) containing highly active proteolytic and lipolytic enzymes. These enzymes function in unison with the ones normally found in the interior of the cheese and develop the typical flavor of the cheese in from three to six months.

Three Soft Cheeses

Limburger is one of a number of reddish cheeses ripened by surface bacteria. Among the others are brick, Liederkranz, Saint Paulin and Pont l'Évêque. The early stages of ripening are marked by the growth of wild yeasts (such as members of the genus

CHEESE	MILK	METHOD OF RIPENING
	HARD CHEESES	
ASIAGO	COW OR EWE	CURED FOR UP TO A YEAR, WASHED AND TURNED FREQUENTLY AND SOMETIMES RUBBED WITH VEGETABLE OIL.
CHEDDAR	COW	CURED AT 36–50° F. FOR 60 DAYS TO 12 MONTHS.
COLBY	COW	CURED FOR 60 DAYS OR MORE.
EDAM	COW	SHELVED IN LAYERS AT 50–60° F. FOR SIX TO EIGHT WEEKS. WASHED, DRIED AND TURNED FREQUENTLY
EMMENTALER (SWISS)	COW	FORMATION OF EYES IN THREE TO FOUR WEEKS AT 72° F. AND 80–85 PERCENT RELATIVE HUMIDITY. RIPENED AT 40° F. AND HIGHER TEMPERATURES FOR TWO TO 10 MONTHS.
GOUDA	COW	CURED AT 50–60° F. FOR TWO TO SIX MONTHS.
GRUYÈRE	COW	FORMATION OF EYES AT 60° F. IN ONE MONTH. CURED FOR 80 DAYS OR MORE AT 50–60° F.
PARMESAN	COW	SHELVED FOR 10 MONTHS OR MORE AT ABOUT 50° F. AND 85 PERCENT RELATIVE HUMIDITY. TURNED, WASHED, SCRAPED AND RUBBED WITH OIL FROM TIME TO TIME.
PROVOLONE	COW	SMOKED AND CURED AT 40–50° F. FOR UP TO 12 MONTHS.
STILTON	COW	MOLD-RIPENED BY *PENICILLIUM ROQUEFORTI* FOR TWO WEEKS, CURED FOR ABOUT SIX MONTHS.
	SEMISOFT CHEESES	
BLEU (BLUE)	GOAT OR COW	MOLD-RIPENED BY *P. ROQUEFORTI*, CURED FOR THREE MONTHS AT 48° F. AND 95 PERCENT RELATIVE HUMIDITY, WRAPPED IN FOIL AND STORED IN A COOL ROOM FOR TWO TO THREE MONTHS.

METHOD OF CURING ripened cheese has a great deal to do with developing the characteristic texture and flavor of the cheese. The chart lists the key steps of curing for 20 well-known ripened cheeses, which are classified according to whether they are hard, semi-

Pichia) on the surface of the cheese. Their enzymes make the mixture less acidic, raising the *p*H to about 5.5, a condition that favors the growth of *Bacterium linens*.

The first Camembert was developed in the tiny French village of that name by Marie Harel in 1791. Since then its manufacture has spread to several other regions of France. Under the old tradition the cheese is made from raw milk; the modern tendency is to use pasteurized milk. Camembert is ordinarily formed into 228-gram (half-pound) wheels.

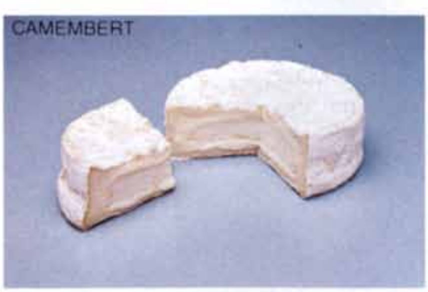
The process for Brie is similar. Both cheeses require the introduction of *P. caseicolum* because of its white mycelia. Camembert and Brie ripen from the surface to the center. The manufacturer must guard against starting with wheels that are too thick; in such a wheel the outer regions ripen excessively before the center turns soft. Moreover, the outer regions develop a higher *p*H, which means that eventually ammonia will be given off there, causing discolorations to appear before the entire cheese is mature.

More an Art

As the reader will have gathered from several of these observations, cheesemaking is still more of an art than a technology. In many countries one can find cheesemakers applying the methods of their ancestors with simple tools and facilities. Yet in the same regions large manufacturing plants may be producing similar cheeses in quantity much as a bakery turns out loaves of bread.

Much cheese is the product of modern technology. Enormous vats, molecular-membrane sieves, continuous conveyors for matting curd, electronic salting devices, block cutters and vacuum presses combine to turn out cheese of surprisingly good quality. The essential fermentation with lactic acid lies hidden under stainless-steel mantles encasing the operational machinery, but its pattern is still the same as that in the copper kettles or 1,000-liter tinned vats of village cheesemaking.

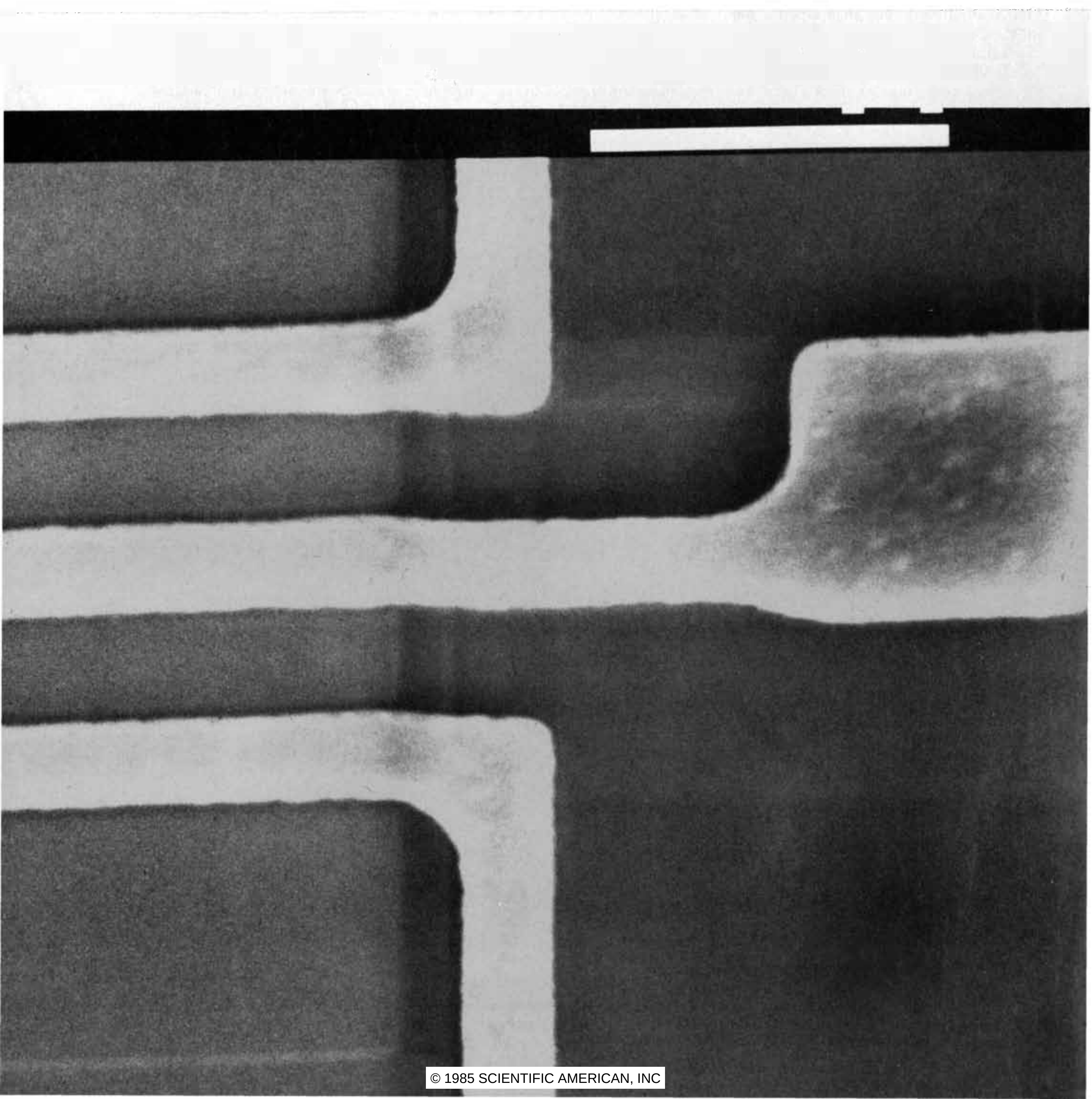
At the other end of the spectrum is the growing number of people in the U.S. who make their own cheese. One result is the recent formation of the American Cheese Society. Its members try their hand at creating a Monterey Jack, a Cheddar, a Brie or some other variety. Just as small wineries are being established by the score in the grape-growing regions of California and New York, so small-scale establishments making specialty cheese can be expected to proliferate.



CHEESE	MILK	METHOD OF RIPENING
SEMISOFT CHEESES		
BRICK	COW	CURED ON SURFACE BY *BACTERIUM LINENS* FOR 14 DAYS, WRAPPED AND STORED FOR TWO TO THREE MONTHS AT 40° F.
GORGONZOLA	COW	MOLD-RIPENED BY *P. ROQUEFORTI*, CURED AT 40–50° F. AND 80 PERCENT RELATIVE HUMIDITY FOR 30 DAYS, THEN AT HIGHER HUMIDITY FOR THREE TO SIX MONTHS.
MONTEREY	COW	CURED FOR SIX WEEKS OR MORE AT 60° F. AND 80 PERCENT RELATIVE HUMIDITY.
MUENSTER	COW	CURED FOR SEVERAL WEEKS AT 50–55° F. AND 80 PERCENT RELATIVE HUMIDITY.
ROQUEFORT	EWE	MOLD-RIPENED BY *P. ROQUEFORTI*. WHEELS ARE SALTED AND STORED IN CAVES AT ROQUEFORT. HELD AT LOW TEMPERATURE AND HIGH RELATIVE HUMIDITY FOR THREE MONTHS.
SOFT CHEESES		
BRIE	COW	RIPENED BY A WHITE MOLD, *P. CANDIDUM*, FOR EIGHT TO 11 DAYS IN CELLAR OR CAVE AT 52° F. AND 90 PERCENT RELATIVE HUMIDITY. DISTRIBUTED WITHIN 14 DAYS UNDER REFRIGERATION.
CAMEMBERT	COW	RIPENED BY *P. CANDIDUM* ON FRAMES OR SHELVES AT 55° F. AND ABOUT 95 PERCENT RELATIVE HUMIDITY FOR 12 DAYS. DISTRIBUTED WITHIN 21 DAYS UNDER REFRIGERATION.
LIEDERKRANZ	COW	RIPENED BY *B. LINENS* FOR THREE TO FOUR WEEKS AT 45° F.
LIMBURGER	COW	RIPENED ON SURFACE BY *B. LINENS* AND CURED ON SHELVES FOR THREE WEEKS AT ABOUT 55° F. AND 95 PERCENT RELATIVE HUMIDITY.

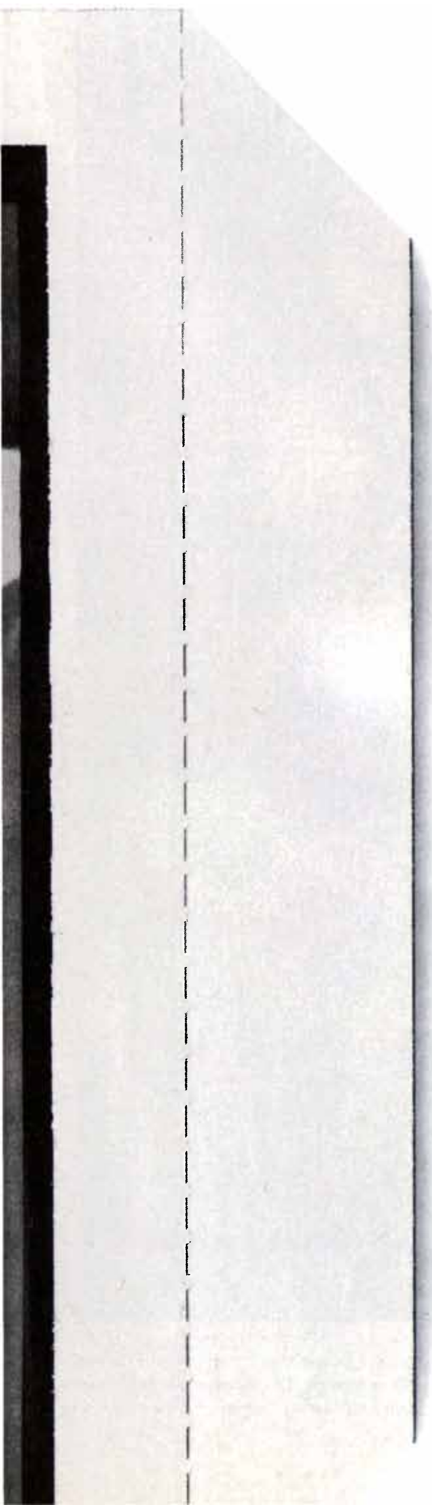
soft or soft. Ripened cheeses make up one of the two major groups of cheese, the other group being fresh cheeses. That group includes cottage cheese, cream cheese, ricotta and mozzarella. Fresh cheeses are not cured and must be eaten fairly soon after they are made.

INTEL TAKES THE LEAD IN CMOS. BY HALF A MICRON.



Low power dissipation of CMOS devices has typically meant low performance, but with Intel's 1.5 μm design rules this limitation is virtually eliminated.

These half-micron-narrower design rules are critical to a process we call CHMOS—a process that yields CMOS products with performance on par with HMOS technology.



What's equally significant is that we provide a complete family of CHMOS 1.5 μm products—not just an isolated product here and there. Including the likes of the highest density DRAM available in CMOS (256K), a high-performance 35ns static RAM, EPROMs and world standard microprocessors and controllers in CMOS.

Intel's CHMOS process is the amalgam of low-power CMOS with high-performance HMOS. Switching speeds of 200 picoseconds at 3 nanowatts power per gate can now be achieved.

Despite the continuous improvement of HMOS by scaling, the delay-power product for CHMOS is more than an order of magnitude lower than its HMOS counterpart in the typical integrated circuit. For example, in a VLSI part with 50,000 gates operating at greater than 10 MHz, the CMOS version would dissipate less than 2 Watts compared to over 25 Watts for an HMOS part.

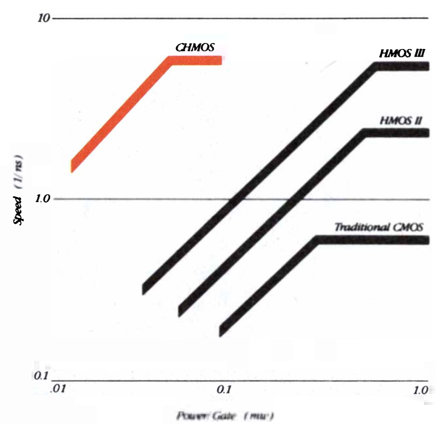
A special use of double-metal layers with CHMOS can produce a large savings in real estate, making package densities of over 500K transistors on a chip possible.

As critical dimensions of the technology are reduced, reliability becomes more difficult to achieve. With Intel's CHMOS technology, reliability is assured by the following:

First, by using epitaxial substrates, the latch-up problem is eliminated over a wide range of operating conditions. For example, no latch-up is observed at supply voltages of 9 volts at 125°C for a part that nominally operates at 5V. And input pins can withstand currents up to 200 mA without initiating latch-up.

Secondly, the lower die temperatures that Intel's CHMOS provides substantially diminish the electromigration problem typical of VLSI devices.

Finally, by locating the storage node in the CMOS well, a natural barrier is provided against soft errors for DRAMs and SRAMs.



Intel's CHMOS gives you the high-performance of HMOS but uses only one-tenth the power.

The result is a low-power, high-performance CHMOS technology with high reliability. And an array of 1.5 μm products unmatched in technical properties.

And that's only part of the story. For more information, call (800) 538-1876. In California, (800) 672-1833. Or write Intel Corp., Lit. Dept. W-203, 3065 Bowers Ave., Santa Clara, CA 95051.

And see just how big half a micron can be.

intel®

FRESHLY MOLTED BLUE CRAB (*bottom*) is shown with the exoskeleton, or hard covering, that it has just shed (*top*). The crab has absorbed water and is 25 percent larger than the old shell. The calcified exoskeleton cannot increase in size and therefore it must be shed intermittently as the crab matures. The blue crab takes its name from the color of the legs and the inner surface of the claws.

Molting in the Blue Crab

The blue crab intermittently sheds its shell and forms a new one. One result is a summer delicacy: soft-shell crabs. The chemistry underlying the formation of a new shell is now being illuminated

by James N. Cameron

Some of the most satisfying of all recipes begin with the instruction "Take a dozen soft-shell crabs..." The basis of these recipes is the familiar blue crab, and whether the recipe calls for deep-frying the crabs in a light batter or simply sautéing, the result is a delicious meal that is associated with pleasant memories of early summer. Soft-shell crabs are most often enjoyed by residents of the eastern seaboard of the U.S., but the taste for them has spread to other parts of the U.S. and indeed to other countries. Underlying the appearance of the soft-shell crabs in the fish markets in early summer is a complex biological cycle, which combines elements of natural history, physiology, ecology and commercial fishing.

The blue crab, *Callinectes sapidus,* is a soft-shell crab only at a particular phase of its life: just after it has shed the old shell and before the new one has begun to harden. The periodic shedding and re-forming constitute a cycle that persists throughout the life of the crab. My scientific interest centers on the chemistry of shell formation. The shell of the adult blue crab is about 7.5 centimeters (three inches) long and 15 centimeters wide. The top of the shell is generally green; the bottom is a whitish shade. The name blue crab comes from the coloring of the legs, of which there are five pairs. The three middle pairs are walking legs. The two front legs end in claws. The hind legs, flattened at the outer ends, are employed in swimming; the blue crab is one of the best swimmers among all crab species.

The main habitat of the blue crab is the Atlantic shoreline of the Americas from Cape Cod to Brazil. Its range extends, however, to the Gulf Coast of Louisiana and Texas and even to the shores of the Mediterranean. In all those places the crab passes through the phases of its life cycle in the shallow waters of estuaries and muddy shores. On the Chesapeake Bay, where the U.S. blue crab industry is centered, the crabs are favored by heavy growths of grasses. Emergent grasses, which protrude from the water, form extensive marshes along the shoreline. Beds of submerged aquatic grasses grow in the shallow water. Together the marsh and sea grasses form the lowest level of the ecological pyramid that supports the crabs. Among the grasses the adult crabs feed on a wide variety of fish, mollusks, crabs and other organisms.

The water of a coastal marsh undergoes wide fluctuations in temperature and salinity, particularly if the marsh is in an estuary or a lagoon. Therefore female blue crabs generally travel to the bay mouth or to ocean water near the shore to release their eggs. From the newly hatched egg emerges the first of the crab's larval stages: the zoea. The zoea is a microscopic form that resembles a shrimp more than it does the adult blue crab [*see illustration on next page*]. Hatching often takes place in June or August, although the timing of events in the crab's life cycle can vary considerably according to such factors as water temperature, salinity and region. After the zoea hatches it floats on currents and tides while undergoing a complicated metamorphosis. There are usually seven zoeal stages (sometimes there is an eighth) and between each stage the larva molts: it sheds its outer covering and forms a new one. At each molt there is a slight increase in size and a slight change in form. These early molts, in which the form as well as the size changes, are called metamorphic molts.

The molting process, which reappears throughout the life of the crab, is an evolutionary accommodation to the increase in size that accompanies maturation. Like the other members of the large phylum of arthropods, the crab has an exoskeleton, or a hard external skeleton. The exoskeleton has several advantages, including increased protection for the internal organs and excellent leverage for some muscle systems. Unlike the bones of a mammal, however, the exoskeleton cannot grow as the crab expands. If the crab were to retain its shell permanently, it would never be able to pass beyond the microscopic zoeal stage. Molting is one solution to this problem and it is the solution observed in most arthropods.

At the end of the zoeal period the crab larva molts again, but this molt does not yield another zoea. Instead the metamorphic molt yields a form known as the megalops. The megalops has eyes on stalks (as the adult crab does), three pairs of walking legs and crude claws. Hence it resembles a blue crab more than the zoea does, but the resemblance is still slight. After about two weeks, however, a final metamorphic molt brings the megalops to the "first crab" stage. The first crab is a blue crab in miniature, identical with the adult in all respects but size and sexual maturity. Under ordinary conditions the period from hatching to the appearance of the first crab is about two months.

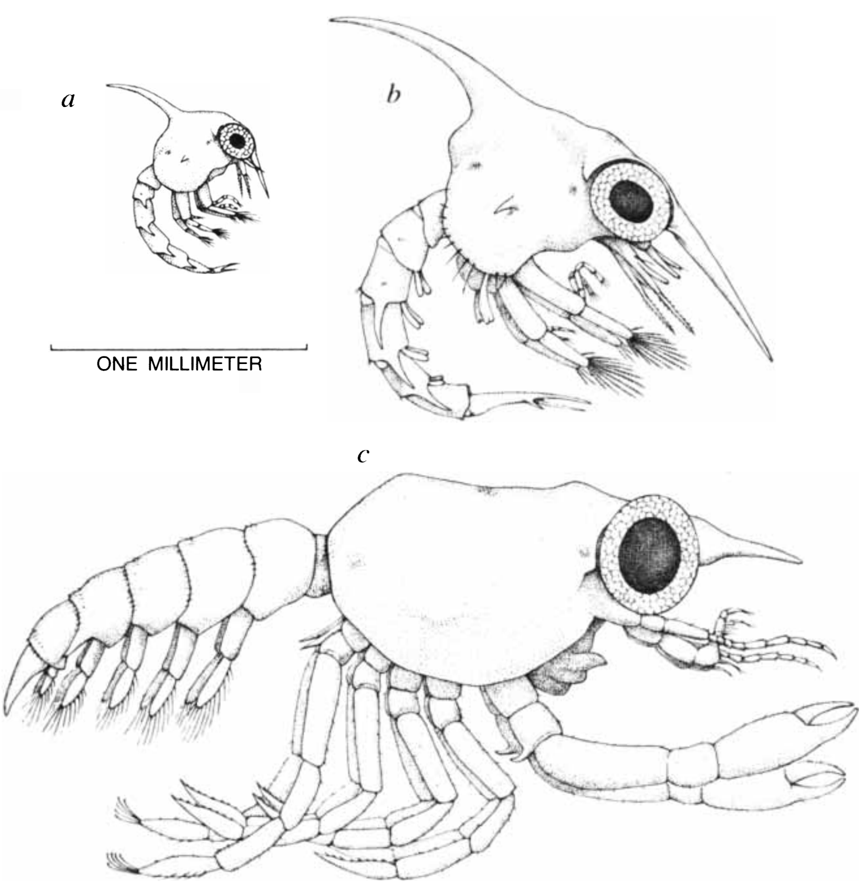
After the transformation into the first crab the successive molts accommodate increases in size without bringing any change in form. A healthy crab may live for about three years and molt some 20 times in that span. As the crab matures the interval between molts increases; by the second or third winter the interval may be many months. The molt cycle has been divided into five stages designated *A* through *E* according to the progressive changes in the exoskeleton and the associated tissues. Stages *A* and *B* are the postmolt phases; stage *C* is the intermolt phase; stage *D* is the premolt phase, and stage *E* is the emergence from the shell, for which the technical term is ecdysis. The neat division into five phases is somewhat deceptive. Based on outward signs the intermolt period, stage *C*, is by far the longest. Internal preparation for the molt, however, begins before there are any obvious external signs, and the physiological recovery from the molt continues after all visible signs have disappeared. Thus the molt is not an abrupt interruption of the crab's life but the culminating event in a continuous process of growth.

Preparation for the molt begins in stage *D* when systems of enzymes cause a separation of the upper and lower layers of the shell and begin to dissolve some of the harder parts. Suture lines, thin separations on the claws, across the front of the outer shell and all around the undersurface of the shell are fully dissolved by the enzyme action. As the dissolution progresses a new shell is forming under the old one. For the most part the forming shell is invisible. The segments at the end of the walking and swimming legs are partially transparent, however, and in those segments the new shell can be seen as a thin border within the margin of the old shell. The double border enables the watermen of the Chesapeake Bay to determine how soon a particular "peeler" (a crab close to molting) will shed its hard outer covering. The double border also gives rise to other colorful terms that are part of the watermens' specialized occupational vocabulary. A week or two before the molt, when the double margin is whitish and indistinct, the term is white-sign crab. Closer to the time of molting the margin becomes quite distinct, taking on a pink color and finally a deep red tone, and the term is red-sign crab. The final preparatory phase begins with the absorption of water into the tissues, which causes swelling and rupture along the weakened suture lines. A crab with a rupturing shell is called a buster.

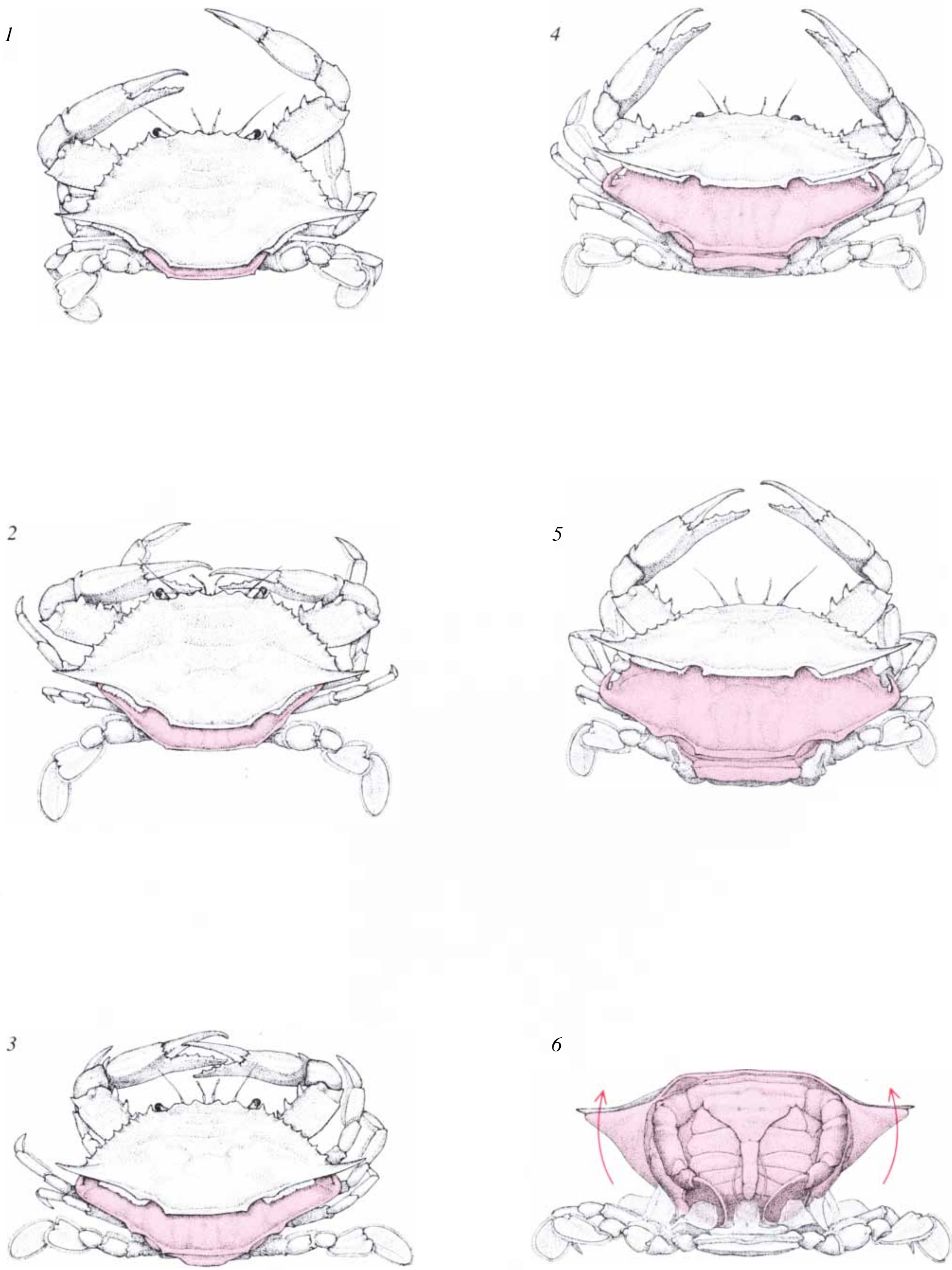
When the physiological preparations are complete, the old shell can be shed quite quickly: we have seen the emergence take 15 minutes or less in the laboratory. The withdrawal is accomplished partly because of the hydraulic pressure of the absorbed water. The crab is also capable of some muscular contraction, and it makes pumping motions of the stomach in the final stages of the molt. All external surfaces are shed in a piece, including the lining of the stomach (which is an external surface although it is rarely thought of as such) and the coverings of the gills. The crab emerges from the old shell backward and upward, withdrawing its flaccid claws, legs and gills from the old, loose exoskeleton [*see illustration on opposite page*]. When the molt is finished, the crab is soft and wrinkled and the shell is brighter colored than it is during the long intermolt phase. Over the next two hours the crab continues to absorb water and soon is about 25 percent larger than the old exoskeleton.

The newly molted crab in stage *A* is the gourmet delicacy that begins appearing on restaurant menus in the late spring. Stage *A* is brief. The exoskeleton hardens quickly. The process begins with the most critical parts, such as the mouth structures that pump water over the gills. Within a few hours the crab has reached stage *B*, the "paper shell" stage. By this time the back has a crinkled texture and the crab is no longer in prime commercial condition. Because the optimal period for shipping the soft crabs is so short, the watermen of the Atlantic and Gulf coasts are by necessity close observers of the molting cycle. They know that as the molt approaches the peelers move from their normal foraging areas to shallow beds of marine grasses near the shoreline. The reason for the move is that freshly molted crabs are quite vulnerable to predators, including other crabs, and the thick clumps of grass offer excellent places to hide. The peelers are collected from their hiding places among the long grass in pots or with trawllike devices called scrapes. Once the peelers are pulled from their hiding places they are held in "peeler floats," large floating containers anchored near the shore. The operator of the float watches the hundreds of crabs within carefully, waiting for the moment of molting. Immediately after molting the now soft crab is removed and shipped in a refrigerated container to a wholesaler and from there to the dinner table.

For the crabs that escape the trip to the dinner plate the completion of the molt is simply a stage of the life cycle. The postmolt stage *B* blends quickly into the intermolt stage *C*, when the shell is hard. The length of stage *B* depends on the size of the crab. In large crabs the process of hardening can be seen for at least 10 days and probably is not finished for two weeks. During the intermolt period,



LARVAL STAGES of the blue crab are the zoea and the megalops. When the egg hatches, the first zoea (*a*) emerges. The zoea molts six times in six weeks, increasing in size and developing in structure until it reaches the seventh stage (*b*). Another molt transforms the zoea into the megalops (*c*). Like the adult crab, the megalops has claws and five pairs of legs. After two weeks a final metamorphic molt yields the "first crab": a blue crab in miniature.



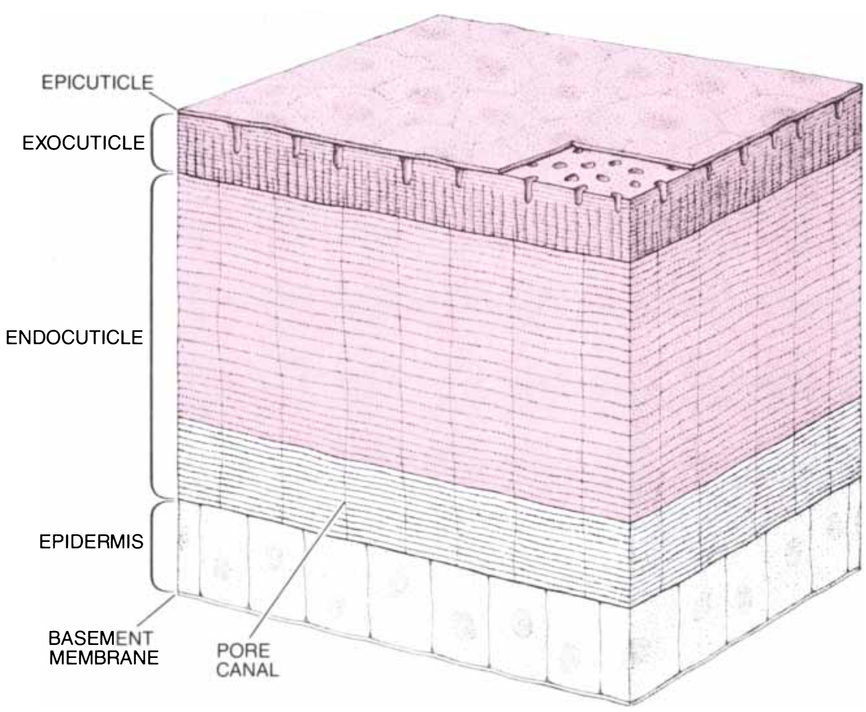
MOLT of an adult blue crab is depicted in drawings based on photographs made over a period of about 20 minutes in the author's laboratory. The molt begins with the rupture of suture lines around the old shell. The upper part of the old shell begins to lift up, revealing the new shell (*color*) underneath (*1–5*). After the upper surface has been shed the crab lifts its body up and back in a maneuver that resembles a back flip in gymnastics (*6*). It then pulls itself from the old leg coverings, the gill coverings and the lining of the stomach.

when no changes are visible, the organic reserves needed for the next molt are being replenished. It is also during the intermolt stage that the majority of tissue growth takes place. Just after the molt there is more fluid than meat within the shell. Toward the end of the intermolt period the carapace is well filled with muscle tissue and the crab is noticeably heavier.

The continuous molting cycle does have an end for each crab. For the female the final molt marks the beginning of readiness for reproduction. The female can mate only after the final molt, which is called the nuptial molt. Copulation can take place only immediately after molting, when the female's shell is still quite soft. Since the period of receptiveness is very short, the mating pair must make contact before the female molts. How the male is alerted remains a mystery, but pheromones probably have a role. After the prospective mates find each other, the male carries the female under him, clasped in some of his walking legs. The joined pair, which are known as "doublers," may hide in the grass for as long as a week awaiting the nuptial molt. After mating they remain doubled until the female's shell hardens. Mating often takes place in the fall and the sperm is stored in the female's body until the following spring, when a mass of fertilized eggs is released to hatch into the tiny zoea.

Thus the natural history of the blue crab includes two superimposed cycles: the long cycle of growth and reproduction that stretches from the zoea to the doublers and the shorter molting cycle that is repeated many times in the life of each crab and forms the basis of the early-summer appearance of soft crabs on the menu. My work has focused on the subtle physiological changes in the exoskeleton and other systems that accompany the briefer of the two cycles. The exoskeleton is not a separate tissue type. It is an epithelium, or skin, that has been strengthened by the addition of inorganic salts of various minerals. The principal mineral salt is calcium carbonate ($CaCO_3$) but there are also small amounts of magnesium, strontium and phosphate.

The toughened epidermis, or outer layer of the skin, is divided into several horizontal layers, which are collectively called the cuticle. The thin epicuticle, or outer layer, contains waxes that reduce the permeability of the shell and afford some protection from bacterial attack. The epicuticle also holds proteins that are tanned by the action of substances secreted by the crab. Tanning links the protein molecules, making the epicuticle tough and leathery. Under the epicuticle are two thicker layers: the exocuticle and the endocuticle. Both include proteins and chitin. Chitin is a natural polymer with properties resembling those of plastic. In many invertebrates, including insects, protein and chitin provide all the structural strength of the exoskeleton. In the crab, however, the deposition of mineral salts adds considerable strength and rigidity.



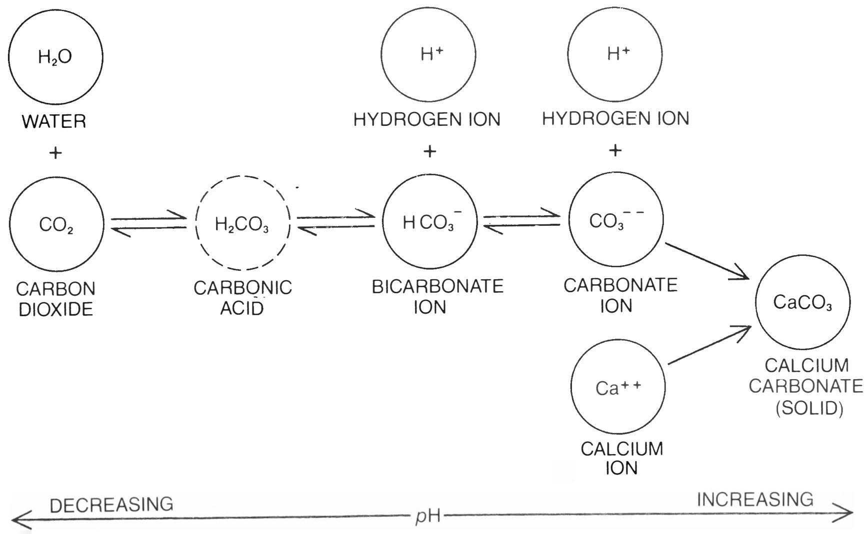
BLUE CRAB CUTICLE is the outer part of the skin, which is made rigid by the deposition of mineral salts. The epicuticle contains proteins and waxes. Horizontal layering in the exocuticle and endocuticle results partly from the presence of proteins and chitin. (Chitin is a tough polysaccharide also found in the outer shell of insects.) These materials provide an organic template for the deposition of the mineral salts, the most abundant of which is calcium carbonate. Just before a molt the upper part of the cuticle (*color*) separates from the lower part. The organic structures in the cuticle then re-form in preparation for the deposition of calcium carbonate. The cells of the epidermis produce a chemical environment that has a key role in the hardening process. The function of the pore canals is not known.

The research in my laboratory began with an investigation of how the blue crab regulates its internal acidity. Acidity and alkalinity are generally expressed in terms of *p*H, or the negative logarithm of the hydrogen ion (H^+) concentration. Low *p*H indicates an acid solution; high *p*H indicates an alkaline one. The level of acidity has a potent effect on the chemical reactions that underlie the formation of the crab's shell. Chris M. Wood of McMaster University in Ontario and I found that 14 percent of the crab's body fluids are contained in the exoskeleton and that the fluid there is much more alkaline than other body fluids. We also found that in molting the crab retains little of the mineral salts from the old cuticle. Those two observations led us to wonder how the mineral salts are built up after the molt and what significance the alkaline fluid of the cuticle has in that process.

Calcium carbonate ($CaCO_3$), the main mineral salt of the cuticle, exists in a delicate balance when it is in solution, including solutions such as the internal fluids of the crab and seawater. The balance is easily changed by fluctuations in *p*H. A slight decrease in *p*H causes the salt to dissolve into its components: the calcium ion (Ca^{++}) and the carbonate ion (CO_3^{--}). A slight increase in *p*H, on the other hand, results in the formation of calcium carbonate. Bicarbonate ions (HCO_3^-) dissociate into carbonate ions and hydrogen ions (H^+). Each carbonate ion joins with a calcium ion, and the solid salt $CaCO_3$ precipitates out of the solution. Clearly the alkalinity of the fluid in the exoskeleton favors precipitation of the solid mineral salt.

The linked chemical reactions in the exoskeleton enable the crab to increase the calcium content of the cuticle very quickly. In a large crab of about 400 grams some 40 grams of calcium carbonate, equivalent in weight to four sticks of blackboard chalk, are deposited during the two weeks after the molt. The new calcium carbonate is formed from calcium (Ca^{++}) and bicarbonate (HCO_3^-) ions in the tissue fluids. The formation of carbonate ions (CO_3^{--}) from bicarbonate ions, however, which precedes the laying down of the mineral salt, generates a surplus of hydrogen ions. If the crab is to maintain the constant internal *p*H needed for life, the hydrogen ions must be transported from the exoskeleton to the blood and from there to the seawater. Furthermore, there must be parallel transport paths in the other direction to supply calcium and bicarbonate ions.

To obtain a better understanding of these processes it would be helpful to measure the rates at which the various ions flow into and out of the crab while the shell is being calcified. There is an experimental difficulty, however, in measuring the exchange of hydrogen and bicarbonate (HCO_3^-) ions with seawater. The difficulty is related to the fact that the ultimate source of bicarbonate ions for the new cuticle is carbon dioxide (CO_2). Carbon dioxide is present in seawater and also in the crab's tissues, where it is a by-product of metabolism. Molecules of carbon dioxide combine with molecules of water to form carbonic acid (H_2CO_3). Each molecule of carbonic acid then dissociates into a hydrogen ion and a bicarbonate ion.



CALCIUM CARBONATE FORMATION depends heavily on a solution's *p*H: the negative logarithm of the hydrogen ion concentration. The *p*H is a measure of acidity and alkalinity. High *p*H indicates an alkaline solution, low *p*H an acid solution. As the *p*H increases, water and carbon dioxide combine to form small quantities of carbonic acid. For the most part the carbonic acid dissociates into bicarbonate ions and hydrogen ions. The bicarbonate ions in turn dissociate into carbonate ions and hydrogen ions. The carbonate ions can react with calcium ions to form calcium carbonate. Most of the calcium carbonate precipitates out of the solution as a solid deposit. Such deposits stiffen the blue crab's cuticle. If the *p*H of the solution falls, the chain of reactions proceeds in the opposite direction.

These reactions, however, can also proceed in the opposite direction: hydrogen ions and bicarbonate ions can join to form carbonic acid, which can then split into carbon dioxide and water. In seawater (indeed in any solution) the ratios of hydrogen ions, bicarbonate ions, carbonic acid and carbon dioxide tend toward a particular level, which is known as chemical equilibrium. If a solution at equilibrium is perturbed by the addition or subtraction of molecules of one of the four substances, the quantity of all four will change until the equilibrium level is regained. Thus it is not possible to distinguish between the addition of a hydrogen ion and the subtraction of a bicarbonate ion because in either case reequilibration will lead to the same change in the concentrations of each ion. It is possible, however, to measure the sum of all the chemical species when equilibrium is reached, which yields a measurement of the net change. To acknowledge the ambiguity of the measurement the term apparent hydrogen ion excretion is employed. We measured the rate of apparent hydrogen ion excretion in freshly molted blue crabs and found that it was very high. Since the hydrogen ions are excreted as part of the process of calcification of the new cuticle, we expected that calcium uptake would also be rapid. Measurement of the rate of calcium absorption showed that it was indeed high.

We then set about uncovering the system that transports the calcium, hydrogen and bicarbonate ions needed for calcification. The transport processes became clearer when we began to investigate the source of the carbon dioxide incorporated into the calcium carbonate. The carbon dioxide could come from the crab's own metabolism. During ordinary metabolism, animals consume oxygen and produce carbon dioxide in roughly equal amounts. Our work showed that some of the carbon dioxide could come from metabolism. The demands of calcification, however, outstrip the crab's capacity to produce carbon dioxide. An adult blue crab produces about 1.6 grams of carbon dioxide per kilogram of body weight. The complete exoskeleton contains about 44 grams of carbon dioxide (incorporated in the calcium carbonate) per kilogram. If only metabolically produced carbon dioxide were used to calcify the new shell, it would take nearly 28 days for the shell to be completely mineralized. Mineralization is completed in far less than 28 days; hence there must be an external source of carbon dioxide in addition to the internal metabolic processes.

The other source of carbon dioxide is the ambient seawater. Carbon dioxide from the environment could enter the crab's tissues in two ways. The first way is direct diffusion of gaseous carbon dioxide dissolved in the water. The second is the absorption of bicarbonate ions (HCO_3^-), which, as we have seen, incorporate molecules of carbon dioxide. In order for a substance to diffuse freely into the tissues from the environment, the external concentration must be greater than the internal concentration. Our measurements showed that there was no such gradient to provide the basis for inward diffusion of carbon dioxide: the concentration of dissolved gaseous carbon dioxide was higher inside the crab than outside. The additional carbon dioxide needed for calcification must therefore be supplied by the absorption of bicarbonate ions from the surrounding seawater.

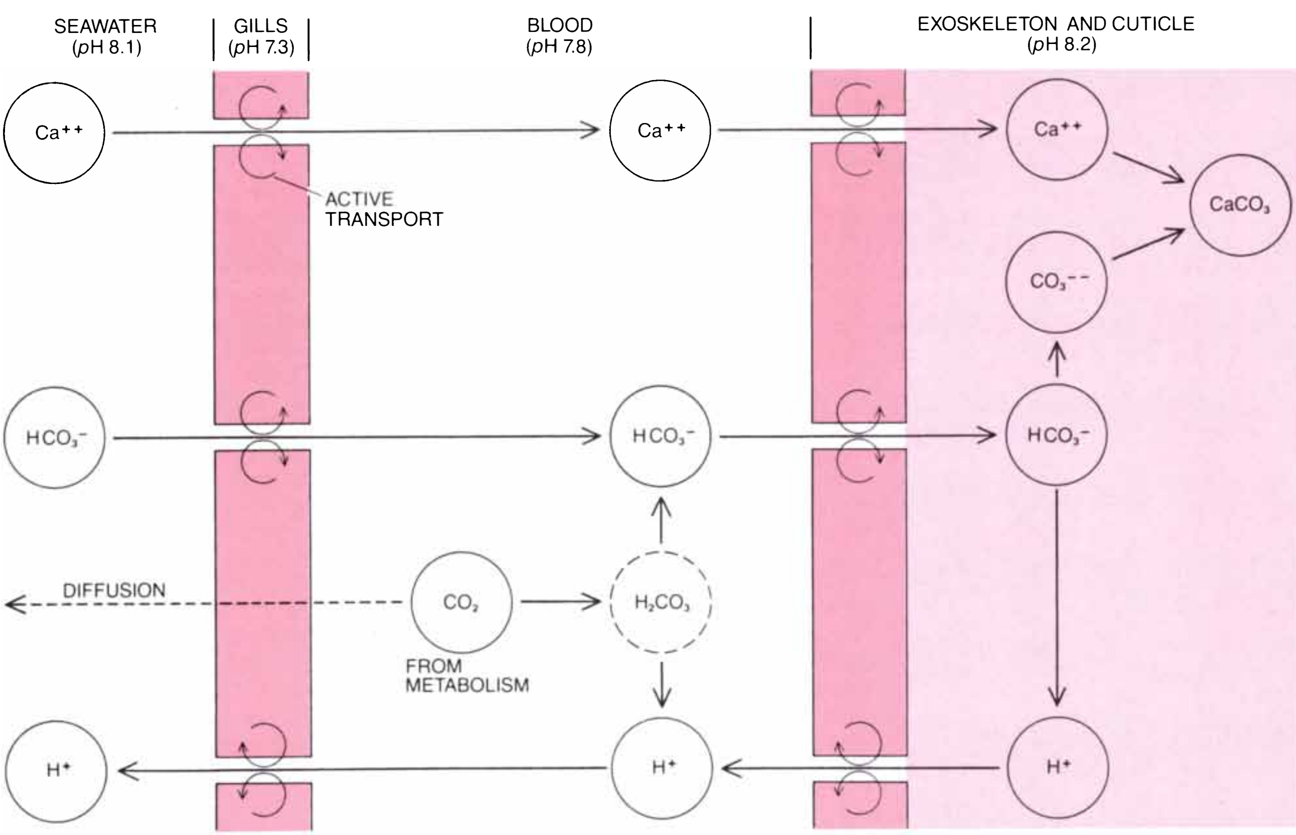
These findings provide a fairly complex picture of the transport processes that take place in the week or so following the molt [*see illustration on next page*]. Calcium ions are transported across the gills to the blood and then carried by the circulation to the inner surface of the forming cuticle. A second transport step moves the calcium ions into the fluids within the exoskeleton. At the same time hydrogen ions are being pumped out of the exoskeletal fluid. The removal of the hydrogen ions raises the *p*H, making the fluid alkaline and thereby favoring the formation of carbonate ions and solid calcium carbonate. Bicarbonate ions are transported from seawater to the blood and then into the exoskeleton. The hydrogen ions resulting from calcium carbonate formation are moved in the opposite direction: from the exoskeleton to the blood and finally into the seawater.

As a result of the work done so far, it is clear that there are three principal ion movements: calcium and bicarbonate ions move inward while hydrogen ions are moving outward. We are now engaged in deciphering the details of the transport system. We recently found that although the rates of transport are particularly sensitive to variations in the calcium and bicarbonate concentrations of the seawater, they are also sensitive to the hydrogen ion gradient between the seawater and the blood. The ion flows are linked but the coupling is not tight. By manipulating the external conditions hydrogen ion movements can be interrupted without completely blocking the movement of calcium; the converse is also true. Each ion flow is affected by changes in the chemistry of the blood, but such effects are not well understood because blood conditions in a live animal cannot be changed very much experimentally.

There are many questions still to be answered about the physiology of molting. A significant area for further work is the relation of the transport processes to the overall endocrine control of the molt cycle. In the premolt period the approach of the molt is accelerated by crustecdysone, a steroid hormone. Crustecdysone, which is secreted by the *Y* organ, a neurosecretory structure in the thorax, was isolated in 1966 by F. Hampshire and Denis H. S. Horn of the Commonwealth Scientific and Industrial Research Organization in Melbourne, Australia. In the intermolt period production of crustecdysone is blocked by a molt-inhibiting hormone (*MIH*), which is secreted by the *X* organ, a neurosecretory structure in the eyestalk. The chemical composition of *MIH* is not yet known. The *X* organ also appears to secrete several other hormones whose roles are not understood in detail.

In view of the great impact of hormones on the molting process it is probable that the calcification of the new cuticle after the molt is under some form of hormonal regulation. Yet when we injected crustecdysone into newly molted crabs, there was no apparent change in the processes of ion transport. Injection of eyestalk extracts, which contain a variety of hormones from the *X* organ, also had little effect. These results do not invalidate the hypothesis that mineralization is hormonally regulated, but they do imply that the regulation is not carried out by crustecdysone or by hormones from the *X* organ. The brain and other nerve tissues of the blue crab, however, also have considerable secretory activity, and experimental attention should probably focus on those structures next.

Solving the intriguing puzzles that remain in the molting cycle of the blue crab could have both scientific and commercial benefits. Calcification takes place in a very broad range of organisms and understanding the process completely would be of great value. The freshly molted crab might serve as a good experimental model for calcification as well as for the study of ion transport and *p*H regulation. In practical terms control of the molting process could increase the value of



TRANSPORT SYSTEM carries the ions needed to form a new blue crab shell. An active pumping mechanism transports hydrogen ions from the cuticle to the seawater. The removal of the hydrogen ions raises the *p*H of the exoskeleton and favors the deposition of calcium carbonate. Other pumps bring calcium and bicarbonate ions into the cuticle from the seawater. At the high *p*H of the cuticle the bicarbonate ions yield carbonate ions, which are incorporated into calcium carbonate. Carbon dioxide from metabolism also contributes to the formation of bicarbonate ions. The carbon dioxide that is not needed in forming the new cuticle diffuses into the seawater.

the crab harvest, make crabbing easier and spread the business of crabbing to new areas.

The harvesting of soft-shell crabs developed, beginning in about 1860, from a colorful local custom into a substantial commercial enterprise. Recipes for soft crabs were created long before 1860, but the shipment of a steady supply of soft crabs to the large cities of the Eastern seaboard had to await the advent of fast railroads and steamships and mechanical refrigeration. When means of rapid shipping became available, the market expanded quickly. Although national statistics are not very reliable, the value of soft-shell crabs on the dock is currently several million dollars per year.

Although hard crabs make up the bulk of the commercial blue crab harvest, individual soft crabs are worth far more than hard crabs. Hard crabs are worth about 20 cents per pound on the dock, whereas soft crabs have brought more than $2 per pound in recent years. If it were possible to control the molting process, the proportion of soft-shell crabs in the harvest would rise and the overall value of the harvest would increase.

In addition, commercial control of molting could make the business of soft-shell crabbing far easier. Operating the peeler floats where the premolt crabs are held until they shed the old shell is arduous, labor-intensive and commercially risky. A reliable method of inducing the molt could make the industry less arduous and help to spread soft-crab harvesting far beyond its current limits. There is now almost no significant commercial soft-shell crabbing outside the Chesapeake Bay region in spite of the fact that other parts of the crab's domain are in some respects more favorable for harvesting. Because of the milder climate in Texas, the molting season there lasts from March through mid-November, almost twice as long as the season on the Chesapeake. If crab harvesting were made easier, the industry might grow rapidly, lowering the price and spreading the soft-shell-crab habit to new palates.

Since the incentives are substantial, perhaps a cheap and reliable molting trigger will soon be found. For the moment, however, the molting of the blue crab must be viewed mainly as one step in a continuous process of growth. Furthermore, it is possible that the savor of a soft-shell crab washed down with a fine white wine is sweetened by the knowledge that the delicacy is available only when nature offers it.

Crystals at High Pressure

X-ray studies of crystals compressed between a pair of gem-quality diamonds reveal a range of responses to increased pressure. Changes in atomic structure are best viewed in terms of polyhedral geometry

by Robert M. Hazen and Larry W. Finger

What happens to the atomic structure of a crystal when the external pressure is increased? In the broadest sense the answer is obvious: the individual atoms move closer together, reducing the crystal's volume. It is only in the past few years, however, that investigators have begun to analyze in detail the various ways in which compression affects particular arrangements of atoms. The gain in understanding has come from the marriage of two experimental tools. One, the diamond-anvil high-pressure cell, is a fairly recent addition to the laboratory; the other, the X-ray diffractometer, has been a mainstay of crystallographic research for decades.

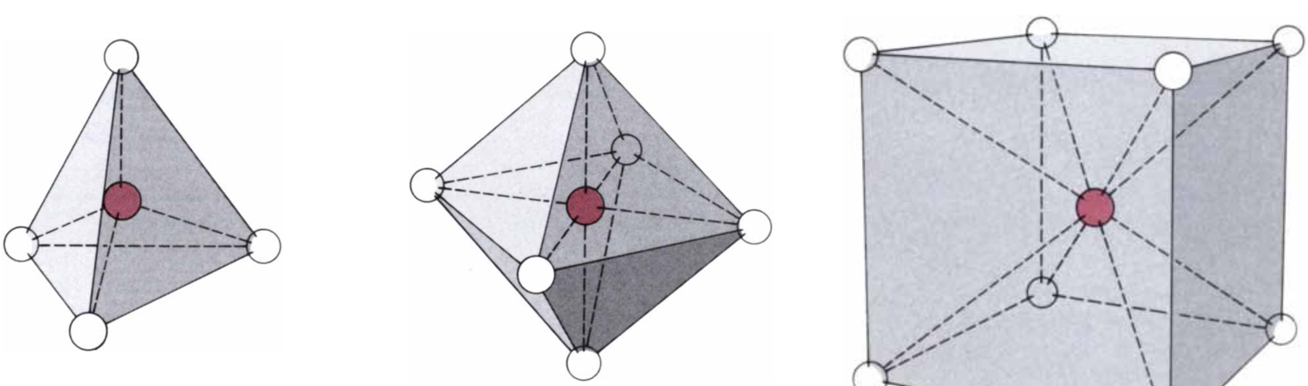
The advent of the diamond-anvil cell has made it possible to achieve extraordinarily high pressures with unprecedented ease. Pressures in excess of two million atmospheres (equivalent to those near the core of the earth) can now be produced routinely at the interface of two small, gem-quality diamonds that are squeezed together in a mechanically levered device similar to a nutcracker [see "The Diamond-Anvil High-Pressure Cell," by A. Jayaraman; SCIENTIFIC AMERICAN, April, 1984].

The tremendous pressure generated in a diamond-anvil cell can be transmitted to a single crystal hydrostatically (that is, uniformly in all directions) by immersing the crystal in a fluid that is confined by a metal gasket between the diamonds. Because diamonds are transparent to X-radiation, a narrow beam of X rays can be focused on the crystal, and the diffracted components of the beam can be measured by standard X-ray detection equipment. Thus the arrangement of the atoms in a crystal under high pressure can be determined in much the same way as it can in a crystal at atmospheric pressure.

Knowledge of the properties of matter at high pressure is crucial to understanding the earth's interior. Accordingly most of the several dozen compounds that have been studied to date with the new high-pressure crystallographic techniques have been minerals. Among them are oxides, silicates, halides and simple molecular crystals. At pressures comparable to those in the earth's lower mantle and core such minerals are reduced in volume by 50 percent or more. What changes in atomic structure can account for such a large loss of volume? What do these changes reveal about the nature of the forces between atoms?

The atomic bonding of most rock-forming minerals can be described in terms of a simple ionic model. Positively charged metallic ions, called cations, are typically surrounded by negatively charged ions, called anions. The cluster formed by a cation and its surrounding anions is usually quite regular in shape, with the anions (most commonly oxygen) corresponding to the corners of a tetrahedron, an octahedron, a cube or some other simple polyhedral form [*see illustration below*].

The representation of ionic clusters as cation-centered polyhedrons simplifies the description of complex crystal structures. Arrangements of many different atoms, which are difficult



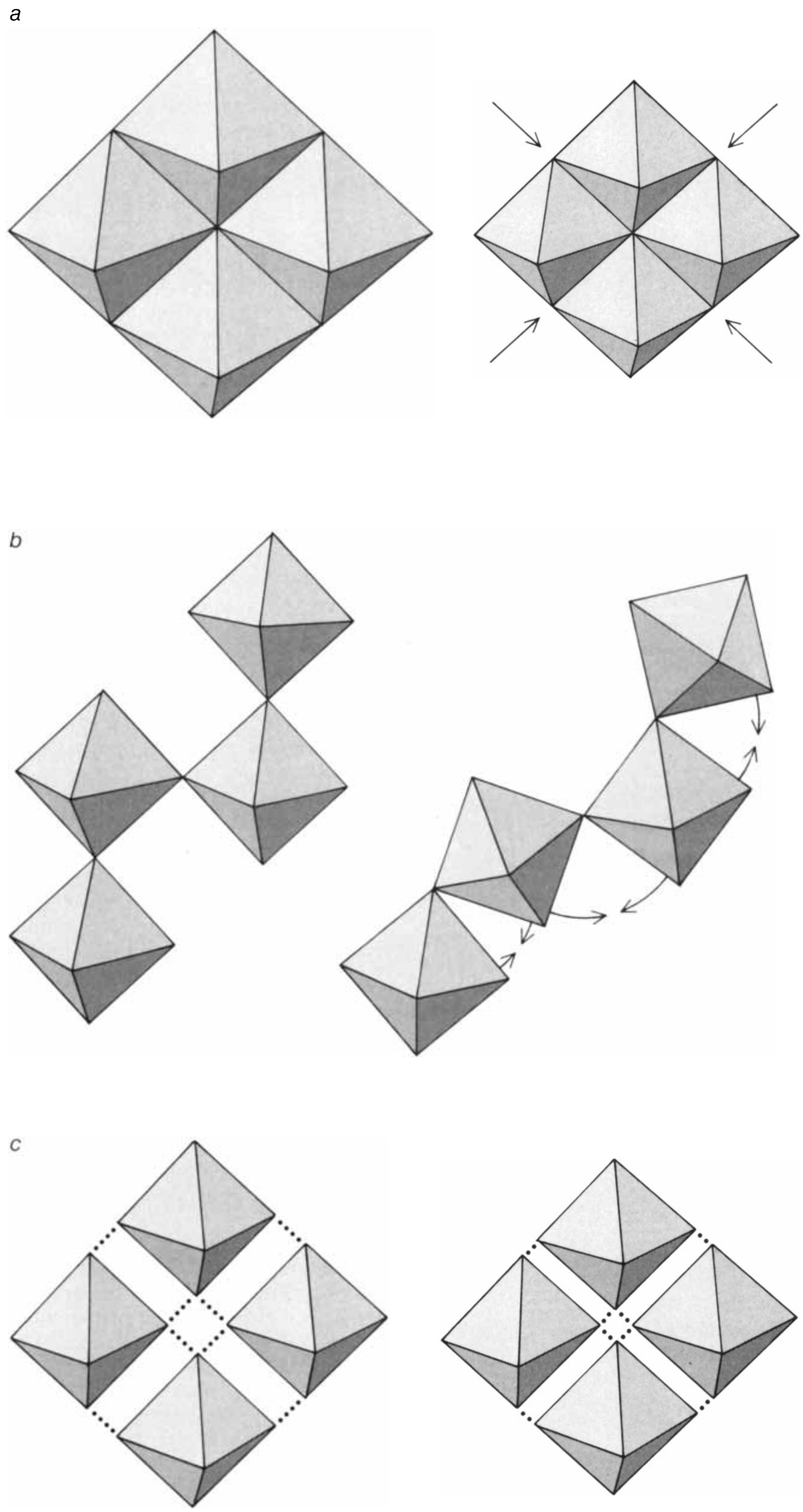
REGULAR GEOMETRIC CLUSTERS of negatively charged anions (*white*) surround a central, positively charged metallic cation (*color*) in many crystalline substances, including most common rock-forming minerals. Ionic compounds of this kind can be conveniently represented as arrays of polyhedrons; at the center of each polyhedron is a cation and at the corners are the anions. Three representative ionic structures are illustrated. In the rest of the polyhedral diagrams accompanying this article the ions are omitted.

to depict if every atom is shown, are reduced to simple geometric forms. Common binary compounds, such as the oxide of a single metal, can be represented by the packing of one type of polyhedron. Silicates and multiple-metal oxides can be similarly treated, although they call for two or more types of polyhedron.

Cation-centered polyhedrons are more than just visual aids. Each type of polyhedron has its own distinctive set of properties, which can be helpful in predicting the behavior of the bulk crystal. The recognition of the importance of these characteristic properties has led to a new approach to the modeling of different forms of solid matter. In this approach, called the polyhedral method, one first identifies the constituent polyhedrons and their properties; one then sums these factors for the bulk crystal. The procedure is often complex and depends to a large extent on how the polyhedrons are linked. In general two polyhedrons can be joined by a shared corner (that is, a single common anion), a shared edge (two common anions) or a shared face (three or more common anions). Alternatively two polyhedrons can be joined by weak molecular forces, in which case no anions are shared. The bulk properties of the crystal depend on the types of constituent polyhedrons and the nature of their linkages.

The polyhedral approach has been particularly effective for studying the behavior of crystals under compression. Each type of cation-centered polyhedron has its own value of compressibility, and the relations between the polyhedral linkages and the compressibility of the bulk crystal are comparatively straightforward. In order to predict how a particular crystal will behave under compression it is first necessary to understand the kinds of change pressure can impose on an atomic structure.

Three kinds of change in structural geometry account for most crystal compression. Bond shortening, which can be modeled as polyhedral compression in ionic compounds, is observed in all substances at high pressure and is therefore always responsible, at least in part, for any reduction in volume. Bond-angle bending, in which the distances between "nearest neighbor" atoms change only slightly while "second-nearest neighbor" atoms move much closer together, dominates compression in crystals in which the atoms are not densely packed. Intermolecular compression is the principal response to increased pressure in condensed molecular substances. Any combination of these three kinds of

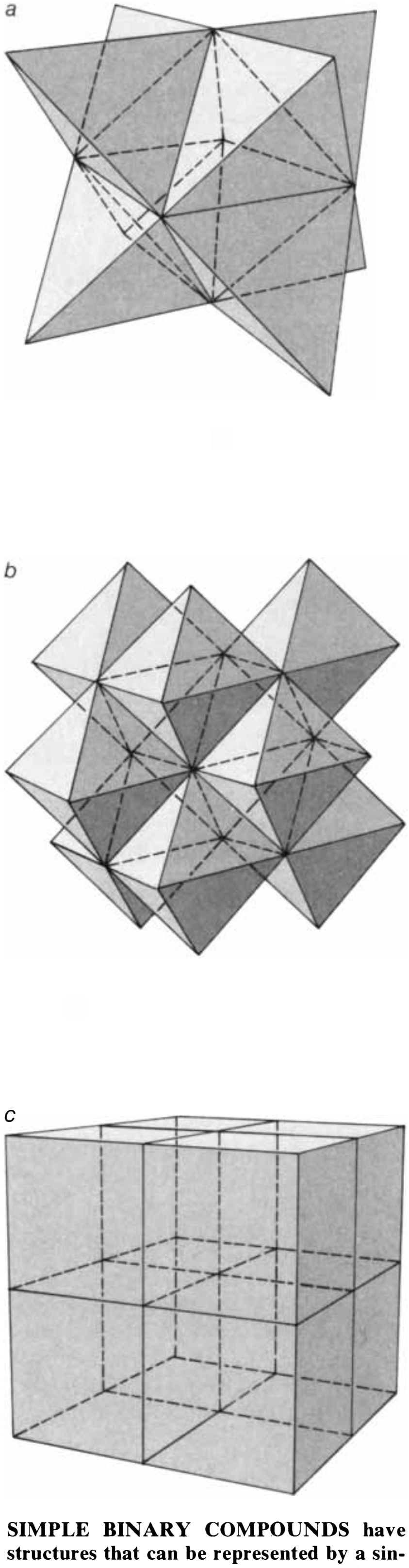


THREE POSSIBLE RESPONSES of a crystal structure to compression are shown in these polyhedral diagrams of typical atomic arrangements. Bond shortening (*a*) reduces the average distance between pairs of atoms. Bond-angle bending (*b*) changes the angles between adjacent bonds without significantly altering the length of the bonds. Intermolecular compression (*c*) is characteristic of solids that are made up of discrete molecules held together by weak bonds (*dotted lines*). Each of the three mechanisms shown reduces the volume of the bulk crystal by decreasing the average distance between the atoms; the magnitude of the changes and the final relative positions of the atoms, however, are different in each case.

a

b

c



SIMPLE BINARY COMPOUNDS have structures that can be represented by a single type of cation-centered polyhedron. The compressibility of such crystals is uniform in all directions and is identical with the compressibility of the individual polyhedrons. In lithium oxide (*a*) each of the lithium cations is surrounded by a tetrahedron of oxygen anions. Sodium chloride (*b*) is composed of an edge-sharing array of sodium-centered octahedrons. Cesium chloride (*c*) consists of a face-sharing array of cubes.

geometric change can contribute to crystal compression.

Although bond shortening is observed in every compressed crystal, the magnitude of the shortening varies for different polyhedrons. Thus some cation-anion clusters in a given structure can be considered "soft," in the sense that they vary significantly in size according to the pressure, whereas others can be considered "hard," in that they show little change in size even at pressures as high as thousands of atmospheres. The range of compressibility from the most compressible bond to the least compressible one is more than 100 to one. In modeling crystal compression it is essential to know the relative compressibilities of the bonds that form the polyhedral clusters.

In spite of the wide range of observed bond compressibilities it is remarkable that any given type of polyhedron—a magnesium cation bonded to six oxygen anions, for example—displays very nearly the same compressibility in all crystals. A value of compressibility that is almost constant from structure to structure can therefore be assigned to each type of polyhedron. What factors govern the relative magnitudes of polyhedral compressibility, and how can these values be predicted?

Ionic bonds, which link cations and anions in the polyhedral cluster, can be described by a few simple parameters, including the length of the bond, the number of nearest neighbors of each ion and the electrostatic charge of both the cation and the anion. Linus Pauling, in his classic monograph *The Nature of the Chemical Bond,* successfully related many subtle features of crystals to these simple variables. We have taken the same approach in modeling polyhedral compression.

Two significant empirical relations have been found between the bonding parameters and the compressibility of a given polyhedron. First, polyhedral compressibility tends to be proportional to the cube of the distance between the ions and hence (roughly) to the volume of the polyhedron. Large polyhedrons, such as those of alkali cations, are more compressible than small polyhedrons, such as those of silicon or aluminum cations. It follows that the polyhedrons responsible for most of the volume of a crystal usually contribute most to crystal compression. Second, polyhedral compressibility is inversely proportional to the charge of the cations and the anions; hence the distance between cation-anion pairs with the strongest electrostatic interaction will be least affected by pressure. These two relations, which were derived from the analysis of hundreds of measurements of bond compressibility, facilitate the prediction of bond and polyhedral compressibilities in oxides and silicates.

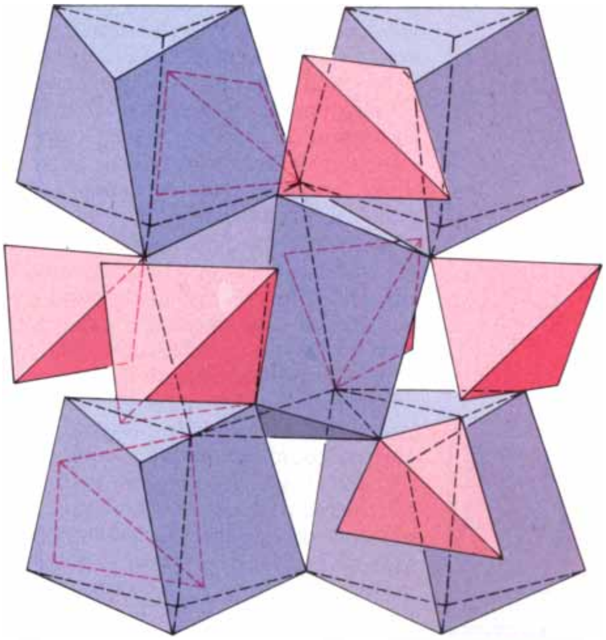
Bond shortening dominates compression in structures in which all polyhedrons are joined to adjacent polyhedrons by shared edges or faces. In these structures there are no discrete molecules for intermolecular compression, and bond-angle bending cannot occur without severe polyhedral distortions, which are not observed. Thus knowledge of the magnitude of polyhedral compression is sufficient to predict the behavior of these crystals under compression.

Simple binary compounds such as lithium oxide, sodium chloride and cesium chloride are among the easiest polyhedral structures to model at high pressure [*see illustration on this page*]. In each of these crystals the bulk compressibility is identical with that of the constituent polyhedrons; a measurement of the macroscopic change in volume with pressure corresponds to a direct measurement of polyhedral compression. When pressure is applied to such structures, it tends to decrease the distance between atoms without resulting in any significant change in the relative positions of the atoms.

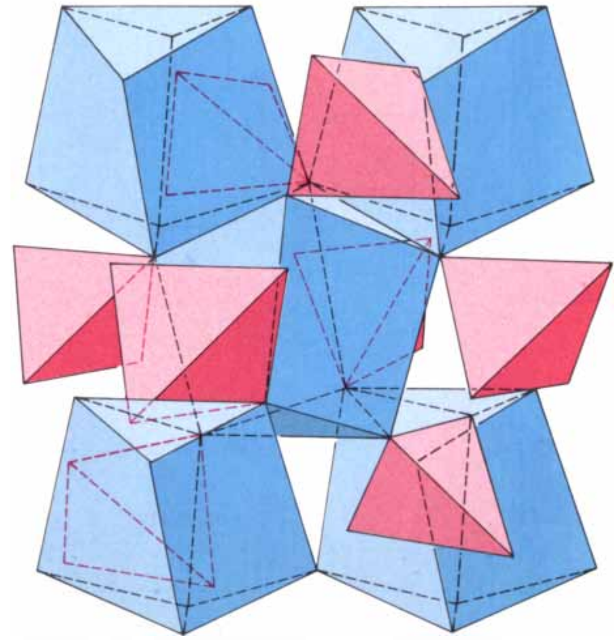
Binary oxides, which have two kinds of cation-anion cluster, often exhibit more complex behavior. Scheelite-type compounds, which have the general chemical formula ABO_4 (in which the arbitrary symbols AB stand for some combination of metallic ions), are particularly interesting because of the great variety of AB cation pairs that are observed to form this structure [*see illustration on opposite page*]. As a rule the compressibility of a scheelite is determined by the compressibility of the large A polyhedron, which in turn depends on the electrostatic charge on the A cation.

Differential polyhedral compressibilities can lead to highly anisotropic, or direction-dependent, behavior in some crystals under pressure. For example, olivine, a common rock-forming mineral with the formula Mg_2SiO_4, has an array of magnesium-centered octahedrons that fill space in much the same way as the octahedrons do in the sodium chloride structure. In olivine, however, rigid silicon-centered tetrahedrons share edges with the octahedrons in one plane of the structure. As a result the compressibility of olivine in this plane is only about half the value observed in the unconstrained perpendicular direction. The small, rigid, silicon-centered polyhedrons, even though they account for only a small fraction of the volume, have an important influence on the bulk properties of the mineral.

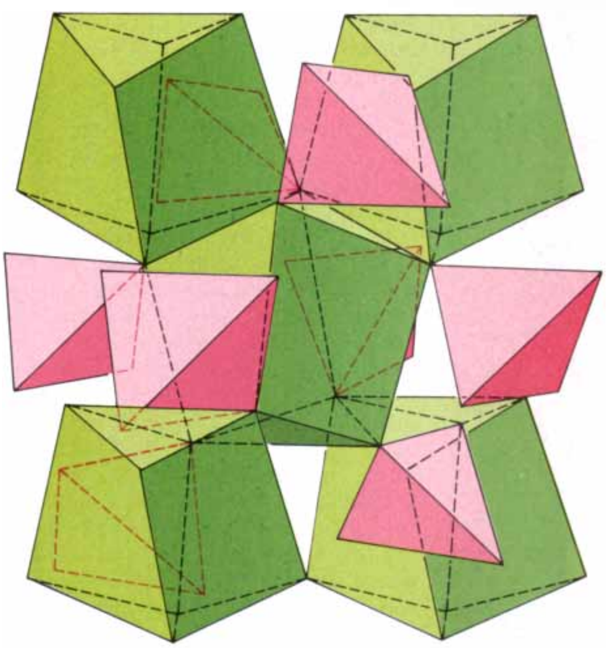
In all the examples considered above the polyhedrons share edges or faces; no bond-angle bending is observed in these structures. If the polyhedrons in a structure share corners, the volume of the crystal is often reduced by changing the angles between the polyhedrons rather than by significantly shortening the ionic bonds within them. The ubiquitous mineral quartz, the commonest oxide of silicon, is perhaps the best-known example of this phenomenon [*see illustration on next page*]. The basic building blocks of the quartz structure are silicon-centered tetrahedrons, each of which shares four corners with four other tetrahedrons to form a contin-



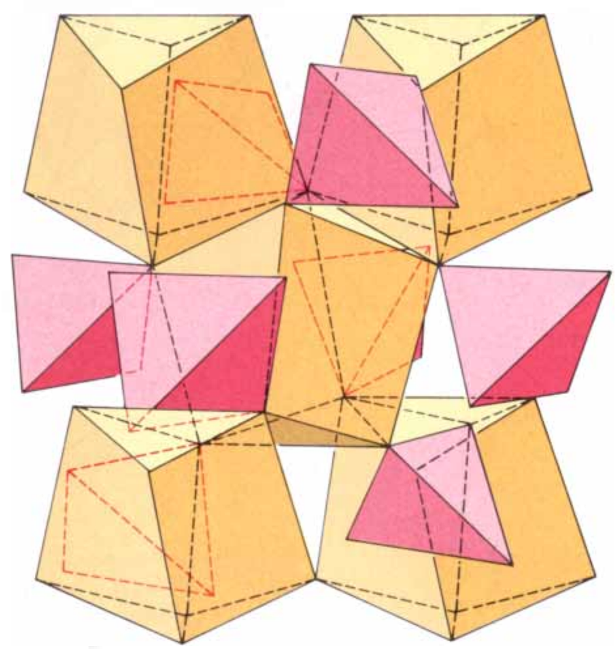
SODIUM IODATE ($NaIO_4$)



CALCIUM TUNGSTATE ($CaWO_4$)



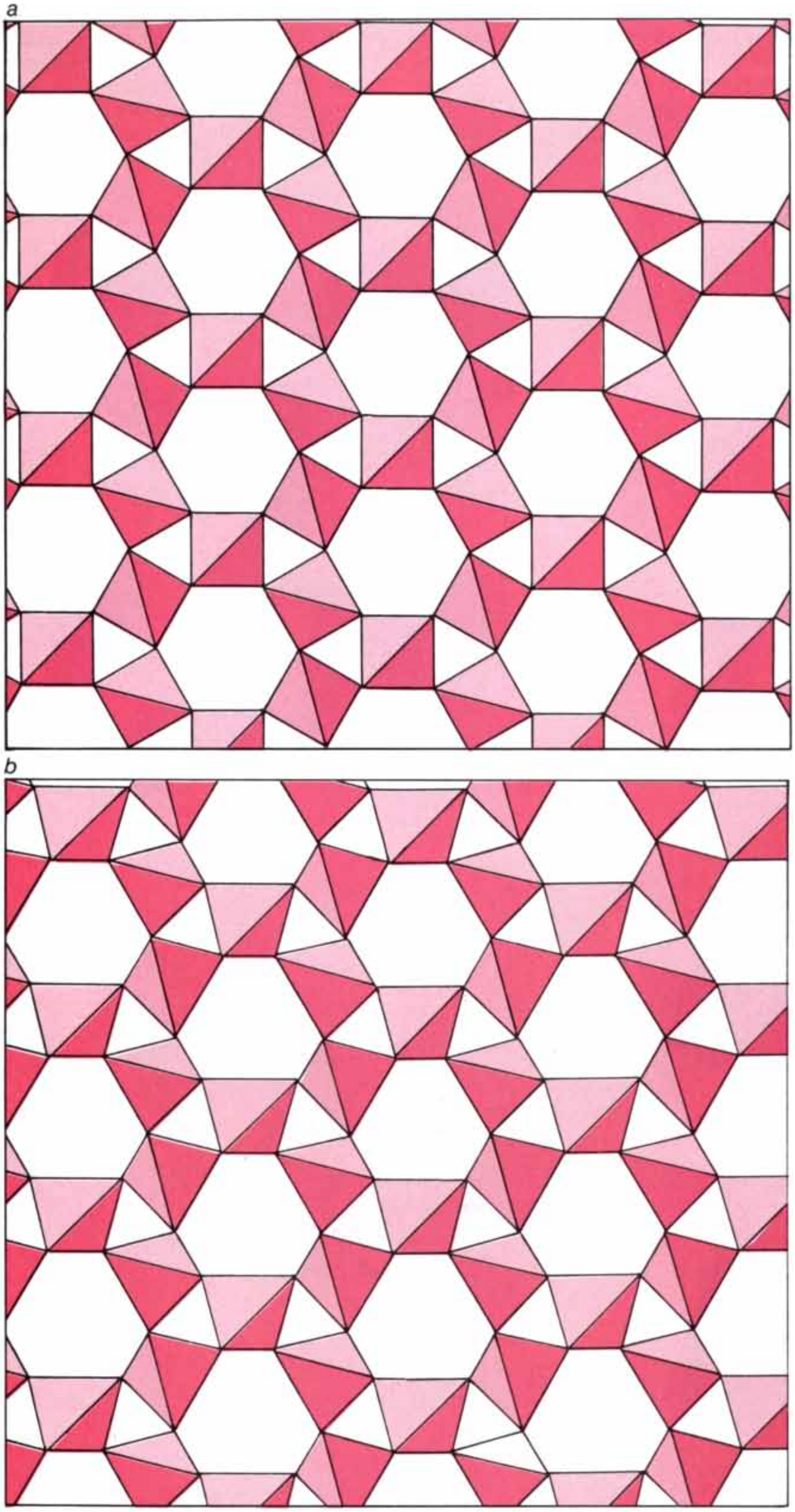
BISMUTH VANADATE ($BiVO_4$)



ZIRCONIUM GERMANATE ($ZrGeO_4$)

SCHEELITE-TYPE COMPOUNDS, which have the general chemical formula ABO_4, are composed of two types of cation polyhedron. The small *B* tetrahedrons, shown in red, are comparatively incompressible and occupy only 5 percent of the volume of the crystal. The large *A* polyhedrons, which can be viewed as distorted cubes, form a continuous three-dimensional network and fill most of the crystal's volume. Thus the reduction in the volume of the crystal under increased pressure is determined predominantly by the compression of the large eight-coordinated *A* polyhedrons. The four representative scheelites shown in this illustration are identical in crystal structure; the bulk crystals differ widely in compressibility, however, because of large differences in the compressibility of their constituent *A* polyhedrons. In this case the range of relative compressibility is indicated by a spectrum of colors: the red end of the spectrum designates the structure with the least compressible bonds and the violet end the structure with the most compressible bonds.

a



b

CRYSTAL STRUCTURE OF QUARTZ (SiO_2), the commonest oxide of silicon, consists of a complex, three-dimensional network of silicon-centered tetrahedrons, all of whose corners are shared (*a*). In quartz compression is attributable primarily to the bending of the bond angles between the polyhedrons (*b*). As a result a bulk crystal of quartz is 10 times more compressible than its constituent polyhedrons. Quartz serves as a counterexample to the intuitive notion that the hardest substances are the least compressible; because of its strong framework of silicon-oxygen bonds, quartz is one of the hardest compounds in nature, and yet it is one of the most compressible of the common rock-forming minerals.

uous three-dimensional framework. The small tetrahedrons with their highly charged cations are quite rigid, changing by less than 1 percent in volume even at pressures as high as tens of thousands of atmospheres. Nevertheless, quartz crystals are 10 times more compressible than their constituent polyhedrons, because the angles between the tetrahedrons are free to bend. The bending of the bond angles between the silicon and oxygen ions requires much less energy than the shortening of the bonds between the ions, and so the "tilting" of the polyhedrons becomes a more efficient compression mechanism.

In some compounds polyhedral tilting is responsible for a dramatic phase transition. For example, rhenium oxide (ReO_3) is composed of extremely incompressible octahedrons, each of which has a rhenium cation at the center surrounded by six oxygen anions [*see illustration on opposite page*]. At normal room temperature and pressure the crystal is cubically symmetrical; that is, all the octahedrons are constrained to align themselves along the axes of a cube. No tilting is possible in this cubic form of the crystal, and the oxide is correspondingly incompressible. At a pressure of about 5,000 atmospheres, however, a striking change takes place. The oxide undergoes a phase transition to a form with the same linkage of octahedrons but with a lower degree of symmetry; as a result the rhenium-oxygen clusters are free to tilt. At the transition point the compressibility of the crystal increases more than tenfold, simply because the dominant compression mechanism changes from bond shortening to bond-angle bending.

The greatest dimensional changes observed in crystals at high pressure are associated with intermolecular compression. The weak bonds between adjacent molecules may shorten by more than 1 percent per 1,000 atmospheres. Such a large effect dominates the behavior under compression of any material in which the bonding is molecular. The simplest condensed substances displaying molecular bonding are high-pressure crystals of elements that are gaseous at normal room conditions. The inert gases neon and argon, for example, can be liquefied at very low temperatures and then loaded into the gasket chamber of a diamond-anvil cell. The elemental fluids crystallize at high pressure, and even at room temperature they retain their crystalline form if they are confined at pressures of several thousand atmospheres. Hence it is possible to study the crystal structures and compressibilities of the crystalline phases.

X-ray-diffraction experiments reveal that at high pressure and room temperature crystals of neon and argon adopt the familiar cubic-close-packed structure, which is also characteristic of many metallic elements. The inert-gas crystals, however, have compressibilities unprecedented in magnitude; changes in volume of more than 1 percent per 1,000 atmospheres are observed in these materials, whereas the cubic-close-packed metals can be compressed by no more than about a tenth of this amount.

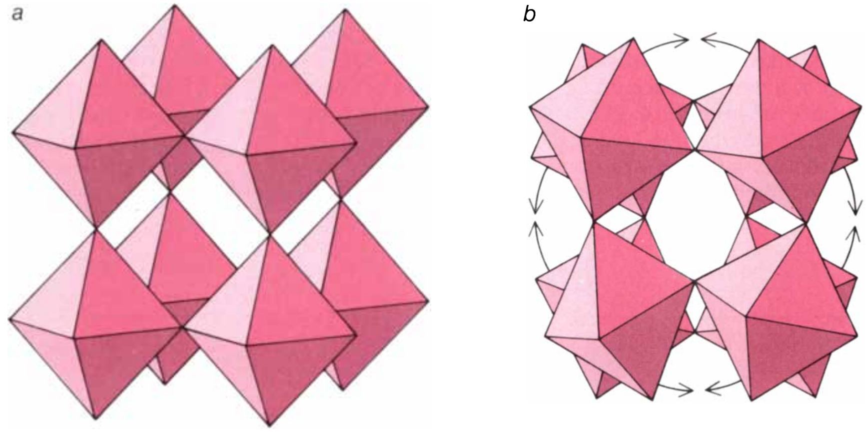
Gases consisting of multi-atom molecules can exhibit much the same behavior. For example, methane (CH_4) crystallizes at a pressure of about 16,000 atmospheres; the structure is also the cubic-close-packed arrangement, but the five-atom methane molecule rather than a single atom of neon or argon forms the close-packed unit. As in inert-gas crystals, the compressibility of crystallized methane is on the order of 1 percent per 1,000 atmospheres. A number of other multi-atom gases, including hydrogen (H_2), oxygen (O_2), nitrogen (N_2) and carbon dioxide (CO_2), also form extremely compressible molecular crystals at high pressure.

Crystals of pressurized gas are among the simplest to display molecular bonding, but they are by no means the only ones. Virtually all organic crystals, from solid forms of alcohols and hydrocarbons with a few tens of atoms per molecule to giant proteins with tens of thousands of atoms in each molecular unit, have highly compressible molecular bonds. It is easy to predict that all these substances will prove highly compressible compared with the inorganic minerals that form the solid earth. It is also easy to imagine the extreme anisotropic compression that must be displayed by some organic molecular crystals such as polymers, in which strong carbon-carbon bonds form continuous chains along one axis of the crystal and are joined laterally only by weak intermolecular forces.

Layered atomic structures exhibit fascinating and revealing behavior under compression. Layered compounds are anisotropic by definition; the bonding within each layer is stronger than the bonding between the layers. These differences in bond strength result in contrasts in compressibility along different crystal directions. Graphite, the commonest form of elemental carbon, is a classic example of a layered structure. Carbon-carbon bonding within each layer is as strong as that in diamond, and yet adjacent layers are bonded by molecular forces so weak that graphite is valued as a lubricant. The contrast between the compressibility parallel to and perpendicular to the layers is also dramatic; compression between the layers is more than 50 times greater than compression within the layers.

Layered silicates, including such common mineral groups as the micas, talcs and clays, can respond to pressure by a combination of all three compression mechanisms: bond shortening, bond-angle bending and intermolecular compression. As a result these minerals reveal much about how pressure affects condensed matter.

One of the simplest layered minerals is brucite, or magnesium hydroxide [$Mg(OH)_2$]; it is the active ingredient in milk of magnesia. The main building block of this structure is an octahedron of magnesium coordinated to six oxygen-hydrogen pairs. The polyhedrons are linked by shared edges to form continuous layers that are one polyhedron thick. Slabs of octahedrons, each with the basic brucite composition, are stacked one on top of another and are linked by intermolecular forces. The compression of magnesium hydroxide within the octahedral layers is controlled by the shortening of magnesium-oxygen bonds and hence is similar to the linear compressibility of magnesium oxide. The compressibility perpendicular to the layers is many times greater.

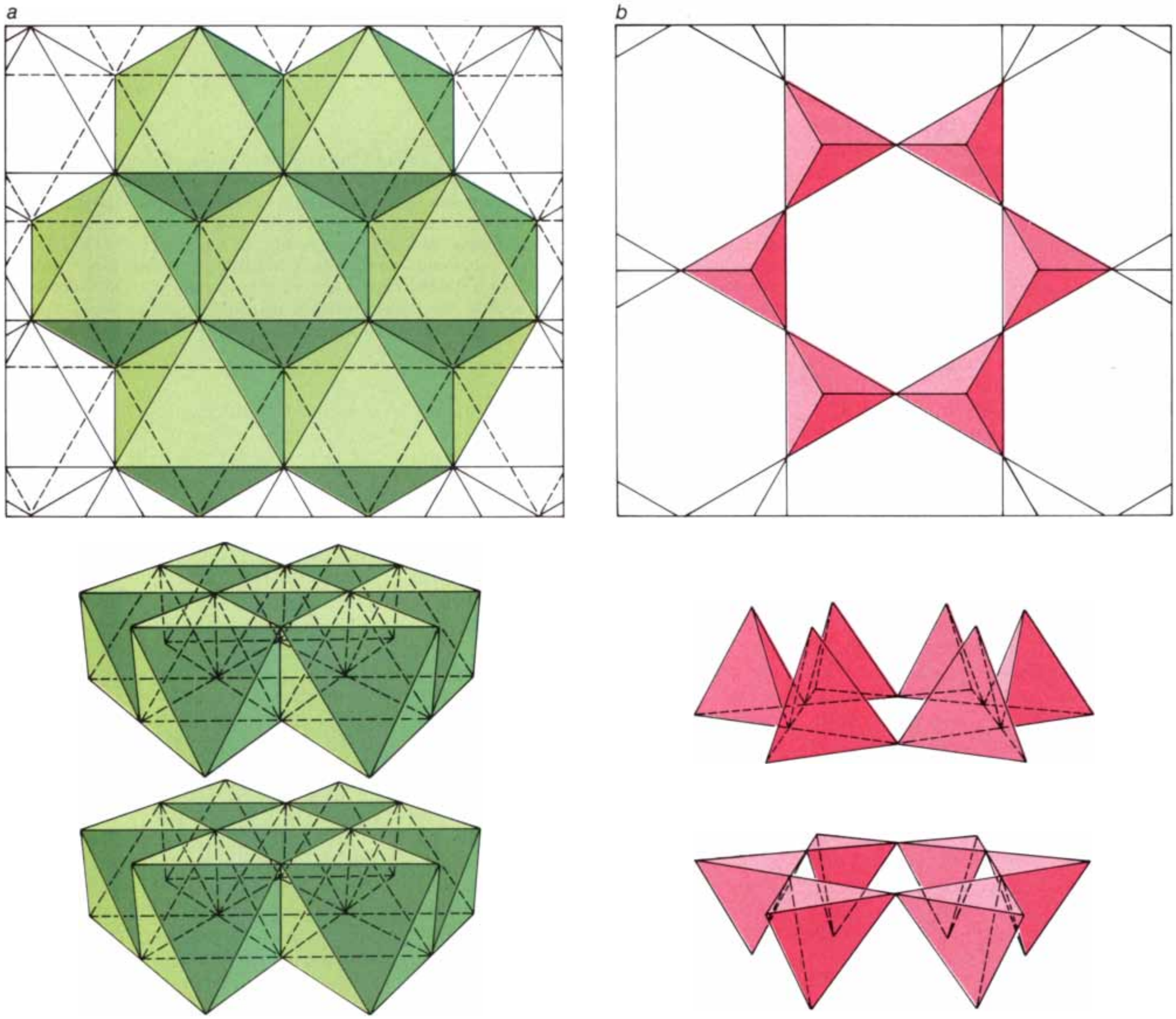


RHENIUM OXIDE (ReO_3) consists of a simple corner-linked array of rhenium-centered octahedrons. Each rhenium cation is positioned at the corner of a unit cube of the crystal structure, and the rhenium-oxygen bonds lie along the edges of the cube (*a*). Under normal room conditions the compound is one of the least compressible oxides known, because compression can occur only by shortening the rigid ionic bonds. At a pressure of about 5,000 atmospheres, however, rhenium oxide undergoes a phase transition to a structure with the same arrangement of octahedrons but with a lower degree of symmetry (*b*). The volume of the high-pressure form is reduced by the more favorable mechanism of bond-angle bending, or polyhedral tilting. At the transition the compressibility increases more than tenfold.

Layered silicates are all composed of several polyhedral layers stacked in some regular sequence. An octahedral layer, like that of brucite, is a ubiquitous feature of layered magnesium silicates; a layer of silicon tetrahedrons is another. The silicon tetrahedral sheet differs from the octahedral sheet in that all the polyhedral units share corners with adjacent units in the same layer. Even though the individual tetrahedrons are rigid, the layers themselves are highly compressible because of bond-angle bending. Bond-angle bending is also important because for a wide range of different compositions it allows the size of the tetrahedral layer to conform to the more constrained octahedral spacing. Accordingly micas, clays and other layered silicates occur in a remarkable range of compositions.

A number of layered silicates are composed entirely of octahedral and tetrahedral sheets. In the mineral serpentine, for example, alternating octahedral and tetrahedral layers are bonded together by intermolecular forces. The anisotropic compression of serpentine is similar to that of magnesium hydroxide: the compression is much greater in the direction of weak interlayer bonding. The most remarkable feature of this structure is related to bonding anisotropies. The two polyhedral layers are joined at flexible shared corners between octahedrons and tetrahedrons. Furthermore, each layer can bend. As a result serpentine layers can curve and form fibrous crystals in which the atomic layers are rolled up like a carpet. This fibrous form of serpentine, called chrysotile asbestos, has a layered atomic structure that is in distinct contrast to many other needlelike crystals, in which long chains of atoms line up along the crystal axis. The crystal form, like the anisotropic compression, is a consequence of the great difference of bond strengths within and between polyhedral layers.

Talclike silicates are composed of one octahedral layer sandwiched between two tetrahedral sheets. Used for centuries in lubricants and as talcum powder, these common layered minerals provide another example of anisotropic properties that result from anisotropic bonding. Talc and serpentine display all three compression mechanisms. Magnesium-oxygen and silicon-oxygen bond shortening occur within the polyhedral sheets. Tetrahedral layers conform to the size of octahedral layers through changes in the silicon-oxygen bond angles. Intermolecular compression accounts for most of the reduction in volume perpendicular to the layers. Without recognizing

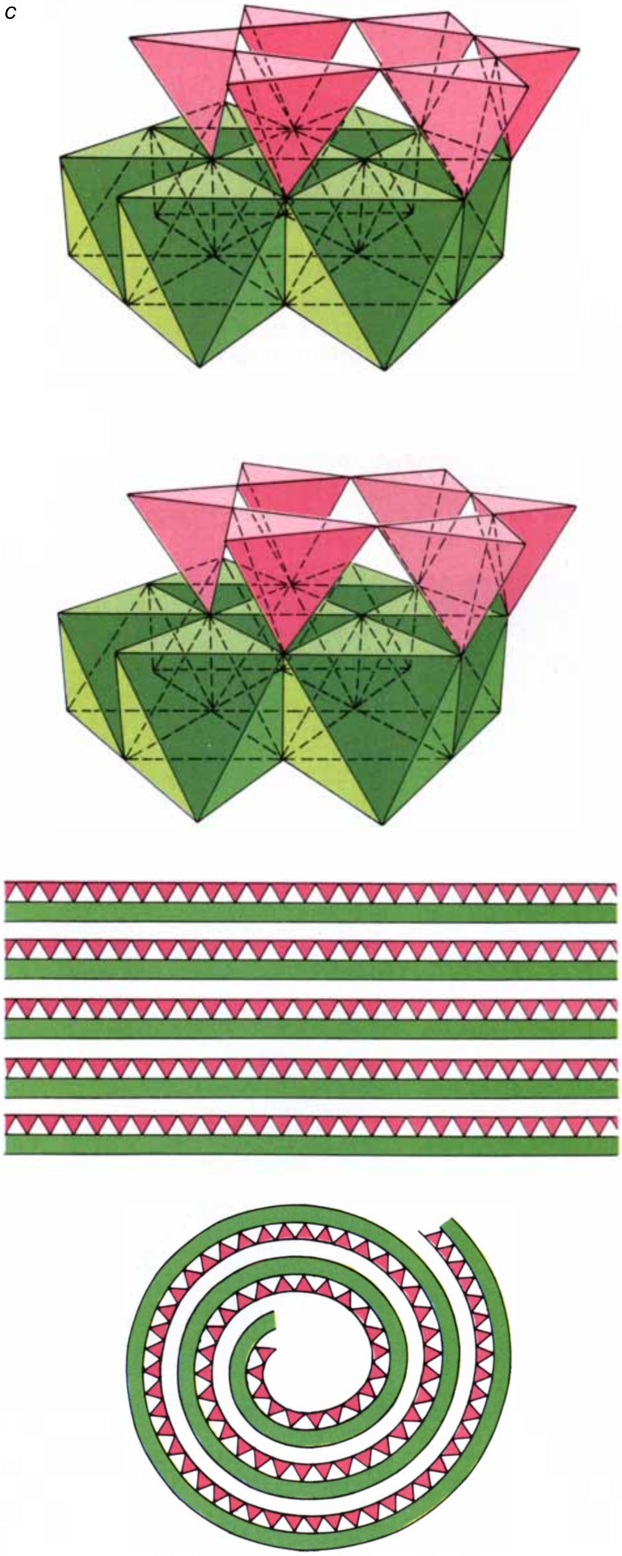


LAYERED SILICATES have a characteristic crystal structure that is based on the stacking of octahedral and tetrahedral sheets. The octahedral sheet viewed from the top and from the side in panel *a*, for example, is made up of a single layer of edge-sharing, cation-centered octahedrons. The simplest structure to incorporate this feature is the mineral brucite, or magnesium hydroxide [$Mg(OH)_2$], in which the octahedral layers are joined by weak molecular bonds. The compressibility of brucite within the layers is determined by the compressibility of the magnesium-oxygen bonds; the compressibility between the layers, however, is many times greater. In contrast, the polyhedral silicon dioxide sheet shown in panel *b* is composed of corner-linked silicon-centered tetrahedrons to form an infinite two-dimensional array of six-member rings. The compression of this layer is facilitated by bond-angle bending, which is energetically more favorable than the shortening of the rigid silicon-oxygen bonds. For example, the serpentine group of minerals is formed from the superposition of one such tetrahedral sheet onto one of the aforementioned octahedral sheets (*c*). In this structure some oxygen atoms are shared by both polyhedral layers. The double polyhedral sheets are also linked by intermolecular forces, which are so weak that many serpentines fail to form regular three-dimensional crystals. The individual layers often bend, curl or even roll up like a carpet to form long, needlelike crystals. In effect the two different polyhedral layers that make up the serpentine sheet behave like the bimetallic strip in a thermostat. One curious result is that the serpentine structure, which is basically planar, produces the most fibrous of all minerals: chrysotile asbestos (*bottom part of panel* c). Layered silicates such as talc, on the other hand, are composed of a stack of triple sheets, each of which consists of an octahedral layer sandwiched between two tetrahedral layers (*d*). Because of the symmetrical sequence of layers, talc is incapable of bending; that is what makes talcum powder a good lubricant. The compressibility of talc within the layers is equal to that of the octahedral layer, but the compressibility between the layers is several times greater.

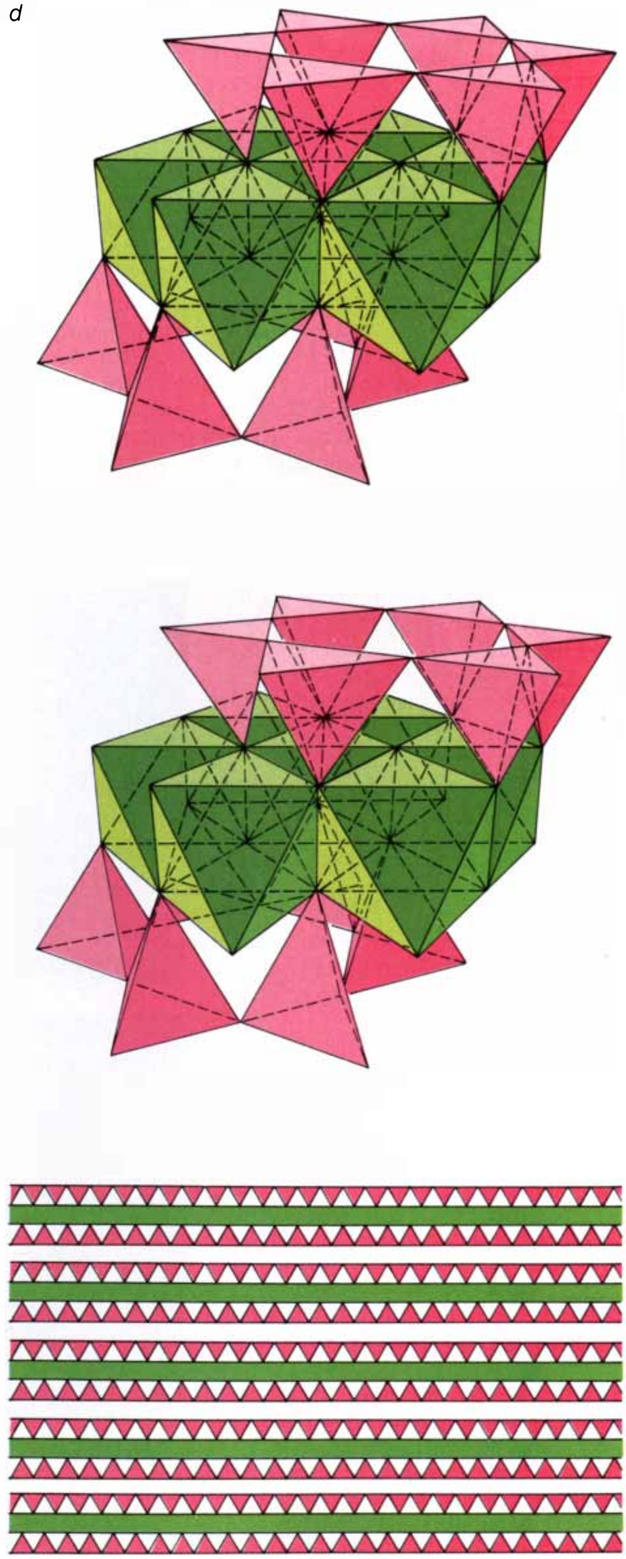
each of these three common mechanisms of crystal compression, it is not possible to understand or describe the response of these structures and others to high pressure.

In sum, the polyhedral approach enhances understanding of crystal compression and other crystal properties by establishing correlations between atomic-scale interactions and macroscopic crystal behavior. The approach is empirical in the sense that it is based on experimental data rather than on theoretical predictions. It is complementary to more rigorous quantum-chemical methods, which can lead to calculations of crystal properties from first principles but which are not yet applicable to the multi-atom structures found in nature. The polyhedral approach therefore greatly facilitates efforts to understand crystal compression and other phenomena in the real world.

c



d



Perceiving a Stable Environment

We perceive our surroundings as stable in spite of the relative motion given the environment by our own movement, because the perceptual system can compensate for such displacements

by Hans Wallach

The world shifts around an observer as he moves through it. As he approaches an object it expands within his field of view; as he passes an object it turns with respect to his changing position. A turn or nod of the head alters the orientation of the surroundings; eye movements shift the image of the world that is projected on the retina. Yet we are ordinarily not aware of the environmental motions precipitated by our own activity. What mechanisms enable us to discount the effects of our own movement and perceive the environment as being stable? In 12 years of research other workers and I have uncovered elegant and precise compensation processes.

An observer sees a stationary scene when he turns his head but would perceive the environment as moving if it actually revolved around him while he was motionless. Conceivably such tendencies reflect a perceptual blockage, triggered by signals that indicate body movement arriving from the tendons, the joints and the vestibular organs. One might propose that any perception of environmental motion is ruled out in the presence of such proprioceptive information.

The notion is easily dismissed. A subject wearing glasses that reverse the direction in which the environment shifts as he turns his head will, at least at first, be acutely conscious of motion in his surroundings. Similarly, a subject who walks forward while facing a large mirror held at eye level will perceive the reflections of the environment behind him as shrinking or receding. The displacements are equal in magnitude, although opposite in direction, to those that ordinarily occur ahead of an observer as he moves forward, but whereas the normal displacements go unperceived, the subject is vividly aware of the abnormal motions of the scene in the mirror.

A converse instance also demonstrates that proprioceptive inputs indicating body movement do not simply block the perception of environmental motion. In an effect familiar to museum visitors the subject of a painting sometimes appears to rotate as the viewer passes it, so that regardless of his actual position with respect to the picture, he seems to stand in the same relation to the scene it depicts; similarly, a head in the painting may seem to turn toward the viewer wherever he walks. If the painted scene existed in three dimensions, it would rotate counterclockwise in relation to an observer passing it on his right (clockwise if he passed it on his left). The configuration of its components would change and individual elements, such as the head, would be revealed from a steadily changing angle. Nevertheless, the observer would perceive the scene as being immobile.

Because the painted scene is flat, however, and the arrangement of its

SEEMING MOTION of a figure in a photograph or a painting is familiar to museumgoers. A three-dimensional object, such as the head of Nefertiti, turns with respect to a passing observer in a direction opposite to that of his own movement. As is shown in the top row, the observer sees the head from a steadily changing angle, just as if it were actually to rotate while he remained still. Yet the relative motion of the figure goes unperceived;

elements is the same whichever direction the painting is viewed from, the normal counterclockwise rotation of the scene does not occur. At the same time it no longer appears immobile but instead seems to turn clockwise. If the perception of environmental motion were simply blocked when the observer himself moves, the absence of normal relative motion would not affect perceptual experience.

Such observations suggest that visual data on the relative displacements of the surroundings is compensated for rather than blocked. The compensation mechanisms compare visual inputs with the proprioceptive data that represent body movement. When the visual and proprioceptive stand in a certain fixed relation, the environment is perceived as being stable. Violations of that relation, in contrast, give rise to an awareness of motion.

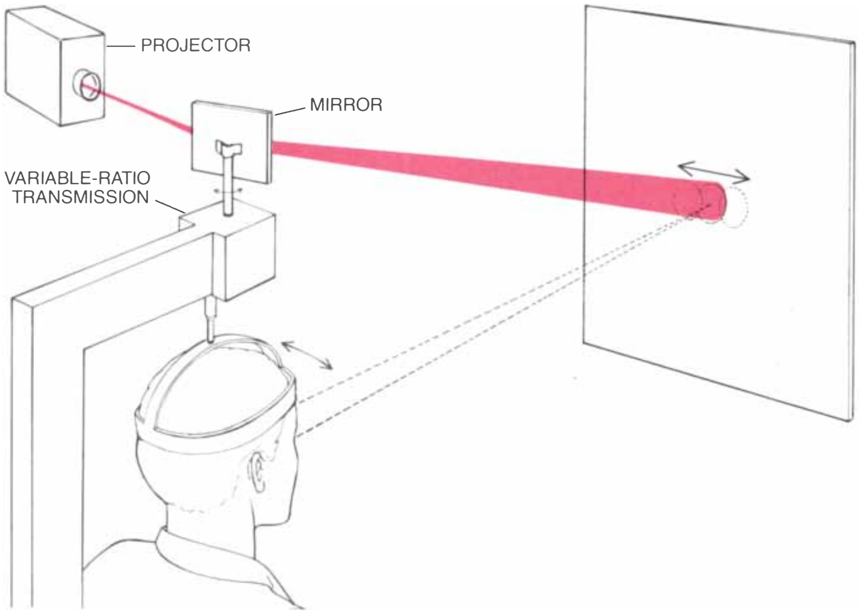
As a first step toward a more detailed understanding of the compensation process Jerome H. Kravitz and I, working at Swarthmore College, investigated its precision. We asked how much discrepancy between body movement and environmental motion is allowed before the observer perceives his surroundings as moving. When an observer turns his head 20 degrees, the visual environment ordinarily is displaced by 20 degrees in the opposite direction and the environment appears to remain stable. Will it still seem stable if it is displaced by 15 degrees during a 20-degree head turn, or by 25 degrees? To use the language of our research, what is the immobility range for head turning, the range of actual environmental displacements through which the surroundings continue to seem immobile?

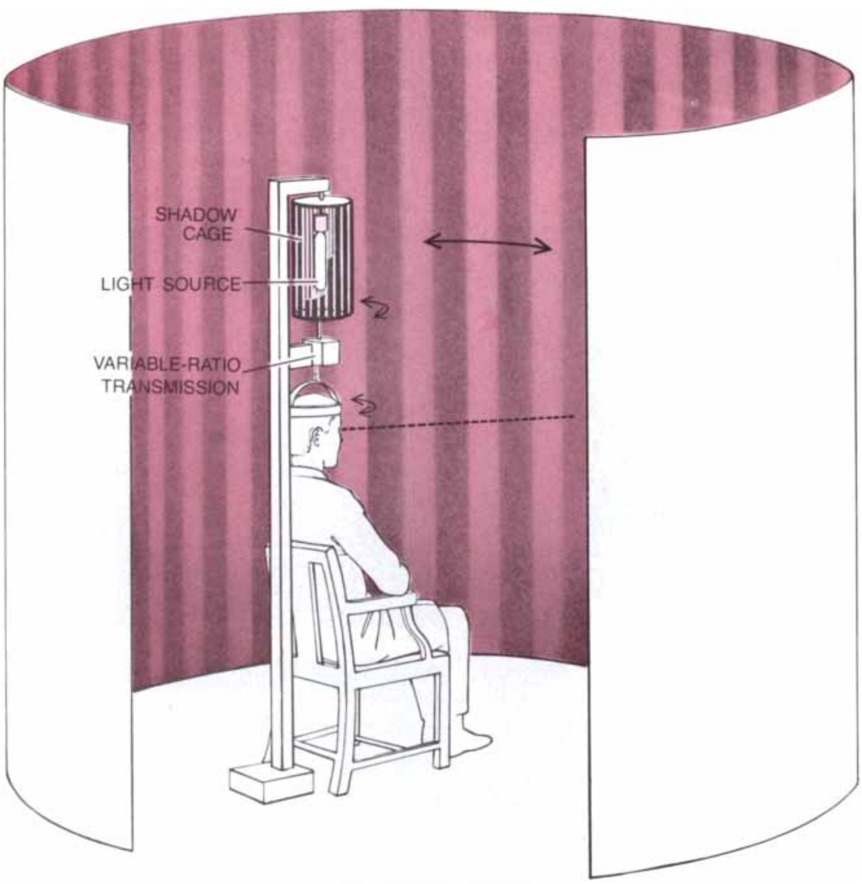
To answer the question we devised an experimental apparatus in which motions of the subject's visual environment could be coupled with head movements. The subject wore a headgear to which a vertical shaft was attached, at a point directly over the head's axis of rotation. The other end of the shaft entered a variable-ratio transmission, mounted above the subject's head. The output shaft of the transmission in turn affected the subject's visual environment.

In some of our experiments a mirror was attached to the output shaft. The beam of a slide projector was reflected from the mirror onto a screen facing the subject. When head movements were transmitted through the output shaft to the mirror, the projected image was shifted across the screen; the setting of the transmission determined the displacement produced by a head movement of a given extent. In other experiments we substituted for the mirror and projector a shadow cage, a cylindrical arrangement of rods with a point source of light at the center. The arrangement cast a pattern of shadows on a large cylindrical screen enclosing the subject. By changing the transmission setting the investigator could alter the ratio of the shadows' angular motion to the angle of head rotation.

In determining the immobility range under these conditions, we chose as a unit of measurement the displacement ratio: the ratio of the actual motion of the environment to the subject's head movement. In a stable environment the displacement ratio is zero; objectively speaking, the surroundings do not move at all during head rotation. We could simulate that condition by adjusting the transmission so that it passed no motion at all from the head to the visual environment. In contrast,

the figure appears immobile. When a three-dimensional object or scene is depicted in two dimensions, however, as in a painting or in the photograph of Nefertiti's head, shown in the bottom row, the same rotation with respect to an observer cannot take place. Although a passerby sees the representation from a range of angles, the angle from which the figure or scene is depicted on the two-dimensional surface does not change. In the absence of the normal contrary rotation, the object or scene that is depicted no longer appears to be immobile but seems to turn with the observer. Thus Nefertiti's gaze seems to follow the viewer as he passes the photograph.





KINETIC RELATION OF HEAD AND SURROUNDINGS was altered in an apparatus designed around a variable-ratio transmission. When the experimental subject turned his head, a shaft connected to the headgear carried the rotation to the transmission. The transmission imparted a proportion of the motion to the visual environment. In some experiments the transmission turned a mirror, which reflected an image from a slide projector onto a screen (*top*). The setting of the transmission determined how far the projected image moved, and in what direction, for a given angle of head rotation. In other experiments the visual environment consisted of a pattern of shadows thrown on a cylindrical screen surrounding the subject by a cage of slender rods enclosing a light source (*bottom*). The setting of the transmission determined the relation between the subject's head rotation and the displacement of the shadows. The apparatus enabled workers to determine how sensitive a subject is to environmental motion during head turning; it also served to show that subjects adapt to abnormal displacements of the environment, eventually failing to perceive them.

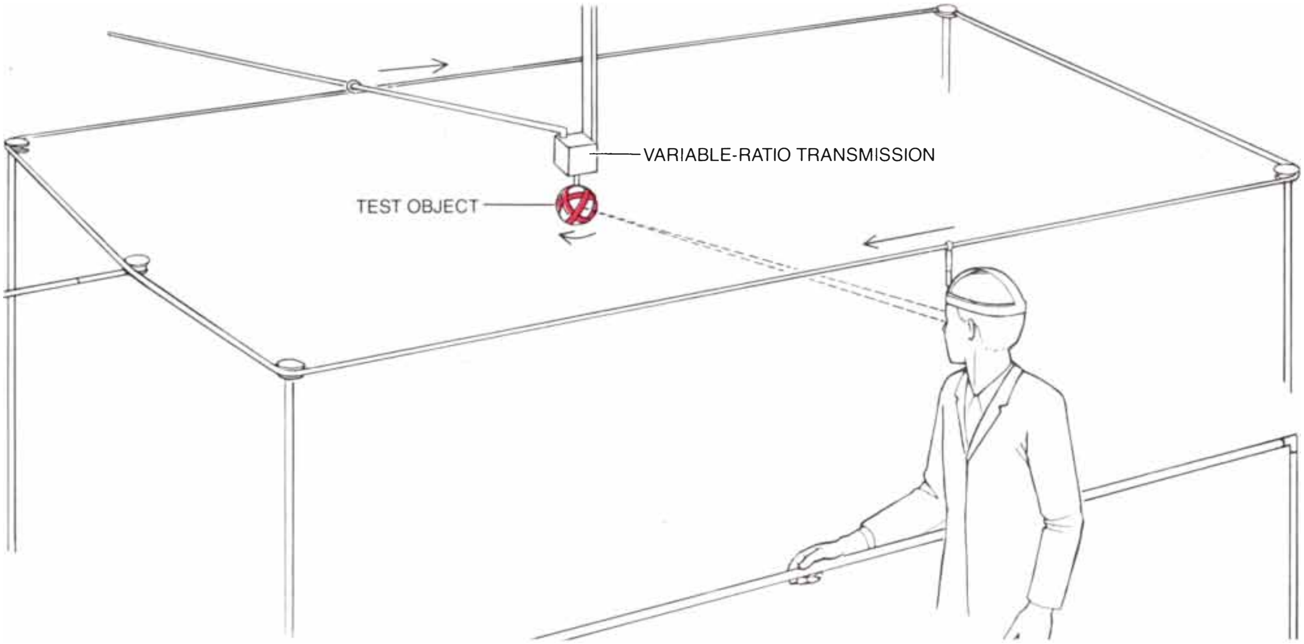
when the apparatus was adjusted so that the visual environment moved by the same amount as the observer's head, whether in the same direction or in the opposite direction, the displacement ratio was 1.

To measure the range of displacement ratios within which a subject perceived the visual environment to be immobile, we first adjusted the transmission so that the environment was displaced in the direction of head rotation in a proportion sufficiently high for the subject to perceive motion. We then readjusted the transmission to reduce the displacement ratio in steps, half a percentage point at a time. At each step the subject sampled the environment by turning his head from side to side. The point on the continuum of displacement ratios at which the subject no longer detected motion marked one boundary of the immobility range. We then repeated the process for environmental displacements in a direction opposite to that of head movement to determine the immobility range's other boundary. The procedure yielded measured immobility ranges extending from .015 to .030 displacement ratios to each side of objective immobility. In other words, our subjects were unaware of displacements that did not exceed from 1.5 to 3 percent of the extent of their head movements.

The immobility range found in other compensation processes is considerably wider. Another relative motion of the environment for which such a process exists is the rotation of objects in relation to a passing observer. To study the compensation mechanism Linda Stanton and Dean Becker, then at Swarthmore, and I constructed an apparatus by means of which an object could be made to turn as a subject walked by it.

Again the setup used a variable-ratio transmission; the subject's transverse movement as he walked past the apparatus was transmitted to the input shaft through a horizontal rod, attached by a system of cables to the subject's headgear. The reference object (a translucent sphere with a dark pattern on its surface or a framework of luminous wires) was suspended from the output shaft of the transmission. The setting of the transmission determined the relation between the subject's movement and the object's rotation. The setup enabled us to determine that for motions of this kind an observer typically does not perceive actual rotations when they amount to less than 40 percent of his own angular displacement.

A relatively broad immobility range is also characteristic of the process that compensates for eye movements,



RODS AND CABLES **transmitted a subject's transverse movement to the input shaft of a variable-ratio transmission, which then turned a patterned globe. The setting of the transmission determined the extent of the globe's rotation. The subject watched the object while he walked past it, guided by a handrail. The apparatus was meant to determine how much of a departure is needed from the normal rotation of a stationary object with respect to a passing observer before the observer becomes aware that the object is turning.**

which displace the image of the surroundings that is projected on the retina of the eye. Separately, Arien Mack of the New School for Social Research and William R. Whipple, who worked in my laboratory, devised experiments in which a subject's eye movements, electronically monitored, triggered the simultaneous motion of a luminous point or of a circle across an oscilloscope screen. The degree of displacement could be varied. Mack found that subjects could reliably perceive a shift only when it subtended an angle greater than 20 percent of the eye movement. Under different conditions Whipple found somewhat greater sensitivity: his subjects perceived motions amounting to 8 percent or more of simultaneous eye movements.

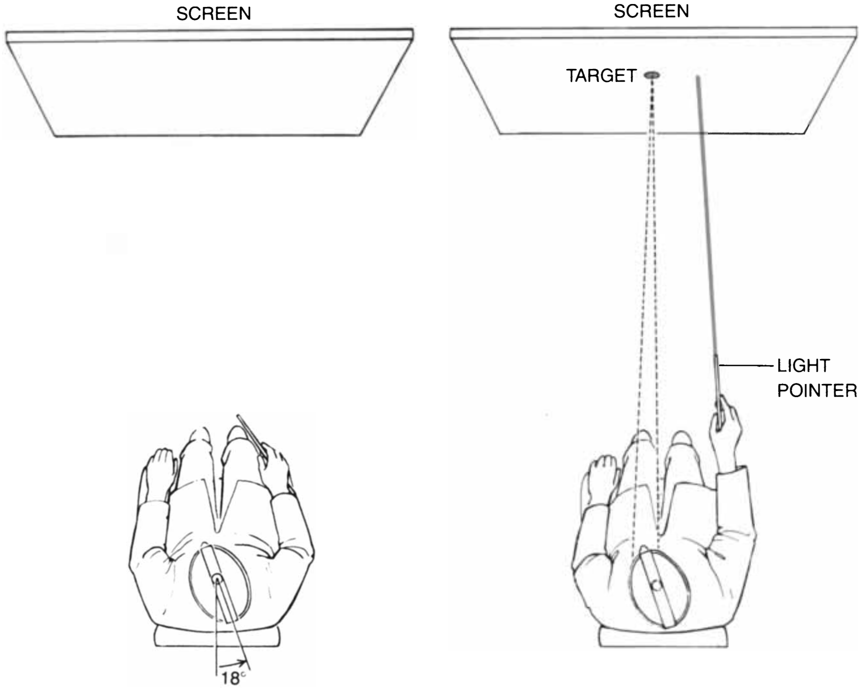
The compensation process for the effect of head movements is thus the most accurate of the mechanisms that have been studied, a characteristic that made it well suited to further experimental inquiry. In an attempt to find out how the process operates we exposed subjects to environments in which the normal one-to-one relation of head movement to the displacement of the surroundings does not hold. We wanted to study the effect on the perceptual system of an environment that moves by disproportionate amounts as the observer turns his head.

It has been known since the turn of the century that the compensation process can adapt to an abnormal relation between head rotation and environmental displacement. In 1896 George Stratton of the University of California at Berkeley wore a system of lenses that inverted his surroundings and caused his visual environment to be displaced in the same direction as his head movements rather than opposite to them, as is normal. At first he was aware of the abnormal motions, but after he had worn the device for two days his environment stabilized. His perceptual system seemed to have adapted to the new relation between his own movements and the resulting motions of the environment—a supposition that was confirmed when he removed the device after eight days and found that with the normal relation restored his surroundings appeared to swing wildly opposite to his head movements. In a process of adaptation to abnormal conditions Stratton's immobility range had shifted.

Our experimental arrangement for measuring the immobility range and our means of characterizing it numerically enabled Kravitz and me to gauge such adaptation effects. In an initial experiment subjects wore wide-angle glasses that caused the visual environment to shift in the direction of head movement by 34 percent of the head rotation's angular extent (a displacement ratio of .34). After six hours we tested the subjects using the procedure described above.

Without the goggles the subjects now perceived the environment to be stable only when it moved in synchrony with head movement at a displacement ratio that averaged .175. Their adaptation hence was partial; their immobility range had shifted enough to accommodate about half of the abnormal displacement they had experienced. The extent of adaptation varied greatly among the 12 subjects, however; one of them adapted fully during his six hours of exposure, perceiving environmental stability at a displacement ratio of .34. In later experiments we encountered faster adaptation. We used the apparatus itself to provide the abnormal environment; we set the transmission to a large displacement ratio and had the subject turn his head from side to side for 10 minutes.

Although such adaptation shifted the immobility range from a position straddling objective immobility to a point elsewhere on the scale of displacement ratios, the width of the range stayed the same. The compensation process remained as accurate as before. It seemed likely that adaptation leaves the process unaltered. It must therefore affect one of the inputs to the compensatory mechanism: either the visual data on the relative motion of the subject's surroundings or the proprioceptive information that



POINTING TEST required a subject to gauge the direction of a target from the extent of the eye movement needed to look at it. The test was done in the dark, so that the subject, who held a pointer, could not aim at the target simply by sighting along the pointer. The subject first turned his head to the left, where at an angle of 18 degrees a stop in the apparatus prevented further rotation (*left*). When a target spot on a screen ahead of his torso was illuminated, the subject moved his eyes to the right to look at the target and aimed the pointer. The investigator then turned on a light source within the pointer and noted where the point of light it cast fell in relation to the target. Subjects who had adapted to a condition in which the entire environment was displaced by less than the normal amount during head turning pointed too far to the right (*right*), suggesting that the extent of the eye movement was misjudged. Apparently adaptation to abnormally short displacements of the surroundings entails an overrating of the eye movements that are needed to track the environment, so that motions of the surroundings are perceived as being longer than they actually are.

indicates the extent of head rotation.

Which category of input does adaptation to a new kinetic relation of head and surroundings affect? Suppose adaptation changes the proprioceptive inputs to the compensation process so that, for example, the nervous system transmits the same signals during a 40-degree head turn that it ordinarily sends to indicate a 20-degree turn. The visual environment will therefore need to rotate in synchrony with head movement by an angle of 20 degrees (equivalent to a displacement ratio of .5) for it to be perceived as stable. Such an adaptation effect should be apparent, however, not only in the visual immobility range but also in the immobility ranges for other kinds of stimuli, since each immobility range reflects a compensation process that acts on proprioceptive inputs.

By using the variable-ratio transmission to control the rate at which a sound was shifted along a bank of speakers as a subject turned his head, Kravitz and I could test subjects' auditory immobility range. We found that adaptation to abnormal shifts in the visual environment was not reflected in the immobility range for sounds, which remained centered on objective immobility. Apparently the adaptation we had induced affected only the visual basis of compensation.

A narrowing of experimental scope was now in order. It was clear that adaptation entails a change in the visual information on the displacement of the surroundings; adaptation experiments therefore promised to serve as an experimental tool for revealing the nature of that information. Joshua Bacon, now at Tufts University, and I varied the conditions of adaptation in order to answer two questions suggested by everyday experience. During a head turn the eyes often track a single point in the surroundings. Do the lengths of the eye movements needed to track a stationary point serve as the visual data on the displacement of the surroundings? Or do the motions enter the compensation process in a more complex form, in which the visual environment is represented as it exists in relation to the head?

Using the variable-ratio transmission, the shadow cage with its light source and a second source of light, Bacon and I separated the subject's visual environment into two components, one that shifted during head movements and one that did not. In the first set of trials the screen surrounding the subject displayed both a stationary pattern of shadows and a point of light that moved with the subject's head movements at a displacement ratio of .4. The subject was instructed to fix his gaze on the light as he turned his head from side to side during a 10-minute adaptation period.

Exposure to these conditions did alter subjects' immobility range, suggesting that the extent of eye movements does play a part in the adaptation process. A single fixation point in the visual environment that consistently required abnormal tracking eye movements, it seemed, was enough to induce the kind of adaptation evoked by abnormal displacements of the environment as a whole. Yet the adaptation effect was smaller than the one we measured for the same period of "normal adaptation," in which, as in earlier experiments, the entire environment moved during the adaptation period. Although the subjects' eyes did not track the stationary background pattern, it apparently was also represented among the visual data on environmental displacement and resulted in a diminished adaptation effect.

To confirm the role of the background pattern we reversed the experimental conditions. In the second set of trials the point of light remained stationary as the subject turned his head, keeping his eyes fixed on the light, whereas the pattern of shadows moved. Again we measured an adaptation effect. Since the eye movements with which the subject tracked the light were of normal length, the adaptation could only have resulted from a representation of the shadows and their motions in relation to the subject's head. Such a representation presumably derived from the slow displacement of the pattern across the retina of the eye as the subject turned his head.

It appeared that the answer to both of our questions was yes. We named the kind of adaptation evoked by the first experiment eye-movement adaptation, because it seemed to be mediated by the extent of eye movements alone. We named the second effect field adaptation; it seemed to reflect displacements of the visual field as a whole, divorced from any abnormality of tracking eye movements.

In further experiments Bacon and I confirmed the distinction between the two kinds of adaptation. The tracking movements of the eyes that compensate for head movement occur both when the environment contains a stationary point on which the eyes can fix and in the dark, when the eyes are not guided by any visual stimulation. The automatic compensatory movements made in the dark are known as the Dodge reflex. The Dodge reflex appears to be precisely tuned to the normal relation between head rotation and tracking movements of the eyes, and therefore to the normal relation of head movement and environmental displacement.

When we measured the extent of the Dodge reflex in subjects who had undergone 10 minutes of eye-movement adaptation at a displacement ratio of .4, however, we found a significant change. The point of light the subjects tracked during their adaptation period moved in the direction in which they turned their head; the tracking movements needed for a given extent of head movement were therefore shorter than normal. The Dodge-reflex movements showed a corresponding shortening, by an average of 13 percent. In contrast, we found normal reflex movements in subjects who had experienced field adaptation, which entails no alteration of tracking movements.

We also tested the group of subjects in whom we had induced normal adaptation. Like the other groups, they had experienced 10 minutes of adaptation at a displacement ratio of .4; their immobility range now fell, on the average, at a displacement ratio of .13. We were intrigued to find that the extent of the Dodge reflex in these subjects was also shorter than normal by 13 percent, on the average—an amount equal to the entire adaptation effect.

In order for a subject to perceive the environment as being stable when the eye movements that track it are shorter than normal, the nervous system must overrate the movements. Bacon and I devised a direct demonstration of such overrating. In a darkened room the subject was seated in an apparatus that limited the movement of the head. He was instructed to turn his head to the left; at an angle of 18 degrees a stop prevented further movement and a target straight ahead of his torso was illuminated. The subject then looked toward the target and pointed at it without straightening his head. In the dark the subject could not see his hand; thus his appraisal of the target's direction depended entirely on a rating of eye movement.

Following eye-movement adaptation subjects pointed an average of 13 percent too far to the right; they overestimated their eye movements by an amount that precisely offset the shortening we had measured in tests of the Dodge reflex. A similar pointing error was found in normal-adapted subjects, but field-adapted subjects, who had tracked a stationary point during the adaptation period and showed a normal Dodge reflex, appraised their eye movements accurately and aimed straight at the target.

The Dodge-reflex and pointing tests detect eye-movement adaptation only; another test Bacon and I devised singles out field adaptation. In the dark, with the head turned 18 degrees to one side, the subject had to move a luminous target to a position straight in front of his torso. Both normal- and eye-movement-adapted subjects performed this "straight ahead" test successfully. Field-adapted subjects erred significantly, however, placing the target off to one side by an amount equivalent to a displacement ratio of .087.

This combination of results suggests that eye-movement adaptation is the process underlying normal adaptation as produced in our experimental apparatus. The tracking of a single point during head movement results in adaptation to abnormal displacements of the entire visual environment. Field adaptation appears to be an entirely distinct phenomenon. It probably takes place at a higher level of perceptual processing, at which eye position and eye movements have been taken into account and displacements of the visual field as a whole are represented.

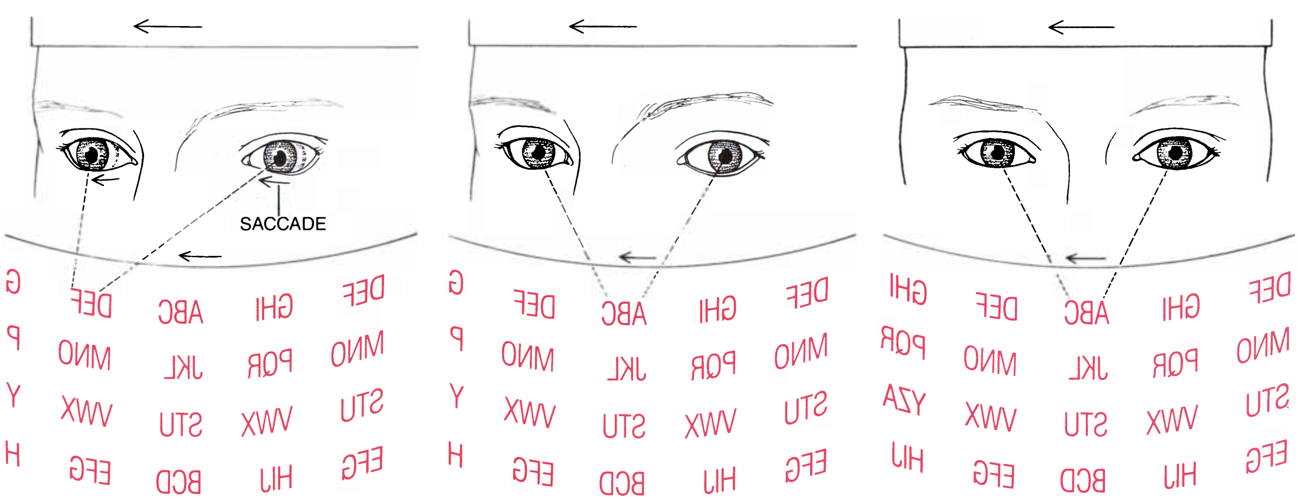
Although field adaptation played no part in the normal adaptation we induced, it does not follow that under ordinary conditions eye movements alone signal the extent of environmental displacement during head rotation. The everyday world differs from the experimental situations my co-workers and I devised. In particular, an observer does not always confine his gaze to a single point in the surroundings while turning his head.

Under ordinary circumstances the tracking of single points is often interrupted by rapid movements known as saccades, in which the eyes shift from one fixation point to another. A saccade brings to the center of the visual field a point in the environment that the eyes had not previously tracked. To find out how the displacement of the surroundings is registered under these more complex conditions we devised an adaptation condition in which a subject was forced to intersperse tracking eye movements with saccades. We used the variable-ratio transmission with a mirror and a projector displaying a slide showing columns of letters, each column consisting of rows of three letters. The image fell on a screen filling the subject's field of view. The array of letters shifted by a displacement ratio of .4 as the subject moved his head.

Each subject underwent a 10-minute adaptation period during which he turned his head continuously from side to side. During each turn to the right he read, from left to right, the letters in specified rows of adjacent columns. Thus the subject had to perform saccades, needed to shift his gaze from column to column, as well as tracking movements, needed to fix his eyes on individual rows of letters. The subject was then given the straight-ahead test, which revealed a substantial measure of field adaptation.

Under these more elaborate (and more lifelike) experimental conditions environmental displacements were registered not only at the level of eye movements but also at a higher perceptual level. It appears likely that in ordinary circumstances the process of compensation (which may in fact consist of two processes operating simultaneously) employs both kinds of input, comparing them with proprioceptive data to yield a perception of stable surroundings. The distinct mechanisms that are altered in field and eye-movement adaptation are both at work in everyday perception.

Through processes of compensaton such as the one my co-workers and I have studied, our perceptual system conveys a world that remains reassuringly stable in spite of our own movement. In everyday life we are not aware of their operation; they simply give us the world as it is. An analysis of what is entailed in the perception of a stable environment, however, makes it apparent that such mechanisms exist. Only through adaptation experiments in which we altered the conditions to which one such process is attuned could we lay bare its workings.



TRACKING AND SACCADIC EYE MOVEMENTS were both evoked in an experiment intended to determine the effect of the combined movements on adaptation to an environment that is displaced by an abnormal amount during a rotation of the head. Using the projector and mirror setup shown on page 120 the investigators displayed a grid of letters on a cylindrical screen surrounding the subject. The variable-ratio transmission linking the subject's head to the mirror caused the array to be displaced in synchrony with head movement but by a smaller angular extent. As the subject turned his head he had to read the columns of letters one after another. The subject thus employed both tracking eye movements, needed to fixate on a single column as his head turned slowly past it (*right and middle*), and quick jumps, or saccades, needed to shift his gaze from one column to the next (*left*). Under these complex conditions of eye movement the abnormal displacement of the visual environment is registered in two forms. One is length of tracking eye movements, which in the experiment were shorter than normal because of the motion of the display. The other, the investigators discovered, is a representation at a higher perceptual level, embodying the subject's visual field as a whole in relation to the head.

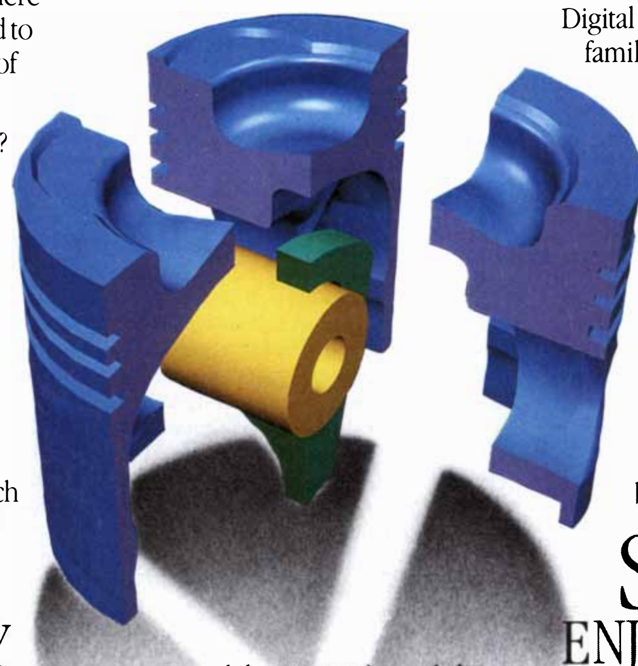
PROFITABILITY

THE INSIDE STORY

Automation has revolutionized engineering. And given rise to an amazing number of computer hardware and software products in the process. There are computer systems of every size and description. There are software programs designed to automate virtually every aspect of engineering.

Can every product be evaluated? Probably not.

Overcrowding in the marketplace can cause real problems for an engineering manager or executive responsible for structuring a long term automation strategy. It is with these courageous individuals, in particular, that we would like to share our very practical approach to automating engineering and increasing profitability.



USE THE VERY BEST SOFTWARE...

Granted that the subject of application software can seem overwhelming. There is, after all, so much of it around. But industry experts agree: There is *nothing* quite like PDA Engineering's PATRAN.®

Put simply, PATRAN is an extraordinarily powerful tool that can help your company solve its most complex engineering problems today, while preparing it for the solution of even more complex problems it will encounter tomorrow.

Here is a CAE application so elegantly designed, so mathematically complete, it models incredibly complex solids as easily as it produces high quality exploded and cutaway views.

As a modeler, PATRAN's capabilities are unmatched: wire frame, surface, solid, even environmental modeling – all with equal facility.

Its applications seem endless: fluid mechanics, thermodynamics, acoustics, electrical and magnetic analyses, radar cross sections, geophysical modeling, camouflage patterns, space planning, medical research, ergonomics (the list goes on and on) – all made possible by PATRAN's inherent completeness. And with an ever-expanding set of analysis modules and full color interactive postprocessing, which brings results evaluation to life, PATRAN provides an extraordinarily powerful and continuingly relevant tool that solves tomorrow's problems as well as today's, at a price few can afford to pass up.

WITH THE VERY BEST HARDWARE...

PATRAN is available on most of the World's finest. Take, for example, Digital Equipment Corporation's VAX™ family of compatible 32-bit computer systems.

Digital has been developing, manufacturing and supporting computer systems for over a quarter of a century now. Digital's VAX computer systems range in size from very large to very small. Each is available with networking, communications and office solutions – virtually *all* of the computing capabilities your company needs.

SIT BACK AND ENJOY THE SHOW!

When you've got PDA Engineering's PATRAN working for you in an interactive computing environment, you've got the means to increase the profitability of your company's operations – easily, economically, yes, even dramatically.

Now that you've got the inside story, we'd be happy to tell you more. Just call PDA Engineering's Software Products Division at 714-556-2800.

PATRAN.............. THE PICTURE OF PROFITABILITY.

pda ENGINEERING

PATRAN is a registered trademark of PDA Engineering.
VAX is a trademark of Digital Equipment Corporation.

The Restoration of Medieval Stained Glass

The victim of its own composition and of modern air pollution, Europe's most radiant art is now threatened with destruction. The efforts at preservation depend on knowledge of the glass

by Gottfried Frenzel

Light has long served religion as a symbol. It has signified creation ("Let there be light" was the first command of the Creator) as well as salvation (John the Evangelist saw the Heavenly Jerusalem illuminated as if made "of jasper" and its walls "like clear glass"). The earthly reflections of such visions, achieved throughout the Middle Ages by means of light, were the period's most brilliant works of art: the stained-glass windows of Romanesque and Gothic chapels, churches, minsters and cathedrals. For almost a millennium, in the case of the earliest stained-glass windows, the glass escaped major damage. Even the catastrophe of World War II inflicted harm that was within bearable limits. In fact, stained glass all over Europe was removed to safety. Today, however, its total destruction is threatened, not by war but by air pollution. If stained-glass windows are kept in situ in their present state of preservation, their total ruin can be predicted within our generation.

A few examples will illustrate the threat. The stained-glass windows of Cologne Cathedral, in the immediate vicinity of the city's main railway station, have been unusually vulnerable. They were endangered by exterior weathering and air pollution as early as the mid-19th century. Seen from outside the building, the windows now look like sheets of chalky plaster. Continuous etching by air pollutants has corroded the exterior surface of the glass, reducing its thickness year by year and giving the decomposed surface a so-called weathering crust. The process of destruction starts anew as each rain washes the crust away. Meanwhile the colored glass itself breaks into tiny particles. The particles fall out of each panel; thus the window disintegrates.

In England stained-glass windows are exposed to heavy smog. Canterbury Cathedral displays the results. The cathedral includes the Trinity Chapel and its ambulatory, or processional aisle, which incorporates the large chapel called the Corona, constructed between 1174 and 1220. In both chapels some of the stained glass has been attacked. Pits have formed, which have now perforated the panels, leaving them quite porous, so that acid rain can reach the inner surface of the glass and eat into the paintwork there.

France is the classic repository of stained glass. A single cathedral, the one in Chartres, is decorated with more than 2,000 square meters of stained glass from the 12th and 13th centuries, the period when the art reached its peak in France. As recently as 20 years ago one could marvel at the glass, and in particular at the richness achieved in the predominantly blue panes of the Romanesque and early Gothic periods: the "blue miracle" of Chartres. Today the contrast is shocking. The blue has not lost all its intensity; indeed, the chemical composition of the blue glass has made it relatively resistant to weathering. (In German-speaking countries the green glass has proved least susceptible to weathering.) The panes of other colors, however, have corroded and turned a mangy brown, rendering the stained-glass images barely recognizable.

Medieval stained-glass windows are constructions of extreme fragility. Each consists of numerous pieces of colored glass of varying chemical composition held together in a firm but elastic network of cames, or thin lead strips with grooves to hold the glass. The cames follow the lines of the composition, producing a unified, mosaic-like image. In most cases the glass was colored with metal oxides, which were put into the molten mass at the time the glass was manufactured. The exception was flashed glass, in which a thin film of color (usually red) was fused onto a clear base glass. Detailed effects could then be achieved by grinding away parts of the colored film. In the early 14th century silver stain was introduced. It consisted of silver nitrate bound in clay or ocher. The stain was painted onto the outer surface of the glass; then the glass was fired, that is, given its final heating. The result was a color ranging from light lemon yellow to deep orange. Finally, in the middle of the 15th century, sanguine was introduced. Sanguine is a pigment containing iron sulfite. Applied to the outer surface of the glass, it takes on a rose to red-brown tint on firing.

The decoration of the colored glass was achieved primarily by means of paintwork known as grisaille, applied to the surface of the glass. As a rule the painting was executed in black or a dark neutral tone. The paint itself was a mixture of copper oxide or iron oxide (which lent the mixture a black, brown or gray-green color), pulverized glass (which allowed the paint to fuse with the surface of the glass when the pane was fired) and a binding agent such as a mixture of wine and gum arabic, from the acacia tree. The paint was applied as opaque lines or as translucent mattes or washes. Dark shading could be reinforced by painting the outer as well as the inner surface of the glass. The washes could then be lightened with delicate dabs of a brush or etched with a needle or a quill. Anyone admiring medieval stained glass at close range for the first time cannot help but be struck by the precision of detail and the subtlety of the means employed in works of art intended to

DAMAGE TO STAINED GLASS from the Middle Ages takes different forms depending on the composition of the glass and the conditions to which it has been exposed. The panel on the left, depicting Adam, is from the great west window of Canterbury Cathedral, for which it was made in about A.D. 1180. Smog and the permanent high humidity have pitted the glass, which in places is perforated like a sieve. The damage is plain on the flesh-tone glass representing Adam's head and body. The panel on the right, depicting three warriors, is from the Church of St. Patrokli in Soest, Germany; it was made before 1166. Over the centuries the surface of the glass has been oxidized. The places with painted linework and halftone were protected for a while. When the paint fell away, however, a negative image remained. The damage is particularly evident on the faces. The green colored panes have resisted damage the best.

LUCKY AND UNLUCKY PANEL were both made by Michael Wohlgemuth, the teacher of Albrecht Dürer. The stained glass at the top, which depicts Lorenz Tucher, the donor, is from the Church of St. Michael in Fürth; it was made in 1485. Exposure to weather did it no harm. In 1815 it was sold and became part of a private collection. Finally, in 1968, the German National Museum in Nuremberg acquired it. The glass is completely free of corrosion; the painted linework is wholly intact. The stained glass at the bottom, which depicts the emperor Heraclius entering Jerusalem, is from the Church of St. Lorenz in Nuremberg; it was made in 1476–77. In the 19th century it was subjected to efforts at restoration that only damaged it further. Many parts of the panel are 19th-century copies of the original pieces of glass. Moreover, the face of the emperor and the face adjoining it have fractured into hundreds of splinters; they are intact because the heads were laminated at the back.

be viewed from a great distance. The final step in the preparation of the pieces of glass was their firing at some 600 degrees Celsius, the temperature at which the surface of the glass would soften and the paint on it would fuse.

From the moment they were in place the panels were in danger. The immediate threats included not only the effects of hail, windstorms and extreme fluctuations of air temperature but also wanton destruction, such as the hurling of stones by vandals. In the Middle Ages religious institutions routinely contracted with glaziers for the maintenance of their glass. The caretaking consisted in cleaning (washing with water, carbonate of soda and a sponge), the repair of cames and the replacement of shattered panels.

Apart from these external threats the glass itself was susceptible to a process of decomposition. Medieval glass was made from local raw materials, usually a mixture of one part sand and two parts ash from beechwood or fern. The mixture had the advantage of being easy to melt. The glass, however, had the disadvantage of being soft, a property that made it susceptible to weathering. The process of decomposition set in as soon as the glass was installed in the form of window panels. Water arriving at the surface of the panel as rain or dew would hydrate the glassy material. In particular, hydrogen ions from the water would take the place of alkali ions in the glass: chiefly potassium and calcium ions. Hydroxyl (OH) ions from the water would then attack the silica (SiO_2) in the glass, turning it from a polymer into a silica gel: an amorphous material consisting of short silica fragments. Eventually, with the alkalis leached out, only silica would remain. The silica layer can be particularly damaging to the appearance of stained glass. The layer can become iridescent, making the panels increasingly opaque.

Since the early 19th century the external threats and the internal ones have been augmented by the hazards associated with industrialization. The chief of these is sulfur dioxide, which is given off into the atmosphere not only by manufacturing processes but also by the burning of coal and oil. Sulfur dioxide combines with humidity to create sulfuric acid, which increases the availability of hydrogen ions. In addition it makes sulfate groups available to react with alkalis such as calcium. The resulting light, chalky layers of sulfates form a weathering crust that can be several millimeters thick. The crust is highly hygroscopic: it absorbs water like a sponge, thus accelerating the destruction of the glass.

DETAILS OF THE DETERIORATION of medieval stained glass appear in three photographs of parts of a representation of Mary, made about 1385, from the Church of St. Martha in Nuremberg. The photographs show, at roughly the actual size, the blue glass of Mary's mantle (*left*), the face of Mary (*middle*) and an ornamental pattern painted on red glass (*right*). In each, corrosion has lacerated the glass, giving it the appearance of a crackled glaze. The corrosion flakes off layers of glass, reducing the thickness of the panes.

Chemical analyses prove that the destruction attributable to sulfur dioxide goes back no more than 10 or 20 years. The rapidity of the destruction can be documented by comparing panels from the same window when some are still in place and others have been transferred to a museum.

The durability of a particular piece of medieval glass depends on a combination of circumstances: the chemical composition of the glass, the metal oxide coloring agent employed, the temperature at which the glass was made and the length of time it was molten during the manufacture. The temperature is crucial. Studies have now established that the melting point of medieval glass ranges from 300 to 900 degrees C. Glass from the Romanesque period (roughly from 500 to 1150) has a fairly high melting point; in glass from the Gothic period (roughly from 1150 to 1550) the melting point is lower. (Glass made still later, in the Renaissance, has the highest melting point.) In general, glass with a high melting point is the most resistant to weathering. The temperature required to manufacture it tends to ensure that its composition is homogeneous and gives it a well-formed, fire-polished surface, which in turn denies footholds to corrosion.

This is not to say that a high melting point is an unmitigated advantage. The glass particles in the paint employed for linework and halftones on the surface of stained-glass panes melt at about 600 degrees C. If the melting point of the pane is substantially greater, the fusion between the two is poor. The result in the course of time is particularly obvious in Renaissance glass. The glass itself is barely corroded, whereas the paint on it is quite poorly preserved. Undoubtedly medieval glass painters were aware of the imperfect fusion between their paint and high-melting-point glass. Sometimes, however, they seem to have wished the problem away. In *Diversarium artium schedula,* a manuscript written between 1110 and 1140, the German monk Theophilus instructed glassmakers removing glass from the firing kiln to "see if you can scrape off the pigment with your fingernail; if not, it is sufficient, but if you can, put it back again."

Certain aspects of the medieval firing process are of some importance for conservators attempting to save the glass today. Pieces of glass ready for firing were often stacked in layers in the kiln; thus the paint evaporating slightly from the surface of a piece in the course of high-temperature firing produced a faint metallic imprint on the piece stacked above or below. The imprint was invisible at the time. Nevertheless, it reduced the susceptibility to corrosion. The Coronation of the Virgin, depicted in the Martyrs Window at Freiburg Cathedral, furnishes an instance. Christ, who is seated next to Mary, wears a crown, which appears faintly, in an imprinted mirror image, on the back of the Virgin's head. The imprinted area is intact, uncorroded glass; the rest is covered by a powdery weathering crust.

Although the danger confronting stained glass today is acute, the problem of its deterioration has a long history. At the time of the Reformation, which rejected sacred ornamentation, the art of stained glass came to a standstill. Some stained-glass panels simply fell into ruin. (As early as 1639 one observer, Adam Gering, was complaining about Freiburg Cathedral: "How terribly the precious windows are already damaged!") Some panels lost so much of their translucency that sections of clear glass were fitted into their midst in order to lighten the interior of the building. The Baroque and the Enlightenment, with their lack of interest in medieval relics, only heightened the neglect. ("Because these painted windows turn everything very dark, heavy and dull, they are disposed of everywhere," a priest of Freiburg Cathedral noted in 1787.)

In the early 19th century interest in stained glass revived. Unfortunately a misguided ambition to surpass the old masters triggered a second wave of destruction. Throughout Europe new generations of glass painters occupied themselves with what they considered "restoration." Damaged panes were replaced by new ones. Damaged grisaille was repainted and refired. In many cases the original panels disappeared, doubtless hoarded by collectors. At the end of the century practices changed. Enthusiasm waned and money became scarce. Hence damaged originals were no longer replaced by copies. Instead original panels were cut up and the pieces were inserted in damaged panels as stopgaps.

The early 20th century brought a number of experimental treatments of stained-glass panels. In the first decade of the century, for example, two panels from the Church of St. Sebaldus in Nuremberg were thinly coated with a low-melting-point overglazing (in particular an enamel) and refired, in order to reattach the grisaille. The damage inflicted by this process was severe. Nevertheless, the treatment was applied to more than 200 stained-glass windows until 1939.

How, then, can medieval stained glass be restored and preserved? The example of the Church of St. Lorenz in Nuremberg demonstrates that in each individual case restoration pos-

es particular problems. There the corrosion caused by air pollution was already under way in late medieval times: adjacent to the church the burghers of Nuremberg treated hops with sulfur. The damage to the glass was apparently not minor. At the end of the 15th century the Council of Nuremberg appointed the famous workshop of Veit Hirsvogel the Elder as the official glazier of the city. Four centuries later the windows of the church were subjected to "restoration." Between 1829 and 1840 the glass painter Johann Jakob Kellner and his four sons removed damaged panes from five choir windows and replaced them with copies or with entirely new creations. The original glass was reduced by 40 percent. The whereabouts of the originals are largely unknown today. Some pieces emerged on the art market and others in the German National Museum in Nuremberg. Examination of the few available specimens shows that they were not damaged sufficiently to have warranted their removal.

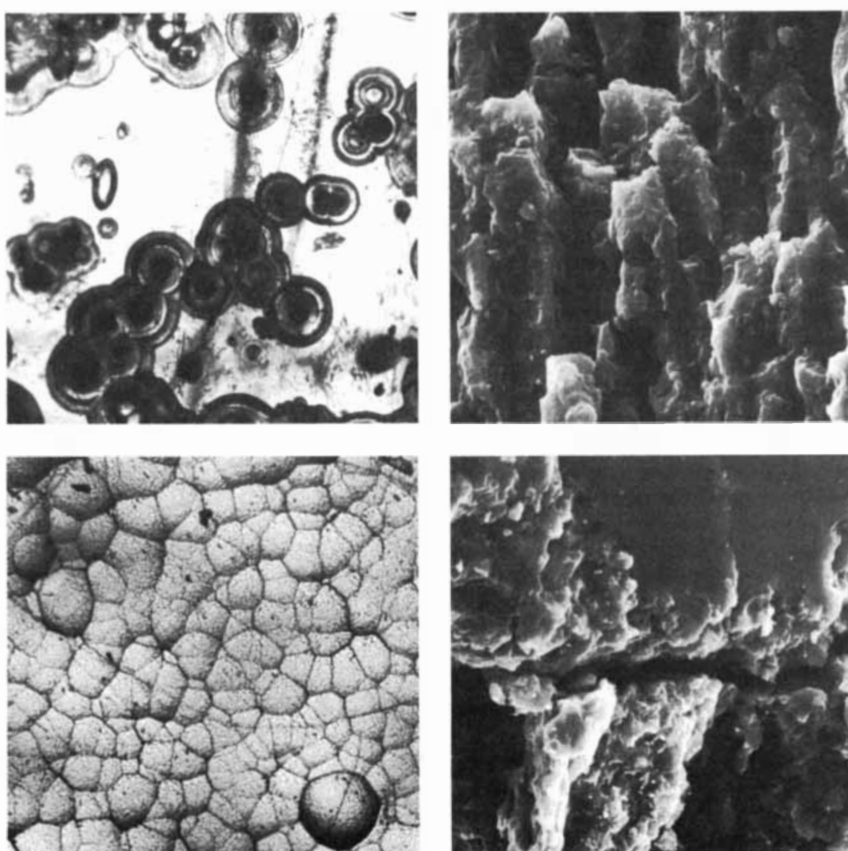
Twentieth-century pollution intensified the deterioration of the St. Lorenz stained glass. Joseph Schmitz, a Nuremberg architect and the son of a glass painter, began to examine methods of conservation. In 1917, after many years of experimentation, test panels were chosen from the Church of St. Sebaldus in Nuremberg. The panels were disassembled; then the pieces were coated with a vitreous dust, which, on refiring, produced an overglazing. Two decades later the Bavarian State Bureau for Conservation decided to employ the method on a larger scale at the Church of St. Lorenz. The outer surface of many types of glass in the St. Lorenz panels was heavily corroded; a thick weathering crust made the panels virtually opaque. The grisaille rested loosely on the inner surface of the glass and fell off in flakes.

The conservators removed the crust by abrasion and applied a low-melting-point overglazing enamel. The inner surface also was cleaned and overglazed, after the loosened grisaille had been carefully pressed down with blotting paper. Efforts began with the Konhofer Window, a work dating to 1477 and created in the workshop of Michael Wohlgemuth, the teacher of Albrecht Dürer. It soon became apparent that the low temperature chosen for the refiring failed to produce a satisfactory refusion of paint to glass. The temperature was increased, whereupon the grisaille fused. In addition, however, a greenish discoloration developed and the paint blurred. At the same time the pieces of glass having a high content of iron and manganese, among them the flesh-tone pieces, turned dark brown. The treatment continued nonetheless, until finally stopped by World War II.

The renewal of conservation efforts in 1968 in my studio at the Institute for Stained Glass Research and Restoration in Nuremberg revealed the true extent of the damage. The refiring and subsequent cooling of the St. Lorenz glass had subjected it to thermal strain, so that it was cracked and in places broken. For its part the overglaze had bubbled, and it was corroding faster than the original glass. The overglaze could be removed with a fiberglass brush, but the exposed paint was unprotected once again.

During the 1950's the church itself was reconstructed. In the course of the effort Richard Jakobi, the former director of the Doerner Institute, undertook to safeguard the church's stained glass. In particular he attempted to sandwich each piece of glass between layers of glass, with a plastic foil separating the original glass from its inner and outer covers. The process had first been tried, with disappointing results, on a panel from Naumburg Cathedral in 1939. Again the individual panes were removed from their network of cames. Splintered pieces in each pane were glued along their edges and reassembled. (Most of the panes were in splinters, each a few millimeters long.) Then a clay-and-plaster mold was made of each side of the pane. Two cover glasses about .8 millimeter thick were cut to the size of the pane, then placed in the molds and settled at a temperature of between 700 and 800 degrees, so that they now had the contours of the surfaces of the original. Finally the original was sandwiched between its new cover glasses. The glass layers were glued together, at a temperature of about 200 degrees, with plexigum foil, a soft acrylic.

The treatment secured the stained glass against air pollution, humidity and even storms and hail. Yet the approach has four serious disadvantages,



MICROGRAPHS of stained glass show further details of deterioration. Pitting (*upper left*) pocks the surface of glass from Augsburg Cathedral. The enlargement is 25 diameters. The bottom of a pit (*upper right*) shows the advance of the decomposition under the influence of humidity. The enlargement is 1,000 diameters. The decomposition of the surface of a pane from the monastery at Lorsch in Germany (*bottom left*) has exposed a scarred understratum of glass. The enlargement is 10 diameters. Deep fissures (*bottom right*) have formed in glass from the church at Lautenbach in France. The smoothness at the upper right is the original surface, preserved in that location. The enlargement is 600 diameters.

which have made it obsolete. First, the application of a cover glass and a plastic foil to the inner surface of stained glass produces a wet-glass effect: a tendency to reflect light, which lessens the visibility of areas painted in subtle halftones. Second, the making of a mold without first reattaching the painted linework is unfortunate: a substantial part of the loose paint inevitably is lost. The subsequent lamination of the panel also takes a toll on the paint. Third, the heating of the sandwich to a temperature of 200 degrees may ultimately damage the glass. It is possible, for example, that internal stress induced by the heating will lead to disintegration. Finally, the use of untested new plastic products in restoration projects is potentially dangerous. In the Church of St. Lorenz damage in fact has occurred. Under the influence of the sun's ultraviolet radiation the plastic employed to join broken edges has turned dark brown.

In 1982 my colleagues and I undid the lamination. Our primary task, during a quarter century of restoration of the stained glass at the Church of St. Lorenz under my direction, has been to use carefully aimed prophylactic measures to create conditions approximating those of a museum.

The first stage in our work is the cleaning of the glass. The purpose of the cleaning is not the improvement of the translucency of the glass. It is the removal of a dangerous source of corrosion: the weathering crust, which attracts moisture. In addition the cleaning exposes what is left of the original paintwork, so that its state of preservation can be accurately judged. Loose parts of the painted trace lines and halftones are reattached to the glass; otherwise they would be lost. The reattachment is done with a nonyellowing acrylic, a substance that can always be redissolved by future conservators. The next stage is restoration. Here no procedure is guaranteed, reliable and universal; each specimen of stained glass requires individual treatment.

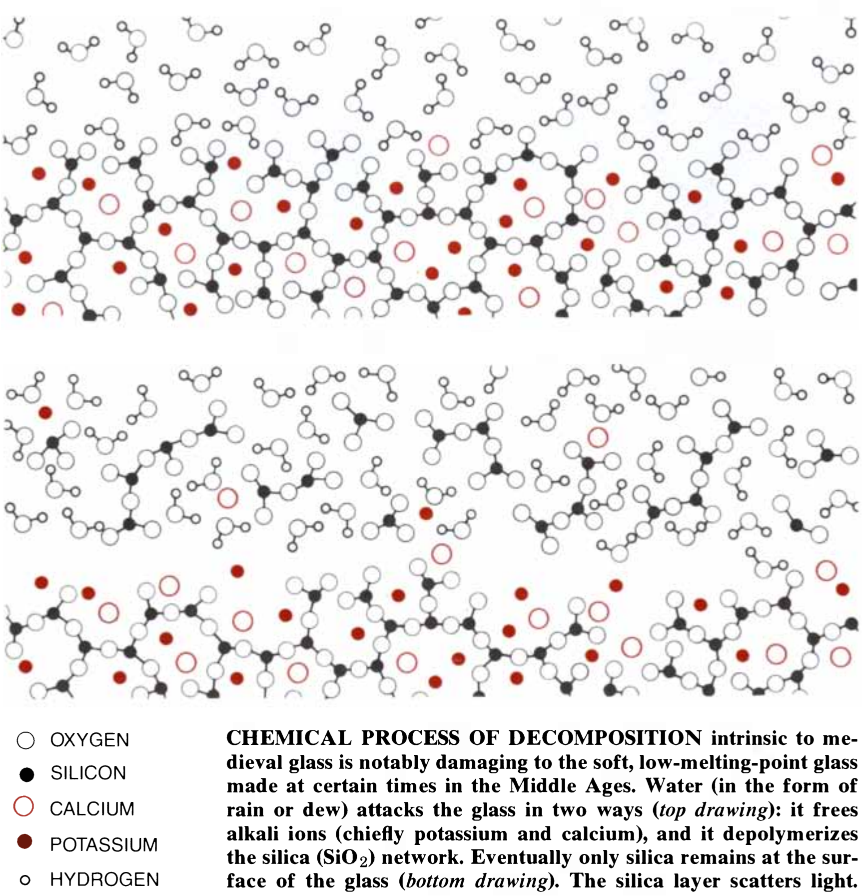
Still, the glass can be protected by two measures. The first is double glazing. A protective pane of glass, not attached to the stained-glass panel, is installed. Second, the temperature and humidity between the stained glass and the protective panel can be controlled, as they would be in a museum. The idea is to superpose an air cushion between the atmosphere and the exterior of the glass. After all, the chief factor that triggers the corrosion of stained glass is humidity. Without humidity even the highest concentration of sulfur dioxide would do no harm.

The oldest double glazing known was installed in England's York Minster in 1861. The intent was simply to improve the building's insulation against the cold. The serendipitous protection of stained-glass windows against exposure to the weather was noted with gratitude later. The protection, however, was unaesthetic. It came in the form of large, greenish sheets of machine-rolled, textured glass; placed between the stained glass and the sky, the new glass not only interfered with the look of the building from outside but also diminished the luminosity of the stained glass viewed from within. Moreover, the sheets were firmly mortared into position, and the tension resulting from their expansion had broken all but one of them after some 45 years.

A second experiment with double glazing came in 1897. Its subject was the Romanesque stained-glass cycle in the small church of Lindenau, now in the German Democratic Republic. Two panels of the cycle have been in the collection of the German National Museum for almost 80 years. Their state of preservation resembles that of the glass still in place in the church. The safeguarding afforded by double glazing evidently corresponds to museum conditions.

Although double glazing protects the exterior of the window, there remains a critical threat to the inner surface of the glass, a threat that arises from the heating of medieval churches. As a rule the churches were not designed to be heated: they have no insulation in their floor, walls, ceilings and windows. The beneficial result is that the relative humidity inside the building has fluctuated little over the centuries. The thick walls have served as a buffer, absorbing moisture or releasing it. Heating, on the other hand, produces a temperature difference between the interior and the exterior of the church, particularly when the heat is turned up quickly in preparation for a service. The humidity in the air then condenses on poorly insulated surfaces, notably the inside of the windows. The humidity traps air pollutants, and so the destruction begins. Grisaille and the glass itself begin to decompose. The painted surface, which for a time had protected the underlying glass, remains discernible in the form of a negative image.

The type of deterioration that produces a negative image is often accompanied by an entirely different process of corrosion, one whose mechanism has recently been clarified by investigations we have carried out in collaboration with the German Museum in Munich. The process affects chiefly



CHEMICAL PROCESS OF DECOMPOSITION intrinsic to medieval glass is notably damaging to the soft, low-melting-point glass made at certain times in the Middle Ages. Water (in the form of rain or dew) attacks the glass in two ways (*top drawing*): it frees alkali ions (chiefly potassium and calcium), and it depolymerizes the silica (SiO_2) network. Eventually only silica remains at the surface of the glass (*bottom drawing*). The silica layer scatters light.

Here are 129 reasons to buy at Elek-Tek, not to mention the fastest delivery anywhere.

Save 30% to 43%
off Manufacturer Suggested Retail prices on
EPSON • Okidata • Star • TOSHIBA • COMREX • TTX

PRINTER ACCESSORIES

1. **FR 192** Epson 80 col. ribbon — $ 3.50
2. **FR 193** Epson 132 col. ribbon — 5.00
3. **FR 153A** OKI & Gemini Ribbon — 2.50
4. **FF 7353** 80 col. printer stand (metal) — 15.00
5. **FF 7354** 132 col. printer stand (metal) — 18.00
6. **CB 5609** IBM PC to Epson cable, 8 ft. — 20.00

star

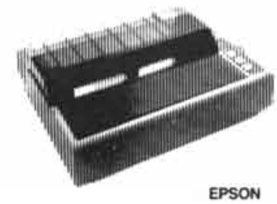
16. SG 10 — CALL
17. SD 10/15 — CALL
18. SR 10/15 — CALL

OKIDATA

19. Oki 84 — CALL
20. Oki 92 — CALL
21. Oki 93 — CALL

EPSON®

22. RX 80/RX 80 FT — CALL
23. RX 100 — $ 350
24. FX 80 + — 345
25. LX 80 — 260
26. FX 100 — 550
27. LQ 1500 par — CALL
28. Comrex II Comriter — 360



EPSON FX 80+

OTHER PRINTERS

48. TTX 1114 — $ 360
49. SCM D200 — 350
50. Toshiba P351 — 1290
51. Toshiba P1340 — 575

PRINTER BUFFERS

7. **DDP-16** 16K Par./Par. (exp. to 64K) — 70

QUADRAM

8. **QRMP8** 8K Parallel/Parallel — 125
9. **QRMPS8** 8K Parallel/Serial — 139
10. **QRMSP8** 8K Serial/Parallel — 139
11. **QRMSS8** 8K Serial/Serial — 139

POWER PROTECTORS

12. **PowerMite PG600-S9** 6 outlet Surge Suppressor — 30
13. **ISOBAR 8** 8 outlet Surge Suppressor — 50

SAFT

Standby Power Systems

Protects your computer *hardware* and *data* when the power goes out

14. **400VA (3.3A)** 475
    - Transfertime—4-6 Mil. Sec.
    - Waveform—rectangular
    - Runtime—up to 24 min
15. **1000VA** 875
    Switches to battery powered back-up in ½ mil. sec.
    - Clean sine wave output
    - AC line synchronized
    - Regulated output voltage
    - Surge Suppression
    - EFI/RFI filter

PRODUCTS FOR IBM-PC

29. **Amdek 310A** Amber Monitor — $ 160
    Other Amdek Mon. ..... CALL
30. **Sakata Monitors** — CALL
31. **AST Six Pak +** Multifunction Board, 64K — 265
32. **AST Megaplus II** Multifunction Board, 64K — 270
33. **AST I/O Plus II** Multifunction Board, 0K — 125
34. **Quadram Quadboard** Multifunction Board, 0K — 210
35. **Quadram Quadboard** Multifunction Board, 64K — 245
36. **Paradise** Modular Graphics Card — 290
37. **Hercules** Monochrome Card — 329
38. **Hercules Color** Color Graphic Card — 170
39. **Novation 4905911** Modem w/software — 320
40. **Novation 4905921** Internal Modem w/software — 320
41. **Hayes 1200B** Internal modem w/software — 389
42. **Hayes 1200** External modem — 425
43. **Hayes 2400** External modem — 620
44. **Anderson-Jacobson 1212-2C01** Internal modem w/crosstalk — 360
45. **Anderson-Jacobson 1212-2C02** Modem w/crosstalk, 2nd ser. port — 400
46. **TEAC FD55B** ½ ht. DSDD Disk Drive — 115
47. **Switchcraft Keyboard** 13 prog. keys, heavy duty — 175

PRODUCTS FOR APPLE

52. **Apple Dumpling GX** Graphics Interface — $ 70
53. **Grappler PLUS** Graphics Interface — 80
54. **Buffered Grappler** 16K Buffer (Expand to 64K) — 150
55. **PKASO** Enhanced Graph. Interface — 120
56. **Kensington System Saver** Fan and Surge Suppr. — 69
57. **Super Cooling Fan** Fan and Surge Suppr. — 29
58. **Rana ELITE 1** SSSD Disk Drive — 175
59. **Rana ELITE 2** DSSD Disk Drive — 335
60. **Rana ELITE 3** DSDD Disk Drive — 420
61. **Rana Controller** 80
62. **ERAM 80** 80 col; 64K RAM for IIe — 115
63. **Novation Apple Cat II** 300 Baud Modem — 200
64. **Novation 212** 300/1200 Baud Modem — 390
65. **Nov 49059112** Mod. & Soft. for MacIntosh — 320
66. **Hayes Micromodem IIe** 300 Baud Modem — 220
67. **Hayes Smartmodem IIc** 300 Baud Modem for IIc — 240
68. **Wico 501030** Analog Joystick — 36
69. **Amdek Monitors** .... CALL
70. **Sakata Monitors** .... CALL

HEWLETT-PACKARD

71. **HP-71B Computer** ..... $ 399
72. **HP-82400A** Series 70 card reader — 125
73. **HP-82401A** HP-IL interface — 95
74. **HP-82402A** 4K memory module — 60
75. **HP-82700A** 8K memory module — 150
76. **HP-82441A** Assembly/Forth — 120
77. **HP-82480A** HP-71 math pac — 75
78. **HP-82482A** HP-71 finance pac — 60
79. **HP-82483A** Surveying Pac — 120
80. **HP-82484A** Curve Fitting — 75
81. **HP-82485A** Text Editor — 60
82. **HP-82488A** Data Com. Pac — 120

SPECIAL PURCHASE

HP-41C $105

While Limited Quantities Last

84. **HP-41CV** $168
85. HP-41CX — 245
86. Optical Wand — 95
87. Card Reader — 145
88. Printer — 283
89. Quad R.A.M. (for HP41C) — 60
90. Ext. Memory Module — 60
91. Ext. Function Module — 60
92. Time Module — 60
93. HP-IL Loop Module — 95
94. Digital Cassette Drive — 335
95. Printer/Plotter (HP-IL) — 335
96. **HP-2225B** ThinkJet Printer — 375
97. **HP-9114A** Disk Drive — 600

SLIMLINE Shirtpocket Styled Power Packed Programmable LCD PROBLEM SOLVERS

98. **HP-11C** Scientific — 58
99. **HP-15C** Scientific — 90
100. **HP-12C** Financial — 90
101. **HP-16C** Programmer — 90

PERSONAL COMPUTATION DEALER

HUGE SAVINGS ON DISKETTES

maxell

102. **MF1** 3½″ SSDD — $ 27
103. **MF2** 3½″DSDD — 37
104. **MD1** 5¼″SSDD — 16.50
105. **MD2** 5¼″DSDD — 21.50
106. **MD2DD** 5¼″DSDD/96 TPI — 35
107. **MD2HD** 5¼″DSDD/IBM AT — 45
108. **FD1** 8″ SSDD — 32
109. **FD2** 8″ DSDD — 37

3M

110. **3M3SS** 3½″ SSDD — 34
111. **3M3DS** 3½″ DSDD — CALL
112. **3M5SSDD** 5¼″ SSDD — 16.50
113. **3M5DSDD** 5¼″ DSDD — 21.50
114. **3M5DSDD96** 5¼″ DSDD/96 TPI — 31
115. **3M8SSDD** 8″ SSDD — 21
116. **3M8DSDD** 8″ DSDD — 29

Dysan

117. **Dys 12** 5¼″ SSDD — 19
118. **Dys 22** 5¼″ DSDD — 25
119. **Dys 33** 5¼″ DSDD/96 TPI — 45
120. **Dys 24** 5¼″ DSDD/IBMAT — 50
121. **Dys 8128** 8″ SSDD — 27
122. **Dys 8228** 8″ DSDD — 32

wabash

123. **WAB 1111** 5¼″ SSSD — 11
124. **WAB 1212** 5¼″ SSDD — 12
125. **WAB 2222** 5¼″ DSDD — 13

DISKETTE ACCESSORIES

126. **RSI 600** 5¼″ Head Cleaning Kit (30 Cleanings) — $ 6
127. **F 320** 3½″ Head Cleaning Kit (20 cleanings) — 22
128. **MM 5** Media Mate Holds 50 5¼″ disks — 10
129. **MM 3** Media Mate-Holds 30 3½″ — 9

DISKETTES *Call for Quantity pricing for 10 boxes or more.*

CALL TOLL FREE 800-621-1269 EXCEPT Illinois, Alaska, Hawaii

ELEK-TEK, inc. 6557 N. Lincoln Ave., Chicago, IL 60645
(312) 631 7800 (312) 677 7660

Corp Accts Invited Min Ord $15.00 Visa or MasterCard by Mail or Phone Mail Cashier's Check, Mon Ord., Personal Check (2 wks to clear) Add $4.00 1st item (AK, HI, P.R., Canada add $10.00 first item) $1.00 ea add'l shpg. & handl. Shipments to IL address add 7% tax Prices subj. to change. WRITE for free catalog RETURN POLICY: Defectives Only: Most products replaced within 30 days of purchase with identical merchandise only Computer and large peripherals replaced only when defective on arrival (within 3 work days of delivery). Other problems covered by mfr. warranty ALL ELEK-TEK MERCHANDISE IS BRAND NEW, FIRST QUALITY AND COMPLETE Delivery subject to availability

25 years ago, the laser was nothing more than an exotic flashlight.

The first working laser—a glass tube coiled around a ruby red rod—was about 3 inches long.

On May 15, 1960, the most intensely pure beam of light the world had ever seen flashed out of a solid ruby crystal.

The laser had been born at Hughes Research Laboratories in Malibu, California.

The Hughes scientists who built that first working laser—and the management behind them—could not have guessed its eventual uses. But now look at what the offspring of their "laboratory curiosity" are doing:

Repairing detached retinas
Reading product codes on groceries
Recording and playing video discs
Cutting fabric for clothes
Drilling holes in metal
Inspecting bottles
Transmitting telephone calls
Surveying roads
Sounding the atmosphere
Dazzling concertgoers
Annealing microcircuits
Welding metal
Characterizing surface roughness
Measuring air pollution
Fingerprinting diamonds
Defining the meter
Slowing atom beams
Cutting cigarettes
Printing computer data
Measuring the earth-moon distance
Cutting airplane parts
Transmitting news wirephotos
Aligning precision machinery
Making three-dimensional pictures
Controlling tunnel machinery
Configuring massive telescopes
Designating military targets
Diagnosing flames
Leveling land
Controlling inventory
Analyzing compounds
Finding impurities
Aligning sawmill cuts
Monitoring polar icecaps
Measuring airplane velocity
Cleaning teeth
Looking for gravitational radiation
Installing acoustical ceilings
Identifying molecules
Aiding robotic vision
Inspecting tires
Positioning medical patients
Probing genetic material
Inspecting textiles
Removing birthmarks
Illuminating fluid flow
Communicating underwater
Enlarging color photographs
Teaching optics
Identifying viruses
Sensing rotation
Peforming microsurgery
Erasing ink
Powering optical computers
Trimming resistors
Altering interconnects
Analyzing materials
Cleaning diamonds
Analyzing auto exhaust
Orienting crystals
Aligning jigs
Ranging targets
Watching continents drift
Sizing dust particles
Cleaning art relics
Tracing air currents
Measuring molecular density
Imploding microfusion pellets
Sensing cloud altitude
Monitoring earthquakes
Gauging fine wines
Testing optical components
Analyzing thin film composition
Drilling holes in diamond dies
Testing relativity
Separating isotopes
Sensing liquid level
Sensing magnetic fields
Programming read-only memories
Counting blood cells
Guiding missiles
Gauging film thickness
Monitoring crystal growth
Aligning large optics
Shaping jewel bearings
Measuring the speed of light
Securing perimeters
Positioning x-y stages
Computing in parallel
Pumping hard-to-pump lasers
Astonishing moviegoers
Creating highly excited atoms
Amplifying images
Cauterizing blood vessels
Diagnosing fusion plasmas
Enhancing chemical reactions
Engraving identification marks
Hardening surfaces
Perforating computer paper
Producing advertisements

Imagine what the next 25 years will bring.

HUGHES
AIRCRAFT COMPANY

glass that has a large content of iron and manganese (glass, as it happens, from the Romanesque and the early Gothic period). First the uppermost, fire-polished layer of the glass is furrowed by crizzling: a great number of microscopic fissures, which allow oxygen, moisture and acids to penetrate. The result (in contrast to the more typical corrosion, which produces a weathering crust) is the stripping of electrons from chemical elements such as iron and manganese. The resulting chemical products tend not to dissolve. Also, they are dark brown. They actually cause stained glass to turn black and opaque. To a degree the process can be reversed (in places where there is no grisaille on the glass) by the use of reducing agents to undo the oxidation.

The ideal solution, impractical in most churches, would be automatic, year-round climate control combined with air purification. A less extreme solution is the provision of a microclimate around the church's stained glass, independent of the conditions elsewhere in the building. Double glazing helps. The protective glass, firmly sealed in the position of the original stained-glass window, absorbs fluctuations in external temperature and provides a cooling surface. (At low exterior temperatures it collects moisture on the inner side; at low interior temperatures the moisture collects on the outer side.) The stained glass, hung in the church next to the protective glass, is exposed to the air in the building. It remains dry on both sides.

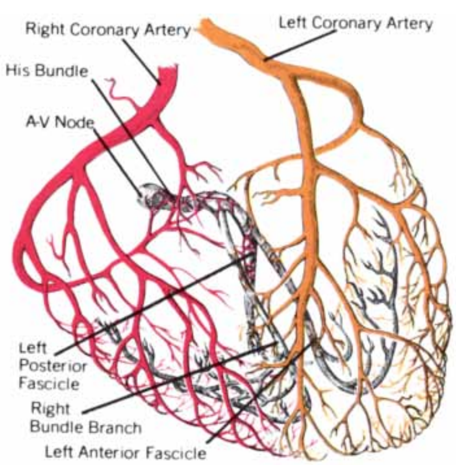
The body of knowledge accumulated so far will be an aid to future efforts. We ourselves, in collaboration with the German Museum, have published a damage atlas for the Federal Republic of Germany, encompassing stained glass from before about 1520. As part of the project we have chosen 30 locations, and in each one we have recorded, over half a year, the local temperature, humidity and air pollution impinging on the medieval stained glass. The intent is to help save a heritage whose loss could otherwise be predicted within our generation.

ST. AMBROSE was depicted by Wohlgemuth in 1477 for the Church of St. Lorenz in Nuremberg. Since then the history of the glass has been particularly unfortunate. In 1836 the glass painter and restorer Johann Jakob Kellner replaced the banderole, or decorative scroll, that bears the name of the saint. A century later it became apparent that all the painted linework on the glass had flaked off; it had been undermined by corrosion. An effort to reattach the paint by refiring the glass served only to melt and blur the hatch lines. In addition white glass took on a yellow-brown hue and glass used for the robe of the saint, which had been blue, turned black. The fine painted detail is now visible only as a negative image.

Physician, did you miss any of these significant developments in medical science?

- **In March 1985, the FDA licensed a test to screen blood products for antibodies to HTLV-III, the AIDS agent; this test should help protect blood product recipients from transfusion-related AIDS.**
- **Oral acyclovir was approved in early 1985 for treatment of initial episodes of genital herpes and for suppression and treatment of recurrent herpes infections.**
- **Several cases of adverse reactions to Fansidar (pyrimethamine-sulfadoxine), some of which were fatal, have been reported among travelers from the United States who took this agent for prophylaxis against chloroquine-resistant strains of *Plasmodium falciparum;* it should not be given to persons with a history of adverse reactions to sulfonamides.**
- **Labetalol, the first approved drug in a new category of agents that combine alpha and nonselective beta blockade, is useful in managing all forms of hypertension, including hypertensive emergencies.**

IF THESE ITEMS are familiar you must be a prodigiously energetic or prodigiously lucky reader. With 2,000 or more journals published each year, information that significantly affects patient management all too easily slips by. Textbooks are out-of-date before they are published.



Branches of the right and left coronary arteries supply blood to the A-V node and intraventricular conduction system.

SCIENTIFIC AMERICAN *Medicine* is lucidly illustrated with drawings and photographs. Some examples are seen here and on the facing page.

SCIENTIFIC AMERICAN *Medicine* is the busy clinician's answer to this problem.

Because its authors update SCIENTIFIC AMERICAN *Medicine* every month, it is always current. Because the new information appears in a single source, it is there when you need it.

This 2,500-page, innovative union of publishing and electronic technology is the work of leading scholar-practitioners from Harvard and Stanford. The editors are Edward Rubenstein, M.D., F.A.C.P., and Daniel Federman, M.D., F.A.C.P.

Each month as authors update their contributions, revisions are entered on the magnetic tape on which the text and index are stored. The tape drives high-speed phototypesetting equipment so that subscribers receive about eight new chapters and a new index every four weeks; a bulletin highlights new developments.

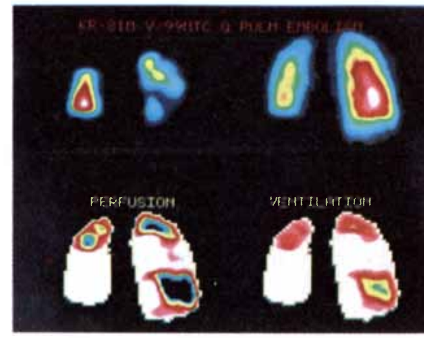
New material replaces old material in the living text, so that the information is there — up-to-date, at your fingertips.

A CME program of eight patient management problems offered over a 12-month period is available at no extra cost. As an organization accredited for continuing medical education, the Stanford University School of Medicine designates this continuing medical education activity as meeting the criteria for 32 credit hours in Category 1 for Educational Materials for the Physician's Recognition Award of the American Medical Association, provided it has been completed according to instructions. This program has been reviewed and is acceptable for 32 prescribed hours by the American Academy of Family Physicians. This program has been approved by the American College of Emergency Physicians for 32 hours of ACEP Category 1 credit.

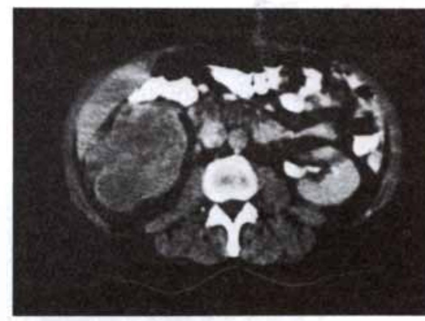
Trial Offer

We invite you to try SCIENTIFIC AMERICAN *Medicine* — for two months at no cost. Send us the coupon and you will receive the two-volume text and two monthly updates. You may also take a CME test for credit. At the end of 60 days, if you decide to continue the subscription, we will bill you for $245 for the full 12 months (renewal is currently $185); otherwise return the two volumes.

Please mail the coupon today and let us take the hassle out of keeping up.



Computerized scintigraphy reveals pulmonary thromboembolism.



Abdominal computed tomogram reveals large renal carcinoma replacing part of right kidney.

THE DISTINGUISHED AUTHORS AND THE FIFTEEN SECTIONS OF SCIENTIFIC AMERICAN ***MEDICINE***

1. Cardiovascular Medicine

Edgar Haber, M.D., Harvard Medical School and Massachusetts General Hospital
E. William Hancock, M.D., F.A.C.P., Stanford University School of Medicine
Roman W. DeSanctis, M.D., Harvard Medical School and Massachusetts General Hospital
Adolph M. Hutter, Jr., M.D., F.A.C.P., Harvard Medical School and Massachusetts General Hospital
Eve Elizabeth Slater, M.D., Harvard Medical School and Massachusetts General Hospital

2. Dermatology

Eugene M. Farber, M.D., Stanford University School of Medicine
Elizabeth A. Abel, M.D., Stanford University School of Medicine

3. Endocrinology

Daniel D. Federman, M.D., F.A.C.P., Harvard Medical School

4. Gastroenterology

Gary M. Gray, M.D., Stanford University School of Medicine

Peter B. Gregory, M.D., Stanford University School of Medicine

John Austin Collins, M.D., Stanford University School of Medicine

Douglas Wilmore, M.D., Harvard Medical School and Brigham and Women's Hospital

5. Hematology

Stanley L. Schrier, M.D., Stanford University School of Medicine

6. Immunology

John David, M.D., Harvard Medical School and Harvard School of Public Health

7. Infectious Disease

Thomas C. Merigan, M.D., Stanford University School of Medicine

Morton N. Swartz, M.D., F.A.C.P., Harvard Medical School and Massachusetts General Hospital

Cyrus C. Hopkins, M.D., Harvard Medical School and Massachusetts General Hospital

Adolf W. Karchmer, M.D., F.A.C.P., Harvard Medical School and Massachusetts General Hospital

Robert H. Rubin, M.D., F.A.C.P., Harvard Medical School and Massachusetts General Hospital

Harvey B. Simon, M.D., F.A.C.P., Harvard Medical School and Massachusetts General Hospital

Peter F. Weller, M.D., Harvard Medical School; Beth Israel Hospital, Boston; and Brigham and Women's Hospital

8. Intensive and Emergency Care

Edward Rubenstein, M.D., F.A.C.P., Stanford University School of Medicine

9. Metabolism

George F. Cahill, Jr., M.D., Harvard Medical School, Howard Hughes Medical Institute; and Brigham and Women's Hospital

10. Nephrology

Roy H. Maffly, M.D., Stanford University School of Medicine and Palo Alto Veterans Administration Medical Center

William Bennett, M.D., Health Sciences University, Portland, Oregon

11. Neurology

Robert W. P. Cutler, M.D., Stanford University School of Medicine

12. Oncology

Saul A. Rosenberg, M.D., F.A.C.P., Stanford University School of Medicine

13. Psychiatry

Ned H. Cassem, M.D., Harvard Medical School and Massachusetts General Hospital

14. Respiratory Medicine

Eugene D. Robin, M.D., F.A.C.P., Stanford University School of Medicine

15. Rheumatology

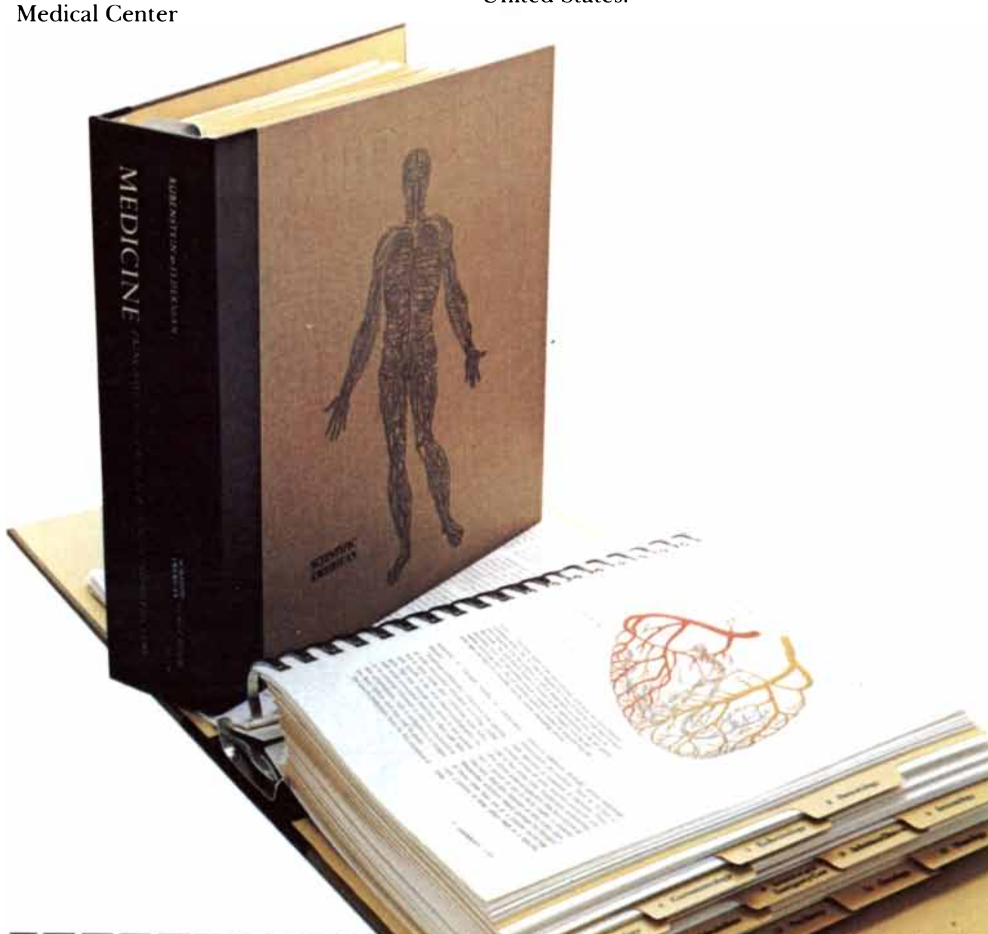
Stephen M. Krane, M.D., Harvard Medical School and Massachusetts General Hospital

Dwight R. Robinson, M.D., Harvard Medical School and Massachusetts General Hospital

Andrei Calin, M.D., M.A., M.R.C.P., Stanford University School of Medicine and Palo Alto Veterans Administration Medical Center

Order by Phone

You can order SCIENTIFIC AMERICAN *Medicine* by telephone. **Please call this toll-free number: 1-800-345-8112 (in Pennsylvania, call 1-800-662-2444);** you will be billed after your subscription begins. Toll-free calls are acceptable only for orders placed in the continental United States.



SCIENTIFIC AMERICAN MEDICINE

415 Madison Avenue, New York, N.Y. 10017

Please enroll me as a subscriber to SCIENTIFIC AMERICAN *Medicine*. On receipt of this coupon you will send me the advanced two-volume text described in your announcement and update it regularly by sending me new monthly subsections. I understand that the price of $245 for first year of service is tax deductible, as is current renewal at $185. If I am not entirely satisfied, I may cancel at any time during the first 60 days, returning all materials for a *complete refund*.

☐ Please enter my subscription for SCIENTIFIC AMERICAN *Medicine*
☐ I shall also enroll in the CME program
☐ I enclose a check made out to SCIENTIFIC AMERICAN *Medicine* for $245*
☐ Bill me ☐ VISA ☐ MasterCard

Expiration Date________ Account Number________

*Please add sales tax for California, Illinois, Massachusetts, Michigan and New York

Name________

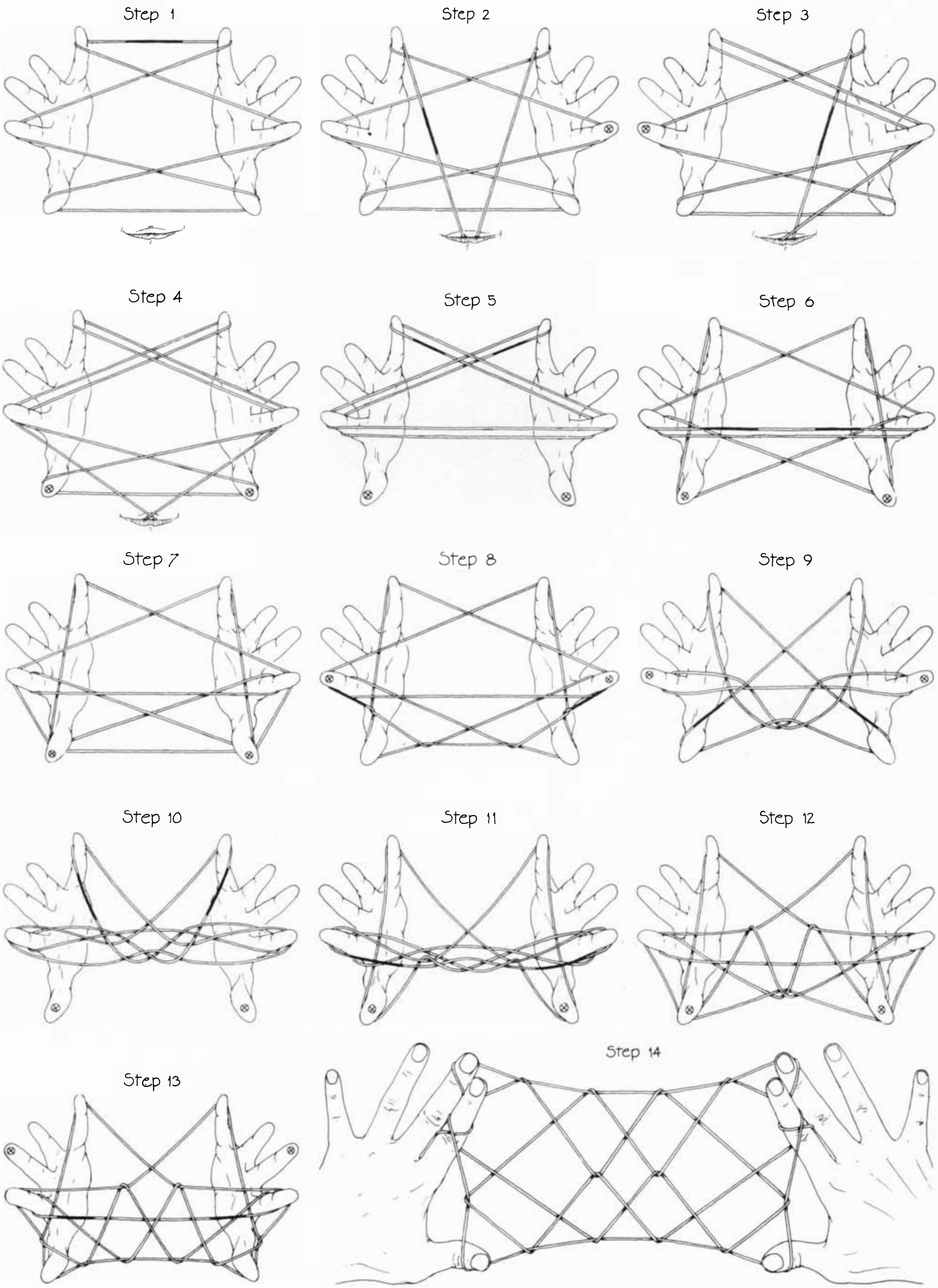
MD Specialty________

Address________

City________ State________ Zip________

Signature________

Please allow six to eight weeks for delivery. All payments must be in U.S. dollars. CME is available outside the U.S. and its possessions for an additional $65. Surface routes used unless airmail delivery is prepaid.



Phylis Morrison's procedure for forming the "ten men" figure

THE AMATEUR SCIENTIST

Cat's cradles and other topologies formed with a two-meter loop of flexible string

by Jearl Walker

The popular game widely known as cat's cradle involves manipulating a loop of string around the fingers in order to create a structure. The exercise invites exploration of the arithmetic rules that seem to emerge from the topology of simple loops.

One begins by wrapping the loop (a flexible string about two meters, or about two yards, long) around a few fingers. The player then uses his or her hands, fingers and even teeth to transform the web of string into the predetermined pattern, often telling a story about the figure. In the standard version a second player takes the structure onto his fingers and forms a new one.

Cat's cradle has been played by children throughout the world for thousands of years. Recently two delightful books inspired me to examine the game and some of its variants. One is *Spiders' Games: A Book for Beginning Weavers* (University of Washington Press, 1979), by Phylis Morrison, who contributes to the "Books" section of this magazine each December. In the book she describes among other things the string figure called "ten men," using it as an example of how one's fingers can serve as a loom [*see illustration on opposite page*]. The second book is *Orderly Tangles: Cloverleafs, Gordian Knots, and Regular Polylinks,* by Alan Holden.

The string in each of these games is a loop in a strict topological sense because the figure collapses to a loop when the fingers are removed. According to the common meaning of knot, however, the string figures are knots. I wondered if the procedures for tying these knots could be catalogued. Surely the figures were devised by people who had mentally constructed their own catalogue of procedures. This month I begin a catalogue by analyzing several string figures to be made by a single player.

Analysis requires terminology: I call the string running between two fingers a section and the string passing around a finger or the entire hand a loop. All string figures have places where one section crosses another; I call them crossings. In the most intriguing string figures some sections cross each other twice. I call them wraps.

To study the wraps in a string figure I mentally follow along the full length of the string from an arbitrary starting point. When I pass through a wrap, I note the sequence of crossings. If the section I am following (the first section) passes over and then under the second section, I describe the relation as an over-under wrap [*see illustration on this page*]. If instead the first section passes under and then over the second section, I describe the relation as an under-over wrap.

I began my study with several questions. How is a wrap created? Are wraps governed by rules? For example, must a string figure have an even number of wraps, as many do? Must it have an equal number of over-under and under-over wraps or can the weaving establish a bias toward one kind? I shall return to these questions after I explain how to begin the game and how to weave the "ten men."

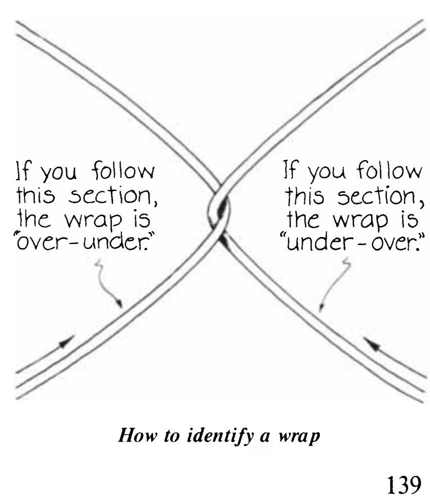
According to W. W. Rouse Ball, whose book is listed in this month's "Bibliography" [*page 144*], most string figures are begun with one of three basic openings: *A, B* and the Navaho. For opening *A* hold your hands with the palms toward each other, as is shown in the top illustration on the next page (which follows the practice of shading the string sections that are to be rearranged and marking with an *X* the fingers that are to be moved). Run the string around the back of the little finger, across the palm, around the back of the thumb, across to the other hand, around the back of the thumb, across the palm, around the back of the little finger and across to the starting point on the first hand. With your right index finger reach below the string section on your left palm and pick it up, hooking the string onto the back of the right index finger. Separate your hands to tighten the string.

With your left index finger pick up the string section on your right palm, hooking the string onto the back of the left index finger. Again separate your hands. You now have opening *A*. Opening *B* is identical except that you start with the left index finger.

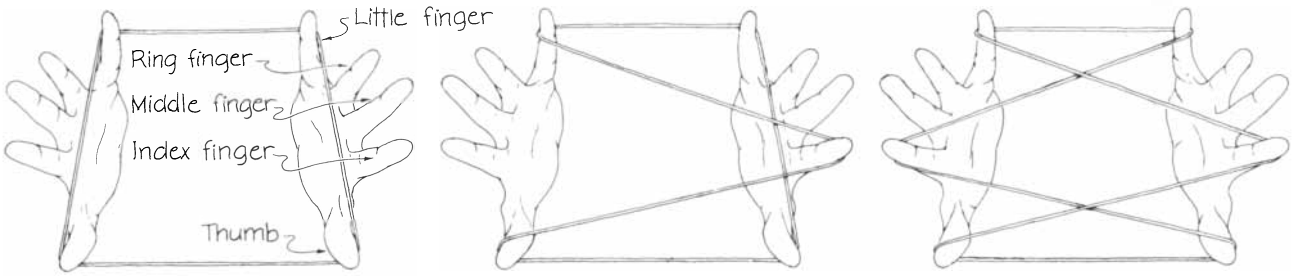
The Navaho opening (named for the Navaho Indians, who were highly skilled at string games) entails a more complicated maneuver. With the string in a figure eight [*see bottom illustration on next page*] insert the index fingers in the top loop of the figure and the thumbs in the bottom loop. (Notice that at the crossing point the top section of string runs from your left to your right.) Turn your palms away from you with the fingers upward. This move catches the top loop of the figure eight on your index fingers and the bottom loop on your thumbs. Rotate your hands to make the palms face.

Now for "ten men." Begin with opening *A* and continue through 14 steps. (1) Seize the far string section with your teeth and pull it over the tops of the other sections. (2) Move your right index finger under the section running from the left little finger to your mouth and pick it up. Keep this captured section near the tip of the index finger to separate it from the string already looped around the finger. (3) Similarly pick up the section running from the right little finger to your mouth. As before, keep this section near the tip of the index finger. (4) Release the string from your mouth and thumbs while you separate your hands to take up the slack.

(5) Bend each thumb away from you so that it passes under four string sections and you can pick up the next section from below (it is the nearer section of the loop around the little finger). Pull the thumbs and the sections they have captured back to you so that the



How to identify a wrap



Starting with opening A

thumbs return to their original orientations. (6) Bend each thumb away from you and pick up from below the nearer section of the string looped around the tip of the index finger. Keep the captured section of string high on each thumb to separate it from the string already looped around the thumb.

(7) With your teeth lift the lower loop at each thumb over the top loop and the thumb tip and drop it in front of the thumb. The procedure is called Navahoing the loops. (8) Release the upper loop on each index finger while you separate your hands far enough to take up the slack. (9) Transfer the loops on each thumb to the top of the adjacent index finger by bending each index finger toward you and then picking up, from below, the loop on the thumb. Keep the captured section at the top of the index finger.

(10) Bend each thumb away from you and under the sections of string on the index finger. Put the back of each thumb under the near section of string on the little finger and pick it up. Return the thumbs to their initial orientation. Keep the loops low on the thumbs. (11) Bend each thumb away from you and pick up from below the nearer string section of the loop around the tip of the index finger. Keep these sections near the top of the thumbs. (12) Navaho the thumb loops.

(13) Bend each middle finger toward you, passing over two string sections. Pick up the next string section from below. Keep these captured sections near the tops of the middle fingers as you return the fingers to their initial orientations. (14) To display the string figure drop the loops around the little fingers, extend your thumbs, keep the index and middle fingers together, separate your hands and turn the palms away from you.

The "ten men" has several crossings and 12 wraps. Although I mastered the construction in about an hour, I spent many hours in gaining an understanding of how wraps are created. In one method that seems to be common to several string figures I add a second crossing to two string sections that already cross. As the situation is depicted in the top illustration on the opposite page a crossing lies in front of a finger around which the string is looped. To create a wrap I bend another finger under the sections and pick up the upper section of the crossing from below. The return of the finger to its initial position pulls the captured string section under the lower section of the crossing, forming a wrap.

I can also form a wrap from the initial crossing if I bend a finger over the upper section and pick up the lower section from below. The return of the finger to its initial position pulls the lower section over the upper one, forming a wrap. I call these moves a hook because a finger hooks one of the string sections in order to pull it around the other section. Sometimes the sections that form the initial crossing loop around one finger. Sometimes the sections are looped around different fingers. In each case the hook adds a second crossing to the sections to form a wrap.

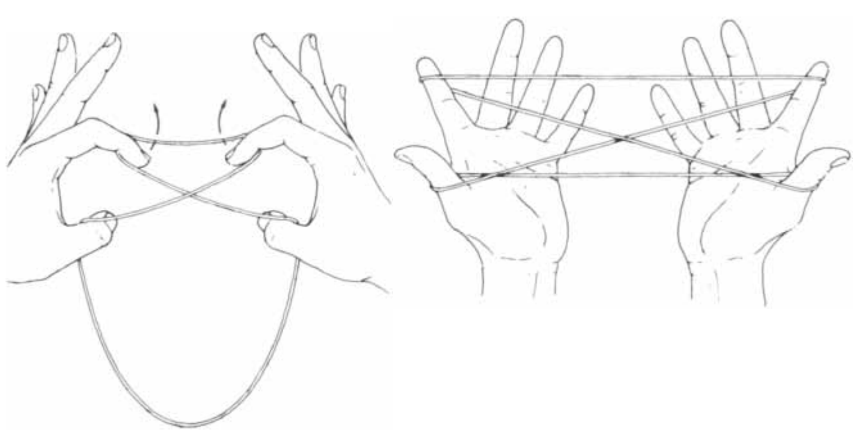
In some string figures a wrap is generated by twisting a finger. For example, run a section across your left palm. With your right index finger pick up the string from below. As you separate your hands rotate your finger so that the captured string twists around itself, creating a wrap. When the finger moves through the first half of the rotation, it crosses the string looped around it. In the second half of the rotation the finger forces the lower section to pass over the upper section, thereby creating the wrap.

I have found another procedure that is essentially a hook [*see top illustration on page 142*]. The string loops around the index finger. The near section of the loop crosses under a string section that loops around the thumb. When the loop around the index finger is expanded to loop also around the thumb, a wrap is created.

A double hook is shown in the middle illustration on page 142. Initially two string sections cross. A third section is pulled under and then over a section on each side of the crossing point, creating two wraps.

The number of wraps that can be produced by Navahoing loops on a finger depends on the shape and orientation of the loops. In the top part of the bottom illustration on page 142 the loop that is low on the finger forms a narrower angle than the high loop. Since the loops have the same orientation (their corner angle opens in the same direction), each side of the bottom loop crosses under the corresponding side of the top loop. When you lift the bottom loop over the fingertip and release it, you create two wraps.

In the middle part of the illustration the top loop forms a narrower angle than the bottom loop. Since the loops



Starting with the Navaho opening

have the same orientation, the sides of the bottom loop do not cross the corresponding sides of the top loop. This time Navahoing the bottom loop over the top loop does not produce a wrap.

In the bottom part of the illustration only the left side of the lower loop crosses under the left side of the top loop. When the bottom loop is Navahoed, a wrap is created on the left side of the top loop. In sum, the number of wraps created by Navahoing loops on a finger is equal to the number of times the sides of the lower loop crossed under the corresponding sides of the top loop before the Navahoing.

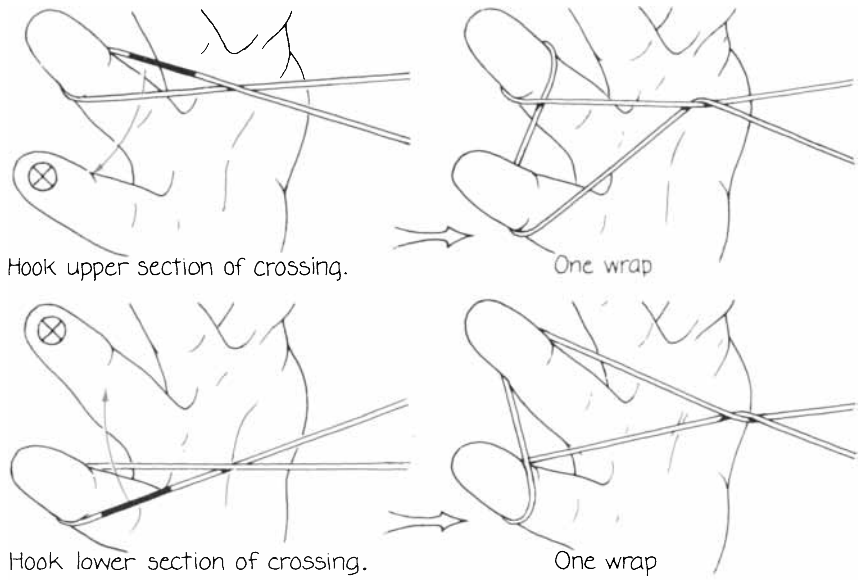
Sometimes making a string figure generates what I call a potential wrap. An example is shown in the top illustration on page 143. If the little finger on the right releases the string looped around it and the hands are separated to take up the slack, the released string wraps around the section that stretches between the index fingers. Many such potential wraps are created during the weaving of the "ten men," but only two of them are converted into real wraps.

Reversing a procedure that creates a wrap eliminates the wrap. In some cases a wrap can also be eliminated by releasing the string from a finger next to the finger you used to Navaho loops. For example, suppose you have just Navahoed the loops on a thumb and one of the loops also passes around the index finger. The wrap between the thumb and the index finger depends on the index-finger loop. If you release that loop, the wrap disappears.

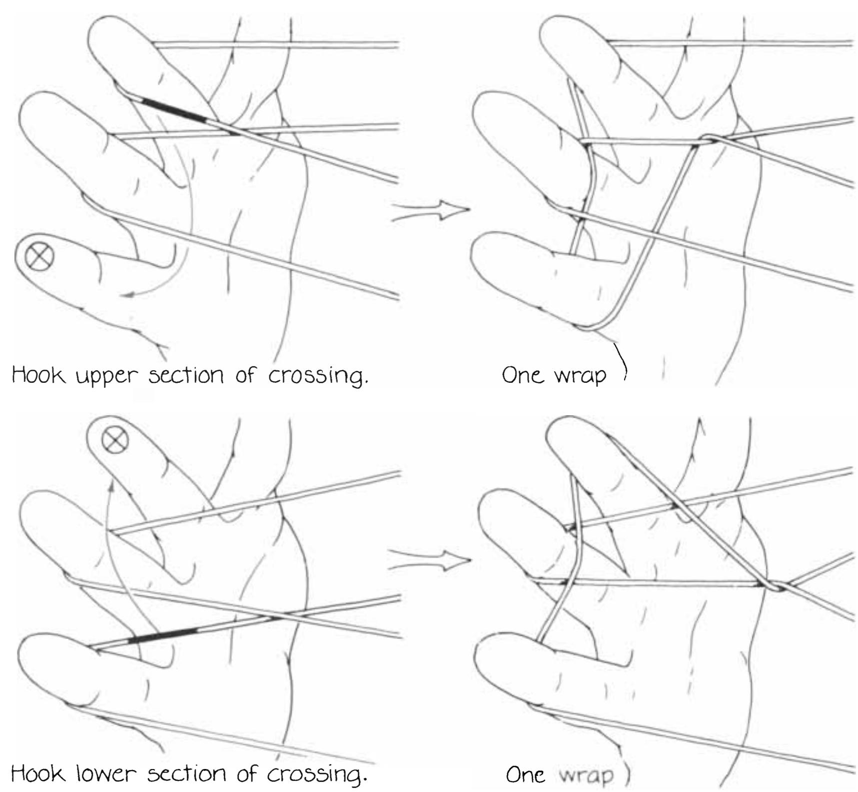
To sum up, there are three basic ways to produce wraps in the weaving of a string figure: the hook, Navahoing loops on a finger and releasing a loop by a finger to convert a potential wrap into a real one. Included in the hook method is the twisting capture by a finger and the expansion of a loop to include two fingers. Wraps can be eliminated by reversing a procedure or by releasing a loop that stabilizes a wrap you created by Navahoing loops on another finger.

Armed with these basic steps, I dissected the formation of "ten men." In the opening position the string has only crossings. Step 1 creates a potential wrap. (If the string around the right little finger were released, a real wrap would appear.) Step 2 creates another potential wrap. Real wraps do not appear until step 7, where the Navahoing of loops at the thumbs creates a total of four real wraps. Step 8 eliminates the wraps that were between the thumbs and the index fingers because they depended on the loops released by the index fingers in this step. Now there are two real wraps.

In step 10 the thumbs hook string

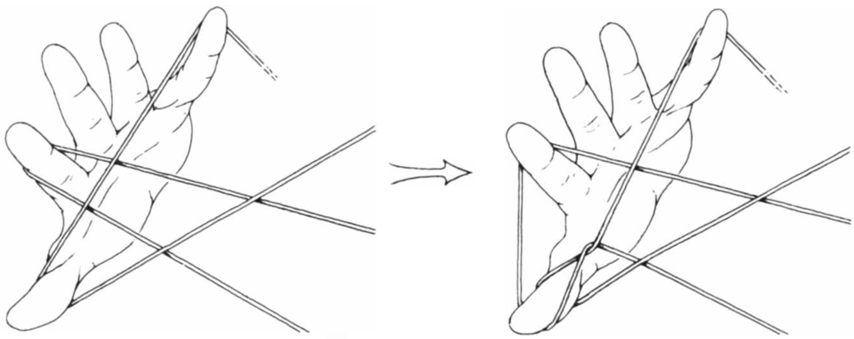


Hooks made on one finger

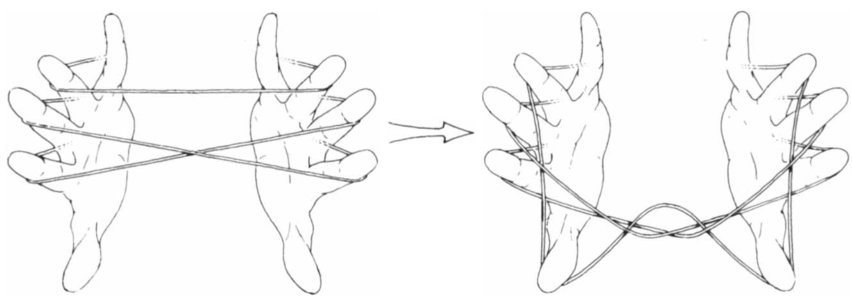


Hooks made on two fingers

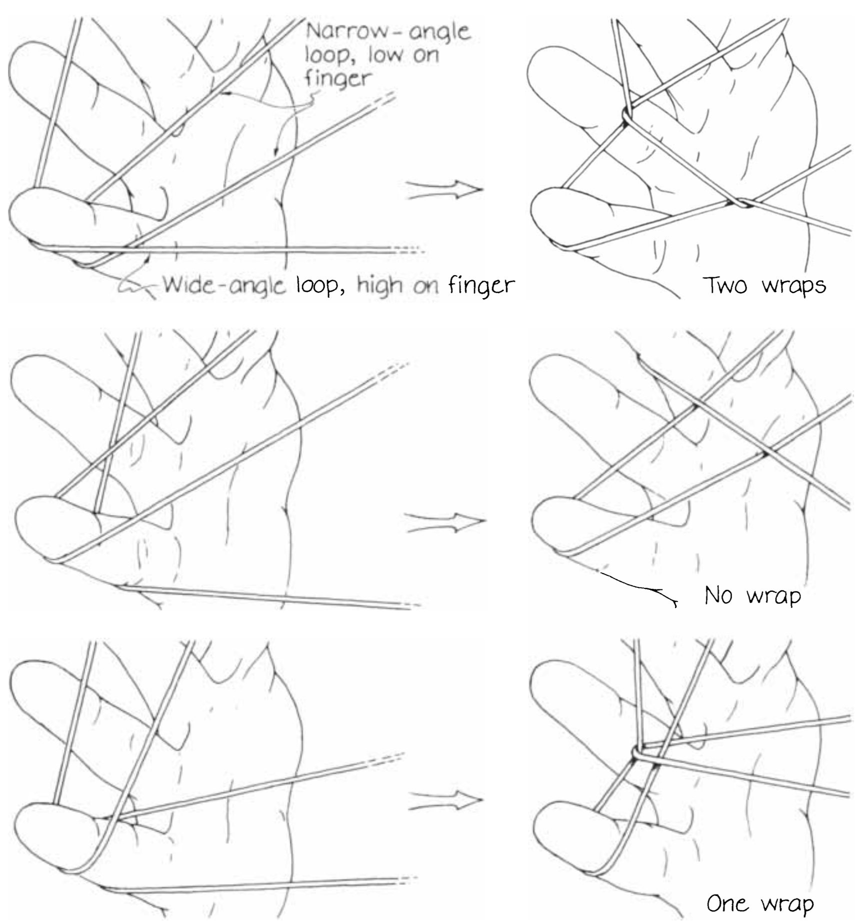
A wrap made by twisting a finger



A loop expanded onto a thumb



A double hook



Various ways to Navaho loops

sections to create two new wraps, again giving a total of four real ones. Step 12 creates another four wraps when the thumb loops are Navahoed. Another two wraps appear in step 13 when the middle fingers hook string sections. (This hook is subtle. The initial crossing is formed by the captured section and a section that emerges from the far side of the index finger at the top.) Finally, the last two wraps are generated when the little fingers release their loops and convert two potential wraps into real wraps, giving a total of 12 real ones.

Can general rules be devised for the wraps of a string figure? I believe they can. For example, if each manipulation of the string is done on both hands, the number of wraps in the figure must be either zero or an even number. The "ten men" is symmetrical in this way because each move on the right hand is also made on the left hand.

Another rule seems to govern the types of wrap in a string figure resulting from symmetrical procedures. From an arbitrary starting point in the "ten men" follow along the string and note the types of wrap you pass. For example, start at the left thumb and move initially toward the left index finger. As you mentally travel along the entire string, returning to the left thumb, you pass through 12 over-under wraps and 12 under-over wraps. I believe a string figure made by symmetrical procedures will always have as many over-under wraps as under-over wraps.

Trace through the "ten men" again, adding arrows along the way to indicate how you enter each wrap. Since during the full trip you pass through each wrap twice (once along each string section in the wrap), each wrap gets a pair of arrows. They point into a wrap either from the same side of the wrap or from opposite sides. In "ten men" 10 of the wraps have a pair of arrows pointing from the same side. Only the central two wraps have arrows pointing from opposite sides. I believe all string figures produced by symmetrical procedures will turn out to have an even or zero number of wraps where the arrows point from the same side and an even or zero number for which the arrows point from opposite sides.

I have also analyzed the production of wraps in the "fishing net." The figure is initiated from opening *A*. Release the loops around the thumbs. Bend each thumb away from you and under three string sections and pick up the farthest string section from below (it is looped around the little finger). Pull the string into place by returning the thumbs to their initial orientation.

You now have one potential wrap formed by the string looped around the left little finger.

Bend each thumb away from you and over one string section. Pick up the next section from below and return the thumbs to their former orientation. This move adds a potential wrap associated with the right little finger. Release the string at each little finger to convert the potential wraps into two real ones. Bend each little finger toward you and over one section. Pick up the next section from below and return each finger to its initial orientation. In this move each of your little fingers makes a hook to add another real wrap. Now there are four real wraps.

Release the string from around the thumbs. Bend each thumb away from you and over two string sections. Pick up the next section from below; it is the near side of the loop passing around the little finger. Return the thumbs to their initial orientation. Using your right hand, pick up the loop that passes around the left index finger. Do not remove the loop from the finger but make it larger so that it passes around both the finger and the thumb. This move is essentially a hook because the section you move to the thumb is forced to wrap over another section from the thumb. Now make the same move on the right hand. You have added two real wraps for a total of six.

Navaho the loops on each thumb. Since the loops do not have the same orientation, this move adds only one wrap on each hand, raising the total to eight wraps. Each thumb now has in front of it a small triangle formed by the string section running from the little finger on that hand. Bend each index finger toward you and pass its tip through the triangle on that hand.

To display the figure rotate your hands to put the palms away from you and the fingers upward. Catch on the appropriate index finger the string of the triangle in which it is inserted. (The loop already on the finger slides off.) Release the loops from the little fingers. Catching a string on each index finger adds one wrap for each maneuver. Releasing the string from the little fingers converts two potential wraps into two real ones. Now you are up to 12 real wraps. Between each thumb and index finger is one wrap and an additional crossing. Ten more wraps are spread through the figure.

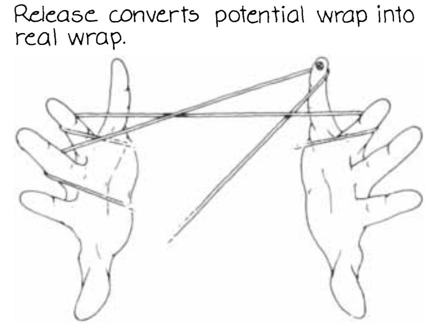
The "fishing net" follows my improvised rules. The total number of wraps is even because of the symmetry of procedures on the two hands. When I follow along the string from an arbitrary starting point, I find that the number of over-under wraps is equal to the number of under-over wraps. When I add arrows to mark how I enter each wrap, I find that the arrows for 10 wraps enter from opposite sides and the arrows for two (the central ones in the figure) enter from the same side. (This result is the reverse of what I found for the "ten men.")

Do the arithmetic relations hold up for other figures? I leave it to you to apply them to the "lightning," a zigzag figure originated by the Navahos. Begin (it is almost needless to say) with the Navaho opening. Bend each thumb away from you so that it passes over two string sections. With the back of the thumb pick up the next section from below. Pull the string by returning each thumb to its original position.

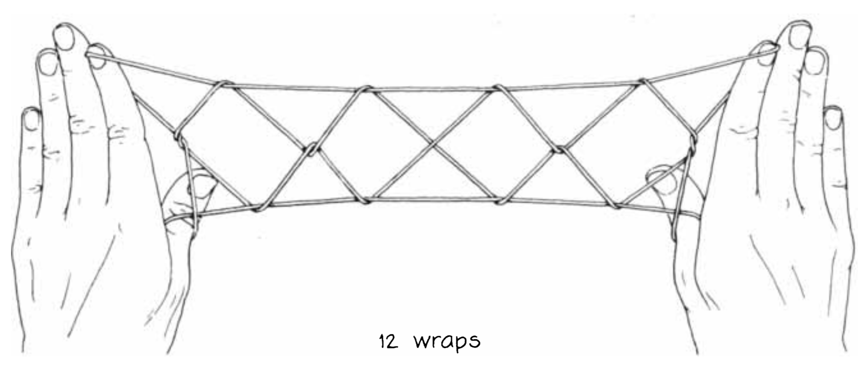
Bend each middle finger toward you, passing it over one string section. With the back of each middle finger pick up the next section from below. Pull the string by returning the middle fingers to their original orientation. Bend each ring finger toward you, passing it over one section. With the back of the finger pick up the next section from below. Pull it by returning the ring fingers to their original orientation.

Now bend each little finger toward you, passing it over one string section, and with the back of the finger pick up the next section from below. Pull on the string by returning the little fingers to their original positions. Bend each thumb away from you and touch its tip on the front string section of the little finger. This move releases the sections wrapped around the thumbs. With a flick of your wrists toss these sections over the other ones. Press your thumb tips down hard on the string sections they touch. To display the figure turn your palms away from you as you spread your fingers and thumbs.

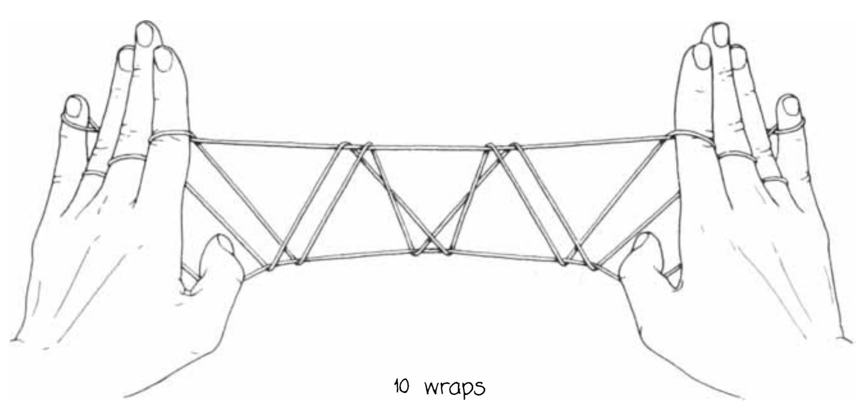
If you find more ways to weave wraps into a figure, I would enjoy hearing from you. Are my rules about the number and nature of the wraps correct? Can you devise general proofs of the rules or find more rules?



A potential wrap



The "fishing net" displayed



The completed "lightning" figure

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

WINNING WAYS FOR YOUR MATHEMATICAL PLAYS. Elwyn R. Berlekamp, John H. Conway and Richard K. Guy. Academic Press, 1982.

THE PLANIVERSE. A. K. Dewdney. Poseidon Press, 1984.

COMPUTER SOFTWARE IN SCIENCE AND MATHEMATICS. Stephen Wolfram in *Scientific American,* Vol. 251, No. 3, pages 188–203; September, 1984.

THE FEDERAL SUPPORT OF MATHEMATICS

FUNDAMENTALS OF NUMERICAL RESERVOIR SIMULATION. Donald W. Peaceman. Elsevier Scientific Publishing Co., 1977.

SHOCK WAVES & REACTION-DIFFUSION EQUATIONS. J. Smoller. Springer-Verlag, 1983.

RENEWING U.S. MATHEMATICS: CRITICAL RESOURCE FOR THE FUTURE. Ad Hoc Committee on Resources for the Mathematical Sciences, the Commission on Physical Sciences, Mathematics, and Resources, National Research Council. National Academy Press, 1984.

MOLECULAR APPROACHES TO MALARIA VACCINES

PROGRESS IN MALARIA VACCINE DEVELOPMENT: CHARACTERIZATION OF PROTECTIVE ANTIGENS. Ruth S. Nussenzweig in *Scandinavian Journal of Infectious Diseases,* Supplementum 36, pages 40–45; 1982.

IDENTIFICATION AND CHEMICAL SYNTHESIS OF A TANDEMLY REPEATED IMMUNOGENIC REGION OF *PLASMODIUM KNOWLESI* CIRCUMSPOROZOITE PROTEIN. G. N. Godson, J. Ellis, P. Svec, D. H. Schlesinger and V. Nussenzweig in *Nature,* Vol. 305, No. 5929, pages 29–33; September 1, 1983.

ISOLATE-SPECIFIC S-ANTIGEN OF *PLASMODIUM FALCIPARUM* CONTAINS A REPEATED SEQUENCE OF ELEVEN AMINO ACIDS. Ross L. Coppel, Alan F. Cowman, Klaus R. Lingelbach, Graham V. Brown, Robert B. Saint, David J. Kemp and Robin F. Anders in *Nature,* Vol. 306, No. 5945, pages 751–756; December 22, 1983.

STRUCTURE OF THE GENE ENCODING THE IMMUNODOMINANT SURFACE ANTIGEN ON THE SPOROZOITE OF THE HUMAN MALARIA PARASITE *PLASMODIUM FALCIPARUM.* John B. Dame, Jackie L. Williams, Thomas F. McCutchan, James L. Weber, Robert A. Wirtz, Wayne T. Hockmeyer, W. Lee Maloy, J. David Haynes, Imogene Schneider, Donald Roberts, Greg S. Sanders, E. Premkumar Reddy, Carter L. Diggs and Louis H. Miller in *Science,* Vol. 225, No. 4662, pages 593–599; August 10, 1984.

HOW A SUPERNOVA EXPLODES

SYNTHESIS OF THE ELEMENTS IN STARS. E. Margaret Burbidge, William A. Fowler and F. Hoyle in *Reviews of Modern Physics,* Vol. 29, No. 4, pages 547–650; October, 1957.

SUPERNOVAE: A SURVEY OF CURRENT RESEARCH. M. J. Rees and R. J. Stoneham. D. Reidel Publishing Co., 1982.

SUPERNOVA THEORY. G. E. Brown, H. A. Bethe and Gordon Baym in *Nuclear Physics,* Vol. A375, No. 3, pages 481–532; February 15, 1982.

SUPERNOVAE, PART I: THE EVENTS. Virginia Trimble in *Reviews of Modern Physics,* Vol. 54, No. 4, pages 1183–1224; October, 1982.

SUPERNOVAE, PART II: THE AFTERMATH. Virginia Trimble in *Reviews of Modern Physics,* Vol. 55, No. 2, pages 511–563; April, 1983.

CHEESE

FUNDAMENTALS OF DAIRY CHEMISTRY. Byron H. Webb, Arnold H. Johnson and John A. Alford. Avi Publishing Co., 1974.

CHEESE AND FERMENTED MILK FOODS. Frank Kosikowski. F. V. Kosikowski and Associates, P.O. Box 139, Brooktondale, N.Y. 14817, 1977.

CHEESE VARIETIES AND DESCRIPTIONS. U.S. Department of Agriculture, Handbook No. 54, Agricultural Research Service, 1978.

MOLTING IN THE BLUE CRAB

BEAUTIFUL SWIMMERS: WATERMEN, CRABS, AND THE CHESAPEAKE BAY. William W. Warner. Little, Brown & Co., 1976.

PEPTIDES IN THE SINUS GLAND OF *CARDISOMA CARNIFEX:* ISOLATION AND AMINO ACID ANALYSIS. Robert W. Newcomb in *Journal of Comparative Physiology,* Vol. 153, pages 207–221; 1983.

APPARENT H^+ EXCRETION AND CO_2 DYNAMICS ACCOMPANYING CARAPACE MINERALIZATION IN THE BLUE CRAB (*CALLINECTES SAPIDUS*) FOLLOWING MOULTING. James N. Cameron and Chris M. Wood in *Journal of Experimental Biology,* Vol. 114, in press.

CRYSTALS AT HIGH PRESSURE

THE NATURE OF THE CHEMICAL BOND AND THE STRUCTURE OF MOLECULES AND CRYSTALS: AN INTRODUCTION TO MODERN STRUCTURAL CHEMISTRY. Linus Pauling. Cornell University Press, 1960.

COMPARATIVE CRYSTAL CHEMISTRY. Robert Hazen and Larry Finger. John Wiley & Sons, Inc., 1982.

GEOMETRICAL AND STRUCTURAL CRYSTALLOGRAPHY. Joseph V. Smith. John Wiley & Sons, Inc., 1982.

PERCEIVING A STABLE ENVIRONMENT

ON PERCEPTION. Hans Wallach. Quadrangle/The New York Times Book Co., 1976.

TWO KINDS OF ADAPTATION IN THE CONSTANCY OF VISUAL DIRECTION AND THEIR DIFFERENT EFFECTS ON THE PERCEPTION OF SHAPE AND VISUAL DIRECTION. Hans Wallach and Joshua Bacon in *Perception & Psychophysics,* Vol. 21, No. 3, pages 227–242; 1977.

A LIMITATION OF POSITION CONSTANCY. Robert Becklen, Hans Wallach and Donna Nitzberg in *Journal of Experimental Psychology: Human Perception and Performance,* Vol. 10, No. 5, pages 713–723; October, 1984.

THE RESTORATION OF MEDIEVAL STAINED GLASS

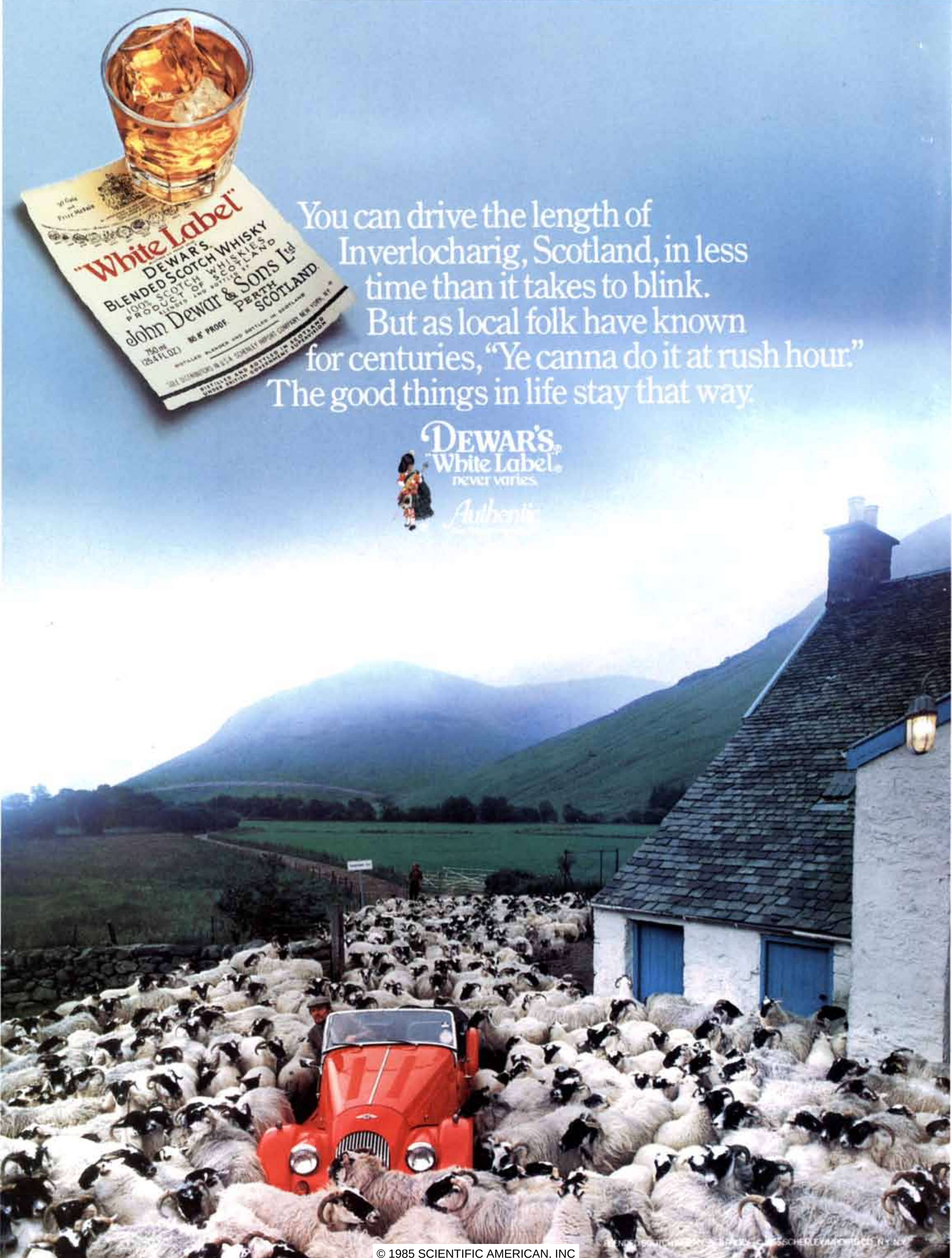
THE DETERIORATION AND CONSERVATION OF STAINED GLASS: A CRITICAL BIBLIOGRAPHY AND THREE RESEARCH PAPERS. R. G. Newton. *Corpus Vitrearum Medii Aevi Great Britain—Occasional Papers 1,* published for the British Academy by Oxford University Press, 1974.

THE EARLY STAINED GLASS OF CANTERBURY CATHEDRAL. Madeline Harrison Caviness. Princeton University Press, 1977.

THE AMATEUR SCIENTIST

FUN WITH STRING FIGURES. W. W. Rouse Ball. Dover Publications, Inc., 1971.

ORDERLY TANGLES: CLOVERLEAFS, GORDIAN KNOTS, AND REGULAR POLYLINKS. Alan Holden. Columbia University Press, 1983.


"White Label"
DEWAR'S
BLENDED SCOTCH WHISKY
John Dewar & Sons Ltd
PERTH SCOTLAND
86.8° PROOF
You can drive the length of
Inverlocharig, Scotland, in less
time than it takes to blink.
But as local folk have known
for centuries, "Ye canna do it at rush hour."
The good things in life stay that way.
DEWAR'S
"White Label"
never varies.
Authentic