SCIENTIFIC AMERICAN

OFFSHORE OIL PLATFORMS

$2.00

April 1982

High Fidelity for Humans:

IT EVEN FINDS
STATIONS WH

Finding your favorite station isn't always as easy as tuning to 123.

For example, now that digital station readouts are standard on most receivers you have to memorize the precise call numbers of all your favorite stations. Not an easy job if you have a dozen or so stations you tune in regularly.

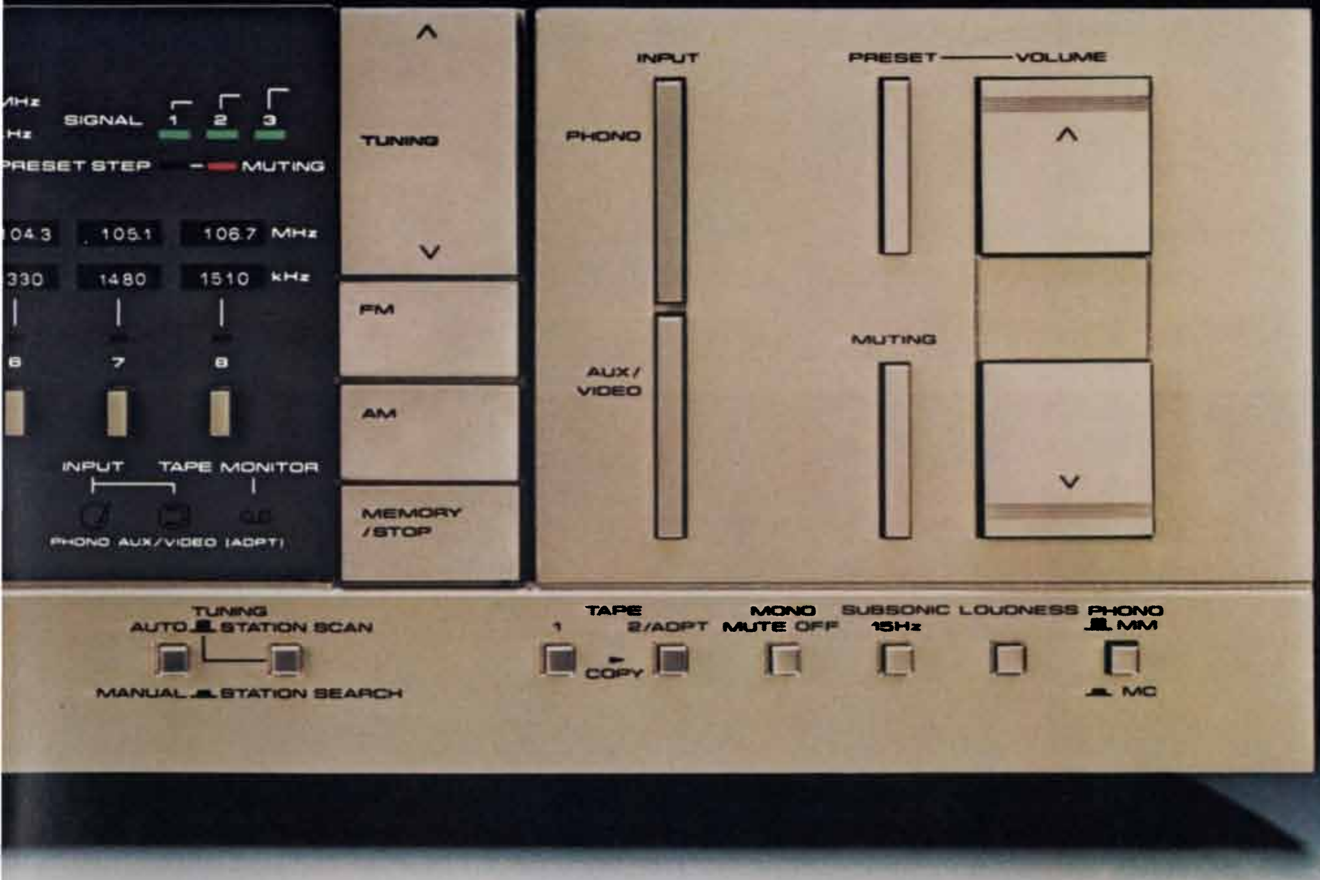
That, however, is just one of the many unpleasantries you have to deal with if you own one of today's conventional receivers. On the other hand, it's just one of the many reasons you should own Pioneer's new SX-7 receiver.

The SX-7 is a product of Pioneer's unique new concept in component design and engineering called *High Fidelity for Humans.* The result is a line of components that are as pleasant to live with as they are to listen to.

For instance, our receiver will commit to memory all your favorite stations. You can preset up to eight AM and eight FM stations. The moment you want to hear one you can recall it instantly.

Should you want to sample a variety of stations without any manual effort, simply press Station Scan. You'll hear five seconds of every strong station on the entire tuning band. If you discover a station you like you simply stop scanning.

Needless to say, not all stations have strong signal strengths. In the past you've had to struggle to tune in those stations with weak signals. The struggle's over. Due to the SX-7's ID Mosfet transistors you can


SIGNAL 1 2 3
PRESET STEP
MUTING
104.3
105.1
106.7 MHz
330
1480
1510 kHz
6
7
8
INPUT
TAPE MONITOR
PHONO AUX/VIDEO (ADPT)
TUNING
FM
AM
MEMORY /STOP
INPUT
PHONO
AUX/ VIDEO
PRESET
VOLUME
MUTING
TUNING
AUTO STATION SCAN
MANUAL STATION SEARCH
TAPE
1
2/ADPT
COPY
MONO
MUTE OFF
SUBSONIC
15Hz
LOUDNESS
PHONO
MM
MC

Seagram's
1857
IMPORTED
BY APPOINTMENT TO HIS EXCELLENCY THE LATE EARL OF ATHLONE
FORMER GOVERNOR GENERAL OF CANADA
Seagram's
V.O.
CANADIAN WHISKY
A BLEND
OF CANADA'S FINEST WHISKIES
DISTILLED, AGED, BLENDED AND BOTTLED UNDER THE SUPERVISION OF THE CANADIAN GOVERNMENT
IMPORTED BY SEAGRAM DISTILLERS COMPANY, NY, NY
THIS WHISKY IS SIX YEARS OLD
BLENDED AND BOTTLED IN CANADA BY
JOSEPH E. SEAGRAM & SONS, LIMITED
WATERLOO · ONTARIO · CANADA · DISTILLERS SINCE 1857
86.8 PROOF
750 ML
BOTTLED IN CANADA
eppes
Schweppes
Club Soda

Established 1845 **SCIENTIFIC AMERICAN** April 1982 Volume 246 Number 4

ARTICLES

38 **ADVANCED OFFSHORE OIL PLATFORMS, by Fred S. Ellers**
Structures that can resist 100-foot waves in waters 600 feet deep call for audacious engineering.

50 **NUTRIENTS THAT MODIFY BRAIN FUNCTION, by Richard J. Wurtman**
The dietary precursors of neurotransmitter molecules sometimes act like drugs rather than foods.

60 **A MAJOR EARTHQUAKE ZONE ON THE MISSISSIPPI, by Arch C. Johnston**
Three of the largest quakes in U.S. history once shook it. The area remains earthquake-prone.

84 **GIANT MOLECULAR-CLOUD COMPLEXES IN THE GALAXY, by Leo Blitz**
Consisting almost entirely of hydrogen molecules, they are the most massive objects in the galaxy.

96 **THE GIANT SQUID, by Clyde F. E. Roper and Kenneth J. Boss**
This deep-sea cephalopod seems almost mythical, but many specimens have now been studied.

106 **SUPERHEAVY MAGNETIC MONOPOLES, by Richard A. Carrigan, Jr., and W. Peter Trower** If isolated magnetic poles exist, they may be exceedingly heavy, slow-moving and rare.

120 **VARIATIONS IN MEDICAL CARE AMONG SMALL AREAS, by John Wennberg and Alan Gittelsohn** They seem to be explained by variations in the preferences of physicians.

136 **THE SHORT LIFE OF ÉVARISTE GALOIS, by Tony Rothman**
Contrary to legend, Galois did not write down group theory the night before he was shot in a duel.

DEPARTMENTS

8 LETTERS
10 50 AND 100 YEARS AGO
15 THE AUTHORS
16 METAMAGICAL THEMAS
32 BOOKS
70 SCIENCE AND THE CITIZEN
150 THE AMATEUR SCIENTIST
162 BIBLIOGRAPHY

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: $21 PER YEAR IN THE U.S. AND ITS POSSESSIONS; $27 PER YEAR IN ALL OTHER COUNTRIES. POSTMASTER: SEND ADDRESS CHANGES TO SCIENTIFIC AMERICAN, 415 MADISON AVENUE, NEW YORK, N.Y. 10017.

A car that "weighs" its own passengers and luggage... then tells the brakes how much power is needed. Such a car is the Peugeot 505 STI.

Test driver demonstrates emergency-braking power of Peugeot 505 STI. Driver maintains control, even though car is braking and turning at the same time.

The car hurtles towards the obstacle. It is a double row of pylons forming a lane 120 feet long and only 12 feet wide that hooks severely to the right.

The driver must bring the car to a halt from an indicated 55 miles an hour within this narrow, curving lane ... *while turning*.

There is the hideous sound of tortured rubber. Acrid tire smoke billows in the air. It is frightening to see and hear.

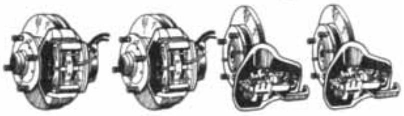
When it is all over, almost two tons of car and passengers have pulled to a stop with several feet to spare.

Brakes that "read"

The car is the Peugeot 505 STI.

It is fitted with an ingenious device that "reads" the weight of the passengers and luggage, then instantly calculates the division of braking power that is needed between the front and rear brakes.

The reason: if a car is lightly loaded, it has a tendency to lock its rear brakes as the weight shifts forward. The result could be a dangerous skid. When the device senses such a light-load condition, it will automatically shut down some of the power to the rear brakes to help keep them from locking. But if a car is *heavily* loaded, as is the case in this test, full braking power is sent to all four wheels.



Peugeot equips every 505 with disc brakes on all four wheels. They shed water and dissipate heat faster than drum brakes.

The sensation in a panic stop has been likened to running into a giant pillow.

Tires that cling like limpets

Powerful brakes are of little use if the carmaker scrimps on tires; it is the tires, after all, that have the final word about stopping the car.

The Peugeot Lion

The lion has represented Peugeot since 1858. It was Peugeot that built the first car ever sold commercially (1891) ... the world's first diesel-powered car (1922) ... and the first high-rpm diesel engine (1967).

You will never find mere letter-of-the-law tires on a Peugeot.

Michelin TRX® low-profile radial tires, mounted on alloy wheels, are standard. Note wide, flat footprint for sure-footed handling.

Every tire is a Michelin steel-belted radial. Special to the 505 STI is the Michelin TRX® tire, mounted on low-profile alloy wheels. The flat, wide footprint of this tire applies a limpet-like grip to the road.

Coupled with stabilizer bars that are thicker than on other 505 models, the tire edges the already-surefooted handling of a Peugeot up another notch in the new STI.

© 1982, Peugeot Motors of America, Inc.

A most-complete car

So lavishly equipped is the Peugeot 505 STI that there is only one option, a 3-speed automatic transmission ($370) in place of the 5-speed overdrive manual gearbox.

Here is a partial list of the equipment that is standard on this $14,990 car:

- Air conditioning, factory-installed
- Power-assisted steering
- Power-assisted 4-wheel disc brakes
- Cruise control
- Electric sunroof
- Electric windows
- Central locking system—a twist of the key locks all four doors and fuel-filler door
- Digital AM/FM 40-watt stereo radio with cassette player, scan tuning, and four speakers
- Automatic electric antenna
- Leather upholstery
- Multi-adjustable driver's seat
- Tachometer
- Limited-slip differential

Unhysterical seats

Click open the door of the Peugeot 505 STI and your nose will immediately tell you what the seats are made of; the car's interior is redolent of leather. What none of your senses will detect, however, is what lies beneath that soft, tanned hide...

Peugeot has banished springs from its seats. Springs have a natural tendency to flutter and bounce hysterically when a car meets with rough roads. In their place, Peugeot has used polymerized foam.

It is more costly to make all 5 seats of the Peugeot in this manner—yes, this is a true 5-seater—but the expenditure has led to seats that have been described as almost sinfully comfortable.

The strokes of a master painter

Pass your hand over the paintwork of the Peugeot 505 STI and you will make a further discovery about the car.

No bored, "just-doing-my-job" worker has created this smooth finish. Instead, the critical part has been done by a computerized robot, taught to imitate the strokes of a master painter.

The STI is the only model of the Peugeot 505 available in a metallic claret or classic black. The black is a wet-on-wet process. First comes the enamel itself. Then, while the black is still glistening wet, it is overlaid with a clear coat of lacquer, sprayed to a thickness of 30 microns.

A final search for faults

Once built, every single car is taken to a special track and road-tested by one of Peugeot's 65-member team of expert drivers. Only when it has met Peugeot's standards is the car permitted to leave Europe for the U.S.

Does this mean that nothing will ever fail on a Peugeot? That would be wishful thinking. The perfect, indestructible car does not exist, no matter what its price. But one can usually count on a car to reveal its little glitches and foibles during the first year of driving.

You will be reassured to know that Peugeot does not care how many thousands of miles you put on your car during the first year you own it. The Peugeot new-vehicle limited warranty is for 12 months... *with unlimited mileage.**

Peugeot has little tolerance for poorly made cars.

*See your dealer for details of Peugeot's limited warranties.

Based on P.O.E. manufacturer's suggested retail prices. Actual prices may vary according to local dealer. Destination charges ($255), state and local taxes, dealer preparation, if any, and license fees are extra.

For name of nearest Peugeot dealer, call **1-800-447-2882** toll free in the continental U.S. In Illinois, 1-800-322-4400.

Please send me the Peugeot Facts Kit, complete with brochures, road tests by leading car magazines, plus list of my nearest dealers.

☐ Please include details on Overseas Delivery.

Name____________________

Address__________________

City_____________________

State____________Zip________

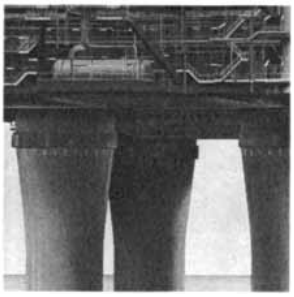
Peugeot Motors of America, Inc.,
One Peugeot Plaza, Lyndhurst,
New Jersey 07071

62-000SA-0400

PEUGEOT

From the second-largest carmaker in Europe comes the new Peugeot 505 STI. Propelled by a fuel-injected gasoline engine that shrugs off tune-ups for 30,000 miles.

© 1982 SCIENTIFIC AMERICAN, INC



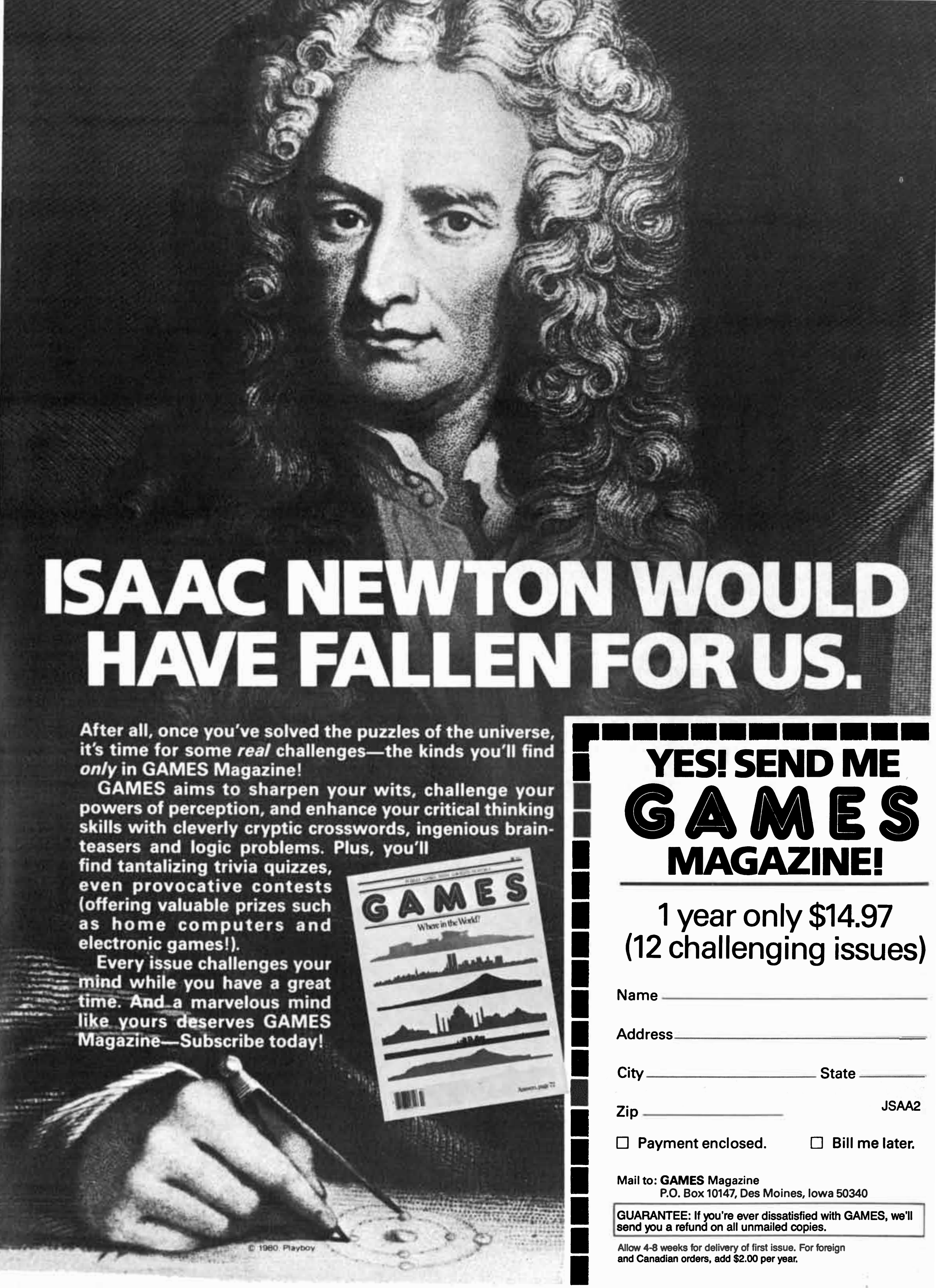
THE COVER

The painting on the cover shows part of an offshore oil platform in the North Sea (see "Advanced Offshore Oil Platforms," by Fred S. Ellers, page 38). The platform, named Statfjord *B*, is the heaviest object ever moved by man. With a weight exceeding 1.2 million tons, it rests in water 472 feet deep 100 miles off the coast of Norway. Later this year the first of 32 wells will be drilled from the structure to tap a reservoir expected to yield 150,000 barrels of oil per day. It is estimated that the finished platform will have cost its owners (principally Statoil, the Norwegian national oil company) $1.8 billion. Statfjord *B* is a "gravity base" platform: its four mighty concrete columns, three of which appear in the painting, rest on a honeycomb concrete structure on the bottom. The entire structure was built in sheltered water and towed to its present site.

THE ILLUSTRATIONS

Cover painting by Ted Lodigensky

Page	Source
16–24	Donald Byrd, Indiana University
27	Courtesy of G. Schirmer, Inc., New York. Frédéric Chopin: *Études,* Op. 10, No. 11; Op. 25, Nos. 1, 2, 3, 6 and 12. Edited by Arthur Friedheim. Copyright © 1916 G. Schirmer, Inc. (renewed). International copyright secured.
38	Norsk Fly og Flyfoto a.s.
40–48	Walken Graphics
49	International Software Services
51	George K. Aghajanian, Yale University School of Medicine
52–59	Alan D. Iselin
61	National Aeronautics and Space Administration
62	Andrew Tomko
63	U.S. Geological Survey, EROS Data Center (*top*); Andrew Tomko (*bottom*)
64	Andrew Tomko
65	Leon Cash and Susan Nava, Memphis State University
66–68	Andrew Tomko
85	© National Geographic Society–Palomar Sky Survey
86–89	George V. Kelvin
90	William J. Welch and John W. Dreher, University of California at Berkeley
92	Philip M. Solomon, State University of New York at Stony Brook; David B. Sanders and Nicholas Z. Scoville, University of Massachusetts at Amherst
94	George V. Kelvin
96–98	Tom Prentiss
99	Tom Prentiss (*top*); Clyde F. E. Roper, National Museum of Natural History (*middle*); L. L. Robins, from *Animals without Backbones,* by Ralph Buchsbaum (*bottom*)
100	Carla Skinder
104–105	Tom Prentiss
108–118	Ian Worpole
121–132	Jerome Kuhl
137	David A. Johnson
138	Jean Dubout, courtesy of Bibliothèque de l'Institut de France, Paris
139–144	Ilil Arbel
149	Jean Dubout, courtesy of Bibliothèque de l'Institut de France, Paris
151–156	Michael Goodman
158	Theodore C. Pickett (*top*), Michael Goodman (*bottom*)


GAMES
Where in the World?
Answers, page 72
© 1980 Playboy

LETTERS

Sirs:

The first complete sentence in the "Books" column for the year 1982 [Scientific American, January] is in error. I hope this is not an omen.

The date at which the earth we ride reached its minimum distance from the sun in 1982—measuring from earth center to sun center—was not January 2, as printed, but in fact January 4, around dawn (Eastern Standard Time).

The source of the error may be instructive. The false estimate was made late last summer, before the authoritative result was in hand. The estimate was got by a seductively simple but naive method: the mean time it takes to go from perihelion to perihelion was simply added to the correct time of perihelion for 1981, which was late on January 1. How could that go wrong by two whole days?

The perihelion is a minor feature of the orbit of the earth, in the sense that the earth's orbit is so nearly circular. The entire variation from the shortest earth-sun distance to the longest amounts to only about three million miles out of nearly 100 million; even Giotto could hardly draw so true a circle freehand. As perihelion day arrives the orbit radius is slowly decreasing; after the instant of zero change the radius grows again. But the speed of change in orbit radius around perihelion is always very small compared with the tangential part of the earth's motion, so nearly perpendicular to the radius is the earth's way during those days. That radial speed component one day from perihelion time itself is only 20 miles per hour, whereas the speed overall amounts to more than 60,000 miles per hour. Any comparably small radial motion of the earth present in addition to its simple Keplerian solar orbital motion is able to shift the sensitive zero in the radial speed component and hence to shift the time of perihelion.

The earth has a companion, the moon. It is not the center of the earth that flies in an ellipse around the sun, neglecting the small pulls of the other planets. It is rather the center of mass of the entire earth-moon system that the sun directs. With respect to that earthy if immaterial point both the moon and the earth are to a good approximation in steady circular motion. Of course, the light moon moves more swiftly, and the heavy earth executes its countermotion reduced by the mass ratio, about one to 80. Even so, that carries the center of the earth around a circle some 6,000 miles across each month. The speed in this circle is nearly 30 miles per hour, clearly enough to shift the perihelion time by a day and more. At full moon and at new moon there is no effect; the earth's motion in lunar orbit is tangential to the solar orbit. But in between the earth's moon-respondent motion does shift the time of perihelion, delaying it on waxing moons, advancing it on waning ones. Moon motion is the source of the larger part of the year-to-year apparent irregularity in the time of perihelion that the carefully made calculations report. The effects of other planets, particularly Jupiter and Venus, are not a great deal smaller.

It is always a surprise to learn anew how rich becomes the texture of even the simplest concepts once they engage not the world on paper but the real one. That this little issue may be cumulatively important to the glaciation of the earth over the longer run is also to be recalled with wonder.

It is a pleasure to thank Owen Gingerich and Brian G. Marsden of the Center for Astrophysics of the Harvard College Observatory and the Smithsonian Astrophysical Observatory, and LeRoy E. Doggett of the Nautical Almanac Office of the U.S. Naval Observatory, for their help, friendly experts all.

PHILIP MORRISON

Book Editor, *Scientific American*
Massachusetts Institute of Technology
Cambridge, Mass.

Sirs:

I read with interest, if not some confusion, "The Anthropic Principle," by George Gale [Scientific American, December, 1981].

It appears that Dr. Gale has finally put Descartes before the horse: "I am, therefore . . . " (I think).

THOMAS HUBER, M.D.

Peoria, Ill.

Sirs:

Your readers might find interest in the following observation:

"The Lining of the Small Intestine," by Florence Moog [Scientific American, November, 1981], presents facts that well exemplify a point that was emphasized by Benoît B. Mandelbrot in his book *Fractals: Form, Chance, and Dimension* (W. H. Freeman and Company, 1977) and was also mentioned by Martin Gardner in your department "Mathematical Games" for December, 1976, and April, 1978. The point is that the result of measuring a length, an area, a volume and so on may depend on the scale of measurement. For example, the length of a country's coastline depends on whether the surveyor uses a "measuring stick" of 10 kilometers, 100 meters, one meter, one centimeter, .1 millimeter and so on. The shorter the measuring stick is, the more detail is detected and the larger is the result obtained for the length of the coastline.

In the case of the adult human small intestine we have, to the roughest approximation, a tube with a length of about six meters and an internal diameter of about two centimeters, and hence a surface area of about a third of a square meter. On the scale of measurement of a few millimeters the "circular folds" are detected and the measured surface area increases as much as threefold, to about a square meter. On the scale of a tenth of a millimeter the villi enter the picture, and the measured area becomes about 15 square meters, a fifteenfold increase. On a scale 1,000 times smaller, a tenth of a micrometer, the microvilli (the villi's villi) are resolved, and the measured area increases some twentyfold to about 300 square meters, 1,000 times larger than the tube approximation.

JOE ROSEN

Department of Physics and Astronomy
Tel-Aviv University
Tel-Aviv, Israel

Editorial correspondence should be addressed to The Editors, Scientific American, 415 Madison Avenue, New York, N.Y. 10017. Manuscripts are submitted at the author's risk and will not be returned unless accompanied by postage.

Advertising correspondence should be addressed to C. John Kirby, Advertising Director, Scientific American, 415 Madison Avenue, New York, N.Y. 10017.

Offprint correspondence and orders should be addressed to W. H. Freeman and Company, 660 Market Street, San Francisco, Calif. 94104. For each offprint ordered please enclose 75 cents.

Subscription correspondence should be addressed to Subscription Manager, Scientific American, P.O. Box 5969, New York, N.Y. 10017. For change of address, notify us at least four weeks in advance. Send both old and new addresses and enclose an address imprint from a recent issue. (Date of last issue on your subscription is shown at upper right-hand corner of each month's mailing label.)

Name ______________________

New Address ______________________

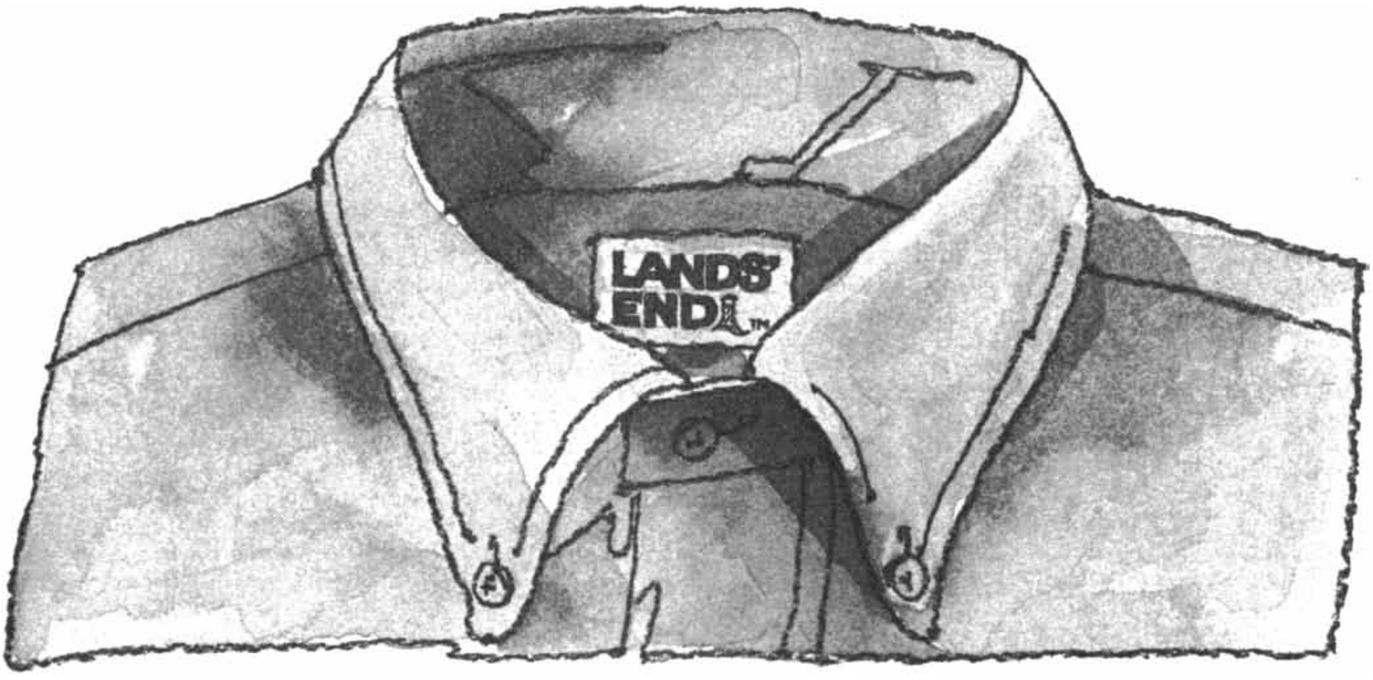
____

____

Old Address ______________________

____

____


LANDS' END

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

APRIL, 1932: "Over the past year or two there has been much comment in the press and from the pulpit concerning the new rapprochement of religion and science. What started it all? The German physicist Heisenberg discovered that it was impossible to determine both the position and the speed of a given electron. This has become the 'principle of indeterminacy,' and it has been construed to mean that, as Sir James Jeans puts it, 'nature abhors accuracy and precision above all things.' Not only Jeans but also A. H. Compton and Millikan and Eddington and others of great fame have given encouragement to this idea, and the clergy have not been slow to seize hold of their encouragement. It would mean the end of the old cause-and-effect principle, so beloved of the hated materialists. But many physicists seriously question the interpretation being given Heisenberg's principle of indeterminacy, and hence the inferences concerning religion and free will to be made from that interpretation. All the principle of indeterminacy means is that man has no way of determining both the position and the speed of a given electron, not that there is no determinacy in either."

"Radio waves represent the first tool with which it may be possible to carry a signal across the great reaches of astronomical space. If among the billions of stars there are some that have a family of satellites such as our own sun has, it is reasonable to expect that some of these satellites would be in a zone of temperatures where life could exist. If life does exist somewhere else, then someday someone is likely to encounter, by means of radio, an extraterrestrial intelligence. What a sublimely dramatic moment it will be for those concerned when this first interstellar contact is made!"

"All previous attempts to obtain extremely high velocities of rotation have been surpassed by J. W. Beams and A. J. Weed of the University of Virginia, who have driven a rotating cone of steel as high as 500,000 revolutions per minute. The rotor resembles a top. Tiny air jets impinge on it and slightly raise it. It finds a position of stable equilibrium very near the stator and continues to rotate at high velocity entirely on 'bearings' of air. The scientist-inventors have perfected a way of introducing and removing liquids without stopping the centrifuge, the rotor being hollow and having small communicating openings near the top center."

"The isolation in pure form of the anti-neuritic vitamin B is the result of a new line of research developed by Professor Adolf Windaus of the University of Göttingen. Vitamin B is now thought to be a complex substance made up of several vitamins. The one investigated by Professor Windaus, called B_1 by some investigators, is the one that protects pigeons and other non-human animals against the disease known as polyneuritis and protects man against beri-beri. Another part of the vitamin B complex is the factor thought to protect man against the disease pellagra."

"It would appear that we have become a nation of cowards. Our pessimism seems to have beclouded our vision and dulled our reason. Businesses become more fearful for their continued existence, retrench, cut salaries, lay off employees. Seeing these things, those fortunate ones with work retrench, cut expenditures, wear worn-out clothes, drive rattly old cars, do without necessities. Demand is lowered and production decreases further. Companies retrench further, lay off more employees. Thus are forged the links of a chain of fear that is binding us down."

SCIENTIFIC AMERICAN

APRIL, 1882: "Charles Robert Darwin, whose influence upon the current of modern thought has been surpassed by that of no other scientific investigator, died at his residence near Orpington, England, on Wednesday, April 19. Mr. Darwin was born at Shrewsbury, England, on February 12, 1809. His early education was received at the public school in Shrewsbury, whence he passed to the University of Edinburgh, where he spent two years. He then went to Christ's College, Cambridge, where he was graduated in 1831. His bent for natural research was not diverted by his schooling; soon after his graduation he read a paper on marine zoology, giving such promise of scientific ability that he was offered the position of naturalist on the now historic H.M.S. *Beagle,* soon to start on a cruise of scientific exploration round the world. Returning from this voyage in 1836 Darwin made ready for publication his *Journal of Researches,* and in 1840–42 he edited *Zoology of the Voyage of the Beagle.* Soon afterward he published his classic works on the structure and distribution of coral reefs. These works were rapidly followed by *Geological Observations on Volcanic Islands* in 1844 and *Geological Observations in South America* in 1846. In 1853 the Royal Society of London awarded him the Royal Medal, and in 1859 he received the Wollaston Medal of the Geological Society. His epoch-marking *Origin of Species by Natural Selection* appeared the same year. The controversies provoked by this work probably did more to attract popular thought to questions of natural science, and to change the popular as well as the scientific mode of regarding such topics, than any other influence of the century."

"During the month of March Dr. Henry Draper succeeded in photographing four times the spectrum of the nebula in Orion. Dr. Draper has also taken photographs of the nebula itself so as to watch for changes in it and observe whether the process of aggregation into stars can be detected. Collated with the photographs of the spectrum, they show clearly, it is said, evidences of such condensations."

"The eyes of capitalists looking for great fields for large investments are turned toward the South. A stranger from any one of the old countries on the other side of the sea, in traveling through the South, is astonished to see the woods and wilds that meet his gaze, even in districts of the country that were among the earliest settlements in the colonies. This long delay in Southern development is attributable to the fact that we have a vast territory with a sparse population, and to the fact that we have always been an agricultural and not a manufacturing people. But manufacturing facilities, especially in this age, afford means for making money and for circulating it among people of all classes and conditions. Manufacturing helps agriculture in immeasurable degree. It attracts population and puts money into the hands of people to buy the products of the farmer."

"The tunnel under the English Channel has finally been begun and is rapidly progressing. The work of tunneling is being executed by means of a Beaumont perforator, which carries two arms having a rotary motion, each of which is provided with seven short steel blades. The frame to which these parts are affixed makes a forward motion of about half an inch at every complete revolution of the steel cutters. In this way a thin slice of the whole surface in front is removed at every revolution of the perforator, and there is thus obtained a cylindrical opening of about seven and a half feet in diameter. The excavating apparatus is making a forward progress at present of about 310 feet per week. If the work on the French side is pushed with the same rapidity as that on the English, it is estimated that the hole may be completed in the space of three and a half years."


The GLENLIVET
AGED 12 YEARS
AGED 12 YEARS
BOTTLED IN SCOTLAND
The GLENLIVET
12 YEARS OLD
Unblended all malt Scotch Whisky
Distilled and Bottled in Scotland
GEORGE & J. G. SMITH LIMITED
THE GLENLIVET DISTILLERY · GLENLIVET SCOTLAND
U.S. IMPORTERS
THE GLENLIVET DISTILLING COMPANY
NEW YORK, N.Y.
750 ML. (25.4 FL. OZ.) 86 PROOF
ESTABLISHED 1824
The GLENLIVET
12 YEARS OLD THE GLENLIVET SCOTCH WHISKY

We're looking for trouble.

Technology should solve your business problems without creating more of them. And Bell has the systems to prove it.

For example, the First Wisconsin National Bank has 1800 data terminals scattered over three states, updating accounts and exchanging information with the bank's computers in Milwaukee. If a terminal suddenly runs wild, part of the network is jammed.

Our trouble-shooting Dataphone® II communications system solves the problem. It not only transmits data over telephone lines, but monitors the flow of data as well, locating and isolating trouble instantly—even zeroing in on the very component involved. For First Wisconsin, that's money in the bank.

If your business has an information management problem, chances are communications products from Bell—like the Dataphone II system—can solve it in trouble-free fashion. One call to your Bell System Account Executive will start putting our knowledge to work for you.

The knowledge business

$750 CASH BONUS

on all new Omegas if you take retail delivery between now and March 31. Participating dealers contribute 25% to these bonuses, so make your best deal. Your actual cost depends on the deal you negotiate.

Omega ES

Available Engine:	2.8L, 2-bbl., V6, high output Bore: 89 mm; stroke: 76 mm
Dynamics:	Horsepower 130 @ 5,400 RPM Torque 145 @ 2,400 RPM High-output cam
Transmission:	4-speed manual
Suspension:	MacPherson strut, front; trailing arm with track bar, rear; firm ride and handling package
Wheels:	Special sport wheels; steel-belted, blackwall, radial tires
Instrumentation:	Rallye cluster: tachometer, voltmeter, temperature gage, trip odometer
Interior:	Individual custom sport bench or available bucket seats in luxury Brougham trim; custom sport steering wheel
Exterior:	"Soft" sloping front end; wraparound bumpers with integral air scoop; wide body side moldings; black-matte styling accents; available in silver, dark blue, dark redwood and charcoal
Models:	Coupe and sedan

For those who believe getting there should be more than half the fun.

Even today, there's still room to do it with style.

Oldsmobile

We've had one built for you.

Some Oldsmobiles are equipped with engines produced by other GM divisions, subsidiaries or affiliated companies worldwide. See your dealer for details.

GM

THE AUTHORS

FRED S. ELLERS ("Advanced Offshore Oil Platforms") is manager of special projects for the Pipeline Division of Bechtel Incorporated. He received his bachelor's degree in 1940 at Oklahoma State University. After graduation he joined the staff of the Tidewater Associated Oil Company, for which he inspected refineries. From 1942 to 1946 he served in the Army Corps of Engineers in the Pacific. He first joined Bechtel in 1946 and stayed until 1948. During most of that period he worked on the development of oil fields in Saudi Arabia. In 1948 he went to work for the Arabian American Oil Company (ARAMCO), leaving in 1961 to rejoin Bechtel. Most of his professional work has centered on the engineering aspects of extracting and transporting petroleum. Ellers worked on developing the first large oil fields in Abu Dhabi in the early 1960's and more recently the Argyll field in the United Kingdom sector of the North Sea.

RICHARD J. WURTMAN ("Nutrients That Modify Brain Function") is professor of neuroendocrine regulation at the Massachusetts Institute of Technology. His bachelor's degree was granted by the University of Pennsylvania in 1956 and his M.D. by the Harvard Medical School in 1960. He was a medical research officer at the National Institute of Mental Health from 1965 to 1967. In the latter year he moved to M.I.T. In addition to the subject of the current article Wurtman's research interests include the catecholamines, acetylcholine, glutamate and the hormones that cause menarche and ovulation.

ARCH C. JOHNSTON ("A Major Earthquake Zone on the Mississippi") is assistant professor of geology at Memphis State University and director of its Tennessee Earthquake Information Center. He got his B.S. at Southwestern at Memphis college in 1967. From 1968 to 1973 he was in the Air Force; he spent much of that time piloting the C130E transport. In 1979 he received a Ph.D. in geophysics from the University of Colorado at Boulder. His main research is on the causes of earthquakes. Johnston writes that his interest in geology was awakened in long flights over the Andes while he was in the Air Force.

LEO BLITZ ("Giant Molecular-Cloud Complexes in the Galaxy") is assistant professor of astronomy at the University of Maryland at College Park. He obtained his B.S. in 1967 at Cornell University and his M.A. (1975) and Ph.D. (1979) in astronomy from Columbia University. His dissertation topic was the giant molecular-cloud complexes that form the subject of his current article. From 1978 to 1981 he was a postdoctoral fellow at the Radio Astronomy Laboratory of the University of California at Berkeley. Last year he moved to the University of Maryland. Blitz writes that he has "been using the giant molecular clouds as probes of the outer portions of the Milky Way. This work has demonstrated that the Milky Way is far more massive than has been thought and that much of the mass is likely to be in an as yet unknown form."

CLYDE F. E. ROPER and KENNETH J. BOSS ("The Giant Squid") are zoologists whose specialty is the anatomy and classification of marine animals. Roper is chairman of the department of invertebrate zoology at the National Museum of Natural History at the Smithsonian Institution. He was graduated from Transylvania University in 1959 with a B.A. and went on to get his master's and doctoral degrees from the University of Miami. In 1966 he joined the staff of the National Museum, becoming curator in 1972 and chairman of the department of invertebrate zoology in 1980. Boss is professor of biology at Harvard University and curator in malacology (the study of mollusks) at the Museum of Comparative Zoology at Harvard College. He received his B.A. in 1957 at Central Michigan College, his M.Sc. in 1959 at Michigan State University and his Ph.D. in 1963 from Harvard. After completing his doctoral degree he worked in a laboratory of the Bureau of Commercial Fisheries at the Smithsonian. He went to Harvard in 1966 as assistant curator in the Museum of Comparative Zoology; he was appointed professor of biology in 1970. Boss writes: "I am suffering from what I believe [Lawrence] Durrell called islomania, an uncommon passion for the escape and isolation offered by islands; to that end I have a book-filled old fisherman's house... on an island in the Gulf of Maine.... It was there that Clyde and I became enthusiastic about the prospect of writing an article on *Architeuthis* [the giant squid]."

RICHARD A. CARRIGAN, JR., and W. PETER TROWER ("Superheavy Magnetic Monopoles") are experimental physicists who are making a joint search for experimental evidence of the existence of monopoles. Carrigan's B.S. (1953), M.S. (1956) and Ph.D. in physics (1962) are all from the University of Illinois at Urbana-Champaign. He served as a research physicist at the Carnegie Institute of Technology from 1961 to 1964 and as assistant professor there from 1964 to 1968. In 1968 he moved to the Fermi National Accelerator Laboratory (Fermilab), becoming assistant head of the research division in 1977. In 1967 and 1968 he was a senior Fulbright fellow at the Deutsches Elektronen-Synchrotron (DESY) in Hamburg. Trower is associate professor of physics at the Virginia Polytechnic Institute and State University. His degrees are also from the University of Illinois: B.S. (1957), M.S. (1963) and Ph.D. in physics (1966). In 1966 he joined the faculty at Virginia State University. In addition to their work on magnetic monopoles both men do research on well-established elementary particles.

JOHN WENNBERG and ALAN GITTELSOHN ("Variations in Medical Care among Small Areas") have collaborated since 1967 in studying health care in New England. Wennberg is professor of epidemiology at the Dartmouth Medical School. He received his B.A. at Stanford University and his M.D. from the McGill Medical School. After finishing his residency he got a master's degree in public health from the Johns Hopkins School of Hygiene and Public Health in 1966. From 1967 to 1972 he was assistant professor of medicine at the University of Vermont and director of the state's Regional Medical Program. In 1972 he was appointed associate professor of epidemiology at Dartmouth, a post he held until he became professor in 1980. From 1973 until 1978 Wennberg was also assistant professor of preventive and social medicine at the Harvard Medical School. Gittelsohn is professor of biostatistics at the Johns Hopkins University School of Hygiene and Public Health. His bachelor's degree and doctorate in biostatistics are both from the University of California at Berkeley. After working in the health department of both California and New York he joined the Johns Hopkins faculty in 1964.

TONY ROTHMAN ("The Short Life of Évariste Galois") is currently a visiting fellow in astrophysics at the University of Oxford. He received his undergraduate education at Swarthmore College, from which he was graduated in 1975. In 1981 he got his Ph.D. in physics from the University of Texas at Austin. His research interests include black holes, the formation of baryons early in the history of the universe and the synthesis of atomic nuclei in stars. He is also a professional writer, having published a novel and a number of popular articles on scientific subjects. He is at work on a second novel. Rothman writes: "My interest in Galois actually stemmed from a play...I wrote a few years ago. It concerns the Russian poet Pushkin as well as Galois. In the course of my background research I discovered that the standard accounts of Galois's life in English were, to say the least, inaccurate."

METAMAGICAL THEMAS

The music of Frédéric Chopin: startling aural patterns that also startle the eye

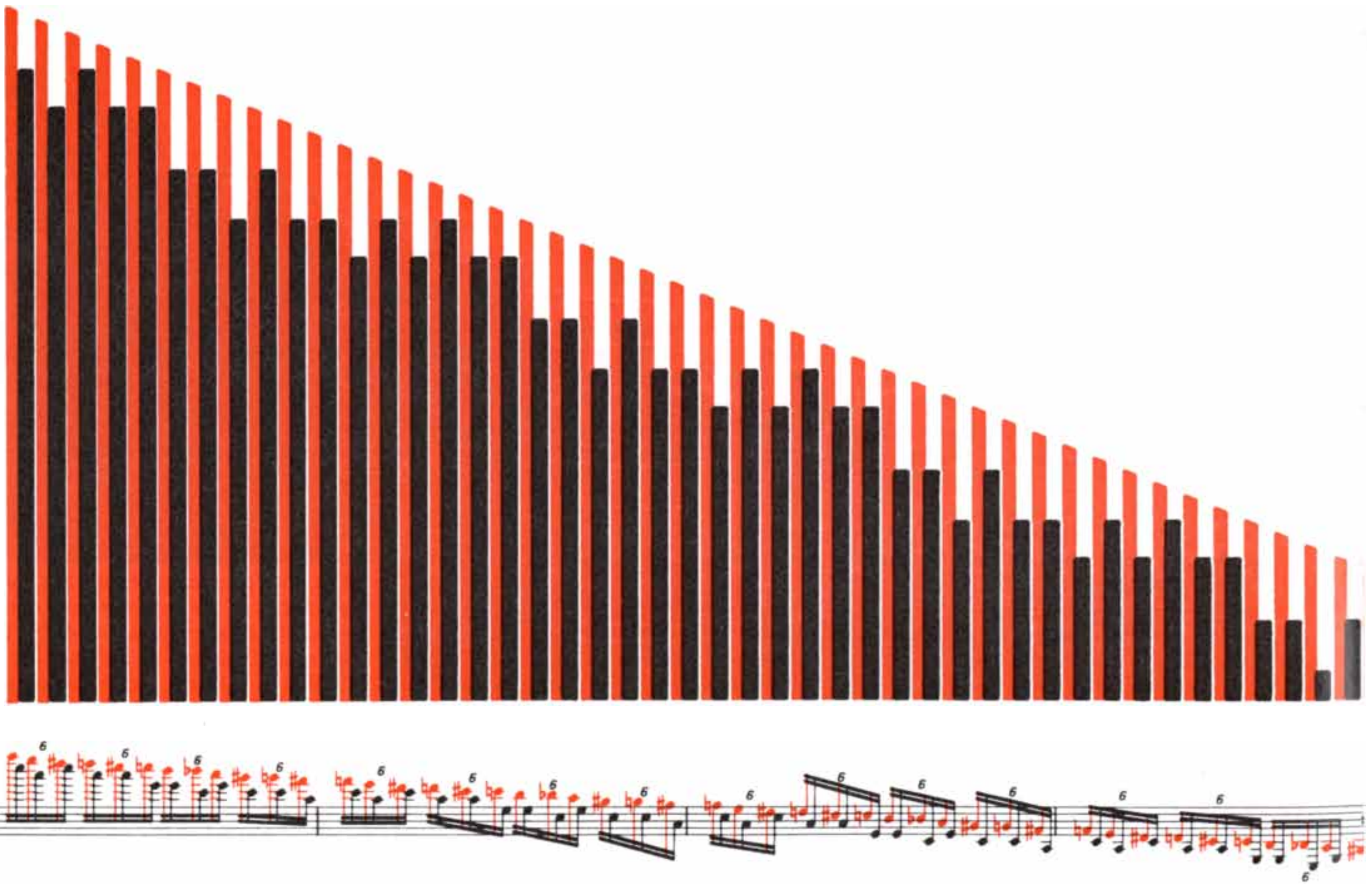
by Douglas R. Hofstadter

The abstract visual pattern of the music running across the bottom of these two pages is the first eight bars of one of the most difficult and lyrical pieces for piano ever composed, namely the 11th étude in A minor from Frédéric Chopin's Opus 25, written in about 1832 when he was in his early 20's. As a boy I heard the Chopin études many times over on my parents' record player, and I quickly grew to love them. They became as familiar to me as the faces of my friends. Indeed, I cannot imagine who I would be if I did not know these pieces.

A few years later, as a teenage piano student, I wanted to learn to play some of these old friends. I went to the local music store and found a book of them. I shall never forget my reaction on opening the book and looking for my friends. They were nowhere to be found! I saw nothing but masses of black notes and chords: complex, awesome visual patterns that I had never imagined. It was as if, expecting to meet old friends, I had instead found their skeletons grinning at me. It was terrifying. I closed the book and left, somewhat in shock.

I remember going back several times to the music store, each time pulled by the same curiosity tinged with fear. One day I worked up my courage and actually bought that book of études. I suppose I hoped that if I simply sat down at the piano and tried playing the notes I saw, I would hear my old friends, albeit a little slowly. Unfortunately nothing of the kind happened. In general I could not even play the two hands together comfortably, let alone re-create the sounds I knew so well. This left me disheartened and a little frightened at the realization of the awesome complexities I had taken for granted. You can look at it two ways. One way is to be amazed at how human perception can integrate a huge set of independent elements and "hear" only a single quality; the other is to be amazed at the incredible skill of a pianist who can play so many notes so quickly that they all blur into one shimmering mass, a "co-hear-ent" totality.

At first it was bewildering to see that "friends" had inside them anatomies of such overwhelming complexity. But looking back, I don't know what I expected. Did I expect that a few simple chords could work the magic I felt? No; if I had thought it over, I would have realized this was impossible. The only possible source of that magic was in some kind of complexity—patterned complexity, to be sure. And I think the



The opening of the 11th étude from Chopin's Opus 25 is printed out **(bottom)** *by a computer program developed by Donald Byrd.*

experience taught me a lifelong lesson: that phenomena perceived to be magical are always the outcome of complex patterns of nonmagical activities taking place at a level below perception. In other words, the magic behind magic is pattern. The magic of life itself is a perfect example, emerging as it does out of patterned but lifeless activities at the molecular level. The magic of music emerges from complex, nonmagical—do I mean "metamagical"?—patterns of notes.

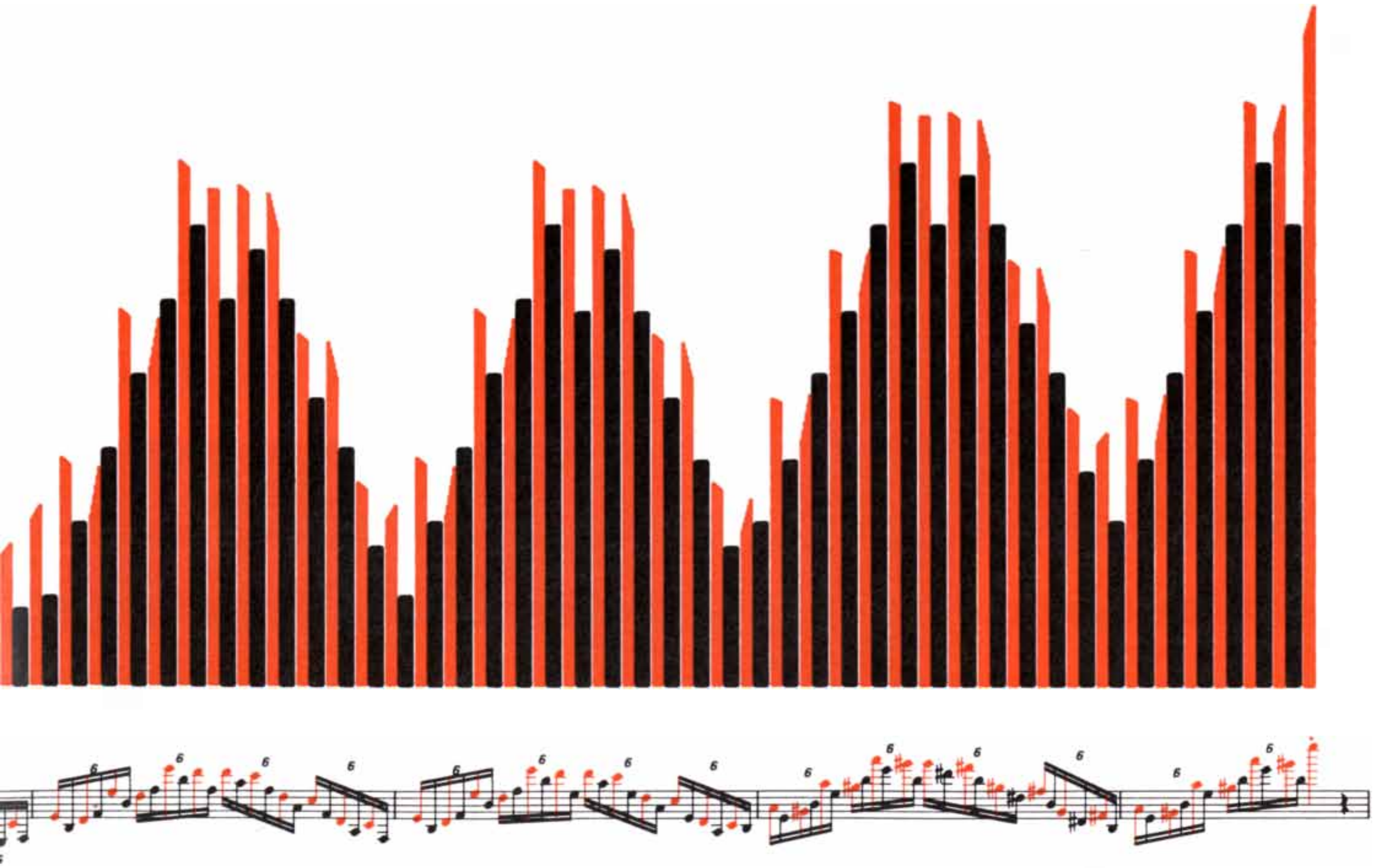
Having bought the volume, I felt drawn to it, wanted to explore it somehow. I decided that, hard work though it might be, I would learn an étude. I chose the one that was my current favorite—the one pictured—and set about memorizing the finger pattern in the right hand, together with the patterns that follow it, making up the first two pages or so. I played the pattern literally thousands of times, and gradually it became natural to my fingers, although never as natural as it had always sounded to my ears—or rather to my mind.

It was then I first observed the amazing subtlety of the lightning flash of the right hand, how it is composed of two alternating and utterly different components: the odd-numbered notes (shown in color in both the bar graph and the musical notation below it) trace out a perfect descending chromatic scale for four octaves, while the even-numbered notes (shown in black), wedged between them like pickets between the spaces in a picket fence, dictate an arpeggio with repeated notes. To execute this alternating pattern the right hand flutters down the keyboard, tilting from side to side like a swift in flight, its wings beating alternately.

A word of explanation. On a piano there are 12 notes (some black, some white) from any note to the corresponding note one octave away. Playing them all in order creates a "chromatic" scale, as contrasted with the more familiar diatonic scales (usually major or minor). These latter involve only seven notes apiece (the eighth note being the octave itself). The seven intervals between the successive notes of a diatonic scale are not all equal. Some are twice as large as others, yet to the ear there is a perfect intuitive logic to it. Rather paradoxically, most people can sing a major scale without any trouble, uneven intervals notwithstanding, but few can sing a chromatic scale accurately, even though it "ought" to be much more straightforward, or so it would seem, since all its intervals are exactly the same size. The chromatic scale is so called because the extra notes it introduces to fill the gaps in a diatonic scale have a special kind of "bite" or sharpness to them that adds color or piquancy to a piece. For that reason a piece filled with notes other than the seven notes belonging to the key it is in is said to be chromatic.

An arpeggio is a broken chord played one or more times in a row, moving up or down the keyboard. Thus it bears a resemblance to a spread-out scale, a little like someone bounding up a staircase three or four steps at a time. Chopin's music is filled with both arpeggios and chromatic passages, but the intricate fusion of these two opposite structural elements in the 11th étude struck me as a masterpiece of ingenuity. And what is amazing is how it is perceived when the piece moves quickly. The chromatic scale comes through loud and clear, forming a smooth "envelope" of the pattern (your eye picks it out too), but the arpeggio blurs into a kind of harmonic fog that deeply affects one's perception, if only subliminally, or so it seems at least to the untrained ear.

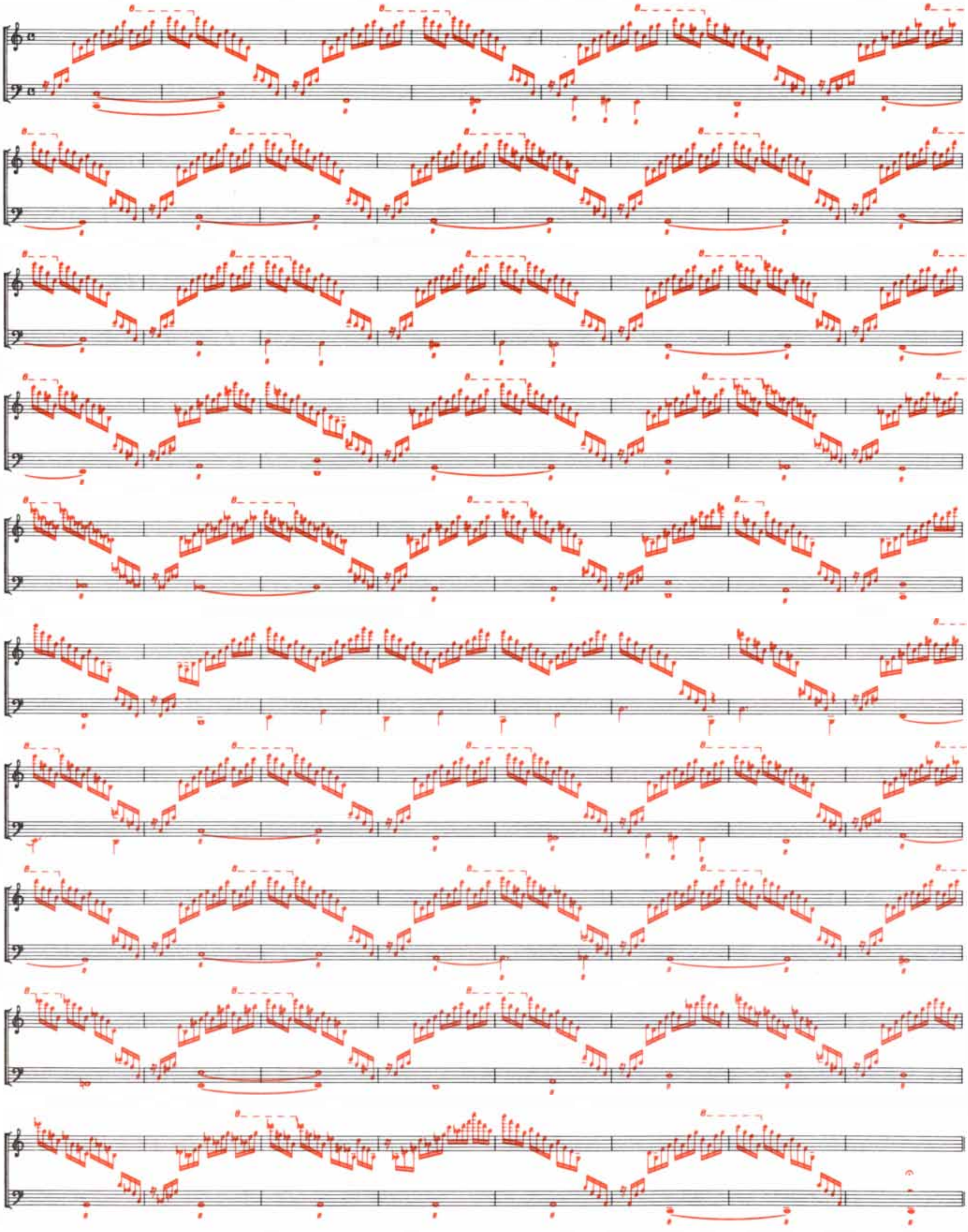
Each étude in the book I bought has a characteristic appearance, a visual texture. This was one of the most striking things about it at first. I was not at all accustomed to the idea of written music as texture; the simple pieces I had played up to that time were slow, so that ev-



The colored vertical bars at the top trace the rise and fall of the odd-numbered notes; the gray bars, that of the even-numbered ones

ETUDE

F. CHOPIN

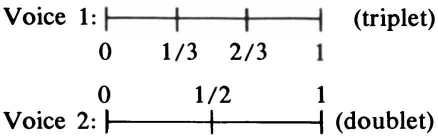


Chopin's étude Op. 10, No. 1, in C major is printed out so that its waves are in phase ery note was distinctly heard. In other words, those pieces were coarse-grained compared with the fine grain of a Chopin étude, where notes often go by in a blur and are merely parts of an auditory gestalt. Conversion of this kind of auditory experience to notated music sheets often yields quite stunning textures and patterns. Each composer has a characteristic set of patterns the eye becomes familiar with, and these études provided for me a stunning realization of that fact.

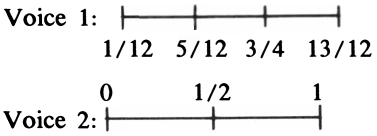
Sadly, I was forced to abandon étude Op. 25, No. 11, after having learned a little more than a page—it was simply too hard for me. James Huneker, an American critic and one of Chopin's first English-language biographers, said of this study: "Small-souled men, no matter how agile their fingers, should not attempt it." Well, whatever the size of my soul, my fingers were not agile enough. For a while that discouraged me from attacking any more Chopin études at all. A few years later, however, when I was working more earnestly on improving my modest piano skills, I came across an isolated Chopin étude in a book of medium-difficult selections. It turned out to be one of three études he had composed later in life, none of which had been on my parents' records. This was a find indeed! Luckily its texture looked less prickly, its pace less forbidding. Somewhat gingerly I played through it very slowly and discovered that it was astonishingly beautiful and not as inaccessible as the others I had tried.

Like the rest of Chopin's studies, this one is centered on a particular technical point, although to think of the études primarily in that way is like thinking of the gymnastic performances of Nadia Comaneci as mere setting-up exercises. Louis Ehlert, a 19th-century musicologist, wrote of one of the most beautiful études in Opus 25 (the sixth one, in G-sharp minor): "Chopin not only versifies an exercise in thirds; he transforms it into such a work of art that in studying it one could sooner fancy himself on Parnassus than at a lesson. He deprives every passage of all mechanical appearance by promoting it to become the embodiment of a beautiful thought, which in turn finds graceful expression in its motion." And so it is for this easier, posthumously published étude in A-flat major, whose chief technical concern is the concept of "three against two," a special case of the general concept of polyrhythm.

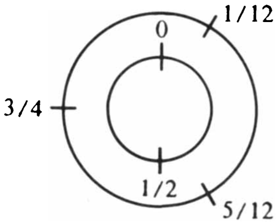
Mathematically the concept is simple enough: play two musical lines simultaneously, one of them sounding three notes to the other's two. Usually the triplet and the doublet are aligned so that they start at the same instant. When they are both plotted on a unit interval, you can see that the doublet's second note is struck halfway between the triplet's second and third notes. Of course, this is simply a restatement of the fact that 1/2 is the arithmetic mean of 1/3 and 2/3:



In theory two voices playing a three-against-two pattern need not be perfectly aligned. If you shift the upper voice by, say, 1/12 to the right, you get a different picture:



Here the triplet's third note starts halfway through the doublet's second. As you can see, the triplet extends beyond the end of the interval, presumably to join onto another identical pattern. We can fold the pattern around and represent its periodicity in a circle:



By rotating either of the concentric circles like a knob we get all possible ways of hearing three beats against two. In Chopin and most other Western music, however, the only possibility exploited is where the triplet and doublet are perfectly "in phase."

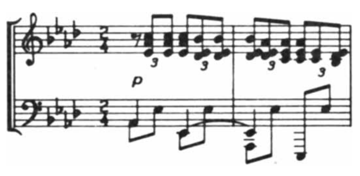
At first I found the three-against-two rhythm hard to perform exactly. One has to learn how to hear the voices separately, to hear the lilt of the three-rhythm weaving itself into the square mesh of the two-rhythm. Of course, it is easy to hear when someone else is playing; the trick is to hear it in one's own playing! In principle the task is not hard, but it is one of coordination that requires practice. I found that once I had mastered the problem of playing the two rhythms evenly and independently I could play the entire étude. To play it—or to hear it—is like smiling through tears, it is so beautiful and sad at the same time.

It is impossible to pinpoint the source of the beauty, needless to say, but it is certainly due in part to the way the chords in the right hand flow into one another. Almost all the way through the piece the right hand plays three-note chords (six to a measure) against single notes by the left hand (four to a measure). The delicacy of the piece comes from the fact that very often when one chord flows into the next one, only a single note changes. And to add to the subtlety of this slowly shifting pattern, usually the steps taken by the shifting voice are single scale steps rather than wide jumps. These "rules" do not hold all the way, of course; there are numerous exceptions. Nevertheless, there is a uniform aural texture to the piece that imbues it with its soft melancholy, known in Polish as *tesknota*.

It is interesting to speculate about the extent to which such formal considerations occurred to Chopin while he was composing. It is well known that Chopin revered Bach's music. "Always play Bach," was his advice to a pupil, and he was particularly devoted to the Well-Tempered Clavier, a paragon of elegant formal structures. Chopin confided to his friend Eugène Delacroix, the painter, that "the fugue is like pure logic in music.... To know the fugue deeply is to be acquainted with the element of all reason and consistency in music." Clearly Chopin loved pattern.

A stunning demonstration of Chopin's extreme awareness of the visual appeal of the textures in his études is provided by the manuscript of his étude Op. 10, No. 1, in C major, one about which Huneker wrote, in his inimitable prose:

"The irregular, black, ascending and descending staircases of notes strike the neophyte with terror. Like Piranesi's marvelous aerial architectural dreams, these dizzy acclivities and descents of Chopin exercise a charm, hypnotic, if you will, for eye as well as ear. Here is the new technique in all its nakedness, new in the sense of figure, design, pattern, web, new in a harmonic way. The old order was horrified at the modulatory harshness, the young sprigs of the new fascinated and a little frightened. A man who could explode a mine that assailed the stars must be reckoned with."

That neophyte might well have been me. Huneker's words form an amusing contrast with what the 19-year-old Chopin himself wrote of this, his first étude, in a letter to his friend Tytus Woyciechowski in 1829: "I have written a large exercise in form, in my own personal style; when we get together, I'll show it to you." A finished copy, believed to be in Chopin's hand, is now in the Museum of the Frédéric Chopin Society in Warsaw. Under the present conditions in Po-

Threes beat against twos in A-flat major étude

It also has front-wheel drive.

The Honda Accord LX Hatchback has everything. It's luxuriously comfortable. It's incredibly well engineered.

Honda has been out in front with front-wheel drive for years. In fact, Honda has always been the only auto manufacturer to sell all of its cars in America with front-wheel drive and a transverse-mounted engine.

Front-wheel drive gives excellent traction because the weight of the engine is over the driving wheels. You corner confidently. And the engine mounted sideways is easy to service.

With front-wheel drive, there's more interior space. As you can see, it has been used to beautiful advantage in the Accord LX.

Everything you see is standard equipment. Including something you also feel.

Air conditioning with new push-button controls and an improved ventilating system.

One thing you hear is a warning chime that sounds when the halogen headlights are left on.

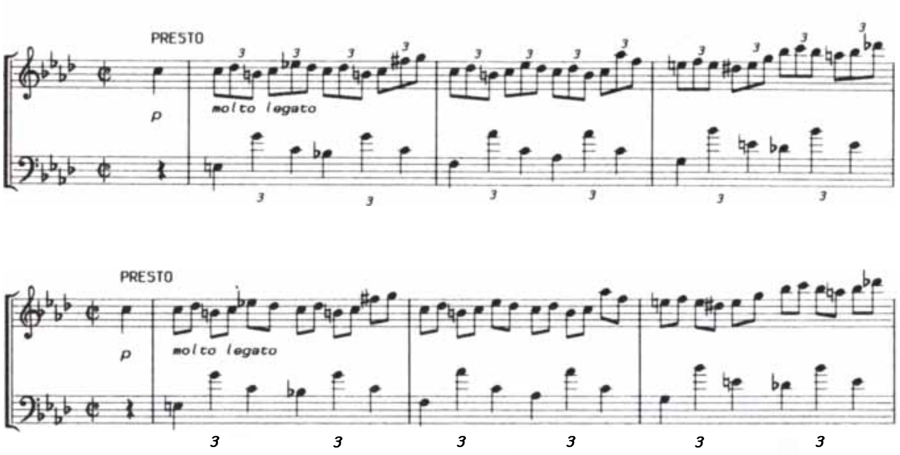
Other features you'll sense are the smoother and quieter ride and better road handling. With Michelin steel-belted radial tires and variable-assist power steering. And ventilated front disc brakes. Again standard equipment.

The Accord LX. It has quality engineering. It has luxury. It simply has all the comforts of Honda.

HONDA

We make it simple.

©1982 American Honda Motor Co., Inc.

© 1982 SCIENTIFIC AMERICAN, INC

The étude Op. 25, No. 2, was written in threes **(top)** *but can be played in twos* **(bottom)**

land it would be difficult to reproduce it directly. Fortunately Donald Byrd of Indiana University has developed a computer program that can print out music according to specification. With his marvelous program, and with the help of Adrienne Gnidec and me, Byrd has printed the music in such a way as to reproduce the large-scale visual patterns of Chopin's own manuscript, in which he took great care to align all the crests of the massive waves. When this piece is played at the proper speed, each sweep up and down the keyboard is heard as one powerful surge, like the stroke of an eagle's wing, with the notes at each crest sparkling brilliantly like wing tips flashing in the sun.

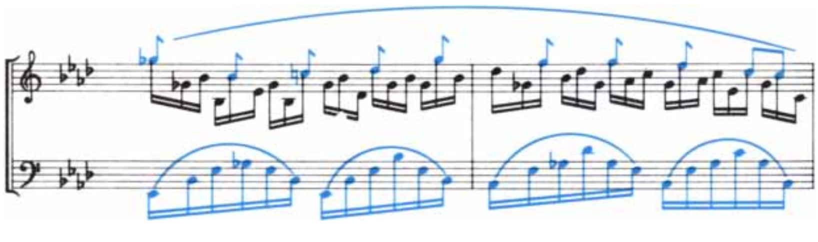
Another interesting feature of Chopin's notation, copied here, is his positioning of the octave whole notes in the bass. Instead of placing them at the very start of each measure, aligned with the sixteenth-note rests, Chopin centers each one in its own measure, thereby creating an elegant visual balance but losing some notational clarity. Musically the centering has no effect. Since a whole note lasts for the duration of an entire 4/4 measure, it must be struck at the start of the measure, otherwise it would overflow into the next measure, and that is impossible. Or rather, it would violate a much more rigid convention of music notation, namely that no note can designate a sound that overflows the boundaries of its measure. Hence the only interpretation possible is that the whole note is to be struck at the outset. In other words, the centering is simply a charming artistic touch with a quaint 19th-century flavor, like the ornaments on a Victorian house. The modern music-reading eye is used to more functional notation; in particular it expects the staff to be in essence a graph of the sound in which the horizontal axis is time. Thus notes struck simultaneously are expected to line up vertically.

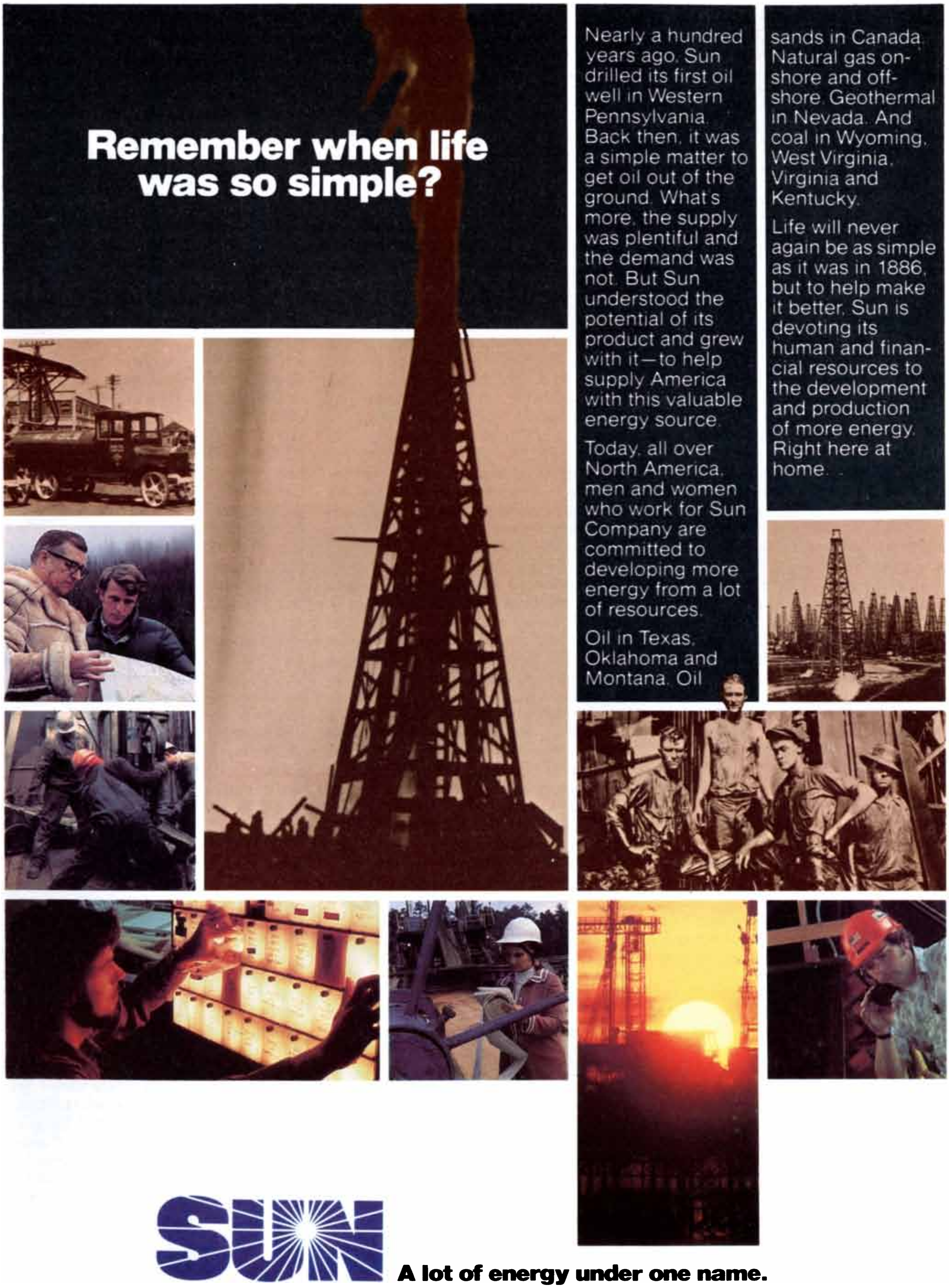
But let us return to the matter of Chopin's preoccupation with form and structure. Few composers of the romantic era have penned such visually patterned pages, have spun a whole cloth out of a single textural idea. With Chopin, however, preoccupation with strict pattern never took precedence over the expression of heartfelt emotions. One must distinguish, it seems to me, between "head pattern" and "heart pattern," or, in more objective-sounding terms, between syntactic pattern and semantic pattern. The notion of a syntactic pattern in music corresponds to the formal structural devices used in poetry: alliteration, rhyme, meter, repetition of sounds and so on.

The notion of a semantic pattern is analogous to the pattern or logic that underlies a poem and gives it reason to exist: the inspiration, in short. That there are such patterns in music is as undeniable as that there are courses in the theory of harmony. Yet harmony theory has no more succeeded in explaining such patterns than any set of rules has yet succeeded in capturing the essence of artistic creativity. There are words to describe well-formed patterns and progressions, but no theory yet invented has even come close to creating a semantic sieve so fine as to let all bad compositions fall through and to retain all good ones. Theories of musical quality are still descriptive and not generative; to some extent they can explain in hindsight why a piece seems good, but they cannot enable someone to create new pieces of quality and interest. It is nonetheless fascinating, if not downright compelling, to try to find certain earmarks of greatness, to try to understand *why* one composer's music reaches in and touches one's innermost core and another composer's music leaves one cold and unmoved. It is a mystery.

After learning the posthumous A-flat étude I felt encouraged to tackle some of the others. One of the ones I had loved the most was Op. 25, No. 2, in F minor. To me it was a soft, rushing whisper of notes, a fluttering like the leaves of a quaking aspen in a gentle breeze. Yet it was not just a scene of nature; it expressed a human longing, a melancholy infused with strange and wild yearnings for something unknown and remote—*tesknota* again. I knew this melody inside out from many years of hearing it, and I looked forward to transferring it to my fingers.

After a couple of months of practice my fingers had built up enough stamina to play the piece fairly evenly and softly. This was very satisfying until one day an acquaintance for whom I was playing it commented, "But you're playing it in twos—it's supposed to be in threes." What she meant by this was that I was stressing every second note rather than every third. Bewildered, I looked at the score, and of course, as she had pointed out, the melody was written in triplets. But surely Chopin had not meant it to be played in threes! After all, I knew the melody perfectly. Or did I? I tried playing it in threes. It sounded strange and unfamiliar, a perceptual distortion the like of which I had never experienced.

I went home and took out my parents' old record of the Chopin études Opus 25 (played by a wonderful but hardly remembered pianist named Alexander Jenner). I put on the F minor étude and tried to hear which way he played it. I found I could hear it *either* way. Jenner had played it so smoothly, so free of accent (as it is said Chopin did), that one really could not tell which way to hear it. All of a sudden I saw I knew *two* melodies composed of the same sequence of notes! I felt myself to be very fortunate, because now I could experience this familiar old melody in a fresh new way. It

Ballade in F minor has flags flying on every fourth note **(color)** *in the treble clef*



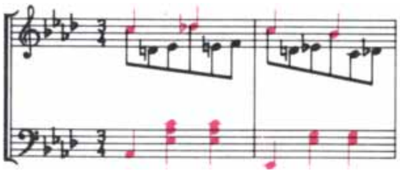
Remember when life was so simple?

Nearly a hundred years ago, Sun drilled its first oil well in Western Pennsylvania. Back then, it was a simple matter to get oil out of the ground. What's more, the supply was plentiful and the demand was not. But Sun understood the potential of its product and grew with it—to help supply America with this valuable energy source.

Today, all over North America, men and women who work for Sun Company are committed to developing more energy from a lot of resources.

Oil in Texas, Oklahoma and Montana. Oil sands in Canada. Natural gas onshore and offshore. Geothermal in Nevada. And coal in Wyoming, West Virginia, Virginia and Kentucky.

Life will never again be as simple as it was in 1886, but to help make it better, Sun is devoting its human and financial resources to the development and production of more energy. Right here at home.

SUN **A lot of energy under one name.**

Melody of A-flat major waltz is in twos (color)

was like falling in love with the same person twice.

I had to practice hard to undo the bad habits of "biplicity" and replace them with the indicated "triplicity," but it was a delight. The hardest part, however, was combining the two hands. With doublets in the right hand this had presented no problem; all the accented notes fell in coincidence with notes in the left hand, moving at exactly half the speed of the right hand in a pattern of widespread arpeggios. But if I were to spread my accents thinner, so that I accented only every *third* note of the right hand, then many of the notes in the left hand would be struck simultaneously with weak notes in the right. This may sound simple enough, but I found it very tricky. The difference is shown in the top illustration on page 22 (which like most of the others in this article was created by Byrd's program at Indiana University).

Even after mastering the right-hand solo in triplets, I found that when I put the parts together, it was at first nearly impossible to keep from softly accenting the melodic notes coinciding with the bass. It was a fearsome task of coordination, yet I enjoyed it greatly. After a while something just "snapped into place," and I found I was doing it. It was not something I could consciously control or explain; I simply was playing it right all of a sudden.

In Huneker's commentary on this etude he quotes Theodor Kullak, another Chopin specialist, on the "algebraic character of the tone-language" and then adds: "At times so delicate is its design that it recalls the faint fantastic tracery made by frost on glass."

Chopin's music is filled with such "algebraic" tricks of cross-rhythm. He seemed to revel in them in a way that no previous composer had. A famous example is his iconoclastic waltz Opus 42 in A-flat major, written in 1840. In it the bass line follows the usual oom-pah-pah convention, but the melody of the first section completely counters this threeness; its six eighth notes, instead of being broken up into three pairs aligned with the left hand's bounces, form two triplets, as in the F minor étude I have been discussing. Here, however, in contrast to the nearly accentless shimmering that is desired in the étude, the initial notes of successive triplets are to be clearly emphasized and prolonged, thus creating a higher-level melody (shown in color in the illustration at the left) abstracted out of the quietly rippling right hand. This melody is composed of two notes per measure, beating regularly against the three notes of the waltzing bass. It is a marvelous *trompe l'oreille* effect and one that Chopin exploited again in his E major scherzo, Opus 54, written in 1842, when he was 32.

In that same year Chopin wrote what some admirers consider to be his greatest work: the fourth ballade, in F minor. The piece is filled with noteworthy passages, but one in particular had a profound effect on me. One day, long after I knew the piece intimately from recordings, a friend told me he had been practicing it and wanted to show me a bit of tricky polyrhythm that was particularly interesting. I was actually not that interested in hearing about polyrhythm at the moment, and so I did not pay much attention when he sat down at the keyboard. Then he started to play. He played only two measures, but by the time they were over I felt that someone had reached into the very center of my skull and caused something to explode deep down inside. This "bit of tricky polyrhythm" had undone me completely. What was going on?

Of course, it was much more than just polyrhythm, but that is part of it. As you can see in the plot of the measures in the bottom illustration page 22, the left hand forms large, rumbling waves of sound, like deep ocean waves on which a ship is sailing. Each wave consists of six notes, forming a rising and falling arpeggio. High above these billows of sound a lyrical melody soars and floats, emerging out of a blur of notes swirling around it like a halo. This high melody and its halo are actually fused together in the right hand's 18 notes per measure. They are written as six groups of three, so that in each half measure nine high notes beat against the six-note ocean wave below—already a clear problem in three-against-two. But look: on top of those flying triplets there are eighth-note flags placed on every *fourth* note! Thus there is a flag on the first note of the first triplet, on the second note of the second triplet, on the third note of the third triplet, on the fourth note of the fourth triplet.... Well, that cannot be. In fact, the fourth triplet has no flag; the flag goes to the first note of the *fifth* triplet, and the pattern resumes. Flags waving in wind, high on the masts of a sailing ship.

This wonderfully subtle rhythmic construction might, just might, have been invented by anyone, say by a rhythm specialist with no feeling for melody. And yet it was not. It was invented by a composer with a supreme gift for melody and harmony as well as for rhythm, and this can be no coincidence. A mere "rhythms hacker" would not have the sense to know what to do with this particular rhythm any more than with any other rhythmic structure. There is something about this passage that shows true genius, but words alone cannot define it. You have to hear it. It is a burning lyricism, having a power and intensity that defy description.

One must wonder about the soul of a man who at age 32 could write such possessed music—a man who at the tender age of 19 could write such perfectly controlled and poetic outbursts as the études of Opus 10. Where could this rare combination of power and pattern, this musical self-confidence and maturity, have come from?

In search of an answer one must look to Chopin's roots, both his family roots and his roots in his native land, Poland. Chopin was born in a small and peaceful country village 30 miles west of Warsaw called Żelazowa Wola, which means Iron Will. His father, Nicolas (Mikołaj) Chopin, was a Frenchman by birth but had emigrated to Poland and had become an ardent Polish patriot (so ardent, in fact, that he had participated in the ill-fated insurrection led by the national hero Jan Kiliński in 1794 against the Russian occupation of Warsaw). Chopin's mother, Justyna Krzyżanowska, was a distant relative of the rich and aristocratic Skarbek family, who lived in Żelazowa Wola. She lived with them as a family member and took care of various domestic matters. When Mikołaj Chopin came to be the tutor of the Skarbek children, he and Justyna met and married. In addition to being a gentle and loving mother she was as fervent a Polish patriot as her husband and also had a romantic and dreamy streak. They had four children, of whom Frédéric, born in 1810, was the second. The other three children were girls, one of whom died young of tuberculosis, which was to finally claim Frédéric as well, at age 39. The four children doted on one another. All in all Chopin had a very happy childhood.

The family moved to Warsaw when Frédéric was very young, and there he was exposed to culture of all kinds, since his father was a teacher and knew university people of all disciplines. Frédéric was a fun-loving, spirited boy. The summer he was 14 he spent away from home in a lilac-filled village called Szafarnia. He wrote home a series of letters gleefully mocking the style of the *Warsaw Courier,* a gossipy provincial paper of the times. One item from his "Szafarnia Courier" ran as follows (in full):

"The Esteemed Mr. Pichon [an anagram of Chopin] was in Golub on the 26th of the current month. Among other foreign wonders and oddities, he came across a foreign Pig, which pig quite specially attracted the attention of this most distinguished Voyageur."

Chopin's musical talent, something he

Hyatt offers a simple basis of comparison: your money's worth.

Hyatt: It's what other hotels try to be. Over the years, and especially recently, other hotels have raised their rates within dollars of Hyatt without justifying that increase with the service and amenities that are, in total, uniquely Hyatt.

The world's most innovative hotel continues to lead the way. Hyatt is always pioneering new standards of contemporary hotel service at 109 prime locations across the U.S. and around the world. Personal attention and special touches are offered as a matter of course. Fresh flowers. Natural greenery. Just-squeezed orange juice. A complimentary newspaper in the morning. Express check-in and check-out.

A touch of Hyatt: you can count on it. Collectively all these special amenities are called *a touch of Hyatt*℠. In terms of value, it simply means getting your money's worth. Every cent of it.

At Hyatt, the difference is everything. Except cost. Without compromising service or amenities, Hyatt strives to maintain standards and set new ones along the way. That's why travelers everywhere are discovering that Hyatt makes all the difference.

Capture the Hyatt spirit at 109 hotels worldwide. Weekend rates are another special touch of Hyatt℠. For reservations, call your travel planner or 800 228 9000.

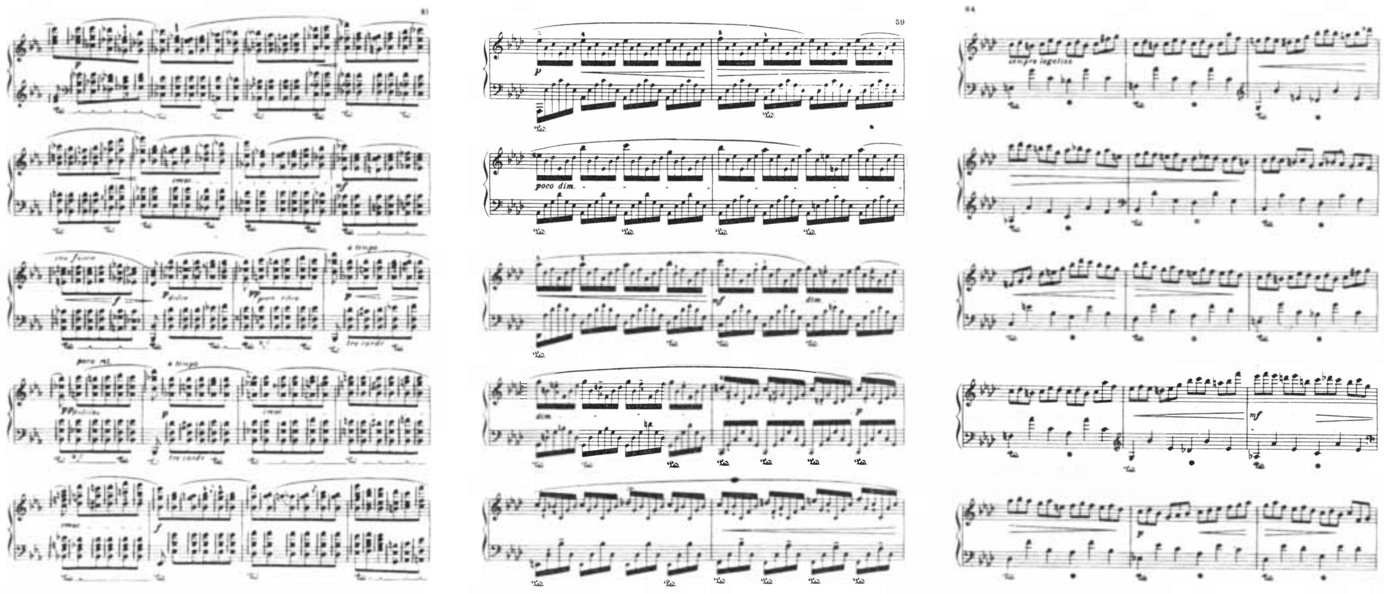
a touch of Hyatt

HYATT HOTELS®

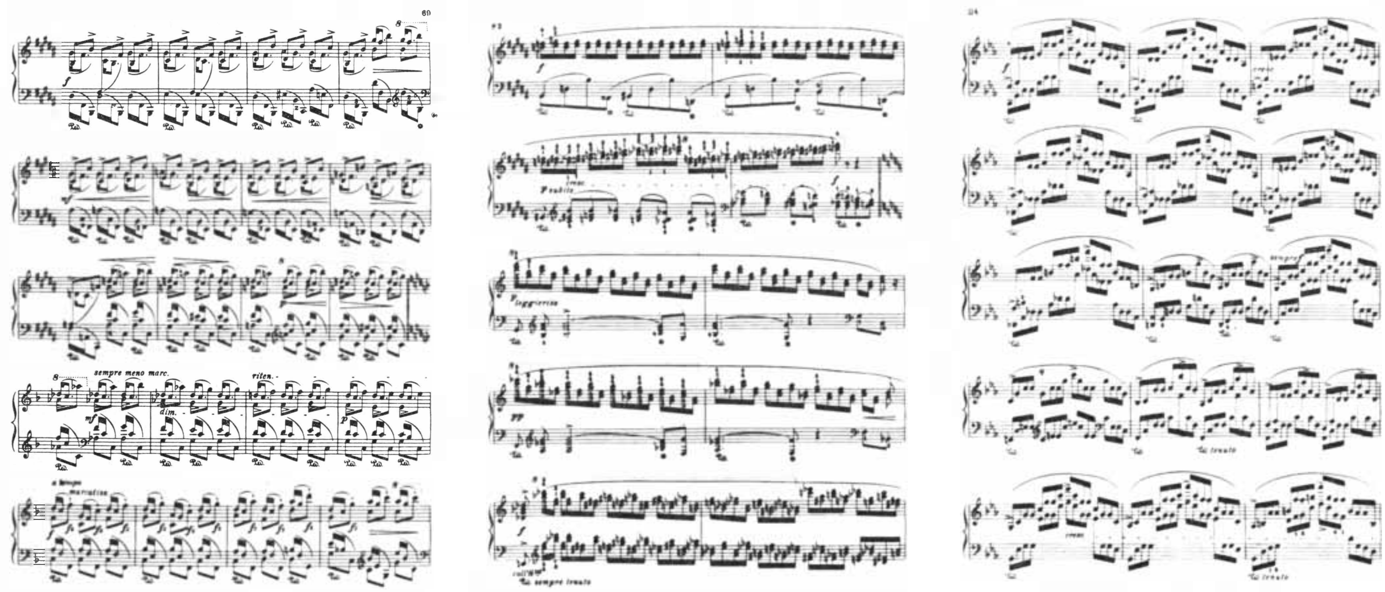

YOUR TIME HAS COME

shared with his mother, emerged very early and was nurtured by two excellent piano teachers, first by a gentle and good-humored old Czech named Wojciech Żywny and later by the director of the Warsaw Conservatory, Józef Elsner.

Chopin grew up in the capital city of the "Grand Duchy of Warsaw"—what little remained of Poland after it had been carved up, in three successive "partitions" in the late 18th century, by its greedy neighbors, Russia, Prussia and Austria. The turn of the century was marked by a mounting nationalistic fervor; in Warsaw and Cracow, the two main Polish cities, there occurred a series of rebellions against the foreign occupiers, but to no avail. A number of ardent Polish nationalists went abroad and formed "Polish Legions" whose purpose was to fight for the liberation of all oppressed peoples and to eventually return to Poland and reclaim it from the occupying powers. When Napoleon invaded Russia in 1806, a Polish state was established for a brief instant; then all was lost again. The Polish nation's flame flickered and nearly went out totally, but as the words to the Polish national anthem proclaim, "Jeszcze Polska nie zginęła, póki my żyjemy." It is a curious sentence, built out of past and present tenses, and literally translated it runs, "Poland has not yet perished, as long as we live." The first clause sounds fatalistic, as if Poland surely *will* perish, but not quite yet! Actually the connotations are not that despairing; a better overall translation would be, "Poland will not perish, as long as we live."

The Poles are a people who have learned to distinguish sharply between the idea of themselves as a people and the idea of the land they live in. The "Polish nation" represents a spirit rather than a piece of territory, although of course the nation came into existence because of the bonds between people who lived in a certain area. It is the fragility of this flickering flame, and the determination to keep it alive, that Chopin's music reflects so purely and poignantly. There is a certain fusion of bitterness, anger and sadness called *żal* that is uniquely Polish. One hears it, to be sure, in the famous mazurkas and polo-

Op. 10, No. 11 (E-flat major)

Op. 25, No. 1 (A-flat major)

Op. 25, No. 2 (F minor)

Op. 25, No. 3 (F major)

Op. 25, No. 6 (G-sharp minor)

Op. 25, No. 12 (C minor)

Six pages from six Chopin études show their remarkable variation in visual pattern naises, pieces Chopin composed in the form of national dances. The mazurkas are mostly smaller pieces based on folklike tunes with a lilting 3/4 rhythm; the polonaises are grand, heroic and martial in spirit. But one hears this burning flame of Poland just as much in many of Chopin's other pieces, for example in the slow middle sections of such pieces as the waltzes in A minor (Op. 34, No. 2) and A-flat major (Op. 64, No. 3), the pathos-filled prelude in F-sharp major (Op. 28, No. 13) and particularly in the middle part of the F-sharp minor polonaise (Opus 44), where a ray of hope bursts through dark visions like a gleam in the gloom. One hears *żal* in the angry, buzzing harmonies of the étude in C-sharp minor (Op. 10, No. 4), and in the passion of the étude in E major (Op. 10, No. 3). In fact, Chopin is said to have cried out on hearing this piece played in his presence, "Ô ma patrie!" ("O my homeland!")

But aside from the fervent patriotism of Chopin's music there is in it that different and softer kind of Polish nostalgia: *tęsknota.* It is his yearning for home—for his childhood home, for his family, for a dream-Poland that at age 20 he had left forever. In 1830, at the height of the turmoil in Warsaw, Chopin set out for France. He had a premonition that he would never return. Traveling by way of Vienna, he made slow progress. When things boiled over in late 1831, when in September, 1831, the Russians finally crushed the desperate Warsaw insurrection, Chopin was in Stuttgart. On hearing the news he was overwhelmed with agitation and grief, partly out of fear for the fate of his family, partly out of love for his stricken homeland. He wavered about going back to Poland and fighting for his nation, but the idea eventually receded from his mind.

It was at about this time that he composed the 12th and last étude of his Opus 10. Of this étude Chopin's Polish biographer Maurycy Karasowski wrote: "Grief, anxiety and despair over the fate of his relatives and his dearly beloved father filled the measure of his sufferings. Under the influence of this mood he wrote the C minor étude, called by many the 'Revolutionary Étude.' Out of the mad and tempestuous storm of passages for the left hand the melody rises aloft, now passionate and anon proudly majestic, until thrills of awe stream over the listener, and the image is evoked of Zeus hurling thunderbolts at the world."

This is pretty strong language. Huneker echoes these sentiments, as does the French pianist Alfred Cortot, who in his famous Student's Edition of the études refers to the piece as "an exalted outcry of revolt...wherein the emotions of a whole race of people are alive and throbbing." I myself have never found this étude as overwhelming as these authors do, although it is unquestionably a powerful outburst of emotion. If someone had told me that one of the études had come to be known as the "Revolutionary Étude" and had asked me to guess which one, I would certainly have picked one of the last two of Opus 25, either No. 11 in A minor, the one pictured at the beginning of this article, with its tumultuous cascades of notes in the right hand against the surging, heroic melody in the left hand, or else No. 12 in C minor, which sounds to me like a glowing inferno seen at night from far away, flaring up unpredictably and awesomely. As for the actual "Revolutionary Étude," I have always found its ending enigmatic, fluctuating as it does between major and minor, between the keys of F and C, like an indecisive thunderclap.

Still, this piece, like the martial A-flat major polonaise (Opus 53), has become a symbol of the tragic yet heroic Polish fate. Wherever and whenever it is played, it is special to a Pole, whose heart beats faster and whose spirit cannot fail to be deeply moved. I shall never forget how I heard it nightly as the clarion call of Poland, when I would try to tune in Radio Warsaw from Germany in 1975. Two measures of shrill, rousing chords above a roaring left hand, like a call to arms, were repeated over and over again as the call signal, preceding a nightly broadcast of Chopin's music. Nor shall I ever forget how the feeble signal of Radio Warsaw faded in and out, symbolizing to me the flickering flame of Poland's spirit.

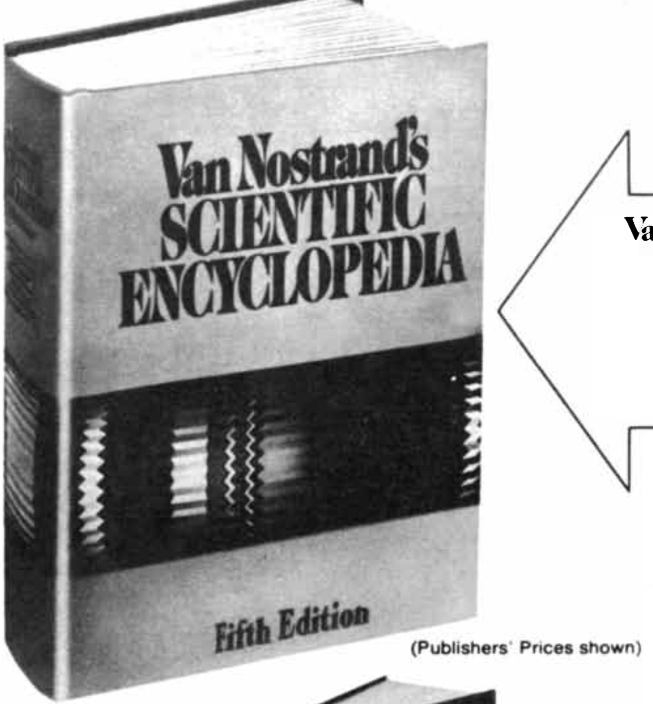
However one chooses to describe it—whether in terms of *żal* and *tęsknota, patriotyzm* and polyrhythm or chromaticism and arpeggios—Chopin's music has had a deep influence on the composers of succeeding generations. It is perhaps most visible in the piano music of Alexander Scriabin, Sergei Rachmaninoff, Gabriel Fauré, Felix Mendelssohn, Robert and Clara Schumann, Johannes Brahms, Maurice Ravel and Claude Debussy, but Chopin's influence is far more pervasive than even that would suggest. It has become one of the central pillars of Western music, and as such it has its effect on the music perceived and created by everyone in the Western world.

In one way Chopin's music is purely Polish, and that Polishness—*Polskność*—extends even to foreign-inspired pieces such as his Bolero, Tarantella, Barcarolle and so on. In another way, though, Chopin's music is universal, so that even his most deeply Polish pieces—the mazurkas and polonaises—speak to a common set of emotions in everyone. But what are these emotions? How are they so deeply evoked by mere pattern? What is the secret magic of Chopin? I know of no more burning question.

As your introduction to

The Library of Science

Take this $69.50 classic for $3.00



(Publishers' Prices shown)

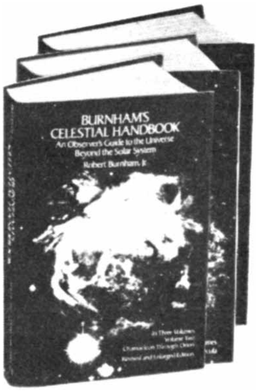
Van Nostrand's SCIENTIFIC ENCYCLOPEDIA

Edited by Douglas M. Considine. Nearly 200 experts have contributed to this thoroughly revised and greatly expanded fifth edition of the most authoritative single-volume source of scientific knowledge ever assembled. Enormous 9¼" x 12" volume contains 2.2 million words, 2382 pages, 2500 photographs, drawings, and charts, and 500 tables. 7200 articles cover everything from the earth and space sciences to energy technology, from the life sciences to mathematics, from information sciences to physics and chemistry. "...an amazing book...for both the general and scientific reader." *–The New York Times*

or 3 other books for only $1.00 each (values to $74.90)

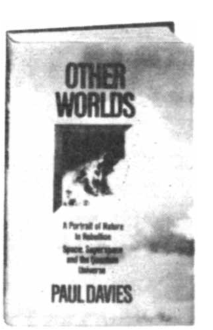
You simply agree to buy 3 more books—at handsome discounts—within the next 12 months.

34440. THE ART OF ELECTRONICS. *Horowitz and Hill.* Emphasizes circuit intuition and non-mathematical design techniques. Packed with diagrams, schematics, and illustrations. **$24.95**

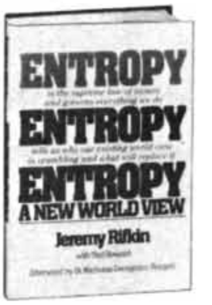


36978-3. BURNHAM'S CELESTIAL HANDBOOK: An Observer's Guide to the Universe Beyond the Solar System. *Robert Burnham, Jr.* 3 volumes containing hundreds of illustrations and photographs. *Counts as 3 of your 3 books.* **$60.00**

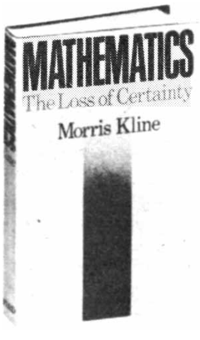
44900-3. ENCYCLOPEDIA OF COMPUTER SCIENCE. *Edited by Ralston.* Monumental 1,500-page volume on every aspect of the computer sciences—software, hardware, languages, programs, systems, mathematics and more. Over 700 charts, tables, graphs and diagrams. *Counts as 3 of your 3 books.* **$60.00**



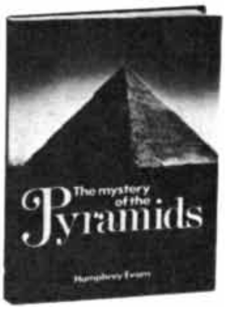
66518. OTHER WORLDS: Space, Superspace and the Quantum Universe. *Paul Davies.* Non-mathematical introduction to the world of subatomic physics and the quantum theory revolution. **$11.95**



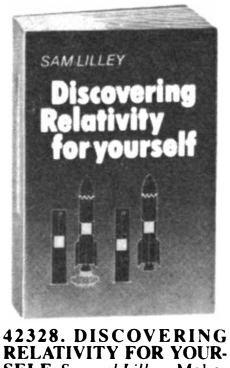
45388. ENTROPY: A New World View. *Rifkin and Howard.* Popular inquiry into the role the Second Law of Thermodynamics plays in energy, economics, politics, health, education, agriculture, urban decay and military activity. **$11.95**



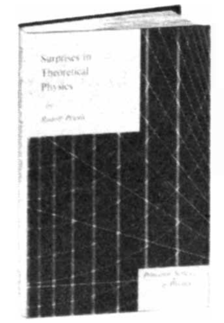
61495. MATHEMATICS: The Loss of Certainty. *Morris Kline.* Traces the history of mathematics' fall from its lofty pedestal. Reveals conflicting concepts and theories. **$19.95**



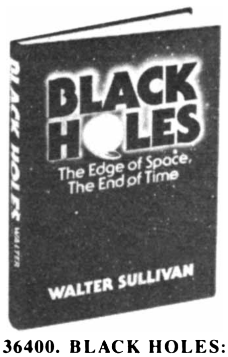
63345. THE MYSTERY OF THE PYRAMIDS. *Humphrey Evans.* A fascinating overview of Egyptian civilization. Covers architecture, statuary, Pharoahs, geographic influence of the Nile and more. Lavishly illustrated. **$19.95**



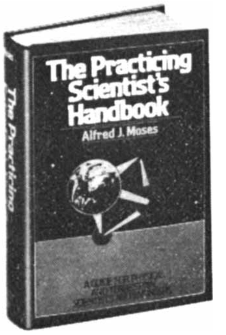
42328. DISCOVERING RELATIVITY FOR YOURSELF. *Samuel Lilley.* Makes Einstein's revolutionary ideas of time, space and motion accessible to everyone. **$19.95**



81477. SURPRISES IN THEORETICAL PHYSICS. *Rudolf Peierls.* Looks at and explains mistakes made by Heisenberg, Einstein and others. "A fascinating book." —J.S. Bell, *Nature.* **$15.00**

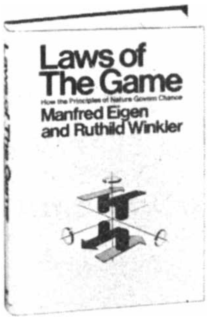


36400. BLACK HOLES: The Edge of Space, the End of Time. *Walter Sullivan.* Bottomless pits into which whole stars can disappear. Concepts discussed include wormholes in space and the end of time. 77 photographs and diagrams. **$17.95**

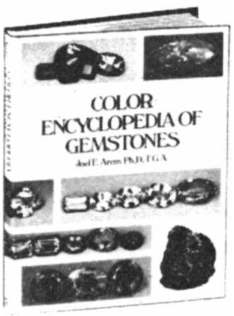


69780-3 THE PRACTICING SCIENTIST'S HANDBOOK. *Alfred J. Moses.* A clear, easy-to-read handbook that provides over 300 figures and tables on the properties of materials ranging from organic compounds through glasses. *Counts as 3 of your 3 books* **$54.50**

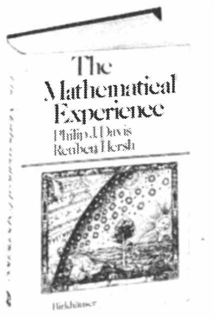
61647-2. McGRAW-HILL ENCYCLOPEDIA OF ENERGY. 6 feature articles on current energy issues; 300 alphabetized articles on energy sources, development and distribution. 838 pages and more than 850 illustrations. *Counts as 2 of your 3 books.* **$34.50**



57922. LAWS OF THE GAME: How the Principles of Nature Govern Chance. *Eigen and Winkler.* A masterpiece of logic which shows how game theory sets the criteria for optimal strategy in many fields. With 60 illustrations including 25 color game boards. **$19.95**



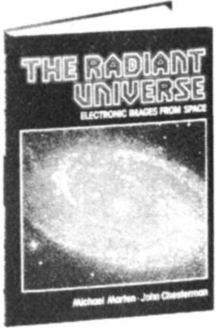
39570-2. COLOR ENCYCLOPEDIA OF GEMSTONES. *Joel E. Arem.* Covers every known variety of gemstones. Photographs. *Counts as 2 of your 3 books.* **$37.50**



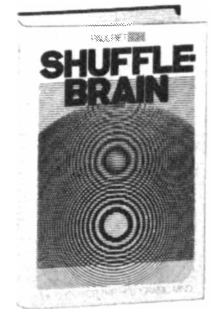
61230. THE MATHEMATICAL EXPERIENCE. *Davis and Hersh.* Captures all the mystery and meaning of mathematics. Filled with charts, diagrams and photographs. 440 pages. **$24.00**

74776. RUBIK'S CUBE/NOTES ON RUBIK'S "MAGIC CUBE." One of the most difficult and popular puzzles *plus* an explanation of how to understand and solve the puzzle. **$15.90**

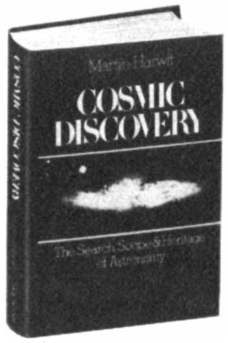
66045. OPTOELECTRONICS. *Robert G. Seippel.* An up-to-date handbook on the devices and systems of fiber optics. With photographs, diagrams and tables. **$21.95**



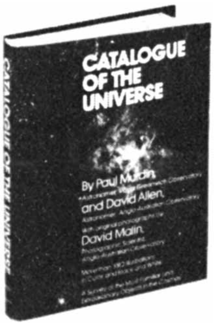
73282. THE RADIANT UNIVERSE: Electronic Images from Space. *Marten and Chesterman.* Introduction to the new tools of visual space science and the use of the computer to solve astronomical mysteries. Over 125 photos. **$17.95**



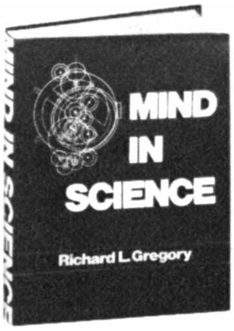
78297. SHUFFLEBRAIN: The Quest for the Hologramic Mind. *Paul Pietsch.* An astonishing report on the provocative hologramic theory of memory. **$10.95**



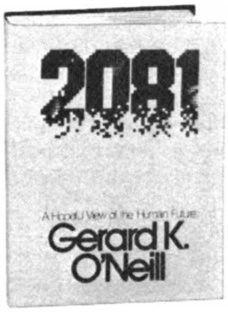
40670. COSMIC DISCOVERY: The Search, Scope and Heritage of Astronomy. *Martin Harwit.* With over 70 illustrations, 512 pages. **$25.00**



37353. CATALOGUE OF THE UNIVERSE: A Survey of the Most Familiar and Extraordinary Objects in the Universe. *Paul Murdin et al.* More than 300 photographs. **$17.95**



62705-2. MIND IN SCIENCE. *Richard L. Gregory.* The evolution of man's understanding of the mind from the earliest myths to the new psychophysics. 627 pages. *Counts as 2 of your 3 books.* **$29.95**



84895. 2081: A Hopeful View of the Human Future. *Gerald O'Neill.* Speculations on the kind of technologies our descendants will be using in such fields as space travel, medicine and communications. **$13.95**

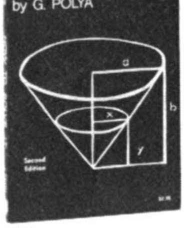
4 Good Reasons to Join

1. **The Finest Books.** Of the hundreds of science books published each year, only the very finest are selected and offered. Moreover, our books are always of equal quality to publishers' editions, *never* economy editions.

2. **Big Savings.** In addition to getting Van Nostrand's Scientific Encyclopedia for only $3.00 or any other 3 books for $1.00 each, you keep saving substantially — up to 30% and occasionally even more. (For example, your total savings as a trial member — including this introductory offer—can easily be over 50%. That's like getting every other book free!)

3. **Bonus Books.** Also, you will immediately become eligible to participate in our Bonus Book Plan, with savings up to 70% off the publisher's prices.

4. **Convenient Service.** At 3-4 week intervals (16 times per year) you will receive the Book Club News, describing the Main Selection and Alternate Selections, together with a dated reply card. If you want the Main Selection, do nothing and it will be sent to you automatically. If you prefer another selection, or no book at all, simply indicate your choice on the card, and return it by the date specified. You will have at least 10 days to decide. If, because of late mail delivery of the News, you should receive a book you do not want, we guarantee return postage.



53854. HOW TO SOLVE IT: A New Aspect of Mathematical Method. Second Edition. *G. Polya.* How you can incorporate heuristics (the study of how problems are solved) into every aspect of your life and work. **$16.00**

More books to choose from...

87112. WHITNEY'S STAR FINDER. Third Edition, Revised for 1982-1985. *Charles A. Whitney.* With a Starlocater Wheel redesigned to include the Southern as well as the Northern Hemisphere. Softbound. **$8.95**

42254. THE FACTS ON FILE DICTIONARY OF PHYSICS. *Daintith and Deeson.* Over 2,000 clear and concise entries. **$14.95**

68715. PHOTOVOLTAICS: Sunlight to Electricity in One Step. *Maycock and Stirewalt.* Everything you want to know about the past, present, and future of solar cell technology. Illustrated. **$19.95**

44780-2. THE ENCYCLOPEDIA OF ANCIENT CIVILIZATIONS. *Edited by Arthur Cotterell.* The lives and cultures of more than thirty ancient civilizations. Over 300 illustrations and 28 maps. *Counts as 2 of your 3 books* **$29.95**

81750. SYNAPSIDA. *John McLoughlin.* Fascinating story of our premammalian ancestors from the first land dwellers to the twilight of the age of dinosaurs. **$14.95**

81435. SUPERPOSITION AND INTERACTION. *Richard Schlegel.* An intense dialogue revolving around two differing philosophical branches of physics: quantum mechanics and special relativity. **$22.50**

70275-2. PRINCIPLES OF REAL ANALYSIS. *Aliprantis and Burkinshaw.* Covers metric spaces, continuous functions, normed and Banach lattices and more. *Counts as 2 of your 3 books* **$27.95**

55001. ILLUSTRATED ENCYCLOPEDIA OF ASTRONOMY AND SPACE. 2nd Ed. **$17.95**

52363. HANDS *John Napier.* Explores the anatomy, function and evolution of this most intriguing human appendage. Outsized. Nearly 60 illustrations. **$13.95**

61944. MEMORY. *Elizabeth Loftus.* Contends that memory is extremely fallible and that people and events can change or even erase it. **$10.95**

62672. MICROELECTRONICS: Processing and Device Design. *Roy A. Colclaser.* Includes pattern generation and photo-engraving, substrate preparation, assembly, packaging and more. **$24.95**

42053-2. DESIGNING A DISTRIBUTED PROCESSING SYSTEM. *Hamish Donaldson.* Includes file strategy, communication networks, and workflow design. *Counts as 2 of your 3 books.* **$34.95**

61540-2. MATRIX COMPUTATION FOR ENGINEERS AND SCIENTISTS. *Alan Jennings.* Practical guide covers the most common matrix methods and shows how they can be applied to the computer. *Counts as 2 of your 3 books.* **$35.75**

If the reply card has been removed, please write to
The Library of Science
Dept. 2-AN6, Riverside, N.J. 08075
to obtain membership information and application

BOOKS

The five kingdoms of life, an atlas of human cultures, animal play, the causes of cancer

by Philip Morrison

FIVE KINGDOMS: AN ILLUSTRATED GUIDE TO THE PHYLA OF LIFE ON EARTH, by Lynn Margulis and Karlene V. Schwartz. W. H. Freeman and Company ($24.95). When Linnaeus published his *System of Nature* he brought boundless classificatory energy to the mapping of three natural kingdoms: the animal, the plant and the mineral. The image has passed into metaphor. In this disciplined and fascinating "catalog of the world's living diversity" the term has a narrow biological meaning: a kingdom is simply the largest category of related living forms. The scheme of the five kingdoms of life used here was developed by the late Robert H. Whittaker, a Cornell ecologist, to reflect appropriately the newer recognition of the profound importance of the host of forms that cannot be regarded as plants or as animals. The oldest of the kingdoms on the earth is the Monera: microbial, usually unicellular forms whose DNA message is not wound onto any proteinaceous chromosomal scrolls. That parts the Monera from all other forms of life; all other cells have chromosomes fit for some more or less complicated ballet of division. The more familiar groupings hold generally multicellular forms: societies of individually larger and more complex cells, all within the three familiar kingdoms, each characterized by a recognizable macroscopic way of life. Fungi are absorbers, Plantae are producers and Animalia are pursuers of nutrient molecules.

The fifth kingdom, necessarily a ragbag since it is defined by exclusion, includes all those forms that have large cells and true nuclei but do not proceed with developmental elaboration as the members of the more familiar kingdoms do. Its variety is wide, although its members are all aquatic, from austerely simple single-cell plankton up to giant kelp. Other authors bent on consistency have carved up to 20 kingdoms out of this domain. Here it is one empire. It has been tagged by a name the Greeks might disown as barbarous: Protoctista. (These Boston biologist authors have their reasons, but the term may not endure.)

This much is mere framework, indispensable but by now familiar to the student of any introductory biology text. What is here is a sparkling string of pictured marvels, examples from about 90 great divisions next under the five kingdoms, each division a broad class of living classes. The category was called a phylum (the Greek for tribe) by Ernst Haeckel a century ago. It includes perhaps 40 phyla that have long been recognized and are widely agreed on, but the authors have added considerably to the list, proposing new phyla to hold many forms of microbial life not widely known. Indeed, such oddities are often half-lost among the bewildering specializations within a study that cannot neglect the botany of the Namibian desert, the sea lilies of the dark Pacific floor, the limestones of the pyramids and the kidney parasites of the octopus.

Naturally microbial life must be characterized more by metabolic style than by visible form; biochemistry is the key. Each phylum is presented by a careful account of its nature and occurrence, amounting to a tentative justification for every category, and a photograph is shown, if possible of the living form, always supplemented by careful annotated diagrams of the structures visible and hidden. A few phyla muster only a handful of known living species; one phylum holds half a million. Arrows on a drawing (a choice among five stock scenes is made) point out the typical habitats of each group. Some phyla are treated more extensively; for our own, the Chordata, the pages display close cousins, a salamander, a sea peach and a swan.

For each kingdom there is offered an illustrated family tree of the descent of the phyla described; necessarily tentative for most of the material, it is full of interest for the reader, particularly because a careful summary of the ideas behind the judgments is given. "Each taxonomy is a theory about the creatures it classifies," we are reminded in a graceful foreword by Stephen Jay Gould. This is not a simple catalogue but a catalogue *raisoné;* we touch not mere variety but a seamless web, the most exquisite and intricate fabric in all the world's weaving. The book is source for innocent and expert alike. Not its conclusions but its factual premises, the diverse organisms we see in these images, make it irrefutable. That is what life has become here on one planet; about how it happened we still have much to learn.

Middle-earth and the tigers of our dreams are fictions less strange than this tapestry. Take *Pelomyxa,* a freshwater cell a fifth of a millimeter long. This little gathered bag has been studied mostly from the muddy bottom of a pond in Oxford, where long ago the discarded carcasses of elephants were thrown by those who prepared exhibits for the University Museum. It has a well-enclosed nucleus but almost none of the complex organelles found in other nucleated cells, no mitochondria, no chromatin, no Golgi bodies, no.... It harbors two distinct types of bacterial symbionts that it seems to feed with its lactic acid waste. They may functionally replace the lacking mitochondria.

Or consider *Trichoplax,* "the simplest of animals." Without tissues or organs, head or tail, it looks like an outsize amoeba, formless and undirected in its motion, except that it keeps topside up. It consists of thousands of cells in two layers; it can divide by mere fission or arise out of an egg, although no fertilization has yet been seen. The beast was discovered in the seawater aquarium of landlocked Graz in Austria a century ago; few have ever seen it who have not searched for it carefully in seawater from Plymouth to the Red Sea.

Here is a tiny spine-studded drifting sphere of the sea, an acantharian, whose radially symmetrical skeleton startlingly consists of crystalline rods of strontium sulfate. Here too are the nitrogen-fixing bacteria; without their enzyme, with an uncommon molybdenum or vanadium atom enclosed, we would all starve. Here are the slime molds; those that form stalks and fruit from a host of aggregating big amoebalike cells; those that thrust up their flowerlike reproductive stages from flat multinucleate masses of protoplasm, mobile only by differential growth; those that send up a similar fruited stalk after the convocation of a vast quorum of tiny bacterial cells, all nucleus-free. Here are the spoon worms, at home from the abyss to tropical shores, which occupy a U-tube burrow a few feet long, ever pumping seawater in a life that is a paradigm of one-way flow.

This gallery begins with the smallest cellular forms, the lipid-walled bacteria (grouped perhaps too simply together), from those that dwell in Yellowstone hot springs to those that are pathogens for a strange pneumonia in human beings and their domesticates. (Viruses, even smaller, are not included; noncellular, they are mere genetic recipes and are not alive.) The pictures close with the lovely purple aster of New England meadows. A thousand genera of organisms are listed at the end in a helpful directory with such common names as exist, each assigned to the right phylum. An indispensable glossary is not forgotten; the nomenclature of these curious forms and habits is a formidable pseudo-Hellenic tongue. There is no other comparable guide to the winding path of organic evolution overall; the work is imaginative and yet circumstantial, a terse visual index to the living library of 10 million species that time has compiled in the script of molecular helixes.

ATLAS OF WORLD CULTURES, by George Peter Murdock. University of Pittsburgh Press ($10.95). This small, dense, personal book by a senior ethnographer is no atlas but an almanac. It tabulates, without a single map or figure but with plenty of numerical tables, the rich diversity not of our biological species, which is securely unitary, but of human cultures, those more swiftly evolving phenomena of an intrinsically social primate, *H. sapiens.* Like any such summary, it is limited to somewhat randomly accumulated evidence. Here that evidence is the entire ethnographic literature: full and careful descriptions of a local culture, detailed enough and yet generalized enough to answer a variety of questions whose coded answers are the substantive component of the book. The author knows that special branch of learning well: "It is probable," he says, "that I have read several times as much ethnographic literature as any other living person." He reads in six languages, those western European tongues that dominate the field, particularly before the mid-20th century, and he has sought out translations from the Russian. Since the spread of modern influence worldwide is steadily reducing the opportunity for ethnographic study of relatively unperturbed cultures, this window seems to allow a good comparative view, although one wonders a little about coverage of the scholarship of China and Japan.

The book divides the world of people into six areas of crudely comparable size: the two New World continents, Oceania, Europe augmented by the circum-Mediterranean fringes of Asia and Africa, and the rest of Asia and of Africa. Professor Murdock has then subdivided these areas into 25 provinces each, using criteria of linguistic similarity and cultural relationship clearly visible to the original reporters. Within each province so found—speaking roughly they are the genera of cultures—he has selected up to six individual cultures, each pinpointed in place and time. These are his species, so to speak: 563 societies, the majority of all those that are well described in the literature he has so carefully examined.

Every societal species is named, numbered and indexed; one or two citations are given for each one as a key to the enormous literature. All the coding decisions are personal; all the data were surveyed in one mind. This provides some kind of sampling stability, even though it has its drawbacks. The literature since 1970 is little covered, although Murdock has tried to avoid blind spots, a task easier for him than for less polyglot scholars. He includes both ancient and modern cultures (the Italians, the Khmers, the English-speaking peoples, the Incas) partly on the basis of historical and folkloristic sources, a practice evidently desirable, if uncommon among anthropologists. Of course, it is starkly skeletal to represent Italian culture, for instance, by just three citations, two on ancient Rome and one on Naples in 1950!

The rich lode of fiction is not represented. Murdock judges that the complexity of Europe and the relative poverty of the literature on societies of northern East Asia and of South America amount to the biggest gaps in what he hopes is a cultural sample representative of the entire world. The codes themselves are interesting. Each culture is assessed for a couple of dozen traits, from the presence or absence of slavery and the mode of marriage to the sex roles in several occupations and terms of kinship. Most of the entries offer more than a single bit of information; subsistence activities are described by five digits coding quantitatively for the relative importance of five ways of making a living, from foraging to agriculture.

So narrow a look at human diversity is bound to put off many; it would appear that Professor Murdock is indeed no favorite of contemporary students of society. Still, as an entry to a treasure house this is an open and inviting little door. Certainly the cross-cultural statistical studies it makes possible need to be understood as tentative at best.

Some leads are worth citing, with all the diffidence implied. It is no surprise that the Kung of the Kalahari Desert are as dependent on foraging for plants and small animals as any other people; it is less expected that in populous southern India there is a forest group, the Chenchu, much like them. It is striking that Africa shows the institution of slavery in some form much more than any other region and that worldwide both pottery and loom weaving are held to be feminine tasks, metalworking masculine and agriculture about neutral. The New World cultures are preponderantly stateless and without social classes; large states circle the Mediterranean, and with them complex social stratification. Here we probably see a mix between space and time. If you seek to know even a little about the sea gypsies (the Bodjau), the Fuegians, the Ibibio, the Todas, the Kapingamarangi or almost 600 other groups, here is an inexpensive and admirable initial guide. (A larger predecessor of this study by the same author has appeared serially in journals and as a book, *Ethnographic Atlas,* in 1967.)

ANIMAL PLAY BEHAVIOR, by Robert Fagen. Oxford University Press ($29.95). The young hippopotamus calf doing a backflip in the warm waters of its lake home, the lion cubs play-wrestling and play-fighting, the play faces of a chimpanzee mother and child and the play-biting of a human partner by a margay kitten are all delightfully seen here in illustrations and closely described. The author, a University of Pennsylvania behavioral scientist out of the Harvard biophysical milieu, has studied animal play in ravens and ponies and in cats domestic and wild. He has plainly lost some of his heart to the Felidae. These carnivores are "creatures of the middle distance" as far as we human beings are concerned. Their nature lies too close for easy objectivity, yet is distant enough to outpace our intuition. Above all, cats are beautiful, particularly in their play.

Let no reader enter this bulky volume to read of such delightful animal behavior. A fourth of the book deals with the natural history of play, tightly organized by group and crowned by an impressive annotated list of all known published accounts of animal play (eagles, rhinoceroses but not aardvarks), apart from the too voluminous work on the three species rhesus, chimpanzee and man. The list is 25 pages long. Of course, it is a useful compilation, but no other part of the work is primarily inductive; most of it is critical, methodological, theoretical, calculational. An early chapter sets out three agreed elements of animal play: play-fighting and its kindred acts, "locomotor-rotational exercise" and diversive exploration with newly examined objects as a base. A sensible view toward definition—let it arise if it can out of our partial conceptions—continues to a discussion of cause and function, close to philosophy.

Professor Fagen regards theoretical foundation as indispensable, and he seeks it in the general attitude of a tentative sociobiology, with emphasis on the adaptive interaction of genes and behavior. That foundation is the study of the costs and benefits of play. Play is costly in time and energy, but it allows the rise of strengths and skills, it is part of an adaptive strategy of social behavior and it patently enters the life history of individual development. Each of these formulations is expanded in a thoughtful chapter, complete with mathematical models based on game theory, recursive relations for gene frequency as affected by various simple cost-benefit interactions and computer simulations along such lines. (Most of the mathematics is safely caged in a series of appendixes.)

A general reader is not much moved by the genetic models, which are patently oversimplified. We cannot do much better now; what is really shown is that the frequency of a play gene can do plenty of odd things, depending on quite subtle changes in the model. Even innovation can be put in, plus cultural transmission. The results are far from simple. It appears that specific scientific questions are hard to answer from the success or failure of such general formulations, absent independent evidence for the precise hypotheses.

One strong statement is supported by all that we see. "Play is a minority phenomenon in nature." It is demonstrable among mammals and birds alone. Reptiles may or may not show it, and for the rest the evidence is negative, if not quite secure. Ants and fishes have complex social behavior but no known play. It would seem that the development of a powerful individually modified system of behavior control—a brain—is the needed condition. The models that stem from the assumption of individual genes for aspects of play may be no more cogent than the same kind of theory to explain why larger animals weigh more. A weightless allele is just not available. It may be that an evolving computer program of any sophistication needs play. In this domain, rather than in behavioral genetics pure, there seems to be a new path to understanding becoming dimly visible. Professor Fagen pays attention to this possibility too, but his own research has struck out more into the world of those unseen but hopelessly self-centered genes. Firm conclusion is surely premature, and it is sensibly avoided.

The volume touches on deep human issues, but it avoids with disdain a grand final "people chapter." That is not to say there are no pages on creativity, on art, on the fine tales of August Kekule and Salvador Luria solving their great problems through play: "unsatisfactory metaphors." The learning is impressive. The initial epigraph is from *The Georgics* (left in the Latin); closing allusions are to Heraclitus and Patrick Henry Pearse. There are nine appendixes and a citation list of some 1,800 items. Professor Fagen is not often at play, although here and there a note of lightness enters. A list includes dragons as one species in which play has been reported, but alas the entry is clearly marked mythical.

THE CAUSES OF CANCER: QUANTITATIVE ESTIMATES OF AVOIDABLE RISKS OF CANCER IN THE UNITED STATES TODAY, by Richard Doll and Richard Peto. Oxford University Press ($16.95). At the request of the Office of Technology Assessment two British epidemiologists here estimate the avoidable risks of cancer in the U.S. today. Theirs is the "black box" approach of the discipline: widespread long-time studies of sizable human populations. "Humans feed themselves, house themselves and arrange their own medical care at no cost to the epidemiologist." Human response to any exposure is exactly what we want, not some overdosed white rat's. How much the experimenter's control is lacking in such spontaneous data is all too clear. It is arguably the worst of methods, but it is all we are likely to have, the only game in town.

The basic ideas are not new; it is detailed sifting of the tables that can persuade. The avoidability of most cancer rests on the fact that worldwide cancers vary in incidence specifically site by site, by one or two orders of magnitude from Iran to Nigeria, or from California to

Japan. The implication is clear: we might come to face only the minimal risk at every site, although that can certainly not happen by chance. It is pretty clear there is no built-in overall rate that simply shifts sites from country to country. The statistics argue against it, and specific exposures that much increase a given type of cancer do not reduce the risk of other types. Nor is the main effect genetic: migrants change their cancers with their residence. It is somehow embedded in the way we live. Of course, no meaningful numbers can be presented without close attention to age distributions; cancer increases at most sites along with age.

The study explicitly considers a dozen avoidable causes in the practical sense, from tobacco and alcohol through diet, food additives, occupation, pollution and on to unknown agencies. Each is examined at length against the abundant but still far from adequate literature. The conclusions are well supported, although these workers do not fail to warn against complacency. The numbers must be kept under scrutiny; there is a great deal we do not know, and here what we do not know can hurt us.

This report appeared in a technical journal in the middle of 1981; it confirms the only two earlier studies (one of them for Britain) the authors know of. Although several dozen types of cancer are reported, more than half of all U.S. cancer deaths result from cancer of the lung, the breast or the colon. That is the central social fact. The chief causes of cancer in the U.S. today are also few: tobacco and food. The tobacco factor is ascribed to smoking, particularly of tarry cigarettes. Each year there die from lung cancer more Americans who would have survived if no Americans had ever smoked than the nation lost in all the years of the war in Vietnam plus the annual deaths in automobile accidents.

That heavy risk is strongly increased by exposure to cigarette smoke in the first decade of adult life. (Young smokers—it is particularly sad that the habit has grown in the 1970's most rapidly among young women—are at serious risk; persuade them to stop!) Details matter; Americans leave longer stubs unburned than the British, as one of the authors can testify from the experience of his early teens, when he scavenged "the Southampton docks for smokable stubs discarded by sailors of various nations." Matters are therefore improving under our new knowledge (which we owe to Sir Richard Doll more than to any other person) but too slowly.

Diet is in principle even more important and almost as modifiable. Here, however, we cannot act with much effect, because we do not know what in the U.S. diet causes the relatively high rate of cancers at half a dozen sites, such as the colon, the breast, the gall bladder and the pancreas. (Stomach cancers are on the decline.) Any number of tantalizing hints are at hand. Underfed mice are sleek, active and long-lived; they have about half the rate of cancers at the major sites. Breast cancer is correlated with consumption of milk, colon cancers with meat. Colorectal cancers are inversely related to a healthy population of certain intestinal bacteria that feed on particular components of dietary fiber, the pentose polymers. The powerful carcinogens our body makes from nitrites in the diet do not show any clear effect, although a test meal of both bacon and spinach resulted in the production of plenty of the hazardous compounds.

There is much more to argue about, but the outcome is that we have good reason to blame the stuff we eat (and how much of it and when) for a third of the avoidable U.S. cancer deaths. But it is not the food additives, not the butter yellow, not the saccharin, not the nitrites, not the antioxidants. Additives can be estimated to contribute "a token

proportion" below 1 percent. It is not hair dyes and not the pill. Nor is it the carcinogens poured out in increasing amounts by industry and transport; the total effect of all pollution cannot be much above 5 percent for air, water or food. There the most uncertain single component appears to be from the halogenated impurities in drinking water. Industrial output as a whole, medical practices, occupational hazards, ionizing and solar radiation, infection by viruses known and unknown may together contribute 20 percent of the avoidable risk. Another 5 percent takes ample account of the indirect effects of infection, such as that of hepatitis on liver cancer.

These estimates are rough, some almost speculative, and Doll and Peto allow for a wide range of error. Still, it remains their conclusion that smoking and some unknown attributes of our diet are the dominant entryways of avoidable fatal cancers, two-thirds of them all. Cancer is not increasing in America, save for lung cancer, clearly the result of increased cigarette use, and for an uncommon form of skin cancer, plausibly (but speculatively) related to increased exposure of bare skin to the sun by those who are not already tanned by steady outdoor work. That overall constancy is a gross argument against attributing much new risk to the flood of novel materials with which our smokestacks and our workplaces assault us; detailed examination confirms that view.

Doll and Peto single out several Government, industrial and private publications that have disagreed with them over the past few years; their critique of those publications seems cogent, both as a matter of explicit argument and as gauged by the concurrence of other studies. They go so far as to characterize one Federal report as written for "political rather than scientific purposes" and another private set of studies as done by "people who wish to emphasize the importance of occupational factors." Overestimates by a factor of 10 follow, they say, from reliance not on observed fatalities but on broad extrapolations from studies of incidence under acute exposure. In one instance the nasal-sinus cancers projected from moderate occupational exposure to nickel would have caused 10 times as many deaths as the national total that was reported.

This is of course no brief for greed or carelessness in the shop, the causes of numerous preventable cruel poisonings of workers and other citizens. It is rather to offer perspective; our experience has not been as bad as our fears. But if we are to progress, we need to know more as well as to act more wisely. These authors recommend a long-term epidemiological investment. Perhaps we can persuade a few hundred thousand healthy people to contribute blood and fecal samples to a bank, along with detailed answers to an open-ended questionnaire on food, smoking, drinking, work, residence, reproductive history, drugs and the like. Epidemiologists could study a set of such samples for speculative inquiry once death from a particular cause took any class of donors. We have a steadily increasing ability to find significant factors in stored blood. Naturally a national index of causes of death would be needed as well; such an index is not available today across the barriers of state record keeping and expense and of our increasing concern with privacy. There are some five or 10 years of life expectancy to be gained by civil victory over avoidable cancers; if we can keep the nuclear peace, we should one day be able to claim that worthwhile trophy for public health.

SCIENCE IN TRADITIONAL CHINA: A COMPARATIVE PERSPECTIVE, by Joseph Needham. Harvard University Press ($12.50). It is 43 years since Joseph Needham, a Cambridge biochemist already distinguished by his literary and philosophical erudition, was inspired by Chinese visitors to his laboratory to learn their "marvellous script or else burst." He learned it all right, and he went to wartime Chungking as Scientific Counsellor to the British Embassy. "Four years in China sealed my fate." What all the world now admires is our share of his fate: the wealth of learning and insight that now occupies eight thick books, his magnificent continuing treatise on the history of science and technology in China. The volumes sail on: they will have traversed the spectrum of all the sciences by the time some 20 of them are on the shelf, at the end of this decade. "I have a sporting chance of seeing the voyage pretty nearly into port," he wrote in this book, a hardy philosopher at 81, about a year ago. After three decades his effort, nearly always joined by one or two colleagues, consists now of a small scholarly industry; some 20 people work in it worldwide. No imaginative person of scientific interests can open any of the volumes for the first time without a sense of discovery and delight. They carry a Western reader into another world of science, manifestly akin to our own and just as manifestly different.

Now, it is no light thing to master those big volumes, to afford them or even to take them down from the library shelf. This small book, a set of five lectures given at the Chinese University of Hong Kong, is a satisfying taste from the kitchen spoon of a great chef as he prepares an august banquet. It is 100 admirable pages, an honest sample of Needham's comparative method that brightens the shadows in our own history with light reflecting from the unexpected Chinese scene. Much of what is here is not yet available full-rigged. An introductory chapter with a personal air and a bent for clockwork, followed by accounts of gunpowder, of the alchemical search for longevity, of acupuncture and moxibustion, and a view of time and change in Chinese thought, the real cycles of Cathay, make this little book a fine day's sail for readers who cannot embark on the long voyage.

The gunpowder chapter is a marvel. There is a Taoist book of the ninth century that warns alchemists against the famous mixture, lest it singe their beard and burn down the lab. Long before that the Chinese were great smoke producers; they fumigated in peace and war by the fourth century. It was villainous saltpeter that was the key to the real stuff. By about A.D. 600 that organically derived encrustation of soil was well recognized. A strange recipe (of course cited here) recommended, for preparation of a rather inert substance, a process that must have turned out to be rather exciting, the first deflagrating mixture in all history. By about A.D. 850 the alchemists were clearly warned what not to do with honey, saltpeter and sulfur. They did it anyway.

Gunpowder is the "fire chemical," *huo yao* in Chinese. A detailed chart maps its progress: first a fuse, then a fire grenade, then a flamethrower. That same fire lance was still in service in this century against wood junks among the pirates of the China Sea. (It is visible here being wielded by a demon in a painting 1,000 years old.) The fire lance was a cylinder—what else?—of bamboo. All gun barrels grew from that stalk of fire. The true propellant mixture—with plenty of oxidant, long evolved beyond the Greek-fire proto-napalm for which fuses were first made—powered rockets for both civil and military purposes by the 1200's. Within a century the true gun and the true cannon were familiars of Chinese battle; every such early bombard of bronze or iron the Chinese archaeologists dig up bears a date earlier than anything of the kind in Europe. The dated relics go back to about the time of the first European description of the same weapon, in a famous Oxford manuscript of 1327. Europe and Islam had learned the art and had become disciples all too apt. If you want to learn how cannon fired against castles remade European polity and yet had little effect in China, read the book.

Those big volumes are luxurious throughout; their magnificent indexes and lists of references are not found in this economy product, and we miss them sorely. But Dr. Needham's own magisterial way with a footnote is fully suggested.

STATFJORD "B" OIL PLATFORM, the heaviest object ever moved by man, is shown being towed out of Stavangerfjord in Norway last August. The structure, which had a displacement weight of 899,000 tons while it was under tow, was brought to rest in 472 feet of water 100 miles off the Norwegian coast in the Statfjord oil field. It is designed to resist North Sea storms by virtue of its immense weight. Statfjord *B* provides living quarters for a crew of 200 in a core structure eight stories high with two seven-story wings. More than 11,000 man-years went into the $1.8-billion platform, which is being commissioned this month. Its rated output is 150,000 barrels of oil per day.

Established 1845 SCIENTIFIC AMERICAN April 1982 Volume 246 Number 4

Advanced Offshore Oil Platforms

Creating structures that can withstand 100-foot waves in waters 600 feet deep calls for audacious engineering and heavy capital expense. Here four innovative offshore platforms are described

by Fred S. Ellers

Over the past dozen years offshore oil platforms have increased an order of magnitude in their size, technical complexity and cost. Today about 12.5 million barrels per day, some 22 percent of the world's oil, first reaches the surface at wellheads on structures surrounded by water, many of them out of sight of land. Some of the earliest platforms were erected in fairly shallow waters, such as those of the Gulf of Mexico, or in lakes, such as Lake Maracaibo in Venezuela, where the technical problems were modest. With the discovery of the great oil fields under the North Sea, between Britain and the Continent, however, platforms have had to be anchored on sea floors 500 feet below the surface and designed to withstand the impact of waves 100 feet high.

Last year one of these great platforms, with an overall height of 890 feet and a total displacement of 899,000 tons, became the heaviest object ever moved by man. Named the Statfjord *B* platform, it rests in water 472 feet deep 100 miles off the coast of Norway. Later this year the first of 32 wells will be drilled from the structure to tap a reservoir expected to yield 150,000 barrels of oil per day. It is estimated that the finished platform will have cost its owners (principally Statoil, the Norwegian national oil company) $1.8 billion. To help build such immense structures semisubmersible derrick barges have been developed that are able to lift 5,000 tons, the weight of two World War II destroyers, nearly 200 feet above the water and to do so in the open ocean.

As examples of the innovative technology that is making it possible to erect platforms in hostile marine environments in even deeper water I shall describe the Statfjord *B* platform in some detail, along with the design of three other types of structure that are extending the ability of the oil industry to tap new submarine resources. The three other platforms I shall describe are the Magnus platform in the British sector of the North Sea, the Hutton platform, also in the British sector, and the Block 280 platform in the Gulf of Mexico. The Magnus platform, the most conventional of the four, will rest on four steel columns held on the sea floor by piles. The buoyant Hutton platform will be held down by vertical steel tubes at each corner, which will remain in tension under all weather and operating conditions. The Block 280 platform can be described as a guyed tower: a slender steel frame held upright in water 1,000 feet deep by a spokelike array of 20 anchor cables. Statfjord *B* will be operational this month; the other three platforms are scheduled for installation over the next two years.

The search for offshore oil and gas in deep water at the four platform sites was first undertaken with temporary floating drilling rigs: either drilling vessels or semisubmersible drilling rigs. After economically significant reservoirs were found oil companies were faced with the problem of designing permanent platforms to optimize the recovery of oil and gas over the economic lifetime of the reserve, in many instances 20 years or more. From such permanent structures wells can be drilled at various angles to the vertical to tap the reservoir over an underground area of nearly four square miles. Depending on the tightness of the oil-bearing rocks, more than 30 wells can be drilled for the production of oil and gas from a single platform. Additional wells can also be drilled to reinject gas associated with oil production and to inject water in order to maintain the pressure of the reservoir.

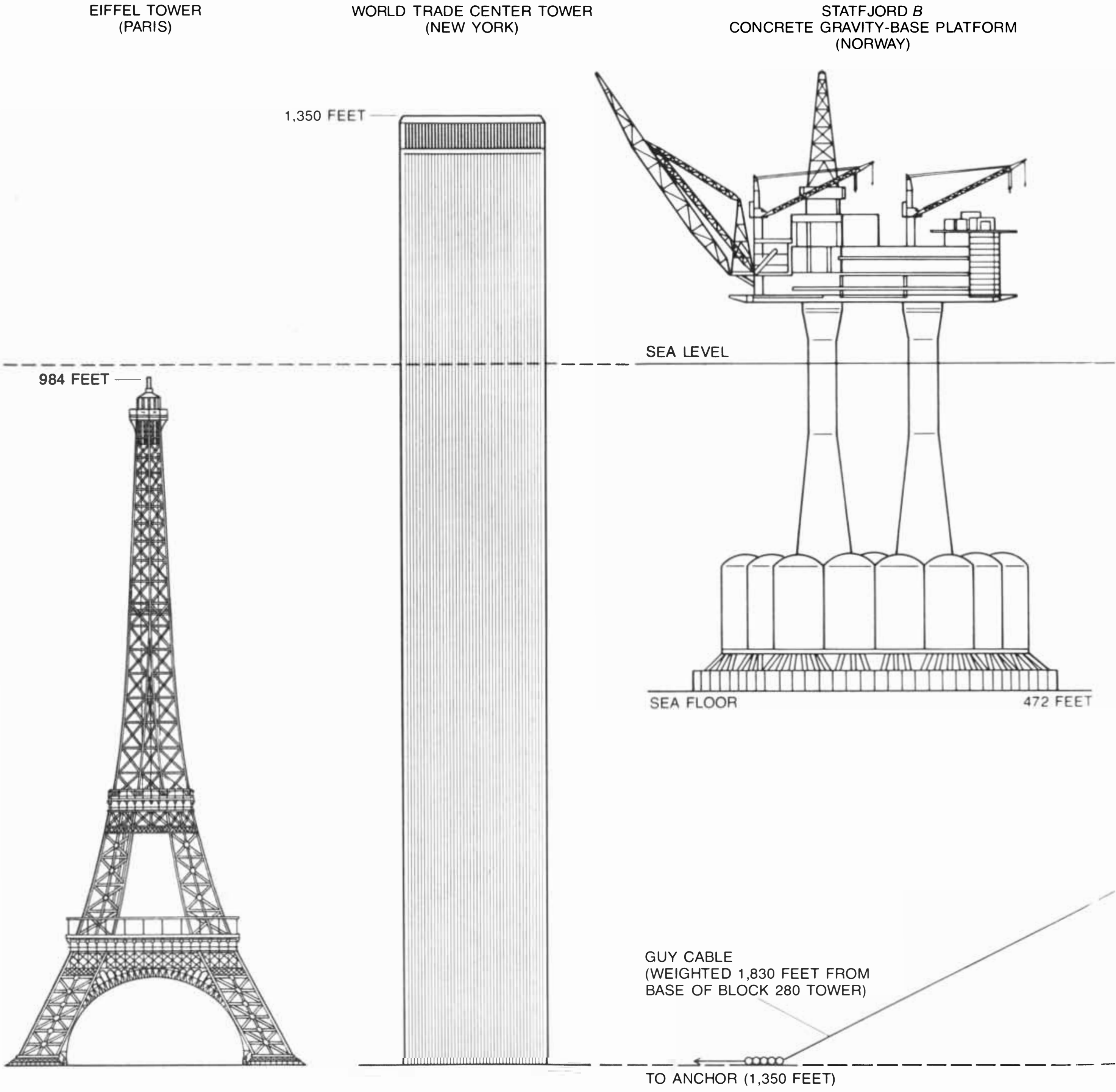
Offshore platforms consist broadly of two components: (1) the drilling and operating facilities, often identified as the topsides, and (2) the supporting structure and its foundations. Included in the topside plant are drilling rigs, oil- and gas-processing equipment, transportation pumps, and utilities and living quarters for as many as 300 workers. All the major platforms now have a landing pad for helicopters. After the oil is processed it is either pumped directly ashore through a submarine pipeline or stored until it can be transferred to tankers. In the latter case sufficient storage is usually provided for the tankers to be loaded intermittently without the need to curtail the flow from the wells.

Topside facilities for the four platforms I have chosen to describe weigh from 16,000 to 40,000 tons and have a total floor space of from 75,000 to 400,000 square feet, much of it dedicated to crew quarters. These massive superstructures must be positioned above the crest of the highest wave expected to come once in 100 years. For the North Sea this calls for positioning the bottom of the lowest deck at least 80 feet above mean water level.

The second component of an offshore platform, and the one that has elicited the most ingenuity, is the supporting structure. Such structures must secure the topside facilities against the assault of winds, waves, currents and in some instances seismic disturbances. The main environmental forces that need to be taken into account are in almost every case those that result from wave impact. (A notable exception is structures that must be built along the West Coast of the U.S., where seismic forces are the major concern.) In establishing the design two wave conditions are critical. The first is the maximum wave that might come as a single event during the lifetime of the project. This predicted wave determines the maximum strength of the structure. The second condition is the cumulative effect of several million waves per year whose period happens to match the fundamental oscillation frequency of the platform itself. Although such resonant forces are individually small, they are amplified dynamically by the structure and therefore determine the fatigue lifetimes that must be designed into the platform.

Finally, the supporting structure must have a secure foundation. In soft sea-bottom soils it may be necessary to drive piles more than 500 feet into the sea floor before adequate support is achieved. The overall length of the pile, including the pile follower between the pile and the pile driver above water, can therefore exceed 1,000 feet and call for handling hundreds of tons of material in one continuous assembly.

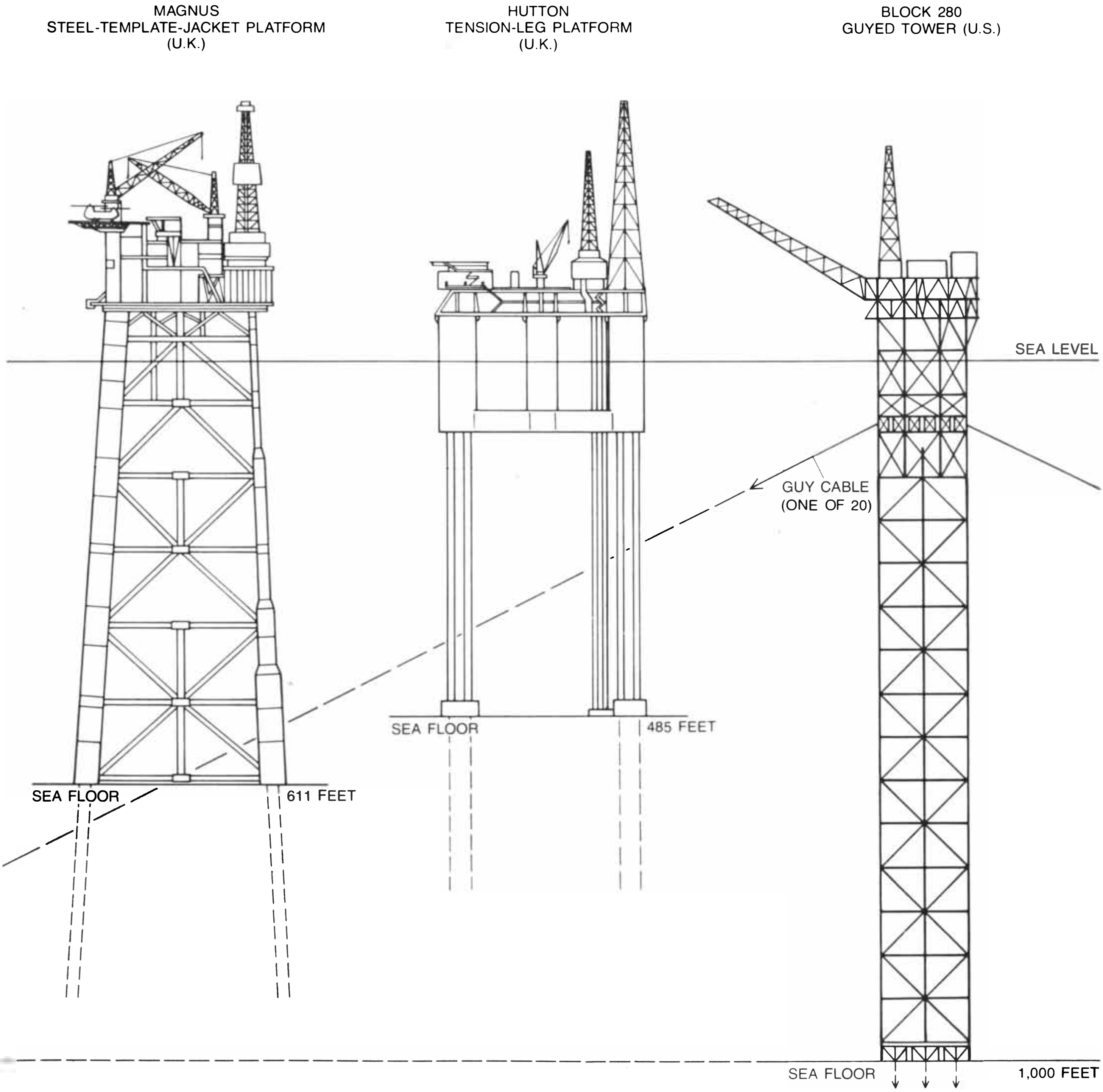
The first fixed structures for drilling and production out of sight of land were installed in 1947 off the coast of Louisiana. These early platforms con-



FOUR OFFSHORE OIL PLATFORMS are compared schematically with two onshore structures of comparable height: the Eiffel Tower, completed in 1889, and one of the towers of the World Trade Center in New York. From the viewpoint of environmental conditions the comparison is highly inadequate. Onshore buildings are subjected to lateral forces no greater than heavy winds. The offshore towers must take the impact of waves as much as 100 feet high. Three of the offshore platforms are designed for neighboring oil fields in the North Sea. The Statfjord *B* platform rests on four massive concrete columns with storage tanks at the base. Since it depends solely on its own mass for stability, it is described as a gravity-base platform. More than 20 such gravity-base structures have been built since 1970. The base of the Magnus platform, known as a steel-template jacket, is the heaviest structure of its type yet fabricated: it will weigh 41,000 tons. Two of the platform's four legs will incorporate flotation chambers, so that later this year the structure can be floated to its site in the United Kingdom sector of the North Sea, about 250 miles northeast of Scotland and about 25 miles north of Statfjord *B*.

sisted of a series of steel frames that were fabricated on shore, transported to the site by flat barges and set in place by derrick barges. They were then pinned to the sea-bottom soil with piles driven through steel sleeves that were part of their structure. Structures pinned to the bottom in this way became known as steel-template jackets.

From 1947 until the mid-1970's virtually all offshore oil platforms incorporated steel-template jackets. As construction companies gained experience they were able to install platforms in deeper water simply by building larger jackets. These structures were fabricated in dock areas accessible to the launch barges. Before the development of the North Sea oil fields got under way, by far the largest concentration of jacket-fabrication yards was along the U.S. Gulf Coast. After each section of the jacket was assembled as a flat panel at ground level it was raised to a vertical position with long-boom crawler cranes that "walked" in side by side as it was being raised. The panels were then welded into



The tower will then be anchored to the bottom by piles. The Hutton structure, known as a tension-leg platform, will consist of a buoyant hull tethered to the sea floor by slender steel tubes at its four corners. The first of its type, the platform will be towed to the Hutton field in the U.K. sector of the North Sea, about 15 miles southwest of Statfjord *B*, next year. The fourth of the platforms, also the first of its type, is the Block 280 "guyed tower." It is designed for service in the Gulf of Mexico in water 1,000 feet deep. It will be pinned to the sea floor by a spokelike array of 20 steel cables, each one more than 3,000 feet long. The tower and its guys will weigh 43,000 tons, slightly more than the Magnus steel-template jacket and more than four times as much as the Eiffel Tower. Of the four platforms the Magnus will be the most expensive, with an estimated cost of about $2.6 billion, or two and a half times as much as the cost of the entire World Trade Center complex, consisting of the two 110-story towers and four smaller buildings. Each of the World Trade Center towers has a dead weight of 325,000 tons, or only slightly more than a third of the displacement of the Statfjord *B* platform when it was being towed.

a single structure, skidded onto a special launch barge, transported to the ocean site and launched over a pair of rocker arms. The structure was made buoyant enough to float horizontally with its top panel at the surface. By controlled flooding it was rotated to an upright position, landed on the sea floor and fastened with piling.

The steel-template jacket launched by barge has many advantages in areas that have shallow-water coastlines, because even the largest launch barges have a modest draft. With careful sea fastening, which calls for welding the jacket to the deck of the barge, jackets much larger than their barges were transported worldwide. Moreover, the most economic sites could be chosen for the fabrication of the jacket. It could then be delivered almost anywhere.

One obvious limitation on this method of platform construction was the availability of barges big enough to handle the high overhanging deck loads when steel-template jackets had to be transported across the open sea. The problem did not become pressing, however, until large oil reservoirs were discovered in the North Sea. By 1972 giant oil fields with recoverable reserves exceeding 500 million barrels of crude oil were being discovered in the Norwegian and British sectors. The initial reaction of the industry was simply to extrapolate practices developed in the Gulf of Mexico. It soon became apparent there were fundamental differences between the two areas.

Platform technology then took two paths, a divergence that continues. In designing platforms for North American conditions one can assume extended periods of calm for fastening the platform to the sea floor and erecting the topsides. Moreover, the platforms can be of modest size, both because the oil reservoirs are fairly small and because until recently the price of crude did not justify large single investments. In designing platforms for the giant North Sea reservoirs, which lie under waters where waves measuring 100 feet from crest to trough are common, major innovations were needed.

Hence one of the biggest differences between the North American and the North Sea platform technologies arises from environmental conditions. In the North Sea as many as 300 workers can be on a single platform with no safe means of departure once a substantial storm has developed. Indeed, high winds and seas are so common that all the platforms in the area are designed with the expectation that their crews will ride out the severest storms. In the Gulf of Mexico, in contrast, hurricanes are carefully tracked and offshore crews are removed before the local conditions become hazardous.

Some 150,000 people are currently employed in the design, construction and management of offshore platforms. This global effort is guiding annual expenditures of more than $15 billion, and nearly all estimates of future growth exceed 20 percent per year. The offshore industry draws not only on such traditional specialties as civil, mechanical, chemical and electrical engineering and naval architecture but also on such newer ones as hydroacoustics, marine geology, fracture mechanics and nonlinear inelastic structural analysis. One measure of this technical diversity is the scope of the annual Offshore Technology Conference, which is sponsored by 11 major international engineering and scientific societies with a membership of more than 400,000 professionals. Some 85,000 attend the annual conferences. Within the past decade several thousand technical papers have been published on topics related to offshore oil platforms.

Three of the four platforms I shall describe will be installed in the North Sea almost within sight of one another. Each platform represents a radically different design concept, yet all have similar duty requirements. The environmental conditions are nearly identical except for water depth and sea-bottom soils. Broadly speaking, any of the three designs might have been adapted to meet the special requirements of any of the three sites. The particular design in each case was dictated in large measure by an initial decision on how the platform was to be fabricated, transported to the site and anchored to the sea floor. I shall therefore discuss the three platforms primarily from the viewpoint of these underlying decisions.

The first of the three projects is the Magnus platform now being fabricated for the British sector of the North Sea by the British Petroleum Company Ltd. It is the heaviest steel-template-jacket platform yet conceived and is therefore an extension of experience gained in this traditional offshore technology. It will also be by far the most expensive of the four projects, with an estimated cost of $2.6 billion.

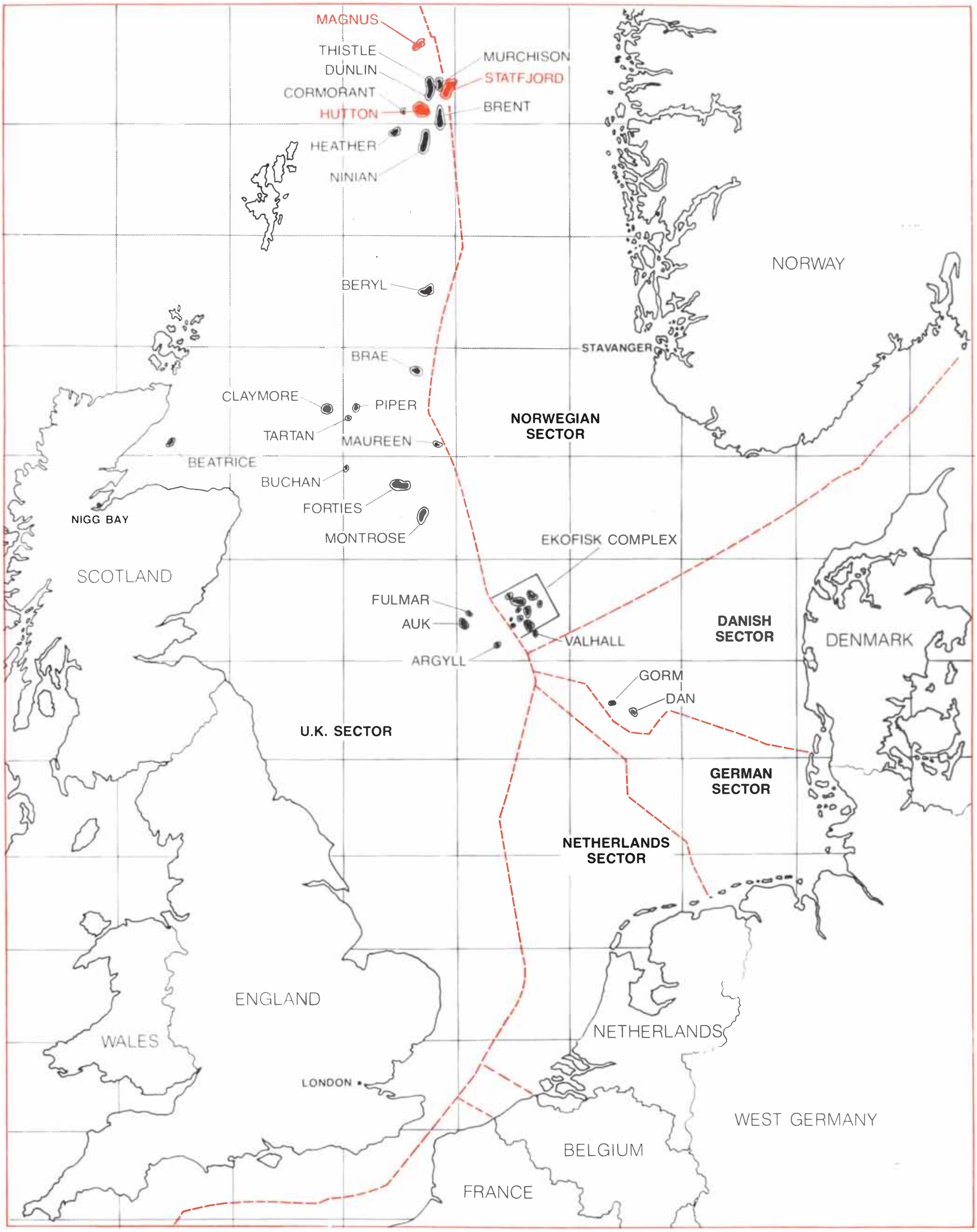
The weight of the Magnus steel-template jacket, 41,000 tons, exceeds the capacity of current launch barges, and so it has been designed with built-in flotation. Two of the corner legs are much larger than the other two. Buoyancy compartments built into the oversize legs will enable the Magnus jacket to float on its side with a modest draft and with most of the frame above water. The jacket is being fabricated on its side in a special dry dock at Nigg Bay in Scotland, a facility large enough for the side panels to be tilted up and welded together in the usual manner. After the Magnus jacket is completed the dry dock will be flooded and the completed jacket will be towed to the Magnus field by an array of tugs. At the installation site, 240 miles northeast of the tip of Scotland and the most exposed location yet scheduled for offshore production, the buoyant legs of the jacket will be flooded, thereby rotating the structure to an upright position. The jacket will then be landed on a "mud mat" (a specially constructed bottom) that will control its penetration into the sea floor.

After the platform is landed it will be leveled and nine primary piles will be driven at each corner. The piles consist of sections of steel tubing welded together. Pile guides are located at regular intervals on each corner column of the jacket so that the piles can be driven by a hammer at the surface rather than one under water. When a pile is driven from the surface, there must be a pile follower between it and the hammer, and the follower must be added to as the primary pile is driven into the bottom 300 feet or so. As the piles are driven they pass through packing glands that expand and seal the space between the pile and its jacket sleeve once the pile has reached its full depth. Cement grout is then pumped into the space, completing the structural connection.

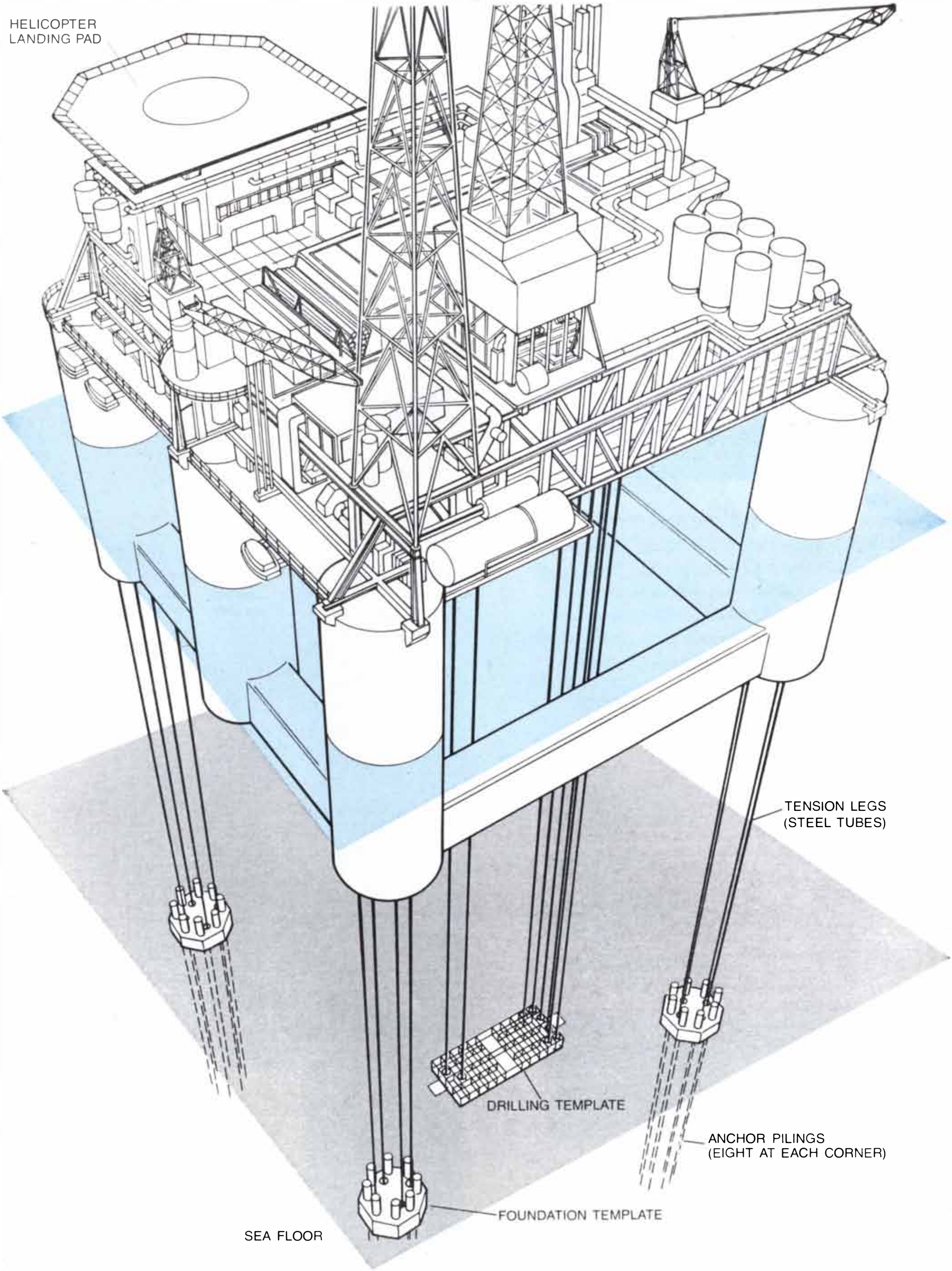
The topside facilities needed for drilling wells, for handling the crude oil as it is brought to the surface and for housing a work crew of 200 are being assembled in the form of 14 modules. The modules will be barged to the site and probably set in place by one of several North Sea semisubmersible derrick barges with a capacity of 3,000 tons. Some of the modules will weigh more than 2,000 tons and must be landed exactly in position 150 feet above the water. Starting with the conceptual phase and continuing through the project schedule, rigid controls on weight and size are needed to ensure that the modules are emplaced safely and quickly.

One might assume that lifting a 2,000-ton module with a derrick whose rated capacity is 3,000 tons leaves capacity to spare, but such is not the case. Derricks and cranes are rated with their boom in a high position, and the boom is not in such a position when modules are being set in the center of the topside deck. Moreover, generous allowance must be made for unscheduled changes in the declared module weight and for various contingencies that may jeopardize the lifting operation. Since the topside structures must be emplaced in reasonably calm seas, the Magnus operation was scheduled for a specific weather "window" some four years in advance. That scheduling is an absolute constraint on the entire project.

After the modules are set in place millions of offshore man-hours are still required for hookup and commissioning before drilling can start. The manpower needed for such activities far exceeds the capacity of the living quarters being put on the platform for the operating and maintenance staff of 200. Supple-



MAJOR OIL FIELDS IN THE NORTH SEA lie in sectors assigned to two countries: the U.K. and Norway. Gas was discovered in the U.K. sector as early as 1965. The first major oil reservoir was discovered four years later at Ekofisk in the Norwegian sector. By 1972 giant oil fields with recoverable reserves of more than 500 million barrels had been identified in both the Norwegian and the U.K. sectors. Denmark also has two small fields. Oil is currently flowing from 21 of the 27 fields named on this map at a rate of more than 2.4 million barrels per day. The three fields whose platform designs are depicted in the illustration on pages 40 and 41 are within less than 25 miles of one another near 61 degrees north latitude on the boundary between the U.K. and the Norwegian sectors. Of the three the Statfjord field was the first to be brought into production (in November, 1979) from a smaller prototype of the Statfjord *B* platform. The platform for the Magnus field is scheduled to be producing by next year. The Hutton tension-leg platform should be producing by 1984.



HUTTON TENSION-LEG PLATFORM illustrates the complexity common to all large offshore oil structures. The topside facilities rest on a buoyant hull designed to yield with the waves. The hull is held down by four groups of highly tensioned tubular-steel tethers anchored to the sea floor by preset foundations at each corner. The tethers pull the hull down so that they will never go slack even in the trough of the maximum expected wave. After the platform is in place wells will be drilled through conductors that will guide the drill pipe through a drilling template on the sea floor. In the heaviest seas the platform may swing as much as 79 feet from the vertical but will at the same time remain level. The Hutton platform is being developed by Conoco, a division of E. I. du Pont de Nemours & Company.

mental housing will be supplied by one or more "flotels" moored alongside the platform. Flotels are generally created by fitting out semisubmersible drilling-rig hulls with living quarters, shops and related facilities. They will be moored close enough to the Magnus platform for it to be reached by gangway.

After the Magnus hookup phase is essentially complete 15 wells will be drilled from the platform. By drilling wells at angles of as much as 55 degrees from the vertical it will be possible to tap parts of the reservoir that lie almost two miles from the platform. Before the platform is in place six vertical wells will be drilled into the underground oil pool at points too far from the platform site to be reached by directional drilling. These wells will be hooked up to the platform by submarine pipelines. Magnus will be in full production in 1985, about 12 years after the discovery of the oil-field. It may well represent the ultimate development of the steel-jacket-platform technology. For many years thereafter it may also remain the most costly offshore oil platform ever built. The capital cost per daily barrel of production will probably exceed $20,000, or as much as 40 times the capital cost for onshore production facilities in the Middle East.

The second offshore structure I shall describe, Statfjord *B*, is the most advanced concrete platform yet constructed. More than 20 such "gravity base" structures have been built by Norwegian, British and French contractors in the past dozen years. The engineering manager of the current project is the Mobil Oil Corporation, which owns a minority share in the giant Statfjord oil field. Some engineers who were originally skeptical of this new technology now concede that the concrete platform has demonstrated important virtues for North Sea conditions.

The base of Statfjord *B* consists of 24 cells in a honeycomb configuration. Each cell is 66 feet across and has walls three feet thick. Construction of the base was begun in a large dry dock at Stavanger, a port on the Stavangerfjord in Norway. When the foundation cells were advanced to the point where they were both buoyant and structurally competent, the dry-dock area was flooded and the partially completed base was moved to sheltered deep water. There the job was completed with the aid of a floating concrete plant. Twenty of the cells were domed over; four were slipformed upward to create the platform columns. As the construction proceeded the platform was gradually submerged in order to keep the floating concrete plant at the most efficient level with respect to the work being done.

The domed cells at the base of Statfjord *B* will serve for the storage of crude oil and diesel fuel. Access to the cells is provided by pumps and piping housed in one of the four columns. Two of the other columns will serve as the conductors through which the wells will be drilled; the fourth column will be given over to various nonhazardous functions. All the columns except the one set aside for pumps and piping will be flooded after installation. Gravity-base structures are designed to remain upright during transit offshore. This makes it possible to install the topside facilities in sheltered waters before the platform is towed to sea.

While the concrete base of Statfjord *B* was still under construction the topside facilities, ultimately weighing 41,000 tons, were being assembled at a nearby shipyard. The various components were built on skids and integrated into a massive steel structure that eventually would be placed on the concrete base. At this stage the supporting frame, weighing 7,000 tons, rested on four mockup columns set in shallow water adjacent to the shipyard. Subsequently drilling modules were installed on top of the frame, and living quarters (eight stories high to accommodate a crew of 200) were set up on the side of the platform farthest from hazardous areas. Further protection was afforded by erecting fire and blast walls between the living quarters and the production facilities. As a result Statfjord *B* is an exceptionally safe installation.

After the topsides of Statfjord *B* were finished early last year four heavily ballasted North Sea barges were moved under the support frame. When the ballast was removed, the barges lifted the entire topside installation as a unit. The assembly was then towed to a deep-water site in Stavangerfjord to which the base had been towed and where it had been submerged until the columns were the same height above the water as the mockup columns. The mating of the two structures was achieved by maneuvering the topsides into position over the four columns of the base. Ballast was then removed from the base until it could bear the weight of its 41,000-ton burden. The two major components were connected with post-tensioned rods and grouted into a monolithic structure.

In August of last year five of the world's largest tugs towed the 899,000-ton Statfjord *B* 230 miles at a speed of 1.7 knots, hugging the Norwegian coast as long as possible, to its final destination in the North Sea. There the platform was landed on a preleveled site 472 feet below the surface and installation was completed without relying on piling. As the weight of the great structure was transferred to the sea bottom it released water in the underlying soil. This free water was then drained by shallow wells drilled from the bottom of the utility shaft. Such drainage also helps to ensure that the soil will not liquefy when storm waves cause the tower to heel slightly and increase the loading on the perimeter of the foundation.

It is clear that a platform such as Statfjord *B* could only be built in certain parts of the world. Without sheltered deep water it would not have been economically feasible to construct such a massive concrete gravity base. It was the juxtaposition of the Norwegian fjords and an industrial center with highly developed skills in the design and fabrication of concrete marine structures that made the project possible.

The concrete base was designed to accept and transport the completed topsides as a unit, which was possible only if the base could remain in an upright position. The sea-bottom soils were such that they could sustain the weight of the platform without deep piling. Having space in the base cells for the storage of oil from the wells also provided surge capacity. Hence the oil could be shipped off by tanker, which made it unnecessary to lay a pipeline to the Norwegian coast. If any of these conditions had not obtained, Statfjord *B* might well have had a different configuration.

The third project of advanced design is the tension-leg platform (TLP). Whereas steel-template platforms and concrete gravity-base platforms are regarded as fixed structures, the tension-leg platform is a compliant structure: it has the ability to yield to the waves in a controlled manner. Tension-leg platforms have two main structural elements: a floating hull similar to a semisubmersible drilling rig but much larger, and an array of highly tensioned vertical tethers at each corner. The tethers, fashioned out of high-tensile-strength steel tubes, pull the floating hull down so far that they never go slack even in the trough of the maximum wave estimated to come once every 100 years. Although the tether system allows a degree of lateral motion, it prevents the heave, or vertical motion, associated with free-floating craft such as drilling vessels.

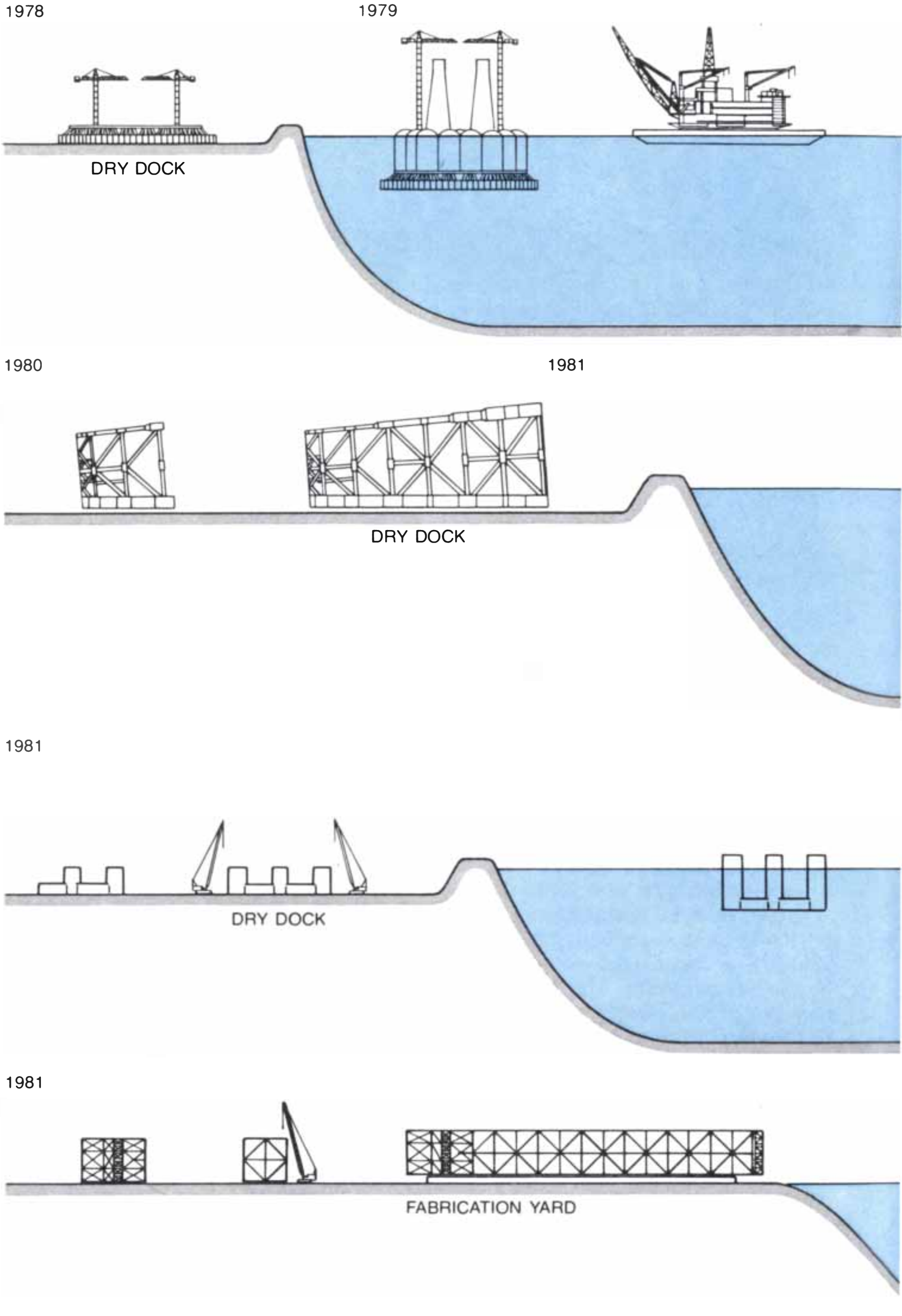
The floating hull must be designed with a careful balance between buoyancy and freeboard (the part of the hull above water) in order to handle extreme troughs as well as maximum crests. The great advantage of tension-leg platforms is their relative insensitivity to the increase of cost with the increase in the depth of the water: other things being equal, only the tethers need to be lengthened. As the offshore industry sensed that it was reaching the economic depth limits for fixed structures its attention shifted toward compliant structures and particularly to the TLP. A secondary but important economic advantage of the TLP is that it can be untethered and anchored at a new site. At the Offshore Technology Conference of last year many more technical papers were concerned with TLP's than were concerned with any other innovative design.

The Hutton platform, being developed and built by Conoco, Inc., for the British sector of the North Sea, is the first commercial tension-leg platform. Like the concrete gravity-base cells, the TLP hull is constructed in a large dry dock and moved to a deep-water mating site, where it can be submerged to receive its topsides. After mating, the hull and topside assembly is towed to the installation site, submerged to allow its tethers to be connected to preset foundations and deballasted to tension the tethers. The dynamic response of a tension-leg platform can be likened to that of an inverted pendulum except that the platform is held level by the pantographlike configuration of its tethers. In the heaviest seas the Hutton platform, riding in 485 feet of water, is designed to swing as much as 79 feet from the vertical and so to diminish wave impact.

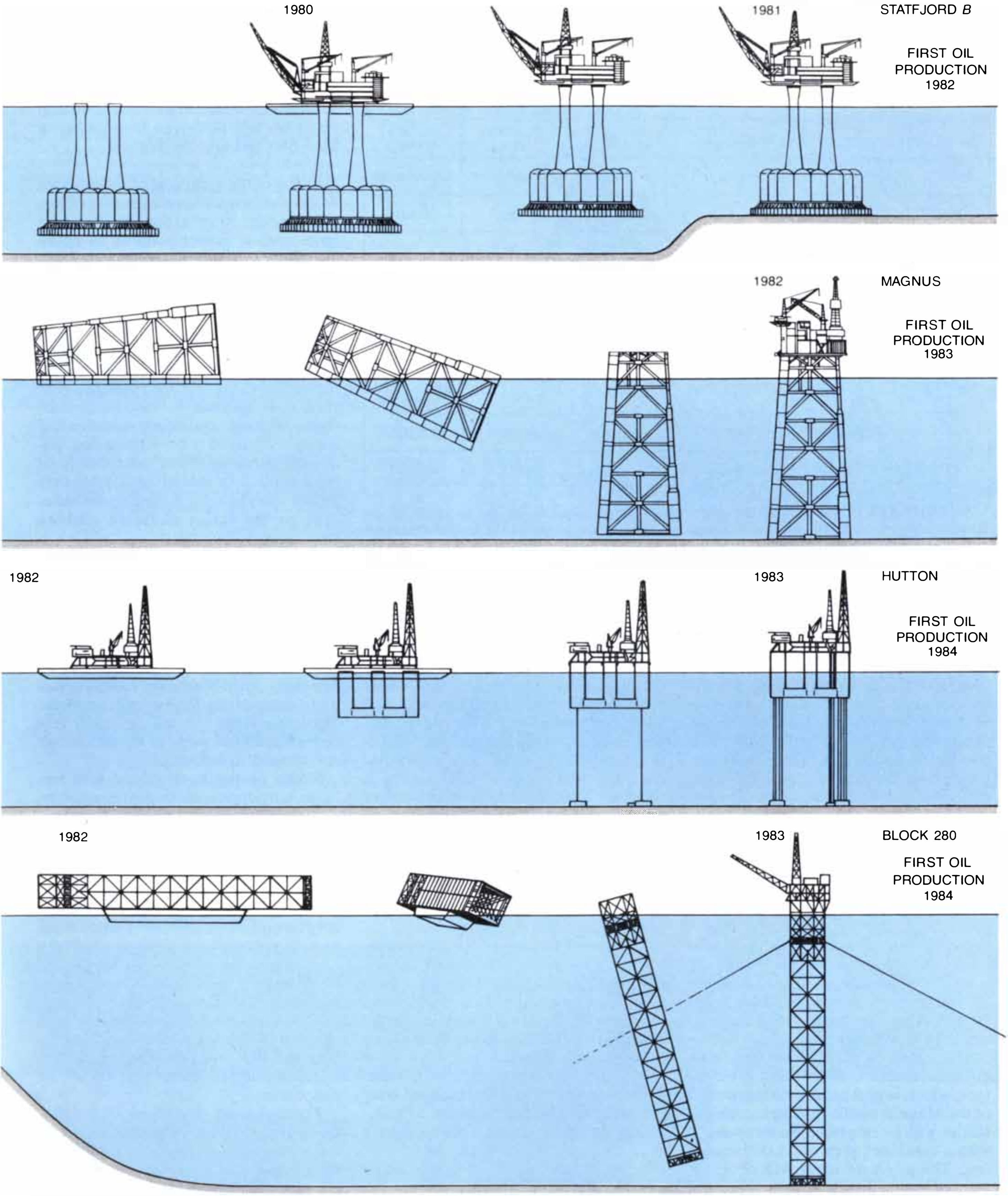
Before the Hutton platform is installed a drilling template will be emplaced on the sea floor within the perimeter of the tether foundations. A semisubmersible drilling rig will then predrill 13 directional wells, which can be connected to the topside piping when the platform arrives. In this way production can start much earlier than it would if all 24 of the projected wells had to be drilled after the platform was in place.

The TLP design was selected for the Hutton platform for three reasons. First, the lifetime of this particular field is expected to be shorter than that of other fields in the North Sea, so that a reusable platform has an extra advantage over a fixed structure. Second, the field happens to be one where little gas is released as the oil pressure in the field is reduced. As a result there will be no need to inject gas back into the field as oil is withdrawn, an operation that adds considerably to the weight the platform must carry. Third, the estimated cost of the tension-leg platform was essentially the same as the cost of a steel-template platform. For these reasons the sponsors concluded that the novel TLP design merited a demanding test. If the Hutton platform can prove its worth in the rigorous environment of the North Sea, much valuable information will be gained for future installations worldwide. The Hutton project may therefore be a milestone in the development of offshore platforms.

The last of the advanced platforms I shall describe is the guyed tower being developed by the Exxon Corporation for Block 280 in the Mississippi Canyon section of the Gulf of Mexico. Although it will be the least expensive of the four platforms, with an estimated cost of about $800 million, it will have by far the highest unit cost: about $32,000 per daily barrel of production. The chief problem facing Exxon was the sheer depth of water at the intended site: 1,000 feet. Moreover, the platform had



METHODS OF CONSTRUCTION AND EMPLACEMENT of the four different offshore platforms are depicted schematically. The design of the Statfjord *B* platform is in some ways the boldest. After the concrete base had been well started it was floated into Stavangerfjord, where the four concrete columns were lengthened by pouring concrete into slip forms from a floating concrete plant. Simultaneously the topside facilities were built on a temporary base and were floated into place over the submerged columns. The integrated structure was then towed some 230 miles to a prepared site in the Statfjord oil field. The steel-template jacket supporting the Magnus platform is now nearing completion in a dry dock at Nigg Bay, Scotland. Flotation compartments in two of the structure's four legs will provide enough buoyancy for



the 41,000-ton tower to float while it is being towed this summer to the Magnus field, the most northerly location of the North Sea yet developed. The tower will be anchored to the sea floor by steel-tube piles driven to a depth of 300 feet. The extensive topside facilities will then be lifted onto the tower by semisubmersible derrick barges. The hull of the Hutton tension-leg platform is also under construction in a dry dock at Nigg Bay. When the hull has been floated into deep water and partially submerged, the topsides will be barged into place as a single unit. The completed platform will then be towed to the Hutton field, within sight of Statfjord *B,* and tethered to its prepared foundations. Fabrication of the Block 280 guyed tower is proceeding at Corpus Christi, Tex. Later this year the tower will be barged 500 miles to a site in the Gulf of Mexico, where it will be launched over the side of the barge. The tower will be sufficiently buoyant for it to float low in the water. By controlled flooding it will be rotated to a vertical position and brought to rest on the sea floor. The array of 20 radiating cables will be fastened to the tower to hold it in place. After foundation pilings are driven topside facilities can be raised into place.

	STATFJORD *B*	MAGNUS	HUTTON	BLOCK 280 (MISSISSIPPI CANYON)
LOCATION	NORTH SEA (NORWAY)	NORTH SEA (U.K.)	NORTH SEA (U.K.)	GULF OF MEXICO (U.S.)
STRUCTURE	CONCRETE GRAVITY	STEEL JACKET AND PILING	TENSION LEG	GUYED TOWER
TYPE	FIXED	FIXED	COMPLIANT	COMPLIANT
WATER DEPTH (FEET)	472	611	485	1,000
100-YEAR WAVE (FEET)	100	102	98	72
PRODUCTION (BARRELS PER DAY)	150,000	120,000	110,000	25,000
INITIAL PRODUCTION (YEAR)	1982	1983	1984	1984
MAXIMUM DECK OFFSET (FEET)	<3	4	79	39
APPROXIMATE TOTAL COST (BILLIONS)	$1.8	$2.6	$1.3	$.8
COST PER BARREL-DAY	$12,000	$21,700	$11,800	$32,000
PROJECT MANAGEMENT	MOBIL	BRITISH PETROLEUM	CONOCO	EXXON

MAIN CHARACTERISTICS of the four offshore oil platforms illustrated on pages 32 and 33 are summarized, together with the estimated costs and the names of the project managers. A fundamental difference in design distinguishes the Statfjord *B* and Magnus structures from the Hutton and Block 280 ones. The first two are rigidly fixed to the sea floor, so that the maximum deck displacement in the heaviest seas will be limited to a few feet. The other two platforms are compliant, designed to yield far more to waves or strong currents. Capital investments as high as $32,000 per barrel-day can be justified now that oil has risen to $34 per barrel.

to be capable of supporting substantial deck loads in order to provide for the development of up to 58 wells. Like the Hutton platform, the guyed tower is a compliant structure. It will consist of a slender steel tower of square cross section, measuring 120 feet on a side. The tower will be held upright by a radial array of 20 anchor cables, each one about 3,400 feet long. Some 1,800 feet from the tower a 180-ton weight will be attached to each cable. Under normal sea conditions these "clump" weights will rest on the bottom and keep the cables taut. Under extreme storm conditions the weights will gradually lift and allow the tower to tilt; the design allows for a maximum departure from the vertical of 39 feet at deck level. At that departure the deck will slope no more than two degrees.

The total weight of the tower and its guys is estimated to be about 43,000 tons, which is comparable to the weight of the Magnus platform. The topside facilities will be emplaced on three decks with a combined area of 75,000 square feet. The guy-wire array will cover an area of more than a square mile, but even where the guys are attached to the tower they will be submerged sufficiently to clear surface craft.

I noted at the beginning of this article that the fatigue lifetime of submerged structures is strongly influenced by the number of waves per year whose period happens to coincide with the platform's natural frequency. An important advantage of the guyed tower over the steel-template-jacket platform is that where a submerged tower 1,000 feet high will have a natural period of about 30 seconds, a submerged steel-template-jacket structure of the same dimensions will have a natural period of five seconds. A structure with a natural frequency of five seconds would resonate with many more waves per year in the Gulf of Mexico than a structure with a much longer natural frequency would. With a natural frequency of 30 seconds the guyed tower will jump over the highest-energy band of the wave spectrum.

Exxon's guyed tower will be assembled in a tidewater fabrication yard at Corpus Christi, Tex., with the tilt-up techniques perfected over many years for the building of steel-template-jacket platforms. When the tower framework is completed, it will be skidded onto a long barge and transported to the site. There the tower will be launched over the side of the barge in order to minimize launching stresses. Ordinarily a steel-template jacket is launched over the end of the barge, a procedure that often creates stresses higher than any exerted when the platform is in service. The tower will have enough buoyancy to float with only its top panel above water. By means of controlled flooding the tower will then be rotated to a vertical position.

Before the tower arrives at Block 280, 65 miles southeast of Grand Isle, La., 20 piles will have been driven 150 feet into the sea floor as anchors for the guy wires. In addition a sea-floor drilling template will have been installed near the base of the tower. With the drilling template as a guide, the tower will be landed on a mud mat and eight foundation piles will be driven from the deck level 560 feet into the bottom.

In view of the great cost of the offshore oil platforms I have been describing, a cost equal to or exceeding that of a large nuclear power plant or an entire oil refinery, it is apparent that the offshore-platform industry has compelling reasons to rely on well-tested designs. On the other hand, the industry is steadily confronted with environmental conditions going beyond its experience. In addition to the four designs I have described many hybrid designs combining steel and concrete in ingenious marriages of fixed, compliant and floating configurations have been examined. Although some of these concepts have been applied in special situations, they do not seem likely to have a major influence on the design of future offshore structures.

One region where new platform concepts are badly needed is the Arctic, where sea ice and icebergs present novel forces of great magnitude. For example, on the North Slope of Alaska oil companies have experimented with gravel platforms built up in water as much as 65 feet deep. Such platforms have worked quite successfully for exploration wells. They can resist the sea ice for a few years but do not seem to be the answer for permanent installations.

In the deeper open Arctic seas icebergs introduce a more capricious type of elemental force. It now appears that small icebergs, which can be moved violently by storm swells, are a greater menace than large icebergs, which can move only as fast as an ocean current can carry them. For dealing with the iceberg hazard there seem to be three main options: concrete structures, which by their sheer mass can absorb the impact of an iceberg, tension-leg platforms designed to break loose under heavy impact and wells embedded in the sea floor coupled to floating production systems. The industry will probably be experimenting with all three systems in the near future.

Turning from the physical installations to the groups of people who create them, one encounters a number of things not characteristic of other areas of engineering. Major offshore-platform projects go through eight readily identifiable phases from the discovery of oil to its production, a development period that lasts for five to 14 years. These eight phases can be consolidated into two basic stages: planning and execution. Substantially different technical skills and attitudes are required for optimum performance at each stage, yet continuity of effort between the two stages is of critical importance.

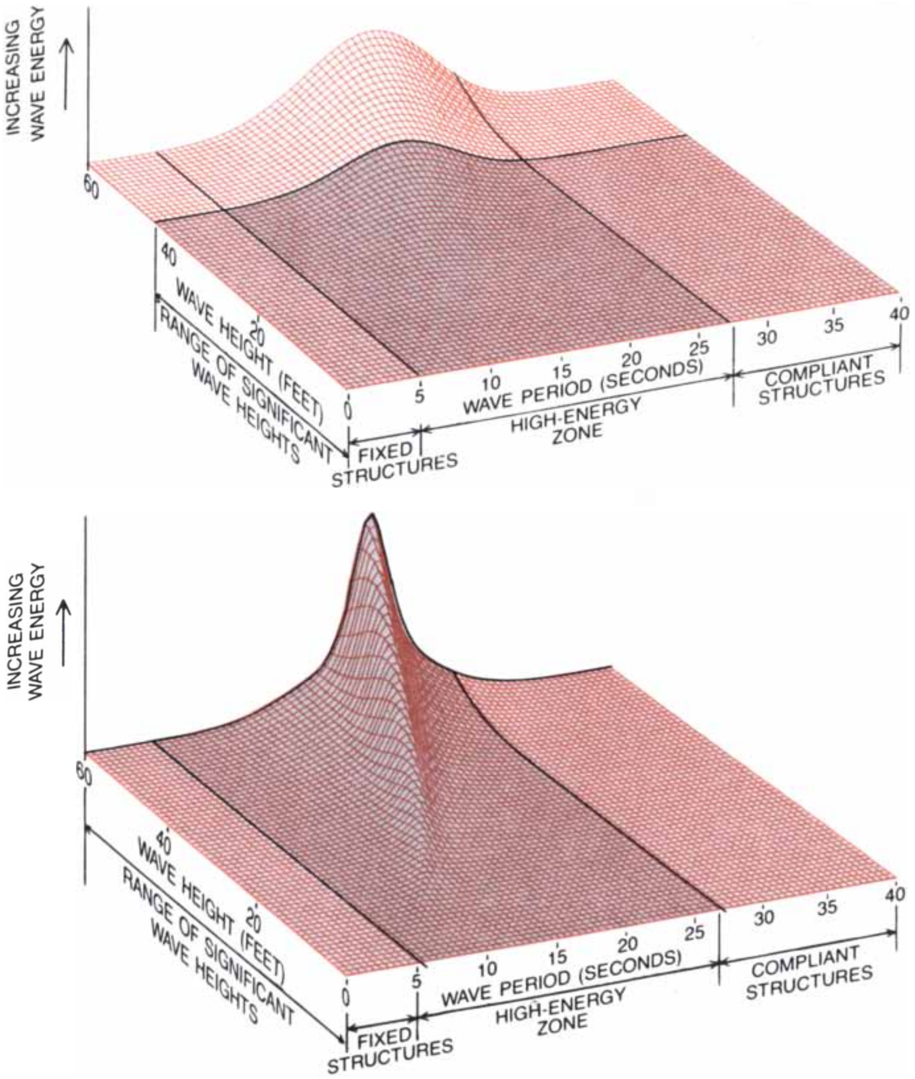
The first stage, deciding what should be built, is a period of intense investigation, drawing on the best experience of a rapidly developing industry. The investigation must be exhaustive and carried out with a completely open mind, uncluttered by proprietary considerations. Properly executed, the effort results in agreement between the project team and senior management on the basic duty requirements of the offshore platform and in key technical decisions on design and construction methods, schedules and costs. As a rule of thumb about 85 percent of the decisions that determine a project's final cost can be made at this stage. The result can best be visualized as an encyclopedia encompassing every significant aspect of the project but lacking the underlying detail. This document (which I prefer to call a project manual), with the transfer of key participants, constitutes the necessary bridge to the second stage.

The second stage in the execution of the project is concerned with detailed design, procurement, fabrication and installation. It is at this stage, after the major commitments have been made, that the offshore-platform industry has encountered its biggest problems. Consider that only 10 years have elapsed since the first giant oil fields were discovered in the North Sea. In that time the price of oil has jumped from about $2 per barrel to $34. Inflation has soared worldwide. Tax structures have been continually modified. Regulatory requirements have become increasingly strict. Meanwhile the industry has had to absorb scores of technological advances. The rate of change has far exceeded anything previously known in the petroleum industry.

It is obvious that however well an original plan is made, it will have to be modified almost immediately to meet unanticipated challenges, technical, financial and political. It is in this difficult period, stretching over five to seven years, that final costs can balloon astronomically unless the decision making and project discipline are of high quality. There is a constant battle between those who seek perfection in detail by trying to incorporate the latest technology and those who try to balance such benefits against their less obvious unwanted side effects. Experience has shown that there is relatively little tolerance for change in this second period.

Looking ahead, of all the structures being developed for offshore oil fields the tension-leg platform seems to have the brightest future in very deep water. Since the design can also be applied over a wide range of water depths, it seems to hold greater promise for standardization than any other type. Since it is a compliant structure, the tension-leg platform is not affected by seismic activity. Perhaps its only weakness arises from its having to be buoyant under all weather conditions, which somewhat limits its deck-load capacity. The design of tension-leg platforms, like the design of guyed towers, is still at the toe of the learning curve and will undoubtedly go through several generations of improvement in the near future. In arctic areas, subject always to the availability of appropriate sea-bottom soils, gravity-base structures with deeply penetrating steel skirts may have the strength and shear resistance that will be needed to offset the massive onslaughts of sea ice, ice ridges and icebergs.

In parallel with the development of platform structures is a trend toward greater reliance on related undersea installations. One example is the six undersea wells that will be tied into the Magnus platform. Another is the undersea drilling template under the Hutton platform, which will be used to predrill wells. This trend will almost certainly continue as more experience is gained with undersea facilities. Nevertheless, the number of air-breathing components, human and mechanical, essential for offshore drilling and production activities guarantees that facilities above the surface will continue to play a key role.



SEA-SPECTRUM DIAGRAMS show the important relation between sea energy and the fundamental vibrational period of offshore oil platforms of different designs. The design must withstand not only the impact of the largest wave expected as a single event over the lifetime of the structure but also the cumulative effect of several million waves per year with a period matching the fundamental period of the tower. Offshore structures are designed to have periods that fall outside the high-energy region of the wave-energy spectrum. If this were not done, dynamic amplification would result in a very short fatigue lifetime. The upper spectrum is based on conditions in the Gulf of Mexico, where the Block 280 guyed tower will be emplaced. The lower spectrum reflects typical conditions in the North Sea. These "blanket" diagrams are prepared by transforming complex records of random waves into a great number of simple regular waves of small amplitude and random phase. "Significant waves" are an average of the highest third of the waves in any given sea state. Thus they are much less than the maximum wave that is expected to occur in the same sea. In these diagrams a 60-foot significant wave can be interpreted as a 100-foot maximum wave. The computer-generated graphic displays were prepared by Ross Cowan of International Software Services of San Mateo, Calif.

Nutrients That Modify Brain Function

They are the precursors of neurotransmitter molecules. Increasing their level in the brain amplifies signals from some nerve cells. In effect they act like drugs, and one day they may serve as drugs

by Richard J. Wurtman

A nutrient is different from a drug, most people would agree. A nutrient is a food substance that in most cases supplies either the energy or the molecular building blocks the body requires. A drug is a substance given for its effect on a specific organ or type of cell. Whereas all healthy people need essentially the same nutrients, a drug would ordinarily be recommended only for people with a particular disease or condition. In this article I shall tell about three nutrients that, when they are administered in the pure form or simply ingested in food, can act like drugs. They give rise to important changes in the chemical composition of structures in the brain. The changes can modify brain function, particularly in people with certain metabolic or neurologic diseases.

Two of the nutrients are the amino acids tryptophan and tyrosine. Amino acids are the building blocks of proteins, and so tryptophan and tyrosine are present in most foods. The third nutrient is choline, a component of lecithin; egg yolks, liver and soybeans are notably rich in lecithin. The composition and function of the brain can be altered by tryptophan, tyrosine and choline because they are the precursors of neurotransmitters: substances that are released from a neuron, or nerve cell, when it fires. The neurotransmitter thereby conveys the nerve impulse across a synapse to either another neuron, a muscle cell or a secretory cell. Tryptophan is converted in the terminals of certain neurons into the neurotransmitter serotonin. In other cells choline is converted into the transmitter acetylcholine. In still another population of cells tyrosine serves as the precursor of dopamine, norepinephrine and epinephrine, which are collectively called the catecholamine transmitters. An increase in the brain level of a precursor enhances the synthesis of the corresponding neurotransmitter product. The enhanced synthesis can in turn cause the neuron to release more transmitter molecules when it fires, amplifying the transmission of signals from the neuron to the cells it innervates.

My associates and I at the Massachusetts Institute of Technology and other investigators have been exploring the interactions that relate the amount of a nutrient administered or ingested to its level in the blood plasma, its level in the brain and its effect on nerve transmission. The interactions are not simple. The conversion of tryptophan into serotonin is influenced by the proportion of carbohydrate in the diet; the synthesis of serotonin in turn affects the proportion of carbohydrate an individual subsequently chooses to eat. In the case of choline and tyrosine the effect on a neuron of an increased supply of the nutrient both varies with the neuron's firing frequency and can lead to changes in that frequency. Choline and tyrosine can therefore amplify neurotransmission selectively, increasing it at some synapses but not at others. It may be possible to exploit such selectivity to develop novel therapeutic agents for several diseases, including hypertension, some forms of depression, Parkinsonism and some memory disorders of old people.

The observations relating nutrient intake to neurotransmission originated in studies of a phenomenon seemingly unrelated to the brain: daily rhythms in the metabolism of dietary amino acids. When people eat, the concentration in their blood plasma of most amino acids (and of other food constituents) changes predictably in ways that depend on what foods are eaten. For people who take their meals at the usual times plasma amino acid levels generally exhibit pronounced daily rhythms. For example, among people consuming the high-protein diet typical of the U.S. the plasma concentration of the amino acid leucine is twice as high between 3:00 P.M. and 3:00 A.M. as it is during the rest of the day. If the same people eat protein-free meals, the leucine level instead falls by half during these hours of active digestion and absorption. In the first case the increase represents the entry into the bloodstream of some of the leucine in the dietary protein. In the second case the decrease results from the secretion of insulin (induced by ingested carbohydrate), which accelerates the passage of leucine and most other amino acids from the circulation into skeletal muscle.

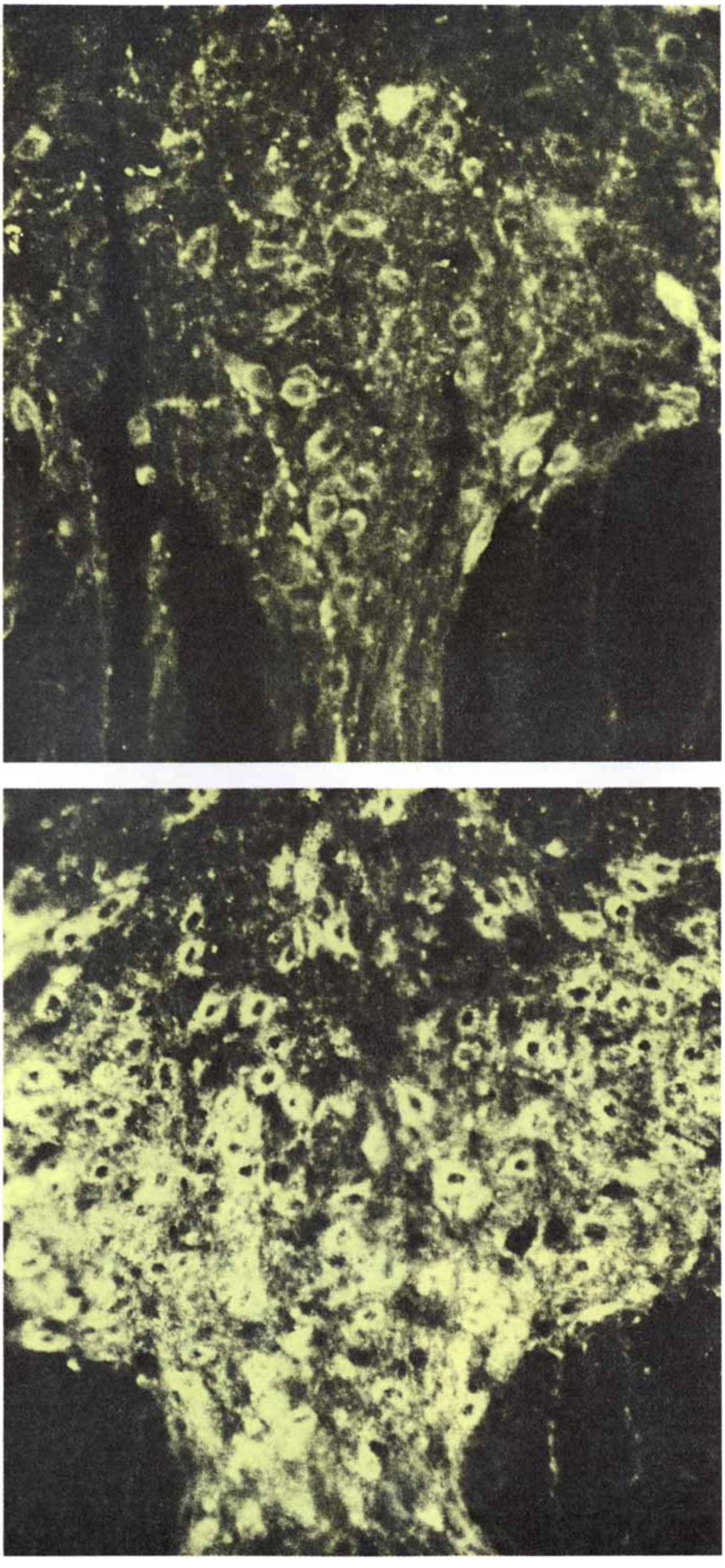
My associates and I discovered these rhythmic, food-induced variations in the plasma level of various nutrients about a decade ago. We wondered whether the changes might have any functional significance. In particular we wondered whether the fluctuations in the concentration of circulating nutrients had any effect on the rate at which the nutrients are converted into cellular constituents. In order for changes in the level of a nutrient to influence the rate of conversion, the enzyme catalyzing the conversion must have a certain property. The enzyme's ability to bind the nutrient preparatory to changing its chemical structure must be relatively poor, so that at the usual nutrient concentrations each enzyme molecule is less than fully saturated with the nutrient and functions at less than peak efficiency. In this situation the quantity of the nutrient available to the enzyme is the rate-limiting element in the reaction, and so an increase in the nutrient's concentration increases the level of enzyme activity: more of the nutrient is converted and more of the product is formed.

We knew that tryptophan is converted into serotonin by just such a low-affinity enzyme. The release of serotonin by neurons originating in the brain stem delivers signals to widely scattered groups of neurons that control such things as sleep, mood and appetite. Other investigators had already shown that the concentration of serotonin in the brain can be increased by giving experimental animals very large doses of pure tryptophan. John D. Fernstrom and I decided to examine the possibility that normal daily variations in the plasma concentration of tryptophan might be enough to alter the rate of serotonin synthesis in the rat brain. We found that even low doses of tryptophan, which raise the plasma concentration of the amino acid but keep it within the normal daily range we had previously established, did indeed enhance serotonin synthesis.

To see whether a reduction in plasma tryptophan had the opposite effect we injected some rats with insulin and gave others a diet of carbohydrates, which induces the secretion of insulin. We expected that the hormone would reduce the plasma level of tryptophan as it does that of other amino acids by moving them out of the bloodstream and into skeletal muscle. To our surprise the insulin did not lower the tryptophan concentration in the plasma, and it actually raised the concentration in the brain, increasing serotonin synthesis instead of reducing it. Feeding the animals large amounts of protein brought another surprise: even though amino acids were plentiful in the diet, both the brain concentration of tryptophan and the synthesis of serotonin were reduced.

The apparent paradoxes were resolved when we found that the amount of tryptophan available in the brain for conversion into serotonin depends not only on the amount of tryptophan in the plasma but also on the ratio of plasma tryptophan to the plasma level of five other amino acids: tyrosine, phenylalanine, leucine, isoleucine and valine. All six of these amino acids are comparatively large molecules, and in a physiological environment most of them are electrically neutral, with about as much positive charge as negative.

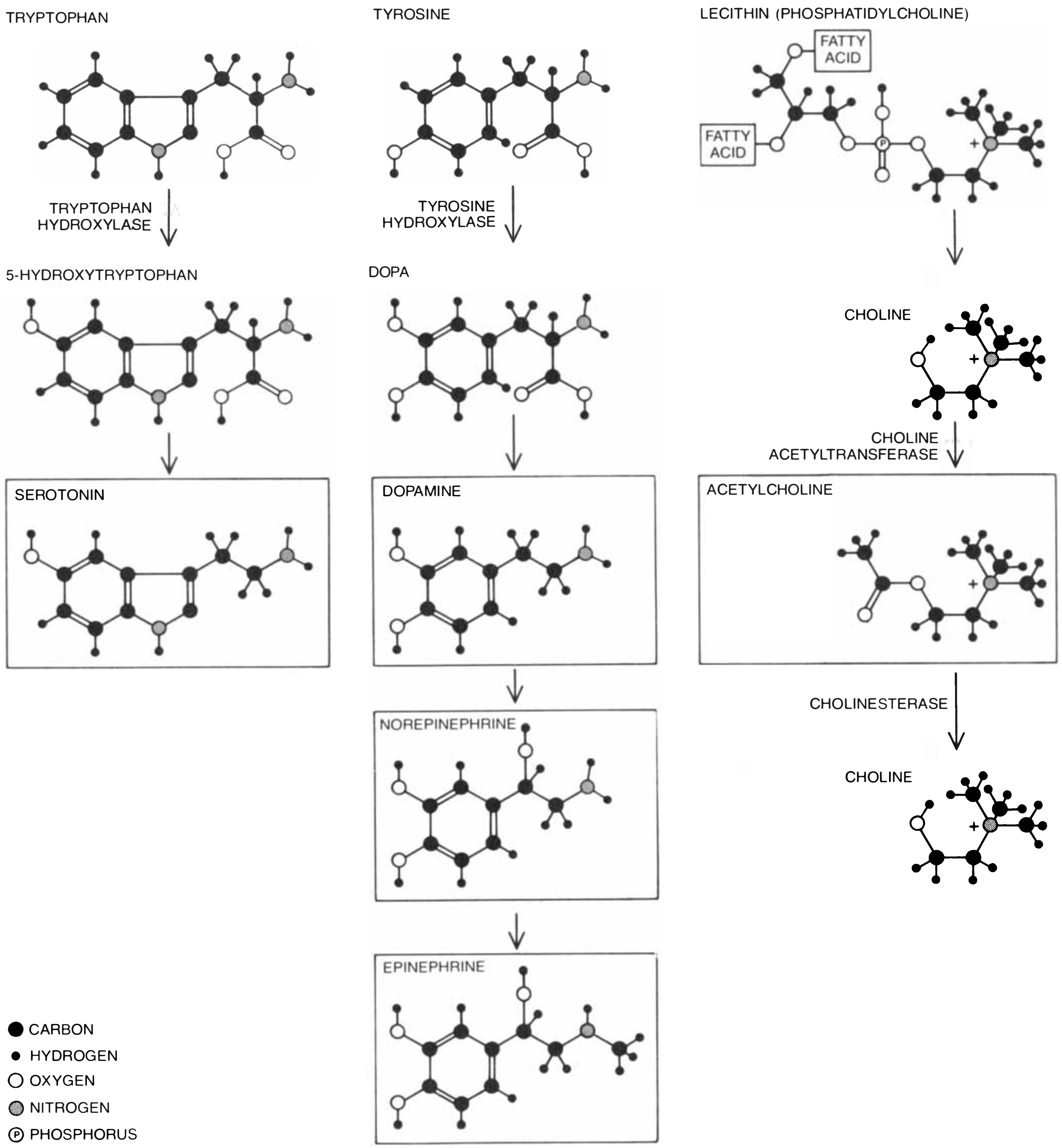
It is very difficult for large, water-sol-



SYNTHESIS OF SEROTONIN, a neurotransmitter, is greatly increased by the administration of the nutrient that is its precursor: the amino acid tryptophan. These photomicrographs, made by George K. Aghajanian of the Yale University School of Medicine, show neurons (nerve cells) in the dorsal raphe nucleus of the rat, one of the structures where most of the brain's serotonin is synthesized and stored. Thin brain-tissue sections were treated with formaldehyde, which reacts with serotonin to form a greenish yellow fluorescent compound. In the control section (*upper photograph*) faint fluorescence reveals the presence of serotonin in the cytoplasm of the cell bodies and along the axons, or nerve fibers. The other section (*lower photograph*) is from the brain of an animal that was injected with a large dose of tryptophan an hour before the tissue sample was taken. There is a dramatic increase in the level of fluorescence, reflecting an increased synthesis of serotonin in the presence of the additional precursor.

uble molecules to diffuse out of the capillaries of the brain and gain access to neurons and other brain cells. Their passage between the blood and the brain is facilitated by carrier molecules present in the endothelial cells lining brain capillaries. A single species of carrier molecule transports all six of the large, neutral amino acids across the blood-brain barrier; the amino acids compete with one another for attachment to the carrier and hence for uptake from the bloodstream into the brain. There is far less tryptophan in most proteins than there is tyrosine, phenylalanine, leucine, isoleucine or valine. A high-protein meal therefore reduces the plasma ratio of tryptophan to the competing amino acids; less tryptophan is carried across the barrier and less reaches the neurons.

A high-carbohydrate meal has the opposite effect because the insulin secreted in response to carbohydrate intake reduces the plasma level of the competing amino acids more than it does that of tryptophan. Whereas the other amino acids circulate as free molecules, most of the tryptophan is bound to the plasma protein albumin; segregated in an albumin reservoir, the tryptophan is essen-



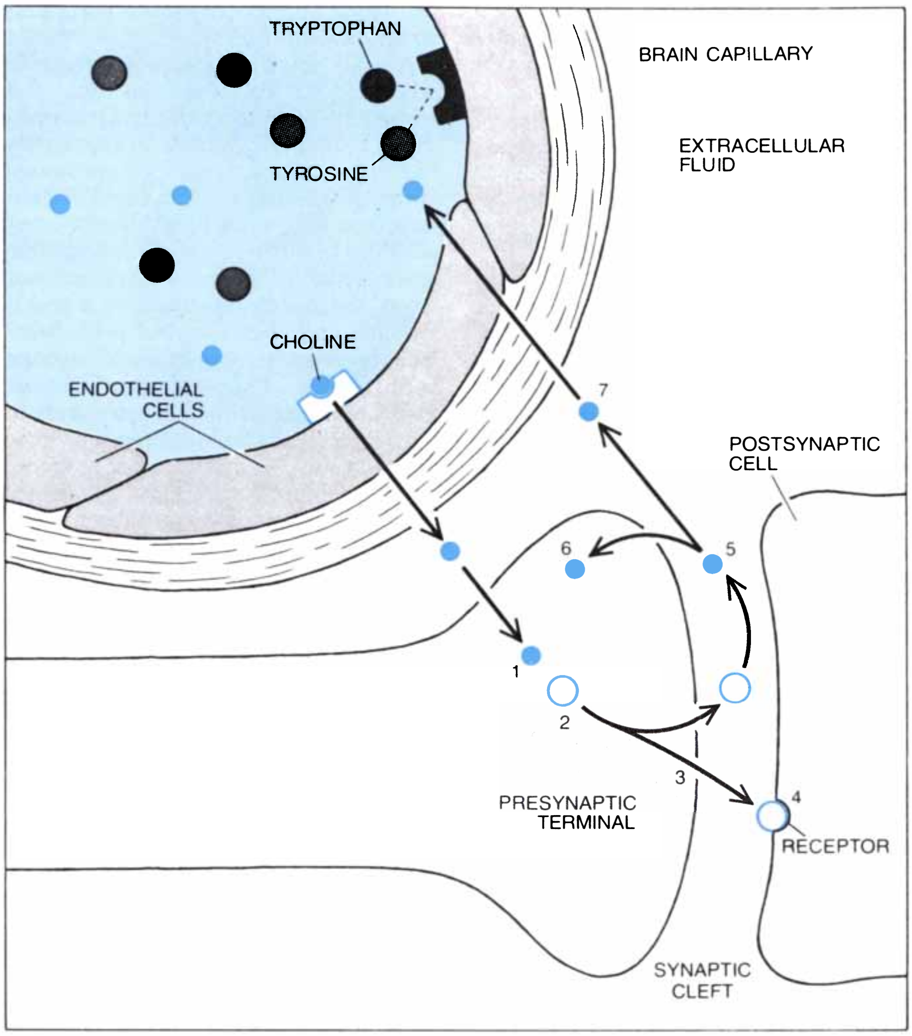
CHEMICAL STRUCTURE of three nutrients and the paths by which they are converted into neurotransmitters are shown. Tryptophan is converted into serotonin in two steps. The three catecholamine transmitters, dopamine, norepinephrine and epinephrine, are formed in successive steps from another amino acid: tyrosine. The food substance lecithin provides choline, the precursor of the transmitter acetylcholine. Unlike other precursors, choline is recycled; it is regenerated when acetylcholine is broken down by cholinesterase.

tially immune to the effect of insulin. The result is that after carbohydrates are eaten the plasma ratio of tryptophan to its competitors rises, causing more tryptophan to reach the neurons [see "Nutrition and the Brain," by John D. Fernstrom and Richard J. Wurtman; SCIENTIFIC AMERICAN, February, 1974].

We proposed that these interactions enable the serotonin-releasing neurons in the brain to serve as sensors of the plasma tryptophan ratio, which increase serotonin release after a carbohydrate meal and reduce it after a high-protein meal. David Ashley and G. Harvey Anderson of the University of Toronto Faculty of Medicine subsequently found evidence that the brain exploits this property of the serotoninergic neurons when an animal chooses one food in preference to another. In an effort to define the specific food constituents whose choice is influenced by brain serotonin, Judith J. Wurtman and I allowed rats to choose between two diets having different proportions of carbohydrate and protein. Various treatments that increase serotonin release in the brain (such as giving the drug fenfluramine) caused the rats to selectively reduce their consumption of carbohydrate. Recently we have shown that serotonin-increasing treatments have a similar effect on obese people with a craving for dietary carbohydrate when they are allowed to choose from a range of snack foods over a period of several weeks.

It appears, in other words, that eating a meal rich in carbohydrate and poor in protein generates a neurochemical change—namely increased serotonin synthesis—that causes the animal to reduce its intake of carbohydrate but not of protein. It seems likely that this control of serotonin release by diet composition and of diet composition by serotonin release evolved because it helps to sustain nutritional balance. Presumably it keeps the bear from eating only honey and keeps human beings from eating sweets and starches to the exclusion of enough protein. Some obese people may suffer from a disturbance of this remarkable feedback mechanism that interlinks nutritional, metabolic, neurochemical and behavioral systems. Unfortunately there is no noninvasive technique for measuring serotonin release in the brain and thus directly establishing the role of serotonin release in man.

In 1975 Edith L. Cohen and I, and independently Dean R. Haubrich of the Merck Institute for Therapeutic Research, showed that the administration of choline increases the synthesis of acetylcholine in the brain. (Consumption of lecithin, the food constituent that provides most of the choline in adult diets, is even more effective.) Soon after that Candace J. Gibson and I observed a similar relation between meal composition and the synthesis of the catecholamine transmitters dopamine, norepinephrine and epinephrine. The administration of tyrosine or the consumption of meals that increase the relative plasma level of tyrosine (the level in relation to the other large, neutral amino acids) raises the tyrosine level in neurons, thereby making more of the amino acid available to the enzyme tyrosine hydroxylase and accelerating catecholamine synthesis.



NUTRIENT MOLECULES must cross the "blood-brain barrier" to get to the brain cells where they serve as precursors of neurotransmitters. The junctions between the endothelial cells of the capillaries in the brain are too tight to allow the passage of tryptophan, tyrosine or choline, and so the precursors must be transported through the capillary wall by carrier molecules. Tryptophan, tyrosine and other large, electrically neutral amino acids compete for the same carrier. The uptake and conversion of choline are shown here. Choline in the extracellular fluid of the brain is taken up by the terminal of a cholinergic neuron (*1*) and is converted into acetylcholine (*2*). The acetylcholine is released into the synaptic cleft when the neuron fires (*3*). The acetylcholine may interact with a receptor and thereby transmit a signal to the postsynaptic cell (*4*). Alternatively, the transmitter may be converted back into choline (*5*), which may be taken up again by presynaptic terminal (*6*) or may enter extracellular fluid and bloodstream (*7*).

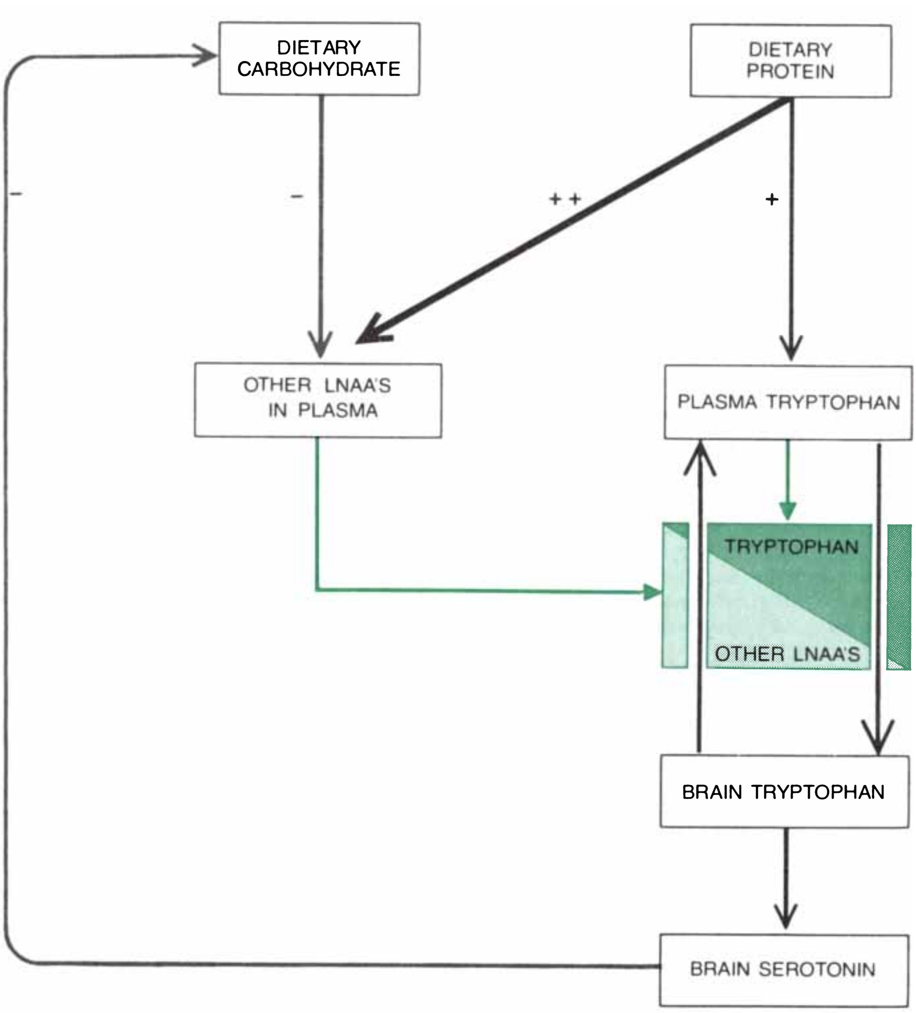
With three nutrients shown to affect the synthesis of their neurotransmitter products, it became possible to formulate some general principles for predicting whether a given transmitter might be subject to such control and to propose some standardized experiments for demonstrating the existence of such relations. I can illustrate the principles by describing the nature of the five biochemical processes that must follow in sequence if the consumption of a meal rich in a nutrient is to increase the synthesis in the brain of the transmitter for which the nutrient is the precursor.

First, consumption of a food that includes the nutrient must significantly elevate the plasma level of the nutrient; the plasma level cannot be held relatively constant by feedback mechanisms like the ones that regulate plasma *p*H, say, or the concentration of calcium. Second, the concentration of the nutrient in the brain must depend on and vary with the concentration in the plasma; there cannot be an impenetrable blood-brain barrier for the precursor. Third, the transport mechanism that mediates the nutrient's movement between the blood and the brain must be of the low-affinity type: it must be unsaturated with its substrate (the precursor), so that it can become more nearly saturated when the plasma level rises. Fourth, the neuronal enzyme that catalyzes the conversion of the precursor into the transmitter must also be a low-affinity one. (As I mentioned above, it was this requirement that led us to the effect of tryptophan on serotonin-releasing neurons.) Fifth, the enzyme must not be susceptible to feedback inhibition when the intracellular level of its product, the transmitter, rises.

All these conditions have now been shown to be fulfilled in the synthesis of serotonin, acetylcholine and the catecholamines. There is some evidence that they are also met in the case of two other transmitters, histamine and glycine, whose production seems to be affected by the availability of their precursors. The fact that so many conditions must be met if precursor availability is to influence neurotransmitter synthesis implies that the precursor-transmitter relation is no biological accident. Rather it seems likely the relation has some adaptive value and has therefore been conserved in evolution.

Are all neurotransmitters subject to precursor control? Probably not. The immediate precursors of some transmitters are usually available in concentrations that fully saturate the neurotransmitter-synthesizing mechanisms and are therefore independent of plasma composition. In other cases it is not possible to assess the extent to which precursor levels control the synthesis of a transmitter, either because the precursor's identity remains to be established or because blood-brain transport systems frustrate the experimenter's attempts to raise the concentration of the precursor in the brain. Even if only some of the 25 or so neurotransmitters known are potentially responsive to nutrient consumption, however, it is interesting that the responsive transmitters include several whose action is thought to be affected by drugs given to treat neurologic, psychiatric and even cardiovascular diseases.

Although the intake of a nutrient may influence the synthesis of a neurotransmitter in the presynaptic terminal of a neuron, it does not necessarily follow that the nutrient alters the transmission of impulses across the synapse. The nutrient must also be shown to increase the number of transmitter molecules released by a given set of neurons per unit time. The number depends mainly on the number of neurons in the tract or nerve being examined, on the total number of synapses they make, on the frequency with which the neurons fire and on the average number of molecules released at each synapse each time the neurons fire. It was theoretically possible that nutrient administration affected only the synthesis of a transmitter and not its release into synapses. This was shown not to be the case in experiments done by George G. Bierkamper and Alan M. Goldberg of the Johns Hopkins University School of Medicine. They measured the release of acetylcholine from motor neurons incubated in various choline concentrations. When they stimulated the nerves electrically, they found a striking parallel between the concentration of choline and the amount of acetylcholine released.



FOOD CONSUMPTION affects the synthesis of serotonin in the brain. Eating protein raises the plasma level of tryptophan, but it raises the plasma level of five other large, neutral amino acids (LNAA's) even further because each of them is more plentiful than tryptophan in proteins. Dietary carbohydrate induces the secretion of insulin, which moves most amino acids out of the bloodstream but has little effect on tryptophan. Because tryptophan must compete with the other LNAA's for transport across the blood-brain barrier, the movement of tryptophan from the plasma to the brain is controlled by the plasma ratio of tryptophan to the other LNAA's (*color*). When the ratio is high, tryptophan enters the brain; when it is low, tryptophan moves from the brain into the bloodstream. The release in the brain of serotonin synthesized from the tryptophan appears to reduce a rat's (or a person's) carbohydrate intake.

Tracing the effect of added nutrient one step further, Ismail Ulus of the University of Bursa in Turkey and I showed that the increase enhances not only transmitter release but also neurotransmission, giving rise to a chemical change in the postsynaptic cells. In these studies we worked with the splanchnic nerve, which runs from the spinal cord to the adrenal glands. There the splanchnic-nerve fibers terminate in synapses with cells called chromaffin cells in the medulla, or core, of the glands. When the splanchnic nerve fires, it releases acetylcholine. This transmitter causes the postsynaptic chromaffin cells to release epinephrine and also to make more tyrosine hydroxylase, the enzyme that ultimately controls the synthesis of epinephrine.

Other investigators (notably Julius Axelrod, Hans Thoenen and Robert A. Mueller, who were all then at the National Institute of Mental Health) had shown that if rats were given a treatment that chronically increased the firing rate of the splanchnic nerve, the consequent release of added acetylcholine increased the level of tyrosine hydroxylase activity in the adrenal gland. We thought that giving choline would increase the amount of acetylcholine released per firing without altering the splanchnic nerve's firing frequency and therefore might similarly increase tyrosine hydroxylase activity. It did. Then we tested animals in which the splanchnic-nerve fibers supplying one of the two adrenal glands had been severed.

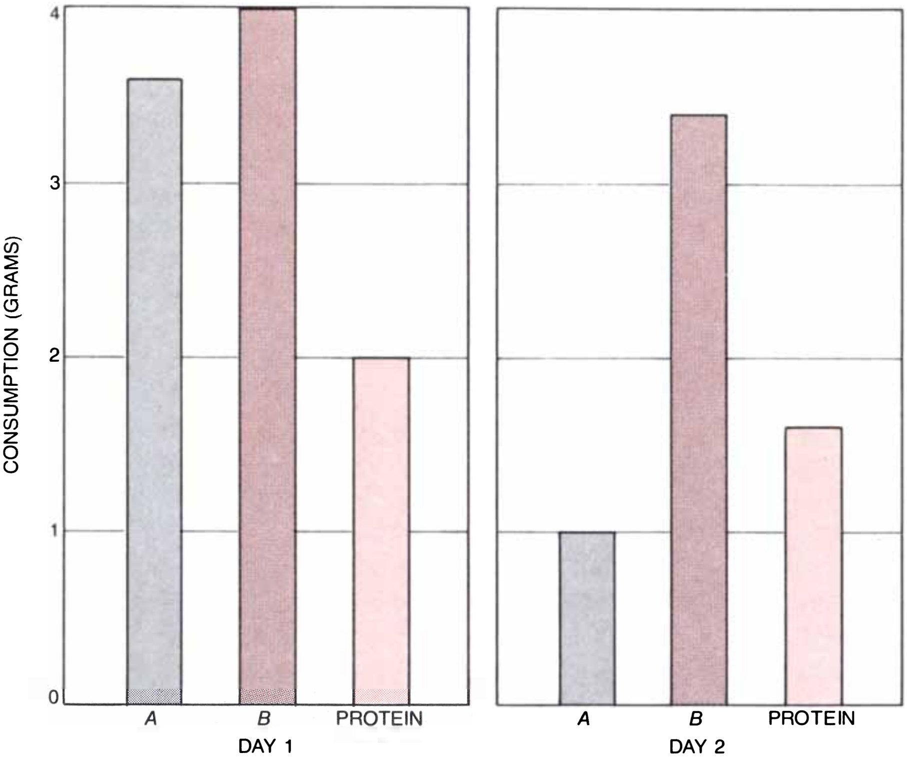
We found that choline induced greater enzyme activity in the intact gland but failed to do so in the denervated one, proving that circulating choline did not act directly. It needed first to be converted into acetylcholine, which was released from the splanchnic nerve.

To confirm that the choline acted by changing the amount of neurotransmitter released per firing (and not by raising the firing frequency) we tested the combined effect of giving rats both choline and a treatment meant to accelerate splanchnic-nerve firing. (The treatments included putting animals in a cold environment, giving them very large doses of insulin and administering drugs that cause prolonged depression of blood pressure.) We reasoned that if choline and the other treatment both acted by increasing the firing frequency, their combined effect on tyrosine hydroxylase would be about the same as the effect of one treatment alone. If, on the other hand, choline increased the amount of acetylcholine released per firing, the combined effect should be multiplicative. We observed that the effects of giving choline and of the second treatment were indeed multiplicative. To our surprise, however, the two treatments also potentiated each other, that is, their combined effect on tyrosine hydroxylase activity was invariably larger than the product of their effects when they were given alone.

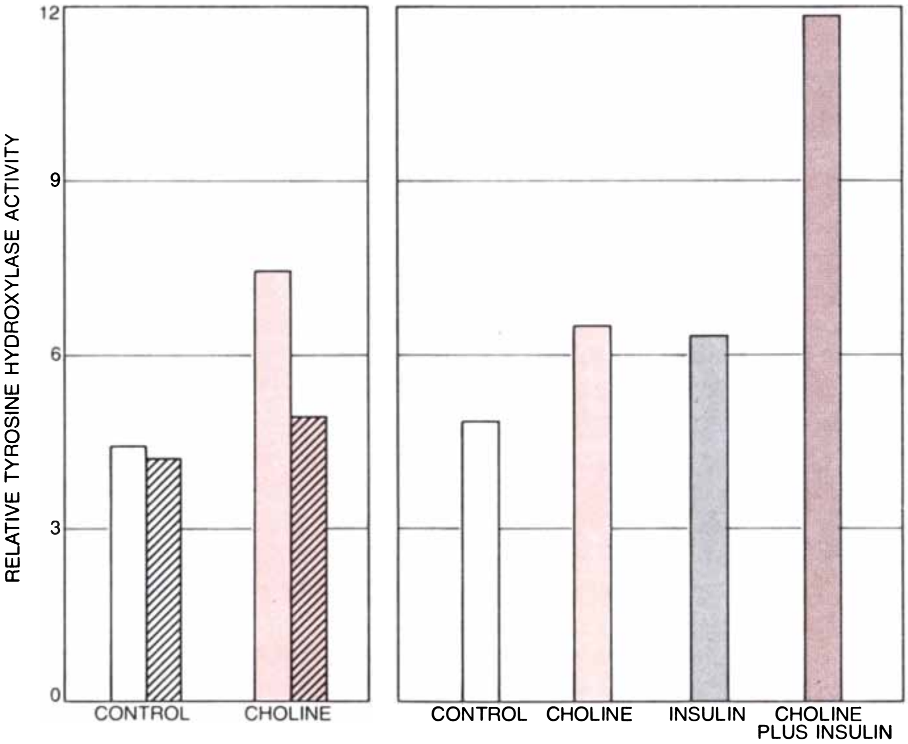
These experiments led to two conclusions. One was that giving a nutrient molecule that acts as a precursor can enhance the release of a neurotransmitter and thereby amplify a neuron's effect on postsynaptic cells. In the instance studied the administration of choline leads to the release of more acetylcholine and hence to greater tyrosine hydroxylase activity. The other conclusion, which had not been expected, was that the magnitude of choline's effect varies with the firing frequency of the neurons on which it acts. The latter finding provided the first suggestion of the important relation, which I mentioned at the outset, between the firing frequency of a neuron and the extent to which it responds to an increased supply of its transmitter's precursor.

Within four months of the publication of our first article describing the increase in brain acetylcholine in rats given choline, the first clinical application of the nutrient-neurotransmitter relation was reported. The report concerned a patient suffering from tardive dyskinesia, a disease characterized by uncontrollable movements of the face and upper body; it is caused (in a substantial fraction of patients) by the prolonged administration of antipsychotic drugs and is considered by many psychiatrists to be the most important side effect limiting the use of those drugs.

Kenneth L. Davis and his colleagues



EFFECT OF SEROTONIN on food choice was demonstrated by an experiment in which rats were offered their choice of two diets. Each diet provided the same amount of carbohydrate and the same number of calories, but diet *A* was low in protein (5 percent), whereas diet *B* was protein-rich (45 percent). On the first day the rats were injected with a placebo. On the second day they were injected with fenfluramine, a drug that increases serotonin release. The additional serotonin caused a selective reduction in the amount of food consumed from diet *A*, so that the rats' carbohydrate intake decreased but protein intake stayed about the same.



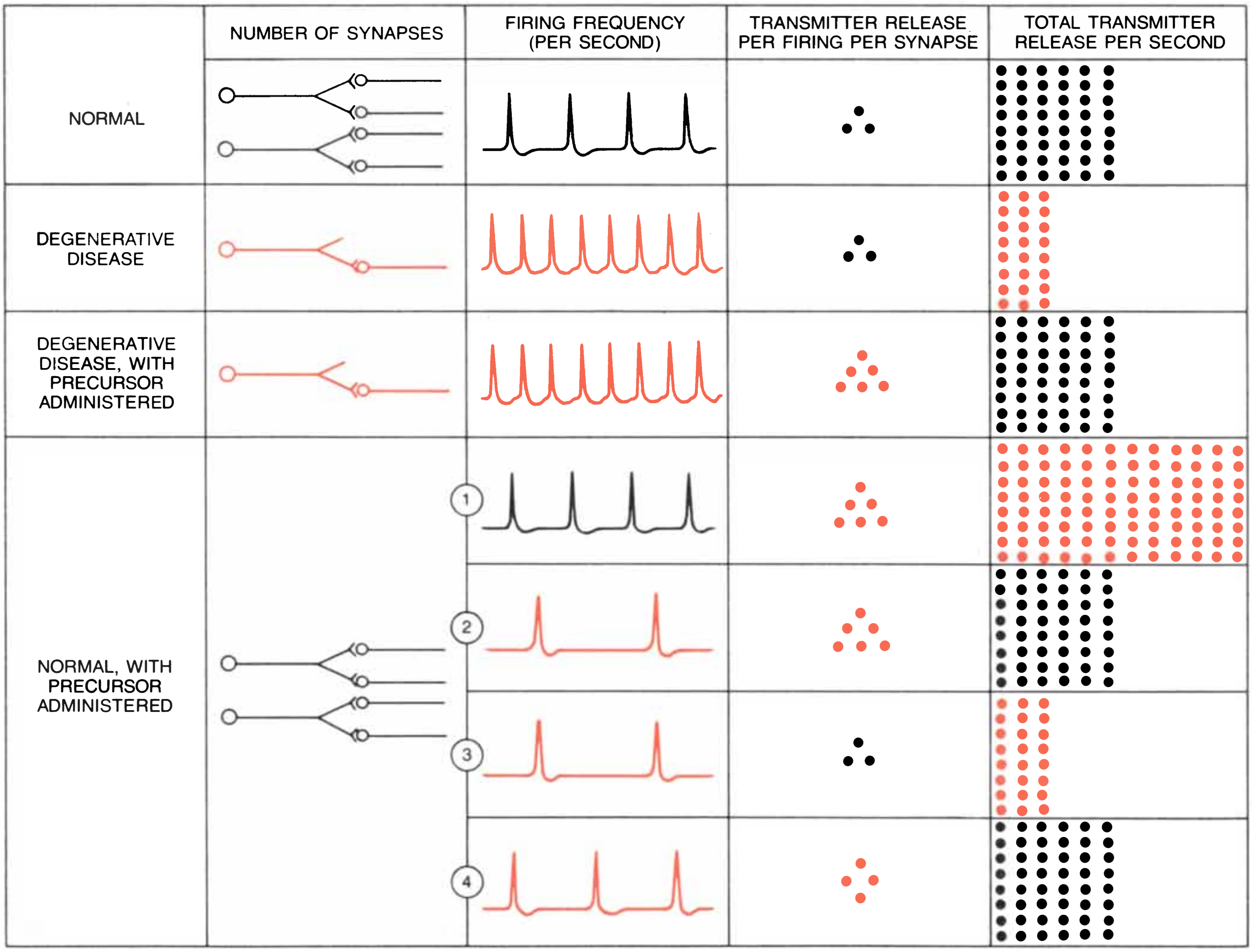
EFFECT OF CHOLINE on signal transmission by acetylcholine from splanchnic-nerve terminals to the chromaffin cells of the adrenal gland was assessed by measuring the activity of tyrosine hydroxylase, an enzyme whose synthesis is augmented by acetylcholine. In one experiment (*left*) one of the two adrenal glands was denervated (*hatched bars*). Giving choline for several days significantly increased enzyme activity in the intact gland but not in the denervated one. In a second experiment (*right*) two treatments were tested in rats, first alone and then in combination. Giving choline increased enzyme activity; so did the injection of insulin, which accelerates the firing of the splanchnic nerve. The two treatments potentiated each other: their combined effect was more than twice as great as the sum of their individual effects.

at the Stanford University School of Medicine had found that giving patients physostigmine, a drug that increases acetylcholine levels (by blocking the degradation of the transmitter by the enzyme cholinesterase), could temporarily ameliorate the abnormal movements. They then showed that increasing brain acetylcholine by giving choline has a similar effect. Other investigators have since confirmed the efficacy of choline in many patients with tardive dyskinesia, including a large group of people treated in a double-blind, placebo-controlled test by my collaborator John H. Growdon of the Tufts University School of Medicine.

Choline has largely been replaced by lecithin as the precursor of choice for raising acetylcholine levels in patients because lecithin is less susceptible to bacterial degradation in the intestine. The degradation wastes the choline and also leads to the formation of trimethylamine, which can impart the aroma of rotten fish to the hapless patient. Unfortunately "lecithin" means a specific chemical compound, phosphatidylcholine, to physicians and biochemists, whereas in the food industry it encompasses an entire family of substances, the phosphatides. Only phosphatidylcholine is an effective source of choline; virtually all the lecithin preparations sold in health-food stores are too impure to be of much value to patients.

The administration of choline or lecithin does not interfere with the therapeutic effects of antipsychotic drugs, only with their side effect, tardive dyskinesia. Moreover, unlike physostigmine and other inhibitors of cholinesterase, choline and lecithin do not give rise to the side effects associated with increased acetylcholine activity itself. The amplified acetylcholine transmission that follows the administration of physostigmine, for example, leads to excessive formation of mucus in the respiratory passages, a very low heart rate, gastrointestinal cramping and numerous alterations of brain function. What explains the absence of such cholinergic side effects after the administration of lecithin or choline?

The answer goes back to the interdependence of firing frequency and neurotransmitter synthesis [*see illustration below*]. Not all neurons that are potentially able to release more transmitter molecules in the presence of more precursor actually do so. The brain can cause particular groups of normally functioning neurons to diminish their firing frequency, so that although they synthesize more transmitter molecules, they do not release more molecules per unit time. The diminished firing frequency has a second effect: it somehow reduces the sensitivity of the slower-firing neurons to the presence of additional precursor. In tardive dyskinesia the neurons that



FIRING FREQUENCY AND PRECURSOR SUPPLY are interrelated. A hypothetical set of four normal synapses is shown at the top; the neurons fire four times per second, and at each synapse three transmitter molecules are released per firing. The result is a total release of 48 molecules per second. A hypothetical degenerative disease reduces the number of neurons and synapses. The remaining neuron fires faster, but transmitter release is still inadequate. The administration of additional precursor increases the number of transmitter molecules released at each firing and thus achieves a normal total rate of release. If additional precursor is supplied to a normal set of neurons, the response is different. First there is a transient increase in total release because more transmitter is synthesized and more is released at each firing (*1*). Soon, however, the neurons' firing rate is slowed, reducing transmitter release to the normal level (*2*). The reduced rate in turn somehow reduces the neurons' sensitivity to the additional precursor; less transmitter is released per firing (*3*). Eventually a combination of firing rate and release per firing is reached that maintains normal total rate of transmitter release (*4*).

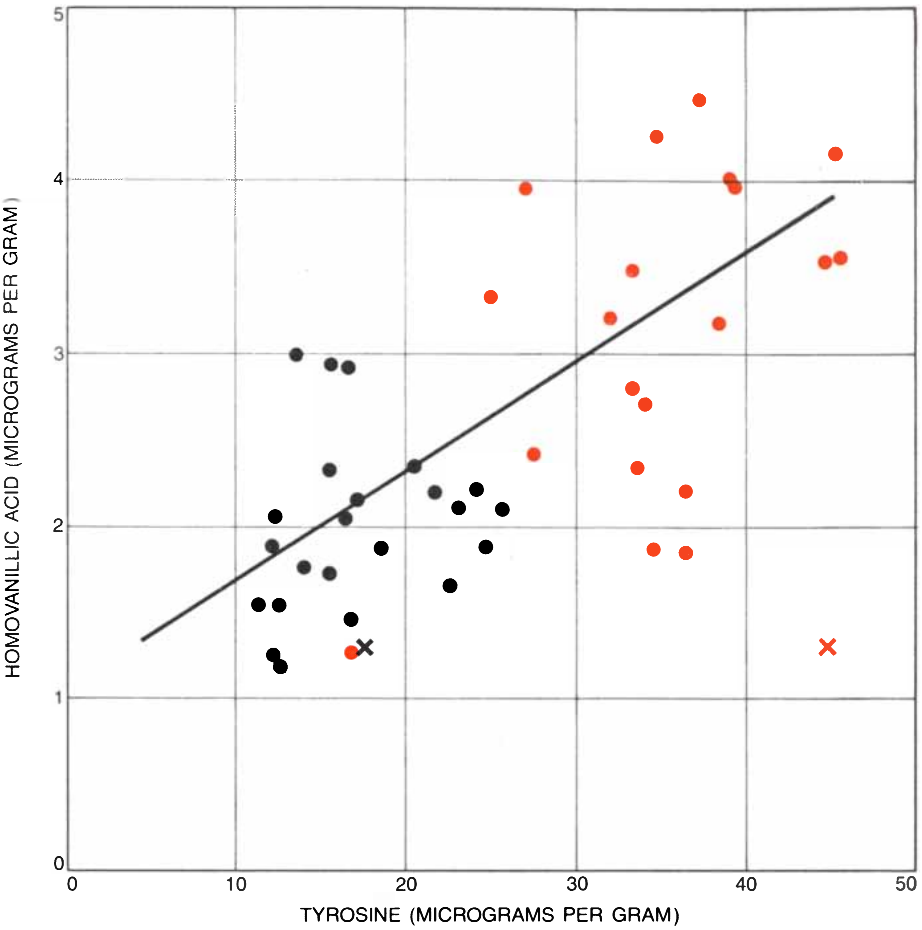
suppress the symptoms are apparently sensitive to the additional supply of choline, and so their action is amplified. Other cholinergic neurons first slow their firing and then become unresponsive to additional choline; as a result their release of acetylcholine does not increase and the side effects associated with overstimulation of the cholinergic neurons are absent.

Any disease state known to result from inadequate cholinergic neurotransmission, in any part of the body, now becomes a candidate for treatment with lecithin given either alone or as an adjunct to drug therapy. A number of brief reports have described improvement after lecithin administration to individual patients suffering from such diseases. Except in the case of tardive dyskinesia, however, too little information is available to sustain even tentative conclusions as to lecithin's therapeutic efficacy.

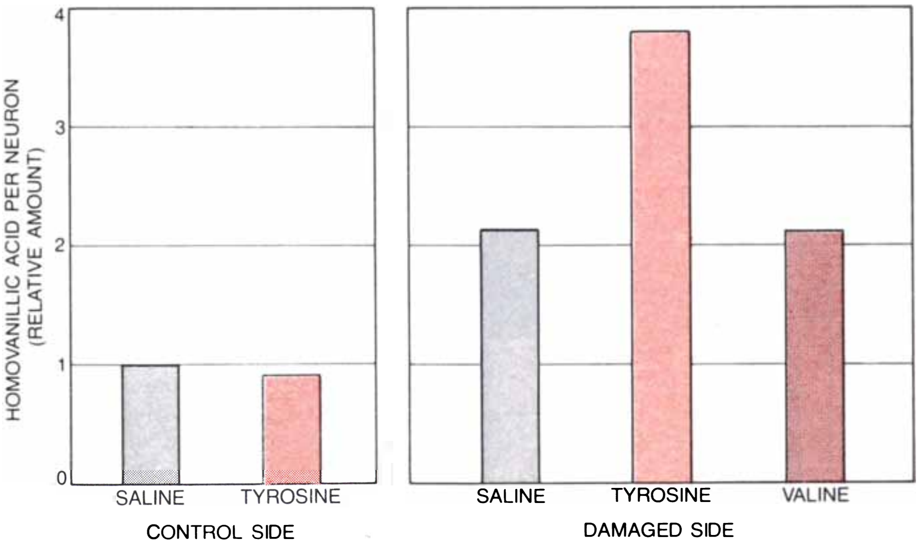
The diseases currently generating the most interest as candidates for lecithin therapy are the memory disorders associated with old age. Aging brings with it a loss of neurons in the brain, and cholinergic neurons seem to be particularly vulnerable. The hippocampus, a region of the brain known to be essential for the formation of new memories, has a particularly large number of cholinergic neurons. The administration to young people of drugs such as scopolamine, which block cholinergic transmission, causes short-term memory impairments similar to those observed in the aged. For these reasons it seems possible that treatments calculated to increase brain acetylcholine may be effective in some patients with memory disorders.

Investigators who would study possible therapies for these disorders are beset by a number of problems. One problem is the lack of a clinical basis for classifying memory-impaired patients into distinct subgroups according to the origin of the disability. It seems likely that the population of patients now said to have senile dementia or Alzheimer's disease will one day be found to consist of people with several different diseases, only some of which may reflect a selective decrease in brain acetylcholine and so perhaps be amenable to choline therapy. For now there is no way to distinguish such patients.

Another problem is the lack of well-validated tests for measuring improvement in memory functions. There are objective tests of memory, but in no case has it been possible to demonstrate that an improvement in a test score presages improvement in the patient's real-life memory functions. The tests cannot be validated because (it is a vicious circle) no treatment has yet been discovered that demonstrably improves real-life memory functions. Furthermore, the substances available for testing are less than satisfactory. Most partially puri-



ADMINISTRATION OF TYROSINE leads to the accumulation of homovanillic acid, a breakdown product of dopamine, in the corpus striatum of the rat, but only when the firing of dopamine-releasing neurons is speeded up. Giving tyrosine alone (*colored* X) increases the concentration of tyrosine in the brain compared with the level in control animals (*black* X), but it has no effect on dopamine release (as measured by the homovanillic acid level). Giving haloperidol, a drug that accelerates the firing of the neurons, enhances the release of dopamine and also makes the synthesis and release of dopamine dependent on the brain level of tyrosine (*black dots*); now administration of tyrosine markedly increases dopamine release (*colored dots*).



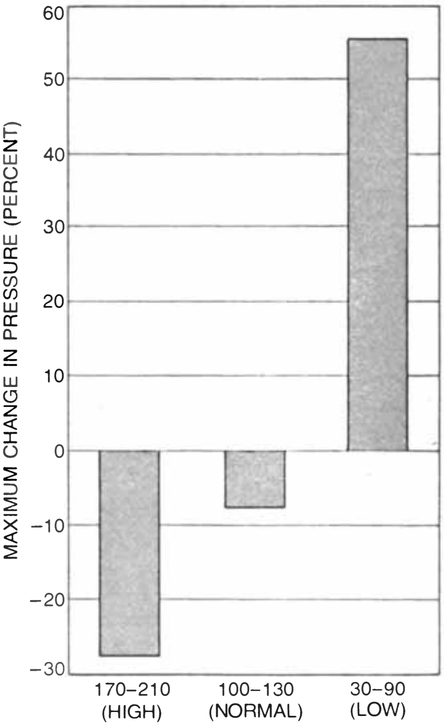
EFFECT OF FIRING FREQUENCY on dopamine release was also demonstrated by an experiment in which 75 percent or more of the dopamine-releasing neurons on one side of the rat brain were destroyed, causing the remaining neurons to fire more rapidly. The injection of tyrosine rather than a presumably inert saline solution does not significantly change the amount of dopamine released per neuron on the intact side of the brain. On the damaged side the increased firing rate of the remaining neurons is manifest: in animals given saline solution the amount of dopamine released per neuron is about doubled. Giving tyrosine leads to a further increase in dopamine release; giving valine, another large, neutral amino acid, does not.

fied lecithin preparations either contain too little phosphatidylcholine or are unpalatable in the large doses that seem to be necessary. The standard drugs with which investigators enhance cholinergic function in animals act too nonspecifically on cholinergic synapses everywhere (and so cause too many side effects) for testing in human beings.

The few reports that have appeared on the memory responses of patients given lecithin suggest that it can be helpful in some cases but is not by any means a fountain of youth. Perhaps combinations of lecithin with drugs that accelerate the firing of cholinergic neurons in the hippocampus or with low doses of cholinesterase inhibitors will be useful in treating patients who have memory disorders attributed to a localized cholinergic deficit. Growdon and Suzanne H. Corkin of M.I.T. are currently testing several such combinations.

The close relation between the firing frequency of a neuron and its response to an increased supply of its transmitter's precursor is best illustrated by a group of dopamine-releasing neurons, those that run from the substantia nigra in the midbrain to the corpus striatum deep within each cerebral hemisphere. These nigrostriatal neurons participate in the control (particularly in the initiation) of movements and are severely damaged in Parkinson's disease.

The release of dopamine from the nigrostriatal neurons can be assessed in animals by measuring the accumulation in the corpus striatum of a dopamine metabolite, homovanillic acid. If otherwise untreated animals receive even a large dose of tyrosine, the nutrient that is the precursor of dopamine, no change is noted in the concentration of homovanillic acid in the corpus striatum. If the animals are pretreated with a drug that accelerates the firing of the nigrostriatal neurons, however, dopamine release becomes sensitive to the tyrosine level.



EFFECT OF TYROSINE on blood pressure depends on the initial pressure. The amino acid lowers the blood pressure of hypertensive rats and raises it in hypotensive animals.

This effect was demonstrated elegantly by Franz Hefti and Eldad Melamed (who at the time were working with me at M.I.T. but who are now respectively at the Sandoz Research Laboratories in Basel and the Hebrew University of Jerusalem Faculty of Medicine). By administering a modified form of dopamine that damages dopaminergic neurons, they disabled more than 75 percent of the nigrostriatal neurons in one hemisphere of the rat brain. The surviving neurons in the damaged hemisphere responded by increasing their firing frequency; they are probably stimulated to do so by a feedback mechanism that enables nigrostriatal neurons to take over one another's functions. (A similar compensatory process keeps most Parkinsonian patients from showing symptoms until at least half of their nigrostriatal neurons have been lost.) The accelerated firing was confirmed by showing that the homovanillic acid level per surviving neuron was much higher on the damaged side of the brain than on the intact side. The animals then received tyrosine. On the damaged side there was a further increase in the homovanillic acid level per neuron; on the intact side there was no increase.

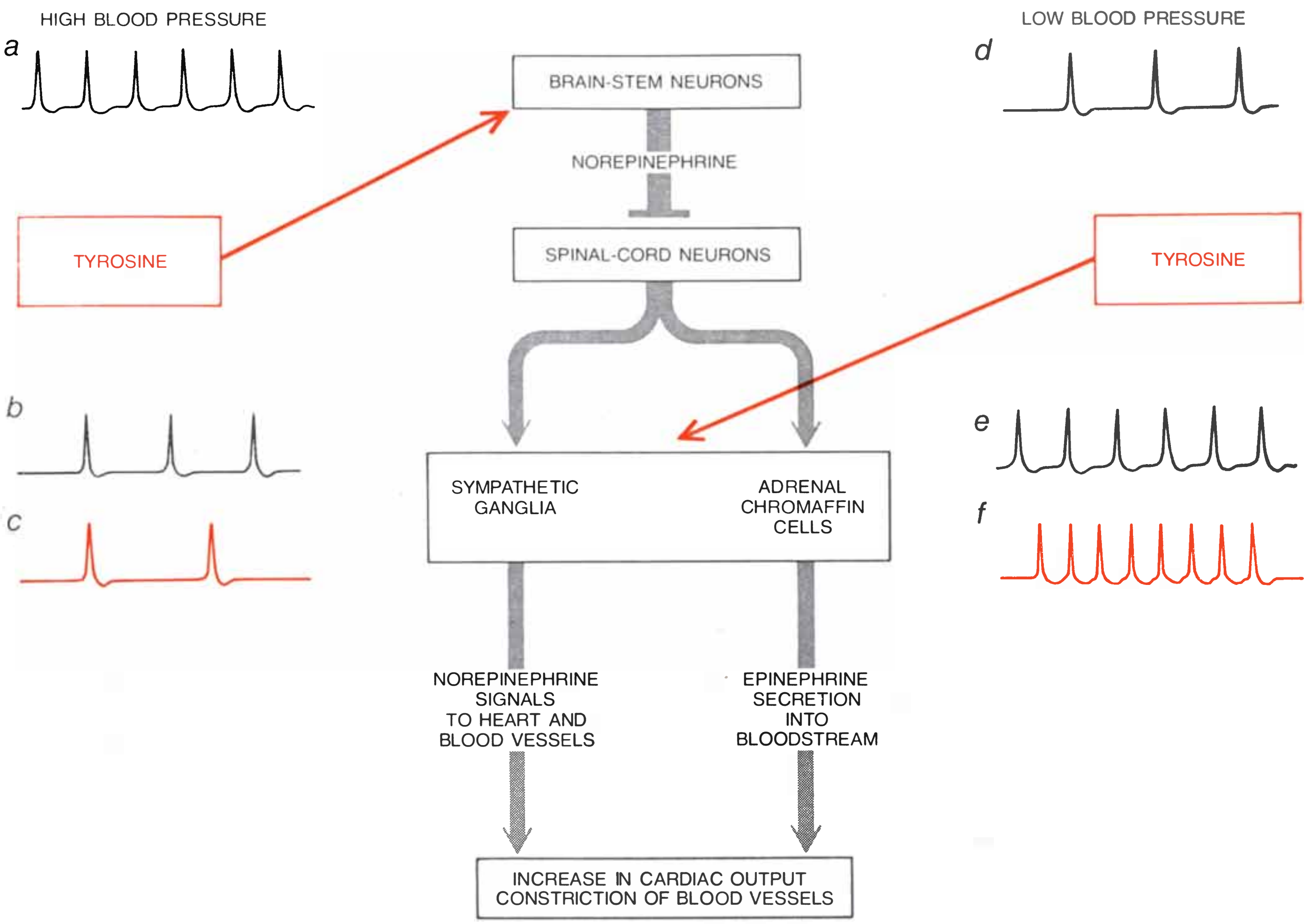
Tyrosine in the diet can also either increase the synthesis of norepinephrine or leave it unchanged depending on the firing frequency of particular neurons. This property explains the remarkable effects of tyrosine on blood pressure. When tyrosine is administered to animals (or people) with normal blood pressure, it has no consistent effect on the pressure. When tyrosine is given to hypertensive rats, however, it markedly lowers their blood pressure; when it is given to animals with hypotension (animals in shock, for example), it raises their blood pressure to near-normal levels. These observations were made by my graduate students Alan F. Sved, Lydia Conlay and Timothy J. Maher, who worked with rats that were genetically hypertensive and with other rats whose blood pressure had been lowered by reducing their blood volume by about 20 percent.

In a hypertensive animal the brain, acting to reduce blood pressure, accelerates the firing of norepinephrine-releasing neurons in the brain stem. At this site norepinephrine acts as an inhibitory neurotransmitter: it suppresses the firing of other neurons, ultimately diminishing the activity of the peripheral sympathetic neurons and the chromaffin cells of the adrenal medulla. The sympathetic neurons and chromaffin cells therefore liberate less norepinephrine and epinephrine, which would otherwise act to raise blood pressure by causing blood vessels to constrict and cardiac output to rise. Since only the brain-stem neurons are firing frequently, only their inhibitory output is amplified by tyrosine administration, and so the blood pressure falls. In animals in shock, on the other hand, the brain acts to raise the blood pressure: the inhibitory norepinephrine-releasing neurons of the brain stem are suppressed, whereas the sympathetic neurons and chromaffin cells are activated to fire more frequently. In this instance tyrosine selectively enhances catecholamine release from those cells, and the blood pressure rises.

Because of this regulatory mechanism tyrosine has a distinct theoretical advantage over many of the drugs now favored for the treatment of circulatory disorders. In theory, at least, it can be expected to act without overshooting the mark because as soon as normal blood pressure is attained, selective changes in neuronal firing frequencies should render the animal insensitive to the effects of additional tyrosine. Tyrosine has yet to be systematically tested in people with hypertension or shock, however, and theoretical arguments are no substitute for well-designed tests of clinical efficacy and safety.

Another human disease state in which the possible therapeutic value of tyrosine is currently being tested at several institutions is depression. Most psychiatrists who seek biochemical explanations of mental illness think that in many patients depression reflects inadequate neurotransmission mediated by either norepinephrine or serotonin. If norepinephrine release is inadequate in certain regions of the brain of some depressed patients, the administration of tyrosine could conceivably be helpful to them. The first evidence that tyrosine can have an antidepressant effect was obtained in studies we did in collaboration with Alan J. Gelenberg of the Massachusetts General Hospital; the findings were quickly confirmed elsewhere. If further clinical testing shows tyrosine can help in treating depression, it will be interesting to learn how it does so. In patients who respond to tyrosine is something wrong with the metabolism of the amino acid (causing the plasma tyrosine ratio to be too low), or is the fault perhaps in the conversion of tyrosine into norepinephrine in brain neurons?

When large doses of a nutrient, separated from the other constituents of foods that are its usual source, are given to people specifically to treat a disease or condition, does the nutrient thereby become a drug? The question is not merely one of nomenclature. The



ABILITY OF TYROSINE to both reduce high blood pressure and raise low blood pressure is explained by the effect of the amino acid on different populations of cells. In response to high blood pressure norepinephrine-releasing neurons in the brain stem speed their firing (*a*). The effect of the norepinephrine released by these neurons is inhibitory: by suppressing the firing of preganglionic neurons of the sympathetic nervous system in the spinal cord it reduces the activity of neurons in the sympathetic ganglia and of the adrenal chromaffin cells (*b*), activity that ordinarily tends to maintain or raise blood pressure. When tyrosine is administered, it affects the rapidly firing brain-stem neurons; those neurons make more norepinephrine, their inhibitory effect is augmented (*c*) and the blood pressure falls. In response to low blood pressure, on the other hand, the firing of the brain-stem neurons is suppressed (*d*); as a result their inhibitory effect is reduced and the activity of the ganglia and the adrenal cells is increased (*e*). Now it is the high-activity sympathetic and chromaffin cells that are sensitive to increased tyrosine when the amino acid is administered. Their activity is further enhanced (*f*) and blood pressure rises.

answer may determine whether these compounds are adequately tested for therapeutic efficacy and safety and, if they pass such tests, whether they become available for clinical applications. Calling them drugs might lead regulatory agencies to require preliminary tests so laborious and expensive that they frighten away some of the most likely developers: the food companies that produce lecithin and protein constituents for inclusion in foods.

Clearly the prescription of these substances by physicians (or their recommendation for use in self-medication) must await the accumulation of evidence that they are safe and effective. Need the evidence be of the same weight for a ubiquitous food constituent as the evidence legitimately required for a drug? It probably need not be, for a number of reasons. Amino acids and choline are rapidly metabolized in the body by enzymes that have been doing the job for as long as animals have existed. How an individual molecule of one of these nutrients is metabolized seems to be largely unrelated to whether it is ingested as the pure substance or as a constituent of a food protein or lecithin. The precursors are water-soluble and are rapidly exchanged between the bloodstream and various tissues by diffusion (or, in the brain, by a transport system); their tissue concentration, unlike the concentration of many drugs, does not continue to rise with repeated administration. Perhaps more important, the brain must acquiesce, so to speak, in most of their effects on neurotransmission: by means of the selective response of neurons to additional precursor the brain modulates the precursor's effects on the brain.

The amounts of these nutrients that must be administered in order to modify neurotransmitter synthesis are relatively large, reflecting the fact that the nutrients are components of a normal diet; their normal blood level and daily intake are not zero, as is the case with a typical drug about to be given to a patient for the first time. Hence the most convenient way of administering them might often be as a constituent of a "special food for medical uses" prepared specifically for that purpose. Such terminology should not imply, of course, that the diseases and conditions in which they might be therapeutic are nutritional in origin. Tardive dyskinesia is not the result of lecithin deficiency, nor is shock the result of tyrosine deficiency. In such conditions the nutrient would be given for its pharmacological effect, as coffee, brandy, bran or prune juice can be given, as if it were a drug.

In a broader sense, however, it may be correct to think in terms of nutrition. Is it possible that old people in whom the brain has lost many of the neurons that release precursor-dependent transmitters might benefit from a routine diet enriched in tyrosine or tryptophan or lecithin? If they might, should their continued reliance on normal but unenriched diets be construed as constituting poor nutrition? I wonder.

A Major Earthquake Zone on the Mississippi

Three of the strongest quakes in U.S. history were not in the West but in New Madrid, Mo. They came in 1811 and 1812, and even now the area is shaken by an average of one small quake every other day

by Arch C. Johnston

At the mention of earthquakes most Americans think of the great San Francisco earthquake of 1906 and the continuing threat to California represented by the San Andreas fault. The fact is that three of the greatest earthquakes in American history occurred 170 years ago not on the Pacific coast but near a placid meander loop of the Mississippi River where the states of Kentucky and Missouri meet. On December 16, 1811, and again on January 23 and February 7, 1812, the river region was struck by earthquakes with respective estimated magnitudes of 8.6, 8.4 and 8.7. The tremors were felt throughout the northeastern U.S. and parts of Canada. They cracked plaster in Richmond, Va., and rang church bells and cracked pavements in the District of Columbia. To this day the region, in the vicinity of New Madrid, Mo., remains seismically active; indeed, it is shaken by small earthquakes on an average of every 48 hours. Why should an area in the middle of the continent be so susceptible to earthquakes? Some decades of research have led to the surprising conclusion that here, near the northern apex of what geologists call the Mississippi embayment, there is a rift in the continental craton: that part of the great North American plate consisting of the oldest and strongest rock in the continental crust. This ancient rift has been reactivated by tectonic forces.

The Mississippi embayment is a wide *U*-shaped expanse bisected by the river. It begins near the confluence of the Ohio and Mississippi rivers and widens southward into the Gulf of Mexico coastal plain. To the east of the Mississippi are gently rolling hills capped by deposits of loess (windblown silt), an indication that the river never flowed there. To a traveler in the western part of the embayment its most arresting feature is its unyielding flatness. Here sand and gravel terraces, natural levees, drainage divides and abandoned meander loops all testify to the fact that the river often wandered far to the west of its present course.

The town of New Madrid lies on the north bank of a Mississippi meander loop known as the Kentucky Bend. It was founded late in the 18th century (in what was then Spanish territory) with the intent of establishing a trade center for the barge and flatboat traffic on the river. By early in the 19th century its hundreds of inhabitants made New Madrid the largest settlement in the region. Thus the town lent its name to the extraordinary events of 1811 and 1812.

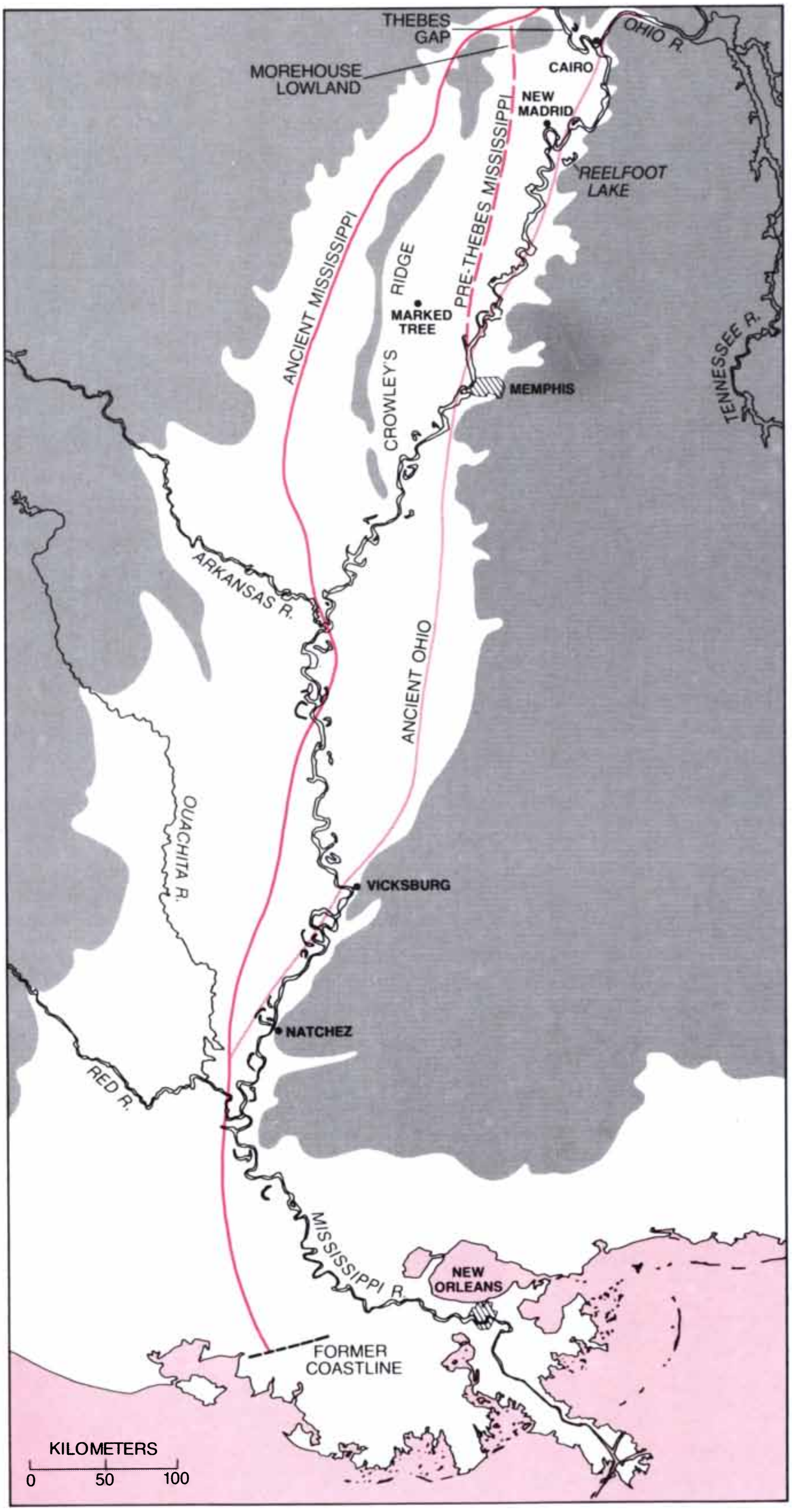
The three massive earthquakes that winter were each followed by a protracted series of strong aftershocks. The effects were awesome. The saturated bottomland soil spurted huge geysers of sand and water. During the main shocks the land heaved and buckled for minutes, trees were splintered, a prolonged roaring was heard and the air smelled of sulfur. According to one account, the aftershocks were a constant trembling, "like the flesh of a beef just killed."

The effect of the third and strongest main shock on the Mississippi itself was even more striking. As an eyewitness wrote, at "about 4 o'clock A.M. a concussion took place so much more violent than those which preceded it that it was denominated the hard shock.... At first the Mississippi seemed to recede from its banks... its waters gathering up like a mountain [only to fall] again with such violence that it took with it whole groves of young cotton wood trees which ledged its borders.... The river was literally covered with the wreckage of boats." The riverbed itself was so disturbed that two waterfalls (or at least rapids) were formed, one upstream from New Madrid and one downstream.

The recent geological history of the Mississippi embayment is closely related to the last of the great glacial advances of the Pleistocene, which in North America ended some 8,000 years ago. At its maximum this advance—the Wisconsin glaciation—covered much of the north-central U.S., extending as far south as southern Illinois. So much water was locked up in the vast ice sheets of the Wisconsin period that the level of the sea was 60 to 80 meters lower than it is today, and most of the continental shelves were dry land. Rivers to the north of the Ohio and the Missouri, which formerly flowed north into Hudson Bay and the St. Lawrence River, were blocked by the ice and were diverted south, augmenting the Mississippi drainage system.

The lowered level of the Gulf of Mexico meant a steeper gradient for the south-flowing rivers. This factor and others (such as flooding) caused both the Mississippi and the Ohio (which then did not join the Mississippi until it reached the vicinity of modern Natchez) to cut deepened valleys through the Mississippi embayment. By the time the Mississippi reached the edge of the continental shelf its channel was more than 100 meters deeper than it is today. The melting of the glaciers brought a rise in the sea level and the release of huge amounts of glacial debris, so that the deeply entrenched river valleys were soon filled with alluvial sediments.

VICINITY OF NEW MADRID is seen in false color in the aerial photograph on the opposite page, part of a survey conducted by the National Aeronautics and Space Administration on behalf of the U.S. Army Corps of Engineers. The Mississippi meander at the far left is the Kentucky Bend; part of the town of New Madrid is visible on the river's north bank. The time is fall and some of the farm fields that make up the rectangular patterns on the fertile bottomland have been cleared of crops (*vegetation appears red*). At the lower-right corner is the northeastern end of Reelfoot Lake, a Tennessee landmark. A map including the area shown in the photograph (which covers part of the Lake County uplift, a seismic zone) appears on page 64.



PRESENT COURSES of the Mississippi and the Ohio, which now meet near Cairo, Ill., are much different from the courses of late Pleistocene times. The Ohio (*light color*) did not join the Mississippi until it reached the vicinity of modern Natchez, and the Mississippi (*dark color*) ran well to the west of its present course for most of its length. The later Morehouse diversion (*broken colored line*) then brought the two rivers together near modern Memphis, rather than below Thebes Gap as at present. Gray areas to the east and west approximate the 250-foot contour line. The low area between the Tennessee River and the western contour line is known as the Mississippi embayment. It widens into the Gulf coastal plain. Crowley's Ridge is the by-product of valley widening and erosion by the ancient rivers on both its east and west flanks.

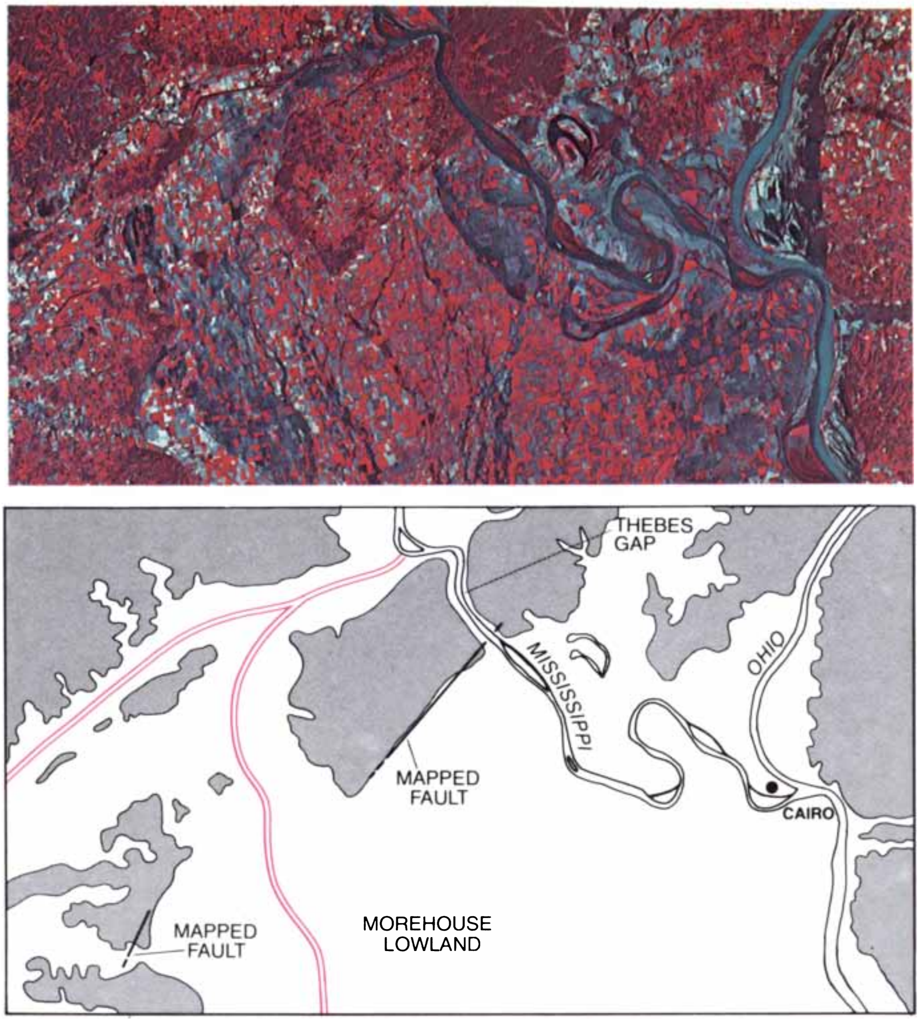
In the 1940's Harold N. Fisk of the U.S. Army Corps of Engineers established a remarkable fact about the Mississippi-Ohio system in late Pleistocene times, some 18,000 years ago. The valley of the Mississippi, Fisk found, was then far to the west of the river's present channel, and the valley of the Ohio passed through what is now eastern Arkansas and western Mississippi. After the retreat of the Wisconsin ice the Mississippi channel was diverted eastward from the western lowlands of the Mississippi embayment until finally it ran through a narrow pass at the apex of the embayment that is today called Thebes Gap.

What caused such a drastic shift in the channel? Fisk considered two possibilities: a gradual deepening of Thebes Gap by periodic erosion during floods and a headward erosion of a tributary of the Ohio that finally resulted in the "capture" of the Mississippi. A third possibility is that the Thebes Gap diversion was tectonic, specifically the result of faulting.

There is little doubt that tectonic influences have controlled the course of the Mississippi in certain places. An example is found in Lake County, Tenn., in the heart of the New Madrid seismic zone. In 1974 Saint Louis University, with support from the U.S. Geological Survey, installed a network of modern seismic instruments in the New Madrid zone. It was augmented in 1979 by the installation of a second network by Memphis State University, with support from the Nuclear Regulatory Commission. The second network is operated by the university's Tennessee Earthquake Information Center. The New Madrid zone is monitored 24 hours per day, and the networks can detect and locate disturbances as small as magnitude 1.0 (an energy release equivalent to the impact of a two-ton weight dropped 12 feet).

Some 700 square kilometers in Lake County are uplifted; the central portion of the zone is from three to 10 meters higher than the surrounding region. This may at first seem insignificant, but it is hard to explain such an uplifted surface in the middle of a level river floodplain. The readings of the seismic network show that the uplifted area overlies the part of the New Madrid seismic zone where the earthquake activity is most intense. Most of the events are microquakes, with magnitudes of less than 2.5 (roughly the threshold of human perception). David P. Russ of the Geological Survey, who has studied the Lake County uplift, considers it responsible for the deflection of the Mississippi channel to the west between Hickman, Ky., and Caruthersville, Mo. He points to anomalous river gradients and channel geometries in the area, and he views the uplift as a surficial response to movement on deep-seated faults. The fact that the area coincides geographically with the zone of intense microquake activity suggests that most of the fault movement is seismic, that is, a sudden displacement that generates seismic waves rather than a gradual release of accumulated crustal strain.

If the Hickman deflection is a product of recent tectonic influences, might not the Thebes Gap diversion be the result of older ones? The gap is about 30 kilometers north of New Madrid. Although the area is not now associated with any marked concentration of microquakes, numerous faults have been mapped in it. Pending further studies such a hypothesis seems plausible albeit speculative.



LANDSAT IMAGE of the area surrounding the confluence of the Mississippi and the Ohio (*top*) shows the narrowness of Thebes Gap. The map of the same area (*bottom*) locates faults on two upland blocks (*colored lines*) traced by the Missouri Geological Survey and indicates two former courses of the Mississippi. The river's shift to Thebes Gap may be the result of flood erosion or stream piracy, but the presence of a fault suggests a tectonic involvement.

An explanation of the seismic activity in the New Madrid region is far more difficult to frame than an explanation of the activity along the San Andreas fault. For example, the California fault has a prominent surface scar with measurable offsets and well-mapped main and auxiliary traces. The opposite is true at New Madrid: the faults are deeply buried under the river sediments and the even thicker underlying marine sediments deposited on the floor of an ancient sea. No permanent surface fault traces, such as the waterfalls or rapids produced by the hard shock of February, 1812, are now found anywhere in the Mississippi embayment; any that may once have existed have long ago been removed by erosion and the reworking of sediments.

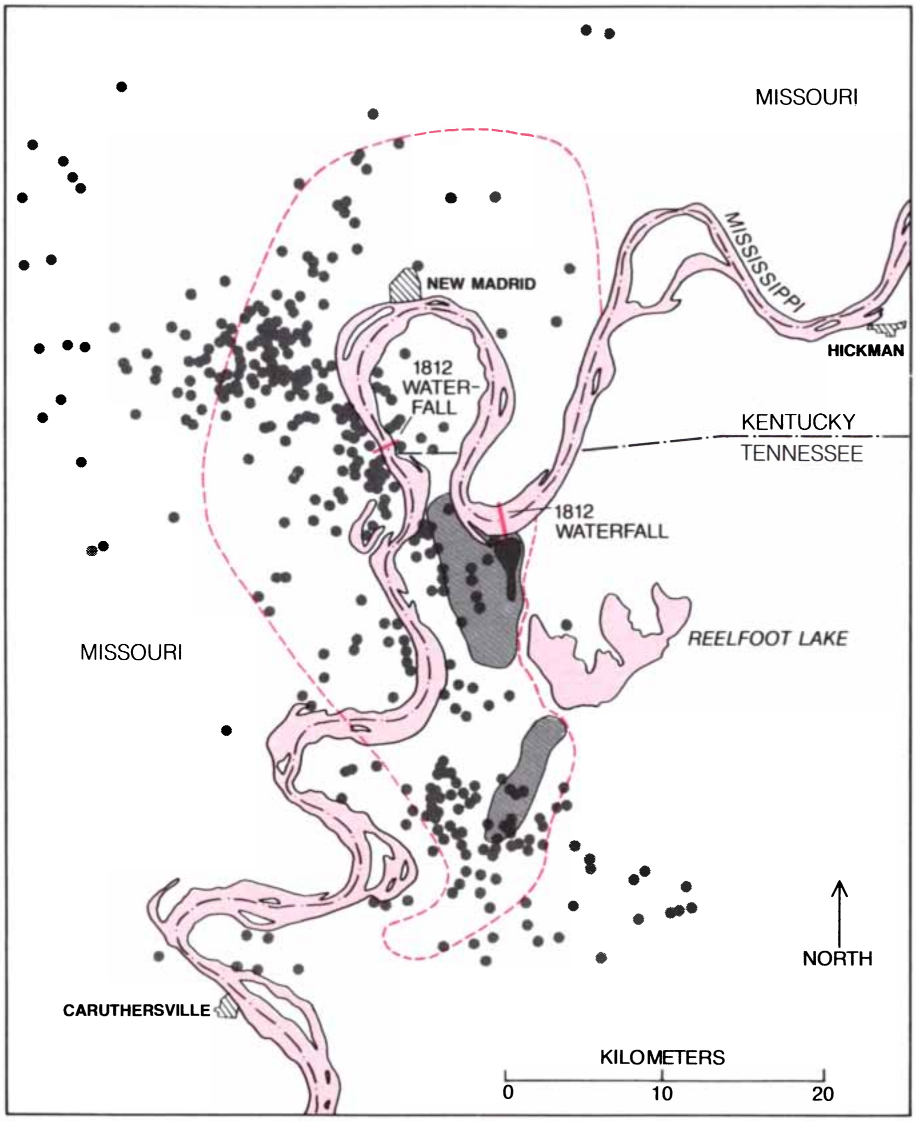
The fundamental difficulty lies in explaining why the Mississippi embayment has earthquakes at all. The San Andreas, after all, is a plate-boundary fault and has been recognized as such ever since the formulation of the theory of plate tectonics in the early 1960's. A plate is a rigid, strong section of the lithosphere: the outermost rock layer of the earth that includes the crust and the upper mantle. It behaves as a single body, and it moves slowly in response to forces generated ultimately by the slow convective release of the earth's internal heat.

In California this movement averages about 5.5 centimeters per year as the Pacific plate slides northwest along the San Andreas fault system, the boundary between the Pacific plate and the North American plate. Where the movement is impeded, as it is today in northern and southern California, the fault is locked and is therefore storing up strain energy that will ultimately and inevitably be released. Given the fact that the average displacement in a major earthquake is from six to eight meters, the recurrence time for a large earthquake along the San Andreas fault is roughly 110 to 145 years, unless smaller and more frequent earthquakes dissipate the energy now stored in the fault.

The New Madrid seismic zone, however, is not a plate boundary, and no such obvious calculation of recurrence time can be made. Indeed, the earthquakes in the New Madrid Zone are of the rare kind known as midplate quakes, a class of disturbances that is of little significance on a worldwide scale. There is still no explanation of midplate earthquakes in plate-tectonic theory. The fact remains that the New Madrid earthquakes of 1811–12 are prominent examples of these generally rare events.

Data from the seismograph networks in the Mississippi embayment yield an informative picture of the continuing New Madrid activity. The seismometers have made possible the precise location of quake epicenters and hypocenters. (An epicenter is the surface projection of the hypocenter, or deep focus, of an earthquake.) Accurate timing is all-important to these determinations; an error of .1 to .2 second in the recorded time of arrival of a seismic wave can shift the calculated location of a hypocenter by several kilometers. As a result of the increased sensitivity of the newer instrumentation what formerly seemed a fairly random scattering of epicenters has now proved to be a much tighter grouping and one that exhibits several distinct linear trends. The epicenters of earthquakes of magnitude 1 or greater can be plotted with an uncertainty of one kilometer or less in latitude and longitude. The uncertainty with respect to the depth of the hypocenters, however, remains a matter of several kilometers.

A word about magnitude scales will be useful here. The first such scale was devised in 1935 by Charles F. Richter of the California Institute of Technology. Since earthquakes vary greatly in magnitude, he found a logarithmic scale convenient. Magnitude is therefore proportional to the logarithm (to the base 10) of the amplitude of a recorded seismic wave. An increase of one unit on the scale represents an increase in wave amplitude by a factor of 10, which in turn represents an increase in radiated seismic energy by a factor of about 32. Large earthquakes, say of magnitude 8 or larger, are thus many millions of times more energetic than an event of magnitude 1. For example, the New Madrid hard shock (magnitude 8.7) released in its seismic waves alone (not



RIVER TERRAIN in the New Madrid seismic zone includes a subtle feature known as the Lake County uplift. The broken line (*color*) outlines the entire uplift feature, mapped by David P. Russ of the U.S. Geological Survey. Gray areas above and below Reelfoot Lake are regions of higher uplift; the darkest zone adjacent to a bar (*color*) is the region of maximum uplift. It once stood more than 10 meters above the level bottomland. The colored bar and a second one downstream and to the west designate places where the earthquake of February, 1812, so disrupted the riverbed that waterfalls were formed. Reelfoot Lake was also enlarged as a result of the earthquakes in 1811–12. The uplift may have deflected the Mississippi westward near Hickman, Ky. Dots indicate some of the earthquakes recorded in the area since 1974.

including either the heat or the mechanical energy released) energy greater than the energy that would be released by the explosion of 150 million tons of TNT.

Today most of the New Madrid events are microquakes. In the period between mid-1974, when the modern seismograph network was established, and mid-1981 an average of more than 150 events per year with a magnitude greater than 1 have been recorded. Of this total of more than 1,000 quakes the local residents have reported feeling fewer than 50. This is about seven per year, or less than 5 percent of the total. Since the large earthquakes of 1811–12 only two events estimated at magnitude 6 or greater have been reported; the first was in 1843 and the second in 1895. There was no seismograph installation in the region until one was set up in St. Louis in 1909. Since then no quakes capable of doing more than minor damage have been recorded. The most recent earthquake, an event of magnitude 5, occurred near Marked Tree, Ark., on March 25, 1976. It was felt in seven states.

Was the 1811–12 series of severe earthquakes a unique event? Recent work by Russ suggests otherwise. Digging a trench at one location in the uplift area of Lake County, he exposed two zones of deformed and faulted sediments. Russ attributes the condition of the sediments to the action of large earthquakes. Fragments of shell from the disturbed sediments have yielded carbon-14 dates of 2,000 ± 250 years before the present. Hence the disturbances, which long predate the 1811–12 events, suggest a possible repeat time for large earthquakes on the order of 600 to 800 years.

The pattern of microquake epicenters seems to reflect the orientation of three deeply buried faults in the New Madrid region. Two main trends are apparent. One fault, possibly segmented, appears to strike southeast from the town of New Madrid into the Lake County region of northwestern Tennessee. The other, a longer fault, runs for 100 kilometers southwest out of the Missouri "boot heel" into northeastern Arkansas, where it ends abruptly at the town of Marked Tree, 60 kilometers northwest of Memphis. A third and less developed fault trend extends northward from the first fault to the vicinity of Cairo, Ill. For all three segments the depth measurements of microquake hypocenters range from five to 20 kilometers below the surface, a depth that places the faults mainly in the granitic basement rock of the upper and middle continental crust.

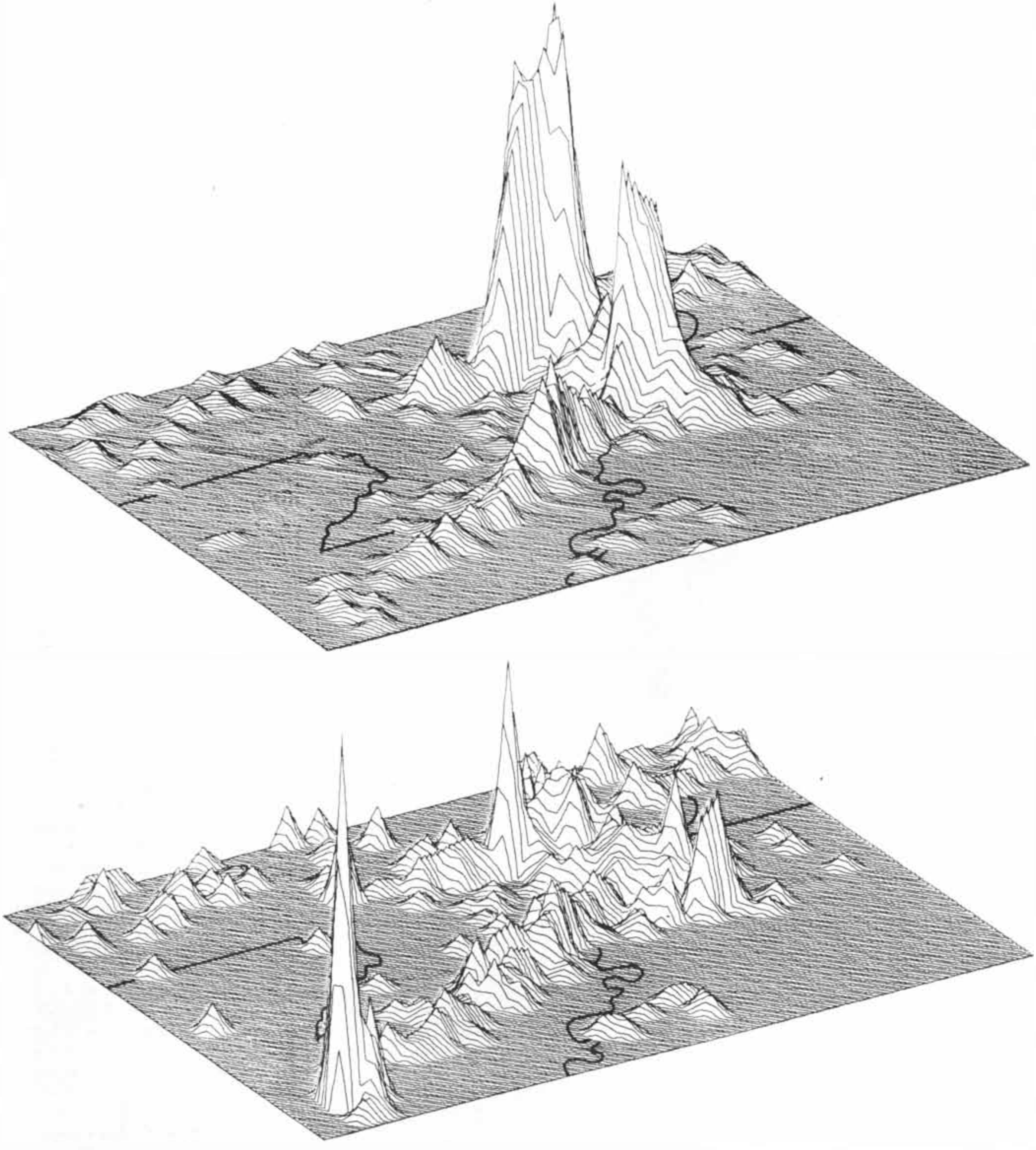
The two main faults differ both in their strain-energy release and in their mode of faulting. The first fault, which I shall call the cross-embayment segment, generates four times as many microquakes as the second fault, or axial segment, but releases only a fourth of the seismic energy. The axial segment produces fewer events, but the shocks of larger magnitude tend to be concentrated along it. This contrast may be due to differences in the stress levels; the rock of the axial segment may be stronger and less fractured than the rock of the cross-embayment segment and therefore able to store more strain energy between releases.

Faults can also differ according to the displacement of one side with respect to the other. For example, a steeply dipping fault that exhibits only horizontal movement is termed a strike slip. If the opposite side moves to the right, the fault is said to have a dextral offset; if it moves to the left, it is said to have a sinistral offset. As another example, if the overriding block of an inclined fault plane moves upward, it is said to be a thrust fault; if it slips downward, it is said to be a normal or gravity fault. These are more than mere classificatory distinctions; different types of faulting require a different orientation of the prevailing stress field. Strike-slip and thrust faulting require a horizontal maximum compressive stress; gravity faulting requires a vertical maximum compressive stress and a horizontal minimum compressive stress or even a tensional stress.

By studying the seismic wave patterns generated by New Madrid earthquakes Robert B. Herrmann of Saint Louis University has determined the dominant faulting types of the main fault segments. He finds that on the axial segment (and possibly on the Cairo segment) dextral strike slip predominates. The results for the cross-embayment segment are less clear but thrust faulting appears to be indicated. The presence of faults of these types is consistent both with the observed uplift in Lake County and with a single overall regional stress field: a maximum horizontal compressive stress with an east-west orientation. This agrees with the extensive North American stress measurements compiled by Mark and Mary Lou Zoback of the Geological Survey; these too indicate an east-west compressive stress field for much of the mid-continent.

The hypothesis is that the compressive stress is transmitted to the plate interior by interactions at the plate boundaries, or possibly by interaction of the rigid lithosphere with the more ductile part of the immediately underlying mantle. In either case the source of the stresses that generate the New Madrid earthquakes is plate motion and the source of plate motion is the heat of the earth's interior.

To know that plate motion is the source of the earthquake-generating stress regime in the New Madrid region

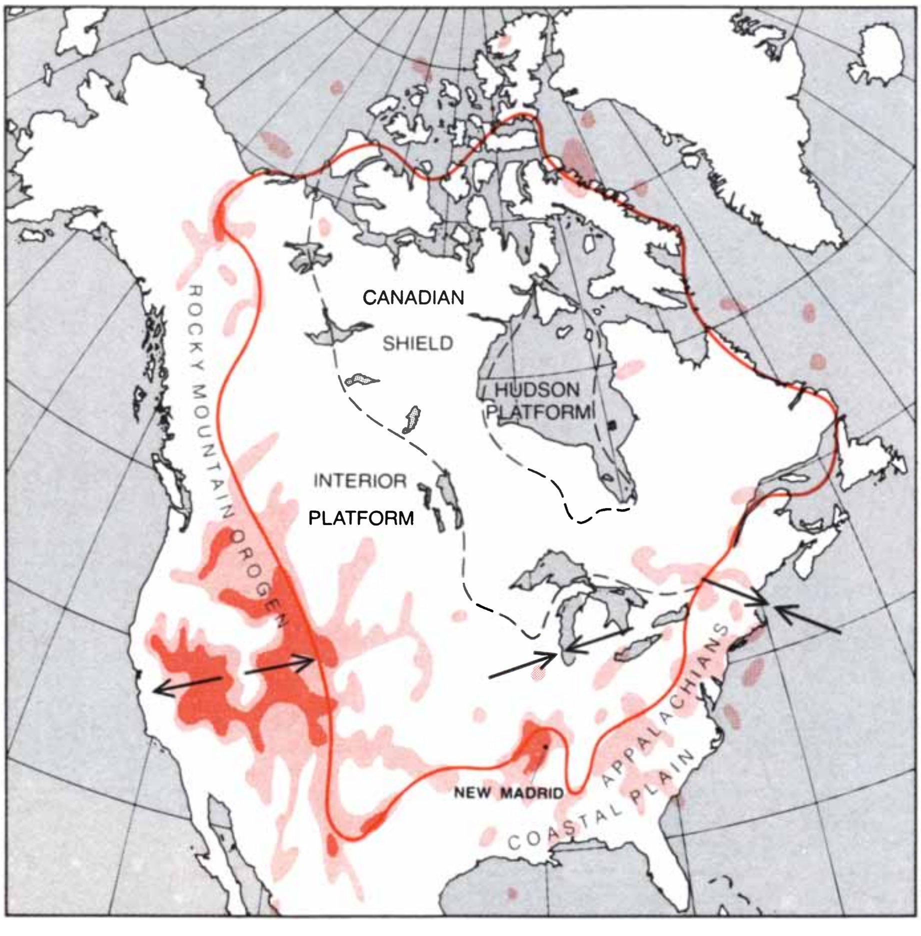


COMPUTER-GENERATED PLOTS measure the earthquake activity of the New Madrid seismic zone. In the upper plot the height of a peak is proportional to the number of earthquakes recorded since July, 1974. The highest peak represents 19 quakes within a zone two square kilometers in extent. This and the next-highest peak overlie fault-segment intersections. In the lower plot peak height represents the cumulative seismic energy released. The vertical scale is exponential; if it were linear, the tallest peak would be more than a million times higher than the low ones. The peak in the lower-left corner represents a 1974 magnitude 5.0 earthquake and its 4.5 aftershock.

is to see only half of the picture. One must also account for the mid-continental faults themselves. Furthermore one must ask why this particular combination of faults and stresses is seismically active when other mid-continental areas subjected to the same regional stresses show little if any such activity. A beginning can be made by summarizing what has been learned about the structure of the earth's crust in the Mississippi embayment.

A wealth of new information concerning the embayment has come from research programs conducted by the Geological Survey and by the New Madrid Seismotectonic Study Group, which is supported by the Nuclear Regulatory Commission. The findings indicate that the crust underlying the embayment is weakened by an ancient rift. The faults associated with the rift may have been inactive for millions of years, but they have now been remobilized. The remobilization is attributed to the present mid-continental regime of east-west compressive stress. Both the stress and the displacements it has induced, however, appear to be quite different from the stress responsible for the initial rifting.

Continental cratons are among the least seismic regions of the earth. The crustal rock underlying the Mississippi embayment, however, appears to represent a "reentrant" into the North American craton; it is not a part of the cratonic block but is rock that is more characteristic of the coastal plain to the south. Such indentations are characteristic of a "failed arm" of three-armed rift systems that are associated with "hot spots" or "plumes" in the earth's mantle. Kevin C. Burke and John F. Dewey of the State University of New York at Albany suggest that such a hot spot underlay the North American craton in this rifted area some 300 million years ago; other estimates suggest that the hot spot was much more ancient, well back in Precambrian times of more than a billion years ago. A hot spot is not the only mechanism that could have produced the crustal rifting, but it is one that serves as a useful working model.



NORTH AMERICAN CRATON is the area inside the heavy colored line. It is the oldest, strongest and most rigid crustal rock of the continent. Most earthquake activity is confined to its periphery (*colored areas*). Earthquake activity along the boundary between the Pacific plate and the North American plate has been omitted from the map. The black arrows show the generalized horizontal stress fields present in the crust: compressive in the east and the mid-continent and tensional in the west. The craton includes both exposed basement rock (shield) and veneers of sedimentary rock (platform); they are outlined by broken black lines. The New Madrid seismic zone overlies not the craton proper but a deeply rifted "reentrant" of coastal-plain rock from the south into the craton. The data were compiled by Steven Wesnousky and Christopher Scholz of the Lamont-Doherty Geological Observatory of Columbia University.

The designation "failed arm" indicates that this particular rift segment did not develop into an active, divergent plate boundary. (The other two arms may or may not have developed into such a boundary.) The failed arm would nonetheless exhibit all or most of the characteristics of an active continental rift: stretched and faulted crust, intrusions of dense magma from the mantle, positive gravity anomalies and an initial doming, followed by subsidence, as the mantle magma intruding into the lower crust cools and contracts. Rivers will often occupy such a subsided rift, and over the thinned and faulted crust thick layers of sediment will accumulate.

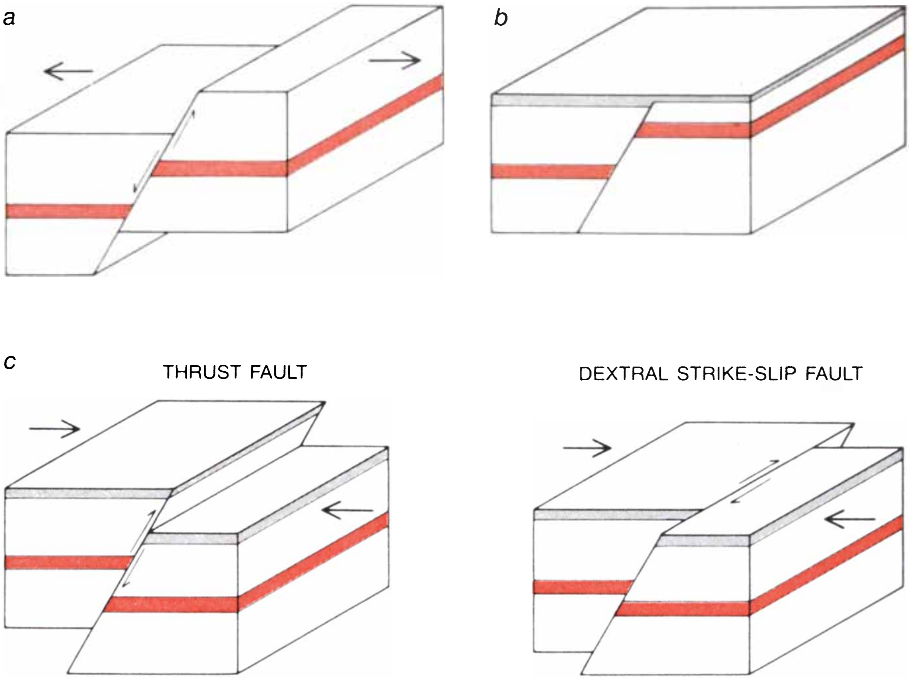
How well does the Mississippi embayment match this model of a failed-arm rift? Recent geophysical investigations that probe subsurface geological structures far beyond the reach of the deepest drill holes support such an interpretation. Gravity measurements have detected plutons (intrusions of dense mantle rock that increase the local gravity field) in the crustal rock of the embayment. The study of seismic waves reflected off subsurface structural discontinuities points to a thin crust within the embayment, underlain by a pillow of anomalous mantle rock.

Still further support for the rift model has come from field work in 1977–78, when Tom Hildenbrand and his colleagues of the Geological Survey carried out an extensive aerial magnetic field study of the region. Like gravity, the magnetic field of the earth is affected by the character of the subsurface rock. For example, the high iron content of igneous rock, such as mantle magma, gives rise to strong magnetic readings and the low iron content of sedimentary rocks gives rise to weak readings.

Hildenbrand's survey made use of a data-enhancement technique that not only revealed igneous intrusions but also located the boundaries of quite small intrusive bodies. A "contour" map incorporating these enhanced data shows a zone of reduced magnetism that runs through the New Madrid region in a northeast-southwest direction and separates flanking areas that have numerous igneous intrusions and are therefore higher in magnetic intensity. The low-intensity zone, roughly 75 kilometers wide, extends from south of Memphis northward to near the apex of the Mississippi embayment. Hildenbrand attributes the reduced magnetic intensity within the zone to a two-kilometer subsidence of basement rock as the failed-arm rift cooled and was loaded with sediments. The axial trend of New Madrid seismic events parallels the ancient rift and is centered between its flanks.

The Geological Survey's magnetic-intensity data have convinced many geologists, including me, of the reality of a subsurface rift in the Mississippi embayment. Hildenbrand's findings have created a situation somewhat analogous to the acceptance of the reality of sea-floor spreading when in 1966 measurements revealed the nearly perfect symmetry of one particular magnetic-intensity profile on each side of the East Pacific rise. New concepts often need the support of a particularly clear piece of evidence in order to turn skeptics into believers.

Evidently at the time the rift was formed the North American craton was subjected to horizontal tensional stress. In today's regime of compressional east-west stress the gravitational faults associated with the initial rifting have been reactivated as faults of the thrust and strike-slip class. Robert M. Hamilton and Mark Zoback of the Geological Survey have recently applied a new technique of seismic analysis to identify for the first time the highly deformed fault zone that generates the axial-trend earthquakes in the New Madrid region. Two other Geological Survey workers, Ernest Glick and Anthony J. Crone, have shown that one fault on the southeast flank of the rift zone has been uplifted 80 meters since the time of the two-kilometer subsidence identified by Hildenbrand.



FAULT REACTIVATION is shown schematically. A gravity fault (*a*) forms when the horizontal component of the regional stress field (*arrows*) is tensional; one block sinks with respect to the other. Erosion and sedimentation (*b*) obliterate the surface evidence of the fault. The fault can be reactivated when compressive stress is applied (*c*). If the overriding block rises, the reactivated fault is of the thrust type, seen at the left. If one block moves horizontally with respect to the other, the fault is of the strike-slip type (motion to the right is dextral and motion to the left is sinistral). Oblique faulting, a combination of the two, is also possible. The gravity faults of the New Madrid seismic-zone rift exhibit both thrust and strike-slip reactivation.

If one could forecast the rate of recurrence of major earthquakes in the New Madrid region, it would be possible to estimate its seismic vulnerability with fair precision. In the region of the San Andreas fault it is possible to make direct fault-slip measurements, but since this cannot be done in the Mississippi embayment, an alternative method is necessary. It stems from the empirical observation that for a given seismic region there is a linear relation between the magnitude of an earthquake and the logarithm of the cumulative number of earthquakes equal to or less than that magnitude.

This is simply a quantitative expression of the fact that small earthquakes are much more numerous than large ones. The slope of the best-fitted line on a graph plotting cumulative earthquake frequency against earthquake magnitude (known to seismologists as the *b* value) is commonly close to unity, so that there is a tenfold increase in the number of quakes for each unit decrease in magnitude. Once the *b* value and the proportionality constant of the equation are specified the average recurrence interval can be calculated for earthquakes of any given magnitude.

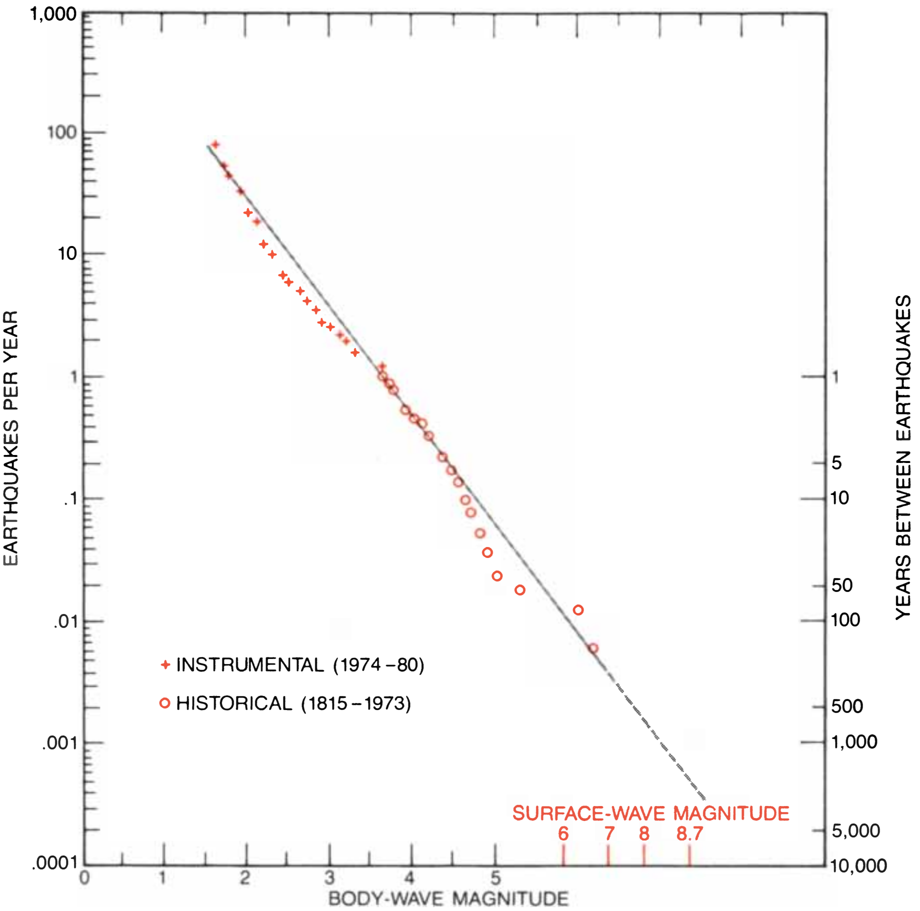
Working with a catalogue of historical New Madrid earthquake data compiled by Otto W. Nuttli of Saint Louis University and with the record of events detected since 1974, I have used the equation to estimate various New Madrid recurrence times. The range of elapsed times is enormous. Whereas, as I have noted, an earthquake of magnitude 1 can be expected almost every other day, an earthquake capable of major structural damage (about magnitude 6) can be expected only every 40 to 80 years. A repeat of an event greater than magnitude 8, such as those of 1811–12, is unlikely more often than at intervals of from 600 to 1,800 years. Obviously the larger the selected magnitude is, the greater the statistical uncertainties are. This results in estimates that have little practical value in the evaluation of risk. Nevertheless, even the empirical estimates are consistent with the results Russ obtained from the disturbed sediments in Lake County.

Even if the frequency of large New Madrid earthquakes cannot be estimated with anything resembling precision, it is still possible to make fairly good estimates of the extent of the damage that would result if such an earthquake occurred. For example, let us assume a recurrence of the 1812 hard shock. The most vulnerable large population center in the New Madrid area is Memphis. The city does not enforce any special seismic provisions in its building code. Moreover, the location of the city's buildings on the unconsolidated sediments of the Mississippi valley makes them far more susceptible to earthquake damage than structures with bedrock foundations would be.

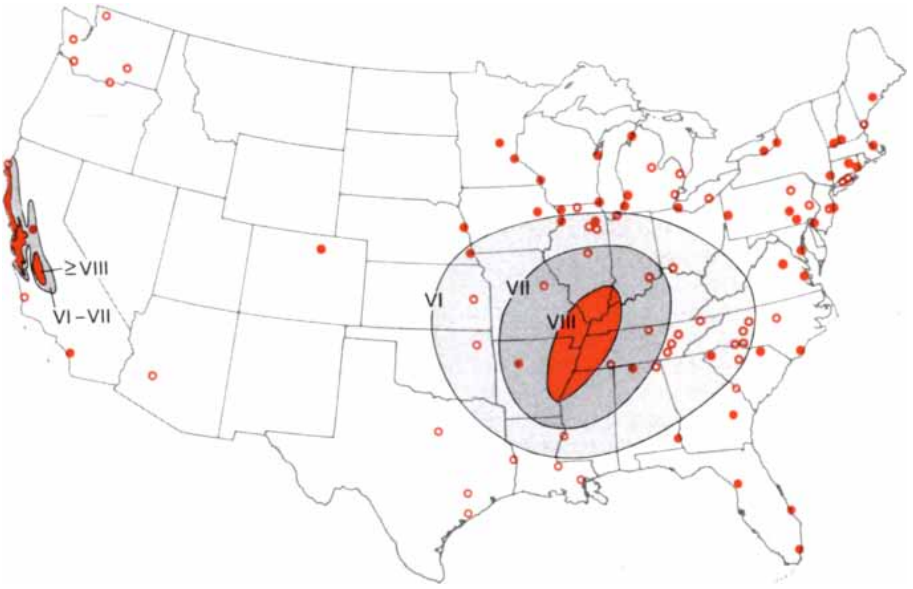
In 1974 Clarke Mann and Warner Howe, consulting engineers in Memphis, joined Frederic H. Kellogg of Memphis State University in an evaluation of seismic hazards in the city on behalf of the Department of Housing and Urban Development. If a maximum New Madrid earthquake were to occur in the daylight hours in the year 2000, they estimated, more than 3,000 people would be killed and the damage to buildings (measured in 1970 dollars) would exceed $1.3 billion.

The damage from such an earthquake would not, of course, be limited to Memphis. In a similar estimate that Nuttli prepared on behalf of the Federal Emergency Management Agency he predicted structural damage to buildings and to "lifelines" such as bridges, utilities and pipelines not only in Tennessee but also in parts of six other states: Arkansas, Illinois, Indiana, Kentucky, Mississippi and Missouri. This is a much larger area of damage than can be expected for an earthquake of similar magnitude in California because the crustal rock of the western U.S. attenuates seismic waves over much shorter distances than the crustal rock of the central U.S. does. The agency's preliminary damage estimate for this extensive region, considering only damage to buildings, came to $13 to $14 billion (measured in 1980 dollars).

Even if the private citizen chooses, in view of the centuries between major earthquakes in the estimates, to ignore the current risk presented by the New Madrid seismic zone, the designers and regulators of large and critical facilities



FORECASTING EARTHQUAKES in a particular seismic region is possible because when the cumulative number of earthquakes per year is plotted logarithmically against earthquake magnitudes, the plot approximates a straight line. The process is applied to New Madrid data in this graph. The graph has a double magnitude scale: surface-wave magnitudes (*color*) for large events and body-wave magnitudes for smaller ones. The plot indicates that an event greater than magnitude 8, such as those of 1811–12, is likely only once every 600 to 1,800 years.



AREAS OF DAMAGE from a recurrence of an earthquake of magnitude 8.6 in the New Madrid seismic zone, as estimated by Otto W. Nuttli of Saint Louis University, are compared here with those of the San Francisco earthquake of 1906. Both quakes are roughly of the same magnitude; the areas of damage differ greatly because the crustal rock in the west attenuates seismic shock over much shorter distances than the crustal rock of the central U.S. does. Colored areas are where buildings were structurally damaged in the San Francisco area and where similar damage can be expected from a New Madrid recurrence. The gray area in California is where the architectural damage ranged from minor to significant. The inner gray area surrounding the central New Madrid damage zone is where significant architectural damage could be expected; the outer gray area is where architectural damage would be minor. Colored dots locate operating nuclear power plants; colored circles, plants planned or under construction.

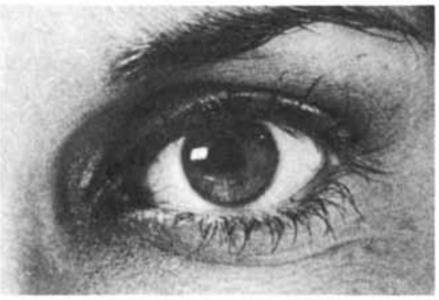
cannot do so. In particular the siting and construction of nuclear power plants are determined in large part on the basis of the estimated earthquake hazard. For example, every nuclear plant built or under construction in the U.S. is designed to withstand two hypothetical earthquakes.

The first of the two is known as an "operating basis" earthquake (OBE), one that would not impair the ability of the structure to continue to operate safely. The second is known as a "safe shutdown" earthquake (SSE), one that would require the shutdown of the plant but would leave functional all systems required for public safety. The hypothetical SSE tremor, normally about twice as powerful as the OBE tremor, represents the maximum potential ground motion estimated for the area concerned and generally has a very low probability of occurrence. In much of the central U.S. assuming a quake equivalent to the New Madrid hard shock of 1812 calls for vibration estimates so large that to satisfy SSE design requirements becomes prohibitively expensive. This is the reason no nuclear plants have been built within the Mississippi embayment and why the design criteria of nuclear facilities nearby include the ability to withstand an imagined maximum New Madrid quake in the section of the rift closest to the plant.

The nuclear power industry faces a dilemma in dealing with earthquake hazards. On the one hand overconservatism in plant design can add tens of millions of dollars to construction costs. On the other hand the Nuclear Regulatory Commission cannot and does not ignore the possibility that the New Madrid seismic zone extends hundreds of kilometers northeast along the Wabash River fault system into Indiana. At the same time until it is better understood why the New Madrid region is active whereas similar continental rifts remain seismically quiescent no one can be sure that what are now inactive zones may not someday generate large earthquakes. For example, dozens of old, inactive faults and rifts are known to exist in the eastern U.S. Should all nuclear facilities in this region therefore be designed to withstand the high energy inputs from possible large earthquakes?

To answer these questions and many others will require a precision of geophysical and geological knowledge that does not now exist. Continuing study of the vast and complex natural laboratory we call the New Madrid seismic zone should nonetheless lead to a deeper understanding of the geological environment in general. The goal is not only to satisfy man's natural curiosity about why things are the way they are; it is also to be better prepared when the next great New Madrid earthquake inevitably comes.

The eyes have it.



In sophistication, they beat anything we expect to offer soon. Therefore we would be foolish to begrudge the salaries and benefits paid for the full-time research of David F. O'Brien, Patricia N. Tyminski, and Richard T. Klingbiel, and the nuclear magnetic resonance (NMR) collaboration of Nicholas Zumbulyadis. At best, the value added to Kodak products by that work can show up only in the distant future and probably in quite unplanned ways. Back in 1912 Kodak realized that research embodied a good measure of risk taking. Actually, it has paid off.

For certain scientific applications, as in astronomy and autoradiography, photography has it over vision by its ability to accumulate energy for hours and days at a time. Furthermore, the resulting physical image can be restudied a century later. However, in everyday photography, the need to adjust, by technology or skill, the intensity and duration of the exposure instead of the sensitivity of the receptor, as the eye does, is a drag. During working hours, we also envy the multimillion-fold range of that sensitivity, which gets us home safely from work in daylight or dark.

Thus motivated, Kodak Research Laboratories have been able to report to all interested the following modest progress:

Initially we demonstrated that some synthetic membranes of rhodopsin, the light-sensitive protein, and the phospholipids, the building blocks of cell membranes, exhibited natural or functional photochemical behavior.*

These observations led to studies of the structure of natural and synthetic rhodopsin membranes by the use of NMR and chemical probes.†

Parallel to the structural studies, we have examined the light-amplification mechanisms. Observations of light-induced ion permeability increases of rhodopsin membranes have shown 10^2 to 10^3 of amplification.‡

Current studies with light-activated enzymes, a phosphodiesterase and a GTPase have produced amplifications of 5×10^5. These enzymes reside on the surface of rhodopsin membranes and are activated by bleached rhodopsin.§

Rhodopsin-activated camera film is most unlikely. But what about lipid bilayers as orienting devices for polymerizations? Maybe they could hold enzymes in place for work they cannot accomplish by random encounters in solution. Nature's way is the best way, it is said.

**Biochemistry* 16 (1977), 1295-1303

†*Biochemistry* 17 (1978), 4186-4192
Biochemistry 18 (1979), 5427-5432
Biophysical Journal 25 (1979), 131a
Biochimica et Biophysica Acta 603 (1980), 313-321

‡*Proc. Natl. Acad. Sci. USA 74* (1977), 5222-5226
Photochemistry and Photobiology 29 (1979), 679-685
Biophysical Journal 21 (1978), 153a

§*Biophysical Journal 33* (1981), 203a

SCIENCE AND THE CITIZEN

State Secrets

Last summer the Department of State issued visas allowing some 7,000 scholars from the People's Republic of China to study at American universities. About 70 of the 7,000 were subsequently singled out by Keith Powell II, the State Department country officer for China. At Stanford University, for example, Powell wrote to Edward J. McCluskey, who was to supervise the work of two Chinese students in computer science, and Bruce B. Lusignan, who was to receive three Chinese scholars studying satellite communications.

The letters were similar. "U.S. government policy," Powell wrote, "encourages the training of accomplished Chinese scholars in modern technology and science." The letters went on, however, to list restrictions the State Department wanted to place on the students. "There should be no access to the design, construction or maintenance data relevant to individual items of computer hardware. There should be no access to design of microelectronics.... This office should be advised prior to any visits to any industrial or research facilities." The letters warned against the "potential transfer of classified technology or of technology which requires an export license."

McCluskey and Lusignan were advised by Gerald J. Lieberman, vice-provost and dean of graduate studies and research, that Powell's letters were "not applicable to Stanford's research programs" because the results of those programs "are generally available to the public.... Indeed, the purpose of university research activities is the discovery and dissemination of new knowledge. By Stanford policy, no secret or classified research may be conducted on campus." The State Department did not press the issue.

In December, Bernard Roth of the department of mechanical engineering at Stanford received a letter from the National Academy of Sciences about a proposed visit by Nikolay V. Umnov, a Russian investigator of robotics. The NAS, which administers exchanges of scholars between the U.S. and the U.S.S.R., advised Roth that the visit would not be approved by the State Department unless Stanford (and three other universities on Umnov's itinerary) agreed to certain restrictions: "*A.* Umnov's program should be restricted to mechanical theory of robotic locomotion. No access should be permitted to control units or programming techniques for robots. *B.* No industrial visits. *C.* Umnov should have no access—visual, oral or documentary—to production, research or other activity funded by Defense Department contracts or grants, classified or unclassified."

On Roth's behalf Lieberman replied to the NAS that Stanford "is willing to host Dr. Umnov as proposed by the NAS, but we are not willing to accept responsibility for Dr. Umnov's actions—either on or off campus—during his visit to Stanford." In February the NAS sent a second letter "clarifying" the restrictions in a way acceptable to the university.

The State Department proposed placing similar restrictions on Mikhail Gololobov, a Russian organic chemist visiting Purdue University, the Massachusetts Institute of Technology and three other institutions. At M.I.T., Gololobov had asked to visit the department of nutrition and food science. Nevin S. Scrimshaw, the chairman of the department, was told that his visitor was to be denied access to work in several areas, including genetic engineering, which is not done in Scrimshaw's laboratory. M.I.T. refused to comply. Gololobov arrived at Purdue in January. His host, Michael Laskowski, Jr., says the State Department's only restriction, that the visitor have no access to defense-related research, does not apply to his laboratory.

The attitude of the Administration was underscored in January by Admiral Bobby R. Inman, Deputy Director of the Central Intelligence Agency, in an address to the American Association for the Advancement of Science. "Make no mistake," he said. "Specific data on technical subjects is high on the wanted list of every major foreign intelligence service, and for good reason.... Restrictions on science and technology should only be considered for the most serious of reasons. But nowhere in the scientific ethos is there any requirement that restrictions cannot, or should not—when necessary—be placed on science."

The legal basis of the proposed limitations of academic freedom is in part the Export Administration Act of 1979, which extends control over exports to the "export of technical data." Such exports include "any release of technical data in the United States with the knowledge or intent that the data will be transmitted from the United States to a foreign country."

More than a year ago the presidents of five institutions (Donald Kennedy of Stanford, Marvin L. Goldberger of the California Institute of Technology, Paul E. Gray of M.I.T., Frank H. T. Rhodes of Cornell University and David S. Saxon of the University of California) joined to protest the application of such export regulations to university activities. "Restricting the free flow of information among scientists and engineers," they wrote, "would alter fundamentally the system that produced the scientific and technological lead that the Government is now trying to protect and leave us with nothing to protect in the very near future. The way to protect that lead is to make sure the country's best talent is encouraged to work in the relevant areas, not to try to build a wall around past discoveries."

Informatization

The government of France has established a world center to promote the social applications of microcomputers and information science. The center is also expected to stimulate the French microelectronics industry and establish French leadership in both containing the threat and exploiting the promise of what is perceived as a fundamental revolution in human affairs: the computerization (or, to use the more evocative French term, *l'informatisation*) of society. The Centre Mondial Informatique et Ressources Humaines will seek to concentrate in Paris the efforts of information scientists from many countries and to disseminate new knowledge and techniques worldwide, with particular attention to the developing countries.

The center was organized, at the behest of President Mitterrand, by the author, journalist and political activist Jean-Jacques Servan-Schreiber, who will be its president. American information scientists will play major roles: the executive director will be Nicholas Negroponte of the Massachusetts Institute of Technology and the chief scientist will be Seymour A. Papert of M.I.T. Alan C. Kay of Atari, Inc., and D. Raj Reddy of Carnegie-Mellon University will be part-time participants. French investigators will constitute a majority of the scientific advisory board, which will be chaired by Jacques-Louis Lyons, president of the National Institute for Computer Science and Automation Research. The board of directors will include nine cabinet ministers. The center's annual budget is expected to be about 100 million francs (about $17 million), some of which Servan-Schreiber hopes to raise from other countries and from international organizations.

The origin of the center can be traced to a concern in France about the economic and social dislocations attending *informatisation,* and in particular the unemployment that is seen as the inevitable first result of the computerization of industry. Fear of unemployment, Servan-Schreiber pointed out in a report to Mitterrand last fall, has blocked needed modernizations. Yet, he argued, the same science that has created powerful information processors and automatic manipulators can, by means of micro-

computers, "regenerate and amplify every person's human faculties, his capacity for new activities, new forms of creativity and new occupations."

In brief, the answer to the threat of the computer is the microcomputer, and more specifically the development and wide distribution of the personal computer. Small computers for the home and the office already exist, Servan-Schreiber acknowledged, but they are far from being true personal computers.

According to Servan-Schreiber, the center's initial activities will be directed along three lines. The first activity will be the development of hardware and software for a truly personalized computer, one about the size of a large book and costing between 500 and 1,000 francs. The second will be the testing of experimental computing devices in developing countries. A number of countries have expressed interest in such projects, which may be financed in part by funds from the Organization of Petroleum Exporting Countries. The third major activity will be experiments in France on the social and educational applications of personal computers, with emphasis on the needs of children, unemployed people and the elderly.

Superstar

Is there any limit to the mass of a star? The question is provoked by a luminous zone of ionized hydrogen atoms in the Large Cloud of Magellan (a companion galaxy to our own), which may harbor a star whose mass is thousands of times the mass of the sun. An increasing number of observations imply the star is there; on the other hand, it is said that a star with much more than 100 times the mass of the sun would destroy itself if it formed, or that it could not have formed in the first place.

The luminous zone, called the Tarantula Nebula, has long been recognized as a place of extreme luminosity. Indeed, the Tarantula is the brightest H II region, or zone of ionized hydrogen, in any nearby galaxy. By 1977 J. V. Feitzinger and his colleagues had begun to report the results of an analysis of photographs of the Tarantula made at the European Southern Observatory in Chile. The workers have found that its central knot of brightness (called R136) has at least three distinct parts. The brightest of the parts (called R136a) may in turn have two components, one much brighter than the other. R136a is about one light-year in diameter, which makes it extremely compact on the scale of the nebula. Feitzinger and his colleagues were therefore led to suggest that the brightness in R136a might indicate the presence of a supermassive star. Its intense ultraviolet radiation would presumably be responsible in large measure for the ionization that makes the Tarantula Nebula luminous.

Feitzinger also suggested that the mass of such a star might be a few hundred times the mass of the sun. That was enough to make the star problematic. No star more massive than a few dozen times the mass of the sun has been found in our galaxy. Moreover, it is calculated that if a star were more than about 100 times the mass of the sun, its self-gravitation (which tends to make the star collapse) would be balanced not so much by the pressure of the gas in the star as by the pressure of the electromagnetic radiation the star emits. In such a circumstance the star is thought to be unstable and hence to pulsate, perhaps wildly.

Other arguments suggest that a star more massive than 100 solar masses might not be able to form. Models describing how dust and gas accrete to form a star imply that the ultraviolet radiation from a newly condensed protostar vaporizes the dust nearby. In this way the protostar gets a dustless cocoon. Gas remaining in the cocoon gives off infrared radiation that sweeps dust from a large volume surrounding the star. The dust pushes gas away with it. The process is thought to limit the amount of matter the protostar can collect.

Only recently has the spectrum of the radiation emitted by R136a been carefully measured. Last year the spectrum at ultraviolet wavelengths was recorded with the aid of instruments on board the International Ultraviolet Explorer satellite. The measurements were made by Joseph P. Cassinelli, John S. Mathis and Blair D. Savage of the University of Wisconsin at Madison, who offer several conclusions. First, they note that the ultraviolet spectrum of R136a includes several lines markedly broadened in a pattern suggesting that R136a emits a "wind" of ionized gas. Evidently the wind is prodigious, with a terminal velocity of 3,500 kilometers per second. The sun's wind is only a tenth as fast, and even the wind emitted by the massive, luminous stars called *O* stars is slower by 1,000 kilometers per second. Moreover, the mass of R136a is being dispersed by the wind at a rate such that in 1,000 years it loses an amount of mass equivalent to the mass of the sun. (In a day it loses the equivalent of the mass of the earth.) The loss of mass for the "windiest" of the *O* stars is 100 times slower, and for the sun it is a billion times slower still.

From the overall pattern of the ultraviolet spectral lines Cassinelli and his colleagues infer that the temperature of R136a is 60,000 degrees Kelvin. The assignment of such a temperature leads to the further conclusion that the intrinsic luminosity of R136a is 100 million times the intrinsic luminosity of the sun. It also yields an estimate of the mass of R136a. A single star with 100 million times the luminosity of the sun must have about 3,000 times as much mass,

DISCOVER ATARI.®

DISCOVER ATARI COIN VIDEO GAMES.

Every so often, a company comes along that literally challenges the way we think. ATARI is just such a company.

Packing a galaxy of motion and logic in dime-sized micro chips, our engineers have developed coin video games like Asteroids, Centipede and Tempest. Stimulating. Involving. Challenging.

ATARI is a worldwide leader of this fast-growing business. But it's just part of a larger world waiting to take you further.

DISCOVER ATARI HOME VIDEO GAMES.

As ATARI technology has grown, we've opened your eyes to the world's most popular home video games. Like Breakout,® Missile Command and Space Invaders.*

Each bringing its own unique excitement to families all over the world. But there is still another facet of ATARI.

DISCOVER ATARI HOME COMPUTERS.

Perhaps this is the most important facet. Transforming the computer into an appliance you can use at home.

The ATARI 400 and ATARI 800 Home Computers are examples of how we've taken sophisticated technology and simplified it so that you can gather information, compose music, play advanced games, and master your finances at the touch of a button.

And this is just the beginning. With ATARI's remarkable ability to humanize technology, we'll continue to broaden and expand. Beckoning you to discover new worlds of fantasy and challenge.

Asteroids, Centipede, Tempest, Missile Command, 400, 800 are trademarks of ATARI, Inc.

*Trademark of Taito America Corporation

A Warner Communications Company

© 1981 ATARI Inc.

ATARI®

DISCOVER HOW FAR YOU CAN GO.

THE LEADING EDGE

#2 in a series of reports on new technology from Xerox

Few inventions have proved more versatile than the laser.

It has been used as a super signal lamp for space communications. It has found a role in exotic metalworking applications, and the incredible precision that it offers has made new kinds of brain and eye surgery possible.

A system for playing TV shows from phonograph-like disks uses a laser to "read" the recorded program material.

Lasers are used in various forms of measurement and information handling systems. For example, one version of the bar-code reader used in supermarkets employs a laser scanner.

The Hologon Laser Scanner is one of the latest developments in practical applications of laser technology. It was invented by Xerox optical physicist Dr. Charles Kramer who wrote this article.

Lasers In Electronic Printing

At Xerox we use lasers in electronic printing systems that are based on xerography. Instead of making copies of existing documents, such printers create documents, drawing on information stored in a computer. In such a system, signals from the computer pulse the laser beam as it scans across a light-sensitive drum or belt that serves as the "camera film" in xerography. The image recorded in this way is then developed and transferred to paper as in a copier.

Laser electronic printers offer the quality of offset printing with significant versatility compared with conventional computer printers. There is virtually no meaningful limit to format or to type style or size. Pictorial or other graphic material can be printed as easily as text. Arabic, Greek or Russian alphabets—even Oriental ideograms—are within its capabilities, provided appropriate programming is fed into the printing system.

Equally significant is the fact that, with electronic printing, documents originated in one location can be printed simultaneously at a number of different locations.

Xerox currently offers three such systems. The 9700 electronic printing system turns out almost two pages per second and has almost unlimited flexibility when it is used with the Xerox Integrated Composition System program. The 5700 electronic printing system is up to 40 times faster than conventional word-processing printers, which it can replace, and it can also be used for electronic mail and remote computer printing. A similar printer is offered as part of the Xerox 8000 network system. Designed for lower-volume applications, it is twelve times faster than a conventional word-processing printer.

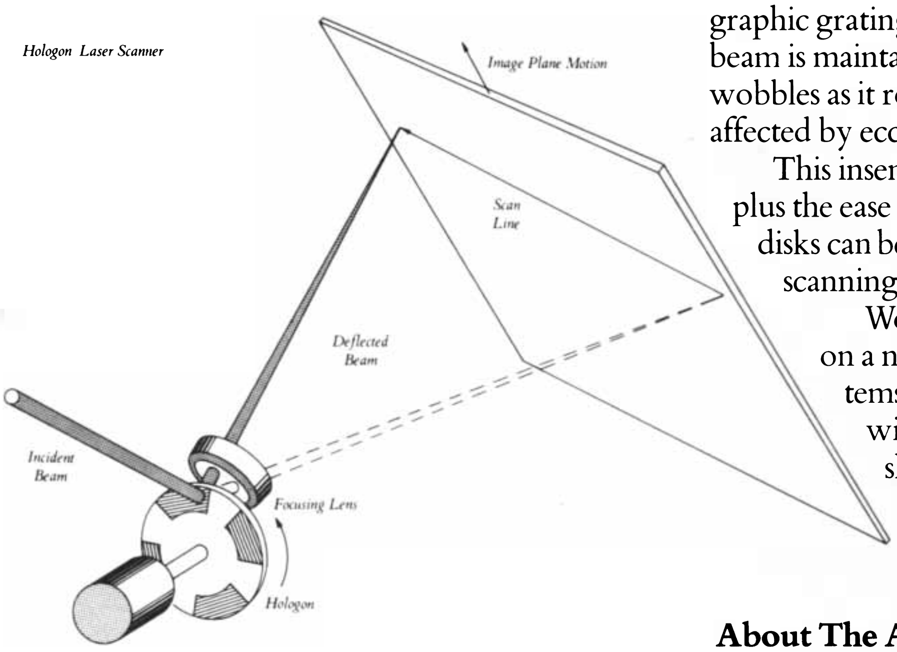
In these printers, the scanning action of the laser beam is created by a relatively complex optical system that is based on a rotating, polygonal prism. Extremely high precision is required in such a system. This complexity and precision make such a laser scanner relatively expensive.

The Hologon Laser Scanner

To simplify laser scanners and reduce their cost, we considered holography to perform the functions of costly lenses and prisms.

Holographic recordings, best known for their reproduction of three-dimensional images, take the form of gratings—corrugated or ridged patterns on a transparent medium, having a spacing of approximately twenty millionths of an inch. In pictorial holograms, these gratings contain the recorded cross section of the wavefront of light that had been reflected from the pictorial subject. When coherent light—usually from a laser—is transmitted through such a hologram, a true three-dimensional image of the subject is reproduced.

Hologon Laser Scanner



However, the holographic gratings used in a Hologon scanning system do not contain pictorial information. Only the optical diffraction properties of the gratings are utilized.

The Hologon System is one of several holographic approaches to scanning. But the others tend to scan in an arc-like pattern which is unsuitable for electronic printing, which requires a straight-line scan, much like the raster pattern used in television to create an image. The Hologon approach gets around this problem through a novel configuration.

In a Hologon, a series of holographic gratings are formed around the circumference of a transparent disk. A laser shines through these gratings as the disk rotates. The gratings diffract the laser light, and the rotating action causes it to scan across the surface on which it is focused, as shown in the accompanying diagram. Focusing is done by a simple, inexpensive lens.

The laser beam in this system is aimed so that it forms a nominal 45° angle to the Hologon's surface as it enters a grating and a 45° angle as it emerges from the grating. In other words, it is diffracted through a right angle by the gratings. This angling results in a straight-line scan.

Because of the optical properties of the holographic gratings, the 90° diffraction angle of the beam is maintained even if the Hologon surface wobbles as it rotates. The beam angle is equally unaffected by eccentricities in the rotating disk.

This insensitivity to mechanical variation, plus the ease and low cost with which Hologon disks can be produced, make a Hologon laser scanning system relatively inexpensive.

Work is currently underway at Xerox on a new generation of laser printing systems utilizing the Hologon laser scanner with all its attendant benefits. This should enable Xerox to make the advantages of electronic printing more widely available than ever before.

About The Author

Dr. Charles Kramer is the inventor of the Hologon Laser Scanner. He is an optical physicist specializing in electro-optical reading and printing devices at the Xerox Joseph C. Wilson Center for Technology in Rochester, New York.

He holds a Bachelor's degree and a Master's degree in Physics from Fairleigh Dickinson University and a Master's degree and Ph.D in Optics from the University of Rochester.

XEROX


Everything fits beautifully.
Including you.

or the pressure exerted by the emitted radiation would make the star explode.

An alternative explanation is that R136a is a cluster of bright stars, but Cassinelli and his colleagues doubt this hypothesis. For one thing, a surface temperature approaching 60,000 degrees K. is attained among known stars only by the hottest *O* stars, the ones designated *O*3. They are extremely rare; only a few have been observed in our galaxy. It would take some 30 of them to emit 100 million times the luminosity of the sun, and in R136a they would have to be found in a volume just one light-year in diameter. Even then the particle wind at 3,500 kilometers per second would be hard to explain.

The most recent contribution to the arguments about R136a was made in January, when two investigators reported their measurement of the spectrum of R136a at visible wavelengths. Peter S. Conti of the University of Colorado at Boulder recorded the spectrum with the four-meter telescope at the Cerro Tololo Inter-American Observatory; then Dennis Ebbets of the University of Wisconsin at Madison subjected it to a technique of image enhancement. The effort revealed emission lines, each one broad but faint. The lines signaled the presence of helium and nitrogen. The nitrogen atoms had been stripped of two, three or even four electrons. An optical spectrum consisting of weak emission lines from a very hot object rich in helium and nitrogen is in many cases the signature of a Wolf-Rayet star of the type designated WN. It is thought such a star begins its lifetime with a mass between 30 and 50 times that of the sun; then, however, it loses enough of its mass to a stellar wind for its core to be exposed. The core is rich in the helium and nitrogen synthesized by nuclear fusion. Some 160 Wolf-Rayet stars are known in our galaxy, and a few dozen have been found in the Large Cloud of Magellan.

Does R136a harbor a supermassive, superluminous, unstable star with a Wolf-Rayet spectrum? Does vigorous convection bring helium and nitrogen to the surface of such an object, so that they are promptly cast off as part of a stellar wind? Conti, in collaboration with Philip L. Massey of the Dominion Astrophysical Observatory in Canada, has been searching for Wolf-Rayet stars in M33, a nearby spiral-arm galaxy viewed almost face-on from the earth. In five bright H II regions they chose 14 luminous stars. Eight of them turned out to have a Wolf-Rayet spectrum.

The Purloined Letters

In the adult human being the main structural components of the hemoglobin molecule are the proteins alpha-globin and beta-globin. A defect in either of the two types of globin chain renders the molecule useless for its function of carrying oxygen. In one form of the genetic disorder beta-thalassemia few functional beta-globin chains are made. About a year ago it was shown that the deficiency results from a point mutation: a defect at a single site in one of the beta-globin genes. Now Meinrad Busslinger, Nikos Moschonas and Richard A. Flavell of the National Institute for Medical Research in England have shown how the minute genetic imperfection leads to the clinical condition.

A point mutation alters one of the pairs of nucleotide bases that make up the DNA double helix. In some cases the substitution of one base pair for another leads directly to the substitution of one amino acid for another in the corresponding protein. The defect responsible for beta-thalassemia, however, acts through a subtler mechanism, and indeed it lies in a region of the gene that does not code for amino acids. The new base pair provides incorrect information to the enzymes that assemble the messenger RNA (mRNA) from which beta-globin is made. As a result an extra piece of RNA is included in the finished messenger molecule. When the beta-globin chain begins to be translated into protein, the extra RNA causes the process to go awry and then stop prematurely, yielding a truncated and nonfunctional beta-globin molecule.

The beta-globin gene is about 1,500 base pairs long. As in all genes the genetic information is encoded in the sequence of the four nitrogenous bases adenine, thymine, guanine and cytosine, abbreviated *A, T, G* and *C. A* can bond only to *T* and *C* only to *G.* Beta-globin mRNA is not formed directly from the beta-globin gene. The reason is that the sequence of nucleotides specifying the protein is split into three coding regions, called exons; between the exons are two noncoding stretches of the gene, known as introns. In the manufacture of beta-globin mRNA an RNA precursor molecule is made first, corresponding to the entire beta-globin gene. A group of enzymes working together cut out the segments of the precursor corresponding to the introns. Other enzymes then splice the three segments corresponding to the exons to yield an mRNA molecule with the code for the 141 amino acids of the beta-globin chain.

In 1980 David Westaway and Robert Williamson of St. Mary's Hospital Medical School of the University of London purified the beta-globin gene from a patient with beta-thalassemia and determined the sequence of nucleotides. The gene included the base pair *A–T* where *C–G* is normally found. The mutation was within the first intron, 19 nucleotides from the junction with the second exon. Busslinger, Moschonas and Flavell combined the gene purified by Westaway and Williamson with a plasmid (a free bacterial gene) called SV40-pBR328. The plasmid provides a means by which a gene can be inserted into a cell where it is not ordinarily found; the inserted gene can then be expressed by the host cell's machinery of protein synthesis. The plasmid with the attached beta-thalassemia gene was put into a colony of the human tumor cells called HeLa cells.

Busslinger, Moschonas and Flavell describe their findings in *Cell.* In the HeLa cells the transcription of the gene to form the RNA precursor was accomplished without incident. When the transformation of the precursor into mRNA began, however, the mutation disrupted the normal flow of events. The assembly enzymes cut the precursor at the point of the *A–T* substitution rather than at the point corresponding to the junction of the first intron and the second exon. As a result when the coding regions were joined, an intervening segment of 19 nucleotides derived from the intron was included in the new mRNA molecule.

The explanation of the incorrect cut and splice is that the mutation creates a group of nucleotides closely resembling those at the end of the intron. The final sequence of the intron is *CCCTTAG.* The correct sequence at the site of the mutation is *CTATTCG,* but the substitution converts it into *CTATTAG.* When the enzymes begin to work on the precursor molecule, they recognize the mutated sequence as a signal to cut the strand. Hence the last 19 nucleotides of the intron are incorporated into the mRNA.

The mRNA with the supernumerary nucleotides enters the cytoplasm of the cell, where a ribosome begins to construct the beta-globin protein. Twenty-nine amino acids are assembled correctly. The ribosome then reaches the point where the additional nucleotides are inserted. The nucleotides derived from the noncoding regions offer incorrect information to the ribosome, and six incorrect amino acids are added to the beta-globin chain. Then a group of three nucleotides at the end of the inserted sequence is reached that is identical with the signal to terminate protein synthesis. Translation ends and the mRNA is freed from the ribosome before the rest of the beta-globin chain can be made.

The defective beta-globin consists of only 35 amino acids (six of them incorrect) instead of the normal 141; such a molecule is not functional. Busslinger, Moschonas and Flavell can thus account for the manufacture of the abnormal hemoglobin. In beta-thalassemia patients, however, a few normal beta-globin chains are also formed, and it is not clear how to explain their presence. About 10 percent of the mRNA molecules observed in the HeLa cells are normal, as are about 10 percent of the beta-globin chains in the cytoplasm. When the RNA precursor is being cut and spliced to yield mRNA, the en-

Speak German like a Diplomat!

What sort of people need to learn a foreign language as quickly and effectively as possible? Foreign service personnel, that's who.

Now *you* can make a start on learning to speak German with the same materials used by the U.S. Department of State — the Foreign Service Institute's *Programmed Introduction to German* and *Basic Course, Continued.*

The FSI spent thousands of dollars and many years developing these materials for use by members of America's diplomatic corps. Today people in all walks of life who need to learn to speak a foreign language are turning to this outstanding audio cassette program.

The Foreign Service Institute's German Course is by far the most effective way to learn German at your convenience and at your own pace. It consists of a series of tape cassettes and accompanying textbook. You simply follow the spoken and written instructions, listening and learning. By the end of the course you'll find yourself learning and speaking entirely in German! A native German speaker, clearly recorded on the cassettes, provides an excellent model to help you develop your skills.

This course turns your cassette player into a "teaching machine." With its unique "programmed" learning method, you set your own pace testing yourself, correcting errors, reinforcing accurate responses.

The FSI's Programmed German Course comes in two volumes: You may order one or both courses:

☐ **Volume I. Programmed Introduction to German,** 9 cassettes, 12 hours and 647-page text, $115.

☐ **Volume II. Basic Course, Continued** 8 cassettes, 8½ hours and a 333-page text, $98.

(Conn. and N.Y. residents add sales tax)

Overseas shipment by air mail only — add $23 for 1 volume, $35 for both.

Your cassettes are shipped to you in handsome library binders.

TO ORDER, JUST CLIP THIS AD and mail with your name and address, and a check or money order. Or charge to your credit card (American Express, VISA, Master Charge, Diners Club) by enclosing card number, expiration date, and your signature.

The Foreign Service Institute's German course is unconditionally guaranteed. Try it for three weeks. If you're not convinced it's the most convenient and most painless way to learn German, return it and we'll refund every penny you paid. Order today!

AUDIO-FORUM®

Audio-Forum
Suite 330
On-the-Green,
Guilford, CT. 06437
(203) 453-9794

Telescopes from Celestron®

Will Take You To Places You've Never Been To Before
• Versatile • World's Largest Selection • Supremely Affordable

With a Celestron, the modern standard of high quality, affordable telescope performance, you can explore the craters of the Moon, the rings of Saturn, twirling galaxies and glowing nebulae. You can easily use a Celestron to photograph and view terrestrial subjects.

All Celestron telescopes have laser-tested, diffraction limited optical systems with hand-figured optics made in the U.S.A.

Celestrons® are available from 1.2- to 14-inches in diameter including the world's most popular modern telescope, the versatile Celestron 8. And, remember, most 35mm cameras easily attach to the Celestron for photography.

Free 8-page brochure or send $3.00 for 52-page full-color catalog on how to select and use a Celestron telescope/telephoto lens.

Celestron® International
2835 Columbia St. (Box 3578-SA2) Torrance, CA 90503, U.S.A. Telephone (213) 328-9560

Dealer inquiries invited.

zymes must be positioned so that it is probable they will make the incorrect cut, but it must still be possible for the correct cut to be made.

Constant Endeavor

Physical interactions outside the nucleus of the atom are governed almost entirely by two fundamental forces of nature: gravitation and electromagnetism. Qualitatively the two forces are similar, but in absolute strength they are profoundly different. Now remeasurements of the numerical constants that define the strength of the two forces have slightly improved the accuracy with which they are known. The new value of *G*, the constant of the gravitational force, is $6.6726 \pm .0005 \times 10^{-11}$ in the standard units of the meter-kilogram-second system of measurement. The corresponding constant of electromagnetism is the dimensionless quantity called the fine-structure constant and designated α (the Greek letter alpha); its new value is $1/137.035965 \pm .000012$.

In order to measure *G*, Gabriel G. Luther of the National Bureau of Standards and William R. Towler of the University of Virginia performed a refined version of the experiment first carried out by the British investigator Henry Cavendish in 1798. Luther and Towler suspended a small dumbbell with a mass of about five grams from a quartz fiber 12 micrometers in diameter and 40 centimeters long. Two larger spherical masses were arranged so that the gravitational force between each large mass and one end of the dumbbell tended to twist the fiber. When the dumbbell was made to oscillate about the axis of the fiber, the torque introduced by the gravitational interaction altered the period of the oscillation. By measuring the period when the masses were at various distances, Luther and Towler were able to calculate their value of *G*.

The determination of α was done by D. C. Tsui and A. C. Gossard of Bell Laboratories and B. F. Field, M. E. Cage and R. F. Dziuba of the National Bureau of Standards. They exploited a novel feature of a well-established physical effect. If an electric current is passed through a thin metal plate in a magnetic field, a potential difference arises along the direction perpendicular to both the current and the field. The induced voltage is called the Hall potential, after E. H. Hall of Johns Hopkins University, who first reported it in 1879. Klaus von Klitzing of the University of Würzburg has recently shown that in certain semiconductors at low temperature the Hall potential is quantized: it changes stepwise instead of continuously as the strength of the magnetic field is varied. From the quantized changes in the Hall potential α can be measured indirectly.

The new determination of α is not in itself superior to earlier measurements.

© 1982 SCIENTIFIC AMERICAN, INC

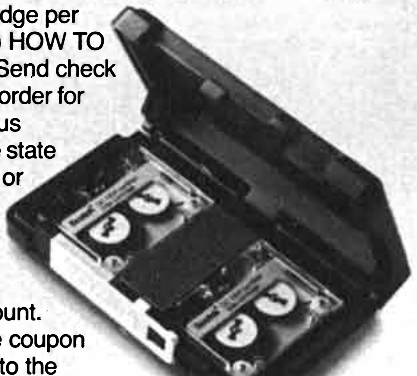
Improve Your Memory With a Free Scotch® Brand Data Cartridge.

Buy One, Get One Free!

We want you to try our Scotch Brand DC100A Certified Data Cartridge. We're so sure you'll like it that we'll send you a free one when you order one data cartridge at the regular price of just $18.00. No strings attached. Two of these superior data cartridges for the price of one.

Scotch Brand Data Cartridges are compatible with all Hewlett-Packard mini-cartridge computers, including the popular HP-85 model. They also fit the Savin Veritext® 950 word processor, the TI-770 terminal, Tektronix models 4041 and DAS-9100 and many others. We *invented* the certified data cartridge and our special back treatment process makes our tape extra durable and longer lasting. Enlarged end-of-tape sensor holes are provided for improved end-sensing reliability. Scotch Data Cartridges are quiet, too, and come with a one year warranty.

Order Now! Get a Free Protective Case.



Order now and we'll send you a rugged, heavy-duty protective shipping and storage case that holds two data cartridges. But don't wait, quantities are limited. (Offer limited to one free data cartridge per customer.) HOW TO ORDER: Send check or money order for $18.00 (plus applicable state sales tax) or charge to your Master Card or VISA account. Fill out the coupon and send to the address shown. Please allow 6-8 weeks for delivery. Offer expires September 30, 1982.

SHIP TO:
Name__________
Company__________ Title__________
Address__________
City__________ State______ Zip______
Phone__________ Computer Type__________
Signature__________

☐ Total enclosed $__________ ($18 plus state sales tax).
Make check or money order payable to 3M.
☐ Charge to ☐ Visa or ☐ Master Card.
Card No.__________. Expires______.

SEND TO: 3M, Dept. CC-W4, DATA RECORDING PRODUCTS DIVISION/3M,
P.O. Box 33133, St. Paul, MN 55133.

3M hears you...

3M

Because the new measurement is independent of the others, however, the result can be combined statistically with the best previous measurements to yield the currently reported value.

Adding another decimal place to the accuracy of constants, even when they are already known with considerable precision, is of great interest to physicists. The value of G determines the orbital periods of the planets and of other orbiting systems, such as binary stars. It can also be employed, as Cavendish realized, to give the mass of the earth and other planets. Similarly, the value of α determines the size, the period and the energy of an electron's orbit in an atom.

Seven Hypotheses

For theories that would explain the apparent mass extinctions at the end of the Mesozoic era, some 63 million years ago, the discovery of sediments enriched in the element iridium has added a new item of evidence. Many marine species of foraminifera seem to disappear from the fossil record at the end of the Mesozoic (which is also the boundary between the Cretaceous and the Tertiary periods); so do certain land plants and many species of reptiles, including all Cretaceous dinosaur species. Walter Alvarez and his associates at the University of California at Berkeley, who discovered the iridium anomaly, proposed in 1980 that the extinctions resulted from a catastrophic collision between the earth and an asteroid some 10 kilometers in diameter. Not all paleontologists agree with this interpretation, but most proposed explanations of the extinctions now take the iridium evidence into account.

At the annual meeting of the American Association for the Advancement of Science in January seven hypotheses about the extinctions were presented. Alvarez was one of the speakers; he reported on a recent finding that appears to weaken his group's hypothesis with respect to the catastrophic extinction of dinosaurs. At Hell Creek, a fossil-rich site in northeastern Montana, Alvarez and William Clemens of Berkeley found an iridium enrichment at the base of a layer of coal dating from the Paleocene, the first epoch of the Tertiary period. The most recent dinosaur fossil at Hell Creek, a *Tyrannosaurus* femur, lay three meters below the iridium layer and fossils of typical Paleocene animals lay two meters below the layer. Thus, on the Hell Creek evidence, the dinosaurs in this part of the New World were gone long before the event that deposited the iridium.

In the Alvarez hypothesis the actual mechanism of extinction was a blotting out of solar radiation by clouds of dust thrown up into the atmosphere. John O'Keefe of the Goddard Space Flight Center of the National Aeronautics and Space Administration favors a similar mechanism but suggests the cause was a volcanic eruption on the moon.

Dewey McLean of the Virginia Polytechnic Institute and State University also argues for a volcanic catastrophe, but one on the earth: the huge Deccan eruption that left 300,000 square kilometers of the Indian subcontinent buried under lava more than two kilometers deep. The massive eruption, McLean suggests, saturated the atmosphere with carbon dioxide and made the surface waters of the oceans acidic, dissolving the foraminiferans' shells and causing a "greenhouse effect" fatal to land organisms. McLean's hypothesis also explains the iridium layer without invoking an asteroid collision: with oceanic populations greatly reduced, the steady rain of iridium-rich micrometeorites would have collected on the sea floor undiluted by the normal and equally steady rain of plankton skeletons.

Stefan Gartner of Texas A. & M. University proposes still another catastrophic event, this one resulting from the reshuffling of the continents by plate tectonics. During the Cretaceous, Gartner contends, the Arctic Ocean was a landbound lake, turned from salt to brackish water by steady rainfall. When Greenland and Norway separated, the Arctic water surged free, forming a brackish layer 150 meters deep over the surface of the other oceans. This sudden environmental change destroyed the plankton, interrupting the ocean food chain and leading to mass marine extinctions. The decline in ocean salinity also gave rise to fluctuations of temperature and rainfall on land that caused the terrestrial extinctions.

Leo Hickey of the Smithsonian Institution supports Clemens' gradualist hypothesis for the land extinctions but agrees that a minor asteroid impact, enough to darken the sky for a short time, could have been responsible for the extinctions in the sea. Finally, Thomas J. M. Schopf of the University of Chicago, another gradualist, suggests that the proposed sudden mass extinctions were neither sudden nor massive. "Extinction is the normal way of life," Schopf stated at the AAAS meeting. As for the dinosaurs, they "had the bad fortune to be in the wrong place at the wrong time, a fate that eventually comes to all species."

Terabits by Soliton

Solitons, or solitary waves, neither crest nor dissipate; instead they propagate without changing their shape. Until recently, however, it was widely thought they were so difficult to generate that they could not be put to work. Now Akira Hasegawa and Yuji Kodama of Bell Laboratories have found a use for solitons: they can modulate a train of light waves into pulses that carry information through optical fibers. Moreover, solitons are not hard to generate: under suitable conditions any input waveform will rapidly alter its initial shape and form at least one soliton. Because a pulse of light modulated by a soliton can be as brief as a trillionth of a second, a communications system based on soliton pulses would have an enormous capacity for transmitting information. Hasegawa and Kodama discuss soliton communication in *Proceedings of the IEEE* (Institute of Electrical and Electronics Engineers).

The classic example of a soliton was an isolated wave on the surface of the water in a channel, observed in 1844 by the British naval architect John Scott Russell. In an optical fiber the kind of soliton being considered as an aid to communications would not be formed by the individual electromagnetic waves that function as carrier waves for the signal. Instead the soliton would be a wave envelope, a waveform superposed on the continuously generated carrier waves in order to modify them and so transmit information. Mathematically, however, the envelope solitons are analogues of the classic solitary waves.

Any complex wave is made up of numerous simple sine-wave components at various frequencies, which add by superposition to create the complex wave. The complex wave tends to disperse because the velocity of the low-frequency components is greater than the velocity of the high-frequency ones. As the wave propagates and its components separate, its initial shape tends to flatten out. In a soliton, however, the dispersion is balanced by the continuous transfer of the energy in the lower-frequency components to the higher-frequency components. Because the frequency dispersion and the energy transfer must cancel each other exactly, it once seemed solitons would be a rare natural phenomenon that could not be reproduced without precise controls. It is now known they can be made readily.

Solitons in an optical fiber were first produced last year by Linn F. Mollenauer, Roger H. Stolen and J. P. Gordon, also of Bell Laboratories. By superposing the soliton wave on a carrier wave at infrared frequencies, the investigators were able to generate pulses of light that lasted for only two trillionths of a second. Hasegawa and Kodama calculate that this result can be improved by a factor of two. They contend that in practical applications a transmission rate of one terabit (one trillion bits per second) could be achieved over distances of 30 kilometers with 10 watts of input power. An input of 10 watts is too much to be feasible, but solitons could transmit 100 billion bits per second with an input power of 10 milliwatts. Such a rate would represent an improvement over the best current transmission rates in optical fibers by a factor of 300.

hexa·photo·cybernetic
The Possibilities are Endless.

Stopped-Down 90

Six-mode exposure control
System versatility. Newer electronics
for wider applications

The Canon A-1 is one of the world's most advanced automatic SLR cameras. Combining the finest in optical and mechanical engineering with the most sophisticated electronics, it's technology applied to give you the ultimate in creative control. At the touch of a button.

Depending on your subject, you can choose from six independent exposure modes to achieve the results you want.

1 Shutter-Priority: You select the shutter speed, to freeze the action and prevent camera shake or create an intentional blur. The A-1 automatically selects the appropriate lens opening.

2 Aperture-Priority: Control the area in focus by selecting the lens opening for the effect you want. The A-1 matches with the right speed.

3 Programmed: When you need to shoot fast, just focus. The A-1 will select *both speed and aperture* for great results.

4 Stopped-Down: For extreme close-up or specialized photography, a bellows, a microscope or almost anything can be attached to the A-1. It's still automatic.

5 Flash: Totally automatic flash photography, of course, with a wide variety of Canon Speedlites to choose from.

6 Manual: Yes. For those times when you absolutely want to do it all yourself. To experiment. To explore the possibilities.

There are over forty fine Canon lenses ranging from Fish Eye to Super Telephoto, plus accessories to meet every need. If you can't photograph your subject with a Canon A-1, it probably can't be photographed.

From the sophistication of its LED viewfinder display, to a ruggedness that allows up to five-frame-per-second motor drive, the Canon A-1 represents an incredible technology. At a price that makes owning one a definite possibility.

Canon A-1

Canon USA, Inc., One Canon Plaza, Lake Success, New York 11042 • 140 Industrial Drive, Elmhurst, Illinois 60126 •
6380 Peachtree Industrial Blvd., Norcross, Georgia 30071 • 123 Paularino Avenue East, Costa Mesa, California 92626 •
Bldg. B-2, 1050 Ala Moana Blvd., Honolulu, Hawaii 96814 • Canon Canada, Inc., Ontario

Chrysler LeBarons as shown: Sedan **$8,712**;* Coupe **$8,618**.*

Chrysler Cordoba as shown: **$9,566**.*

Chrysler New Yorker as shown: **$11,391**.*

THE NEW CHRYSLERS

We've re-engineered the American luxury car.

The Chrysler LeBaron Series. They are like no other cars in America, Europe or Japan.

No other six-passenger car combines the high mileage, luxury and unexpected price of the surprising LeBaron coupe and sedan.** They stand alone.

LeBaron is built in one of the world's most technologically advanced plants. Computer-directed robots assure precision welding, and along with special sound insulation this accuracy helps make LeBaron a car of unusual quality and quietness on the road.

With advanced front-wheel-drive and a refined suspension system, LeBaron assures sure-footed traction, cornering and a sense of confidence behind the wheel. And LeBaron is equipped with a sophisticated electronic fuel system that's computer-directed.

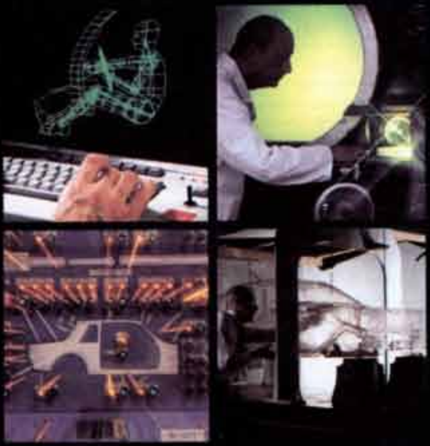
And to further assure quality, the LeBaron Quality Assurance Program of 200 inspections is the most demanding ever for a Chrysler.

CHRYSLER QUALITY

In LeBaron you will find your new car. In your new car you will find the genius of Chrysler engineering. The series includes the ultimate luxury, LeBaron convertible† and the functional luxury of the new Town & Country.

Chrysler LeBaron coupe base sticker price $8,143.*

40 EST HWY 26 EPA EST MPG



Today Chrysler engineering is setting new standards for the auto industry. In Computer Aided Design (CAD) and Computer Aided Testing (CAT) Chrysler is so advanced that two auto companies and Boeing bought Chrysler programs. In Computer Aided Manufacturing (CAM) our automatic welders help build quieter cars with better fit and appearance. And by predicting parts endurance through plastic simulation, we build longer life into our new Chryslers.

The Chrysler Cordoba. Engineering excellence in a personal luxury car.

Here is the sophisticated design and road manners a fine car should possess. Cordoba's riding comfort and privacy from road noise are simply astounding, the result of a refined suspension system created with the aid of computer design.

Power steering, power front disc brakes and automatic transmission are all standard, all matched to a smooth and supple six-cylinder engine that was honed to excellence for drivers who love to drive.

Come. Come to Cordoba. Come for elegance. Come for excellence. Come for Chrysler engineering.

Chrysler Cordoba base sticker price $9,197.*

22 EST HWY 18 EPA EST MPG

The Chrysler New Yorker. The best luxury car value in America.

The sleek, new contemporary-sized New Yorker is crafted in the great Chrysler tradition of personal attention to engineering quality. And it is fitted with so many standard luxuries that it is the best-equipped luxury car value in its class. No other luxury car matches its combination of astonishing equipment and favorable price.

New Yorker standard luxuries include • 60/40 Cloth & Vinyl Seats with Passenger Recliner • Air Conditioning • Automatic Transmission • Electronic Ignition System • Dual Horns • Power Front Disc Brakes • Power Windows • Power Steering • Premium Wheel Covers • WSW Steel Belted Radials • AM Radio • Electronic Digital Clock • Special Sound Insulation • Galvanized Steel for Corrosion Protection • Front Torsion Bar Suspension System • Halogen Headlamps • Rear Window Defroster • Premium Plush Pile Carpeting • Dual Remote Mirrors • Tinted Glass • Trunk Dress Up • Bumper Guards • Warning & Interior Light Packages • And More.

The plush new New Yorker is the flagship of Chrysler luxury, Chrysler value, Chrysler engineering. It drives on pride.

22 EST HWY 18 EPA EST MPG

Chrysler

THE NEW CHRYSLER CORPORATION

*Sticker price excluding taxes, title and destination charges. Cordoba wide WSW tires shown $116 extra. **Use EPA est. MPG for comparison. Actual mileage may vary due to speed, trip length and weather. Highway and Calif. mileage lower. LeBaron ests. for coupe and sedan only. †Convertible available late spring.

© 1982 SCIENTIFIC AMERICAN, INC

Giant Molecular-Cloud Complexes in the Galaxy

Consisting almost entirely of hydrogen molecules, they are the most massive objects in the galaxy. They also give rise to most of the galaxy's stars. A decade ago their presence was unknown

by Leo Blitz

Astronomers who study how stars form are faced with problems similar to those faced in the 19th century by physiologists who sought to study the growth and development of the human embryo. When a star is forming, it is hidden from view; it is surrounded by the opaque cloud of dust and gas from which it is condensing. Hence when the cloud is dispersed and the star is finally seen, the processes that make it shine are already well under way. Like a newborn infant, the star has lost the traces of how it came into being.

One aspect of how stars form is becoming accessible. On a large scale it has been known for some time that the spiral arms of disk-shaped galaxies such as our own are delivery rooms for stars. More recently it has become clear that the specific birthplaces for most of the stars forming today in the spiral arms of our galaxy are giant complexes of clouds composed primarily of hydrogen molecules. These complexes share with certain star clusters the distinction of being the most massive objects in the galaxy. They are, however, much more numerous than the clusters. Indeed, they lock up a sizable fraction of all the gas in the galaxy. Yet little more than 10 years ago they were completely unknown.

Interstellar Gas

A discussion of the giant molecular-cloud complexes should begin with a description of the matter of which they consist, the matter in interstellar space. Looking at the night sky one might suppose the space between the stars is empty, but nowhere in our galaxy is matter completely absent. To be sure, the matter is extremely rarefied: the average density of atoms of hydrogen (the most abundant element) in interstellar space is less than one atom per cubic centimeter. The interstellar matter takes many forms. In some places it is glowing gaseous nebulae, in others it is dark patches in the sky that were first noticed because they block the light from the stars behind them. Wherever astronomers look and apparently no matter what scale they choose for their observations the interstellar matter is distinctly nonuniform. That is, the matter between the stars is organized everywhere into regions of higher and lower density. One speaks of the high-density regions as clouds and of the low-density regions as intercloud gas.

Until recently it was thought the clouds consist of hydrogen atoms with a small admixture (about 10 percent) of helium atoms and trace amounts (adding up to no more than a few percent) of atoms heavier than helium atoms. In addition about 1 percent of the mass of a cloud was (and still is) thought to be solid dust particles of microscopic size. The dust particles attenuate the light of background stars: they absorb it at visible wavelengths and reradiate it in the infrared. As a result the clouds tend to be opaque.

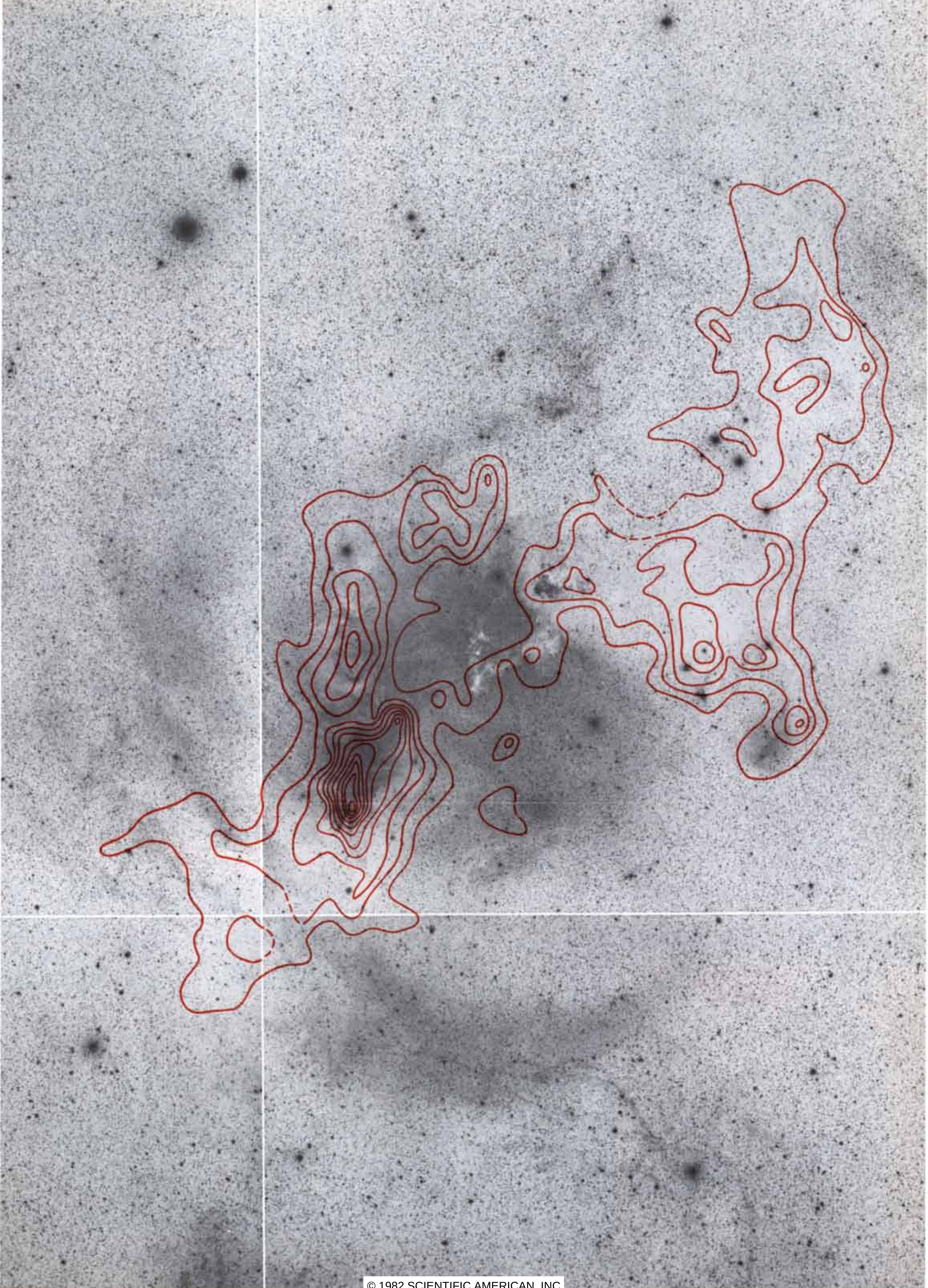
In the past decade this view of the composition of the clouds has changed in one respect: it has been discovered that many of the clouds are made up not of atoms but of molecules. The difference is crucial: molecular clouds are typically much colder and denser than atomic clouds. At present the most widely accepted view is that between 10 and 50 percent of the gas in the galaxy is molecular. An individual cloud is almost entirely molecular or almost entirely atomic, but it appears that the largest molecular clouds are embedded in envelopes of atomic gas.

A molecule of course consists of two or more atoms bound together in a stable unit. It is the smallest parcel into which a compound can be divided and still maintain its chemical properties. Molecules, like atoms, emit and absorb electromagnetic radiation only at specific and characteristic wavelengths. Each such wavelength corresponds to a spectral line: a narrow peak (or a bright line) in the spectrum of the radiation emitted by the molecule or a sharp dip (or a dark line) in the spectrum of the radiation passing through a gas made up of the molecules. For a typical molecule hundreds of spectral lines are detectable, and the wavelength of each line can generally be measured with an accuracy of better than one part in a million. Hence if several spectral lines are detected in the radiation from a particular direction in the sky, and if the lines are known to characterize a particular molecule, the probability of a misidentification is vanishingly small.

Rotating Molecules

This is not to say the identification of the molecules in a giant molecular-cloud complex is completely straightforward. The most notable difficulty is that hydrogen molecules cannot be identified directly from their radiation at the temperature of about 10 degrees Kelvin (10 degrees Celsius above abso-

MOLECULAR COMPLEX in the constellation Monoceros is a clumpy distribution of cold dust and molecular gas some 800 parsecs (2,600 light-years) from the solar system. The dust blocks the light from the stars behind it, but the extent of the complex is best revealed by the electromagnetic radiation that carbon dioxide molecules emit from inside the complex at a wavelength of 2.6 millimeters. Here the pattern of the radiation, as mapped by the author with the 1.2-meter radio telescope at Columbia University, is displayed as a set of contour lines on a mosaic of four plates from the Palomar Sky Survey. The contours suggest that the complex consists of two clouds joined by a slender bridge. The lower cloud lies behind a glowing nebula. (The photographs are negatives, and so the nebula looks dark.) The nebula consists of hydrogen that has been dissociated into atoms and then ionized by the ultraviolet radiation from a single massive star, visible as the black dot at the top of the densest clump in the lower cloud.



lute zero) that typifies most parts of a molecular cloud. Fundamentally the molecules in a giant molecular-cloud complex are rotating, and the rotation (as viewed from a distance) is equivalent to a periodic oscillation of the molecule's electrons. In effect the rotation shuttles the electrons back and forth, much as they would move in a radio antenna. A given molecule rotates only at discrete rates, and the transition from a higher rate to a lower one is accompanied by the emission of radiation by the electrons.

It is thought 99 percent of the mass of a giant molecular cloud is molecular hydrogen. The hydrogen molecule, however, is symmetrical about its axis of rotation. Therefore the average position of the two electrons in the molecule is on the axis of rotation. As a result the rotation of each molecule (as viewed from a distance) is not accompanied by an oscillation of electrons. What is more to the point, a change in the rotational state of the molecule is not accompanied by the emission of radiation.

Fortunately many of the molecules present in trace amounts in a giant molecular-cloud complex are not symmetrical around their axis of rotation. In such molecules the average positions of the electrons are at a distance from the axis. The collisions between molecules (both symmetrical and asymmetric) in a complex are sufficiently frequent and sufficiently energetic to yield numerous transitions between rotational states. The transitions of the asymmetric molecules account for the radiation the complex emits. Typically the radiation has a wavelength between a few centimeters and a fraction of a millimeter; thus it can be detected with radio telescopes on the earth.

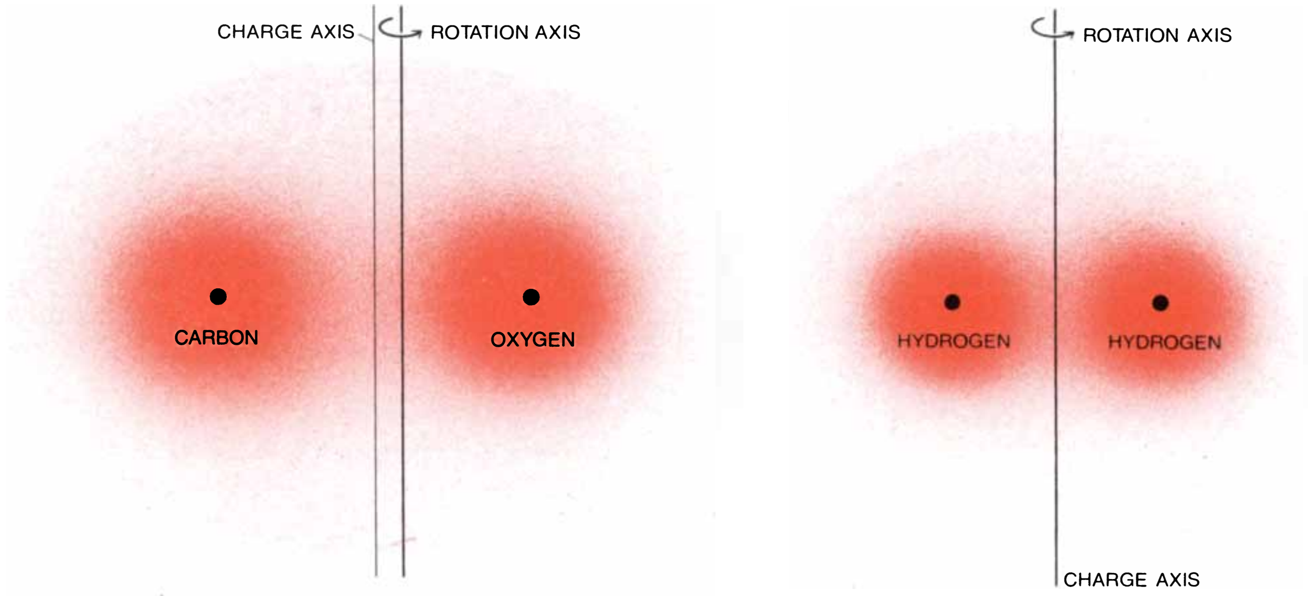
As of now some 53 different molecules have been firmly identified in interstellar space and four more have been tentatively identified. Many of them, including some of the most abundant ones, are simple molecules also found on the earth. Four examples are H_2O (water), CO (carbon monoxide), NH_3 (ammonia) and H_2CO (formaldehyde). Many others, including OH and HCO^+, are known on the earth only in the laboratory; they are too reactive to exist at the densities of matter typical on the earth. Still others, including C_4H, HC_7N and HC_9N, have never been detected on the earth even in the laboratory. These last have been identified in interstellar space on the basis of theoretical calculations. Specifically, the assignment of an otherwise unexplained spectral line to a certain hypothetical molecule enables one to predict the wavelengths of other lines in the molecule's spectrum. The subsequent detection of some of those other lines confirms the presence of the molecule.

One of the more complex gaseous molecules detected so far in interstellar space is CH_3CH_2OH, known on the earth as ethyl alcohol or vodka. According to Ben M. Zuckerman of the University of Maryland at College Park and his colleagues, the amount of the stuff at the center of our galaxy could fill the volume of the earth more than 10,000 times. Estimates of the water content in the molecular clouds containing the ethyl alcohol indicate, however, that the water is 100,000 times more abundant. The alcohol is therefore bottled in the cloud at only .002 proof. In addition it is laced with such unpotables as HCN (hydrogen cyanide).

Of all the gaseous molecules detected so far the large majority are composed of hydrogen, carbon, nitrogen and oxygen. (A few include silicon and sulfur.) This composition mirrors the relative cosmic abundances of the chemical elements. Hydrogen and helium (which is not found in molecules because it is chemically inert) are the principal elements that formed when the universe began, and carbon, nitrogen, oxygen and helium are the principal elements produced by nuclear fusion deep inside stars. The elements synthesized inside stars are released into interstellar space in several ways. Some stars (perhaps all of them) emit "winds" of particles, including the atomic nuclei synthesized by nuclear fusion. Some stars blow off an outer envelope of matter in the process of becoming novas. Finally, some stars explode catastrophically and become supernovas. Most of their mass is expelled into interstellar space.

The Importance of CO

The molecule most responsible for the growing understanding of the large-scale properties of giant molecular-cloud complexes is carbon monoxide, which was first identified in interstellar space by Robert W. Wilson, Keith B. Jefferts and Arno A. Penzias of Bell Laboratories. Carbon monoxide is only about a ten-thousandth as abundant as molecular hydrogen in space, but it radiates strongly at low temperatures. Moreover, more than any other molecule detected in space it appears to be coextensive with molecular hydrogen.



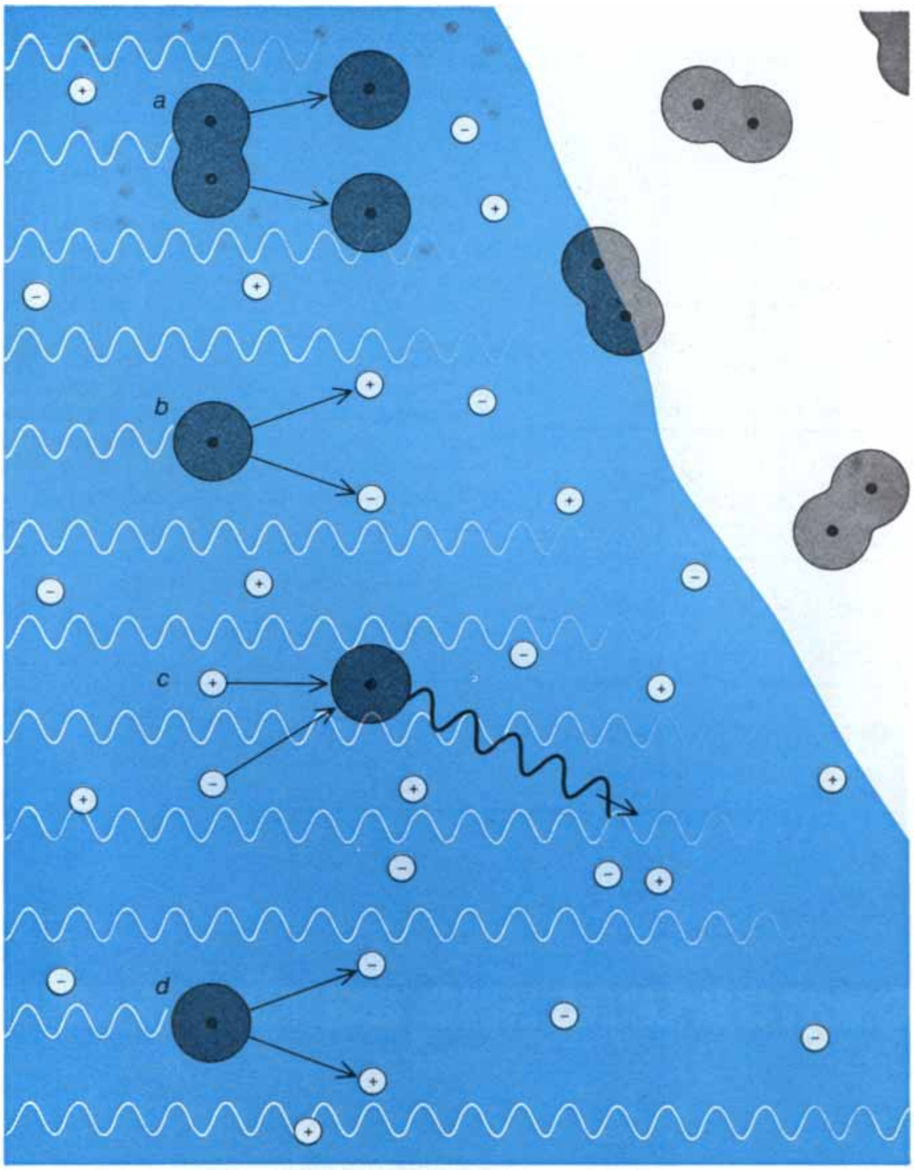
ROTATING MOLECULES are the source of the electromagnetic radiation detected from giant molecular complexes. Carbon monoxide (*left*) is an asymmetric molecule: its oxygen atom is more massive than its carbon atom; hence the molecule spins on an axis that is closer to the former. The resulting asymmetry of the molecule's electric charge around the axis of rotation means that the charge oscillates as the molecule spins. When the spin decreases from one discrete rate to another, the molecule emits radiation at characteristic radio wavelengths. Hydrogen (*right*) is a symmetrical molecule: it does not emit radiation merely because it spins. Therefore molecular hydrogen cannot be identified directly in most of a giant molecular complex even though it makes up more than 99 percent of the mass of the complex.

(It is collisions with molecular hydrogen that give carbon monoxide a variety of rotational states and allow it to emit radiation.) Thus the radiation from carbon monoxide serves as a tracer of the interstellar molecular gas, much as a stain can reveal undetected structure in living tissue.

In 1973 Kenneth D. Tucker and Marc L. Kutner of Columbia University, working with Patrick Thaddeus of the Goddard Institute for Space Studies, attempted to determine the extent of a giant molecular cloud in proximity to the Horsehead Nebula, a bright gaseous patch in the constellation Orion. They found that the cloud spans an area amounting to almost six square degrees of the sky, or 30 times the area covered by the moon.

The great extent of the cloud presented a curious problem. Radio telescopes collect the radiation from an area of the sky inversely proportional to their diameter; hence a large telescope yields measurements of the radiation from small areas. The telescope used by Tucker, Kutner and Thaddeus was the five-meter instrument at the McDonald Observatory in Texas. It can measure the radiation from the carbon monoxide in circular patches of the sky with a diameter of only 2.3 minutes of arc. With a week of observing time the investigators sampled the radiation from only about 150 positions in the cloud; to map the cloud completely in 2.3-arc-minute patches would have taken six months of full-time work. In any case the molecular cloud they studied is now known to be only one of about 4,000 such clouds in the galaxy. Clearly making a complete map of the clouds in the galaxy would have required an unrealistically large amount of time.

The simplest way to circumvent the problem was to design and build a smaller radio telescope, one that would examine a larger patch of the sky with each observation. Thaddeus took that approach: with several graduate students and a few technicians he designed and built a 1.2-meter telescope on the roof of the Pupin Physics Laboratory at Columbia University in the middle of Manhattan. No professional astronomers had attempted observations in New York City for half a century, but at the radio wavelengths at which interstellar carbon monoxide is detected New York remains as pristine as it was when it was populated only by Indians. (The carbon monoxide produced today by urban combustion does not affect the astronomical observations. Its spectral lines are broadened by the pressure of the earth's atmosphere to such a degree that they can be unambiguously subtracted from a celestial signal.) Thaddeus' new telescope could sample 12 times more sky with each observation than had been possible before, and so for the first time a large number of giant molecular clouds could be fully mapped, although at lower resolution. The group that did the work at Columbia included Gregory P. Baran, Gordon Chin, Richard S. Cohen, Hong-Ih Cong, Thomas Dame, George R. Tomasevich and me.



RECOMBINING ATOMS are the source of the electromagnetic radiation from the glowing nebulas (called H II regions) with which giant molecular complexes such as the one shown on page 85 are associated. The ultraviolet radiation from young, massive stars dissociates hydrogen molecules into atoms (*a*) and then ionizes the atoms (*b*), thereby creating a region occupied almost entirely by free protons (+) and electrons (−). When these particles recombine into atoms (*c*), they emit the radiation that makes the nebula visible. Soon the atoms are ionized again (*d*). Most of the ultraviolet radiation inside the H II region escapes from the region only after it has been converted by such a process into electromagnetic radiation at longer wavelengths.

Meanwhile several other astronomers working with large radio telescopes surveyed, although incompletely, the distribution of carbon monoxide in giant molecular clouds. Charles J. Lada and Bruce G. Elmegreen of the Center for Astrophysics of the Harvard College Observatory and the Smithsonian Astrophysical Observatory worked with Anniela I. Sargent of the California Institute of Technology to map the extent of two giant cloud complexes. Philip M. Solomon of the State University of New York at Stony Brook and Nicholas Z. Scoville of the University of Massachusetts at Amherst studied the inner part of the galaxy and concluded that a large fraction of the gas there is contained in giant molecular clouds.

By 1977 it was understood that the largest molecular clouds share a number of characteristics. First, each such cloud has a mass between 100,000 and 200,000 times the mass of the sun. The clouds are thus the most massive objects in the galaxy. The largest globular star clusters have a comparable mass, but there are about 20 giant molecular clouds for each of the galaxy's globular clusters. Second, the density of the gas in each cloud is 100 times greater than the average density of the interstellar medium. Nevertheless, the interior of a cloud is far more rarefied than the best vacuum ever produced in laboratories on the earth. Third, the gas in each cloud is organized into clumps whose density is 10 times greater than the average density in the cloud. In some clouds, clumps with densities much greater than that are observed.

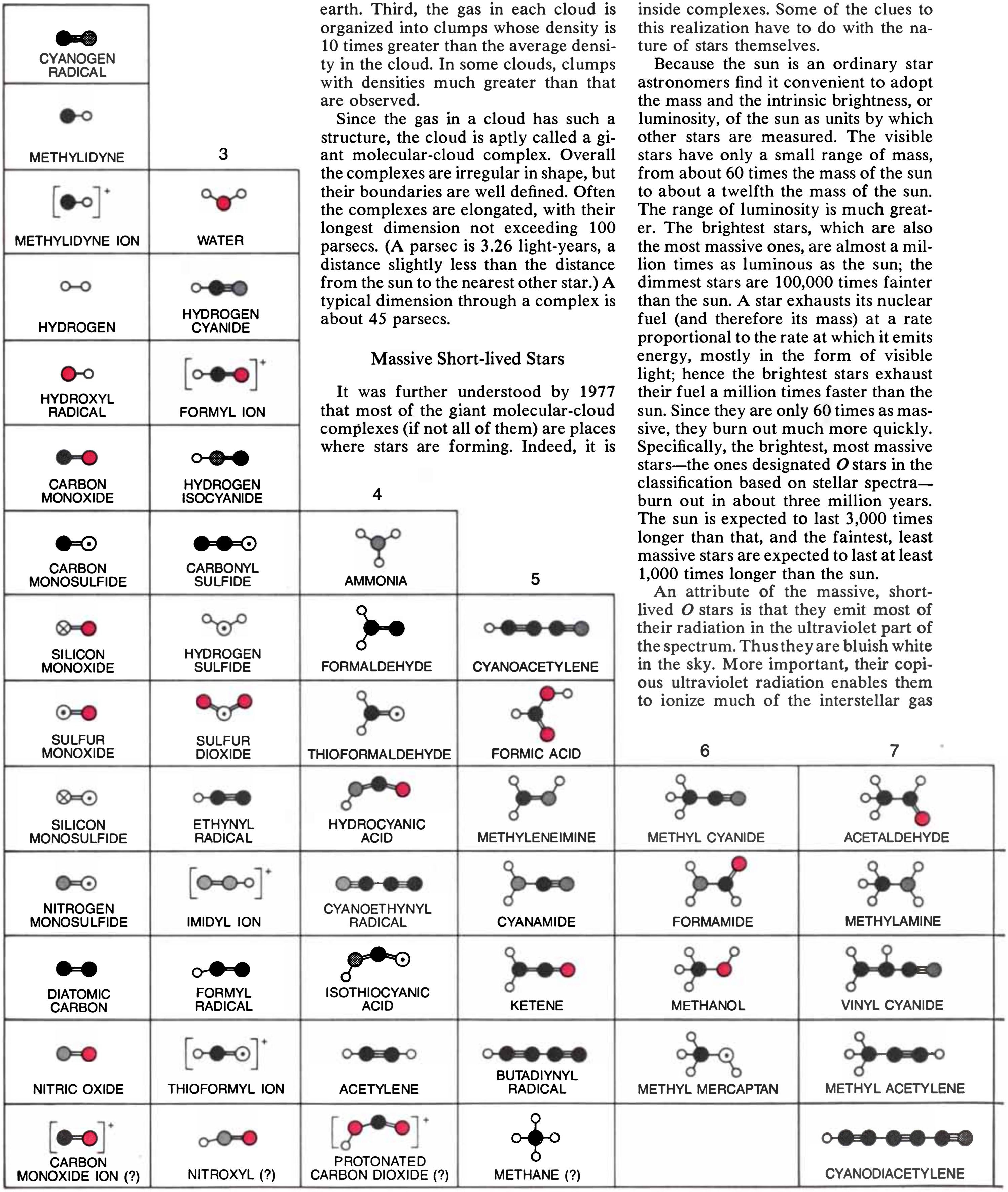
Since the gas in a cloud has such a structure, the cloud is aptly called a giant molecular-cloud complex. Overall the complexes are irregular in shape, but their boundaries are well defined. Often the complexes are elongated, with their longest dimension not exceeding 100 parsecs. (A parsec is 3.26 light-years, a distance slightly less than the distance from the sun to the nearest other star.) A typical dimension through a complex is about 45 parsecs.

Massive Short-lived Stars

It was further understood by 1977 that most of the giant molecular-cloud complexes (if not all of them) are places where stars are forming. Indeed, it is now agreed that most of the stars currently forming in the galaxy are forming inside complexes. Some of the clues to this realization have to do with the nature of stars themselves.

Because the sun is an ordinary star astronomers find it convenient to adopt the mass and the intrinsic brightness, or luminosity, of the sun as units by which other stars are measured. The visible stars have only a small range of mass, from about 60 times the mass of the sun to about a twelfth the mass of the sun. The range of luminosity is much greater. The brightest stars, which are also the most massive ones, are almost a million times as luminous as the sun; the dimmest stars are 100,000 times fainter than the sun. A star exhausts its nuclear fuel (and therefore its mass) at a rate proportional to the rate at which it emits energy, mostly in the form of visible light; hence the brightest stars exhaust their fuel a million times faster than the sun. Since they are only 60 times as massive, they burn out much more quickly. Specifically, the brightest, most massive stars—the ones designated *O* stars in the classification based on stellar spectra—burn out in about three million years. The sun is expected to last 3,000 times longer than that, and the faintest, least massive stars are expected to last at least 1,000 times longer than the sun.

An attribute of the massive, short-lived *O* stars is that they emit most of their radiation in the ultraviolet part of the spectrum. Thus they are bluish white in the sky. More important, their copious ultraviolet radiation enables them to ionize much of the interstellar gas



MOLECULES IDENTIFIED IN INTERSTELLAR SPACE are now 53 in number. (Four more, indicated by a question mark, are tentatively identified.) With the exception of hydrogen each accounts for less than .1 percent of the mass of a typical giant molecular-cloud complex. Many account for less than .0001 percent. Some of them are molecules common on the earth; others, such as the long cyano-

that surrounds them. That is, the radiation dissociates molecules into their constituent atoms. It also strips electrons off the nuclei of atoms. When the free electrons recombine with the nuclei (although they can do so only briefly before they are ionized again), the atoms they form emit radiation at characteristic wavelengths. Most notably, the recombination of electrons with protons to form hydrogen atoms gives rise to an intense red spectral line. In this way groups of young stars come to energize fluorescent gaseous nebulas. Astronomers have given these objects (some of the most beautiful in the heavens) the ungainly name H II regions.

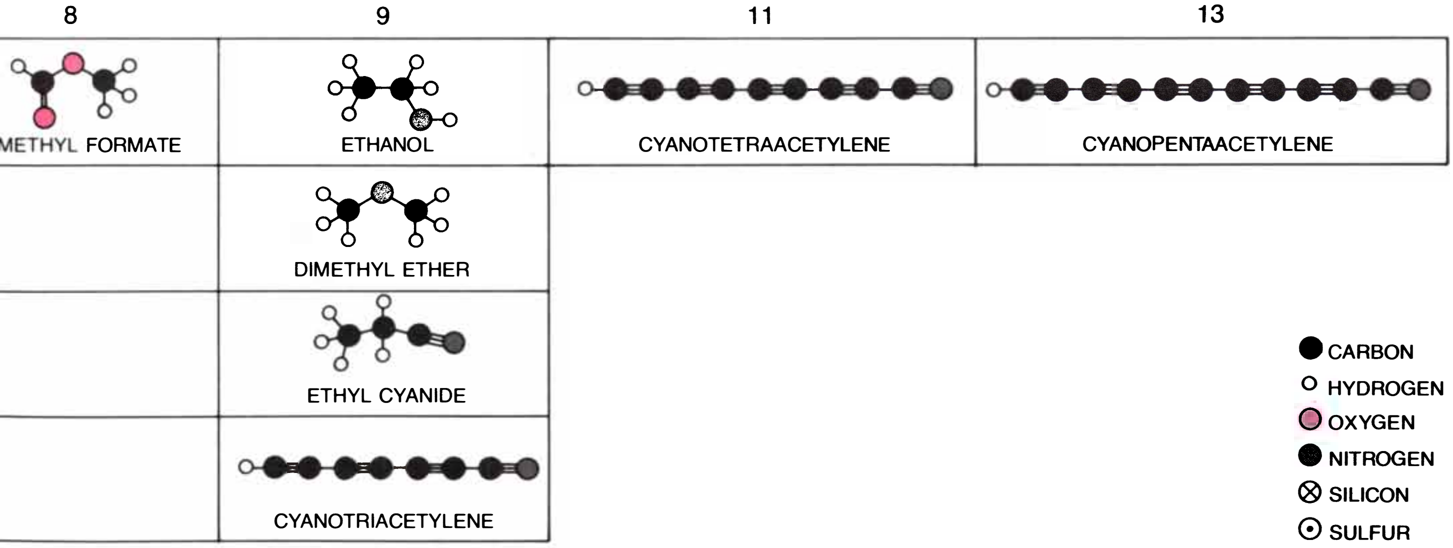
When astronomers began searching for carbon monoxide with radio telescopes in the early 1970's, it soon became clear that the best places to find it were in the vicinity of H II regions. It can be said, therefore, that giant molecular-cloud complexes and H II regions are intimately related. Indeed, a recent survey by Michel Fich of the University of California at Berkeley, Antony A. Stark of Bell Laboratories and me shows that about 80 percent of all the visible H II regions in the galaxy are associated with molecular clouds. The *O* stars that ionize the H II regions have lifetimes so short that they cannot have moved far from their birthplace. (Even at a speed of 10 kilometers per second with respect to the gas of a giant molecular-cloud complex such a star would need an entire lifetime of three million years to travel 30 parsecs, a distance that might leave it still inside the complex.) It seems a safe deduction that the massive, short-lived *O* stars form inside the complexes, and that their ultraviolet radiation ionizes the gas that surrounds them, making it into an H II region.

The radiation from the *O* stars not only ionizes the gas in a giant molecular-cloud complex; it also pushes it. Indeed, the gas can be ejected from the surface of a complex at a velocity of 10 to 12 kilometers per second. At such a speed the gas can escape from the cloud's gravitational influence. Hence the stars born in a giant molecular-cloud complex can help to destroy it. In fact, the rate at which *O* stars ionize and push the gas in a complex is high enough so that if stars arise continuously inside the complex, the complex can be destroyed in a few tens of millions of years.

There is reason to think the stars really do arise continuously. When an H II region and a giant molecular-cloud complex are contiguous, the pressure of the radiation leaving the H II region and the pressure of the particle wind from the *O* stars inside it combine to drive a shock into the complex. The shock is a front along which the temperature, pressure and density of a gas suddenly change. Its effect on the complex is to compress the molecular gas. Under normal circumstances the pressure inside a cloud roughly balances the cloud's self-gravitation, which would tend to collapse the cloud if its action were unopposed. What generates the pressure is a major unanswered question. The pressure arising from the heat in the cloud at a temperature of 10 degrees K. is known to be insufficient. Many investigators suspect that the magnetic fields in the gas are important. In any case it seems possible that when the gas is compressed by a shock, the self-gravitation of a clump of gas in the complex can be increased so much that the clump can no longer resist collapse. The resulting contraction could continue until the clump becomes a star.

Several years ago Elmegreen and Lada developed a theoretical model of the effects of a shock in a giant molecular-cloud complex. They analyzed the interaction of an H II region with a uniform layer of molecular gas and suggested that the shock caused by one generation of stars could initiate the formation of a new generation. More recently Richard I. Klein of the University of California at Berkeley and Maxwell T. Sandford and Rodney Whittaker of the Los Alamos National Laboratory have modeled the interaction of an H II region and a nonuniform cloud of gas, a more accurate representation of the structure of a giant molecular-cloud complex. They have suggested that a rather violent process may be important for forming stars: a clump of gas can be driven inward on itself (in effect, imploded) by the arrival of a shock.

In principle the triggered formation of new generations of stars by the older generations in a giant molecular-cloud complex could continue until all the molecular gas in the complex has been transformed into stars or dispersed into the interstellar medium. The presence of several generations of *O* stars in giant molecular-cloud complexes such as the ones associated with the H II regions known as the Orion Nebula and the Rosette Nebula suggests that this is the case. There is also observational evidence that stars can form spontaneously in a giant molecular-cloud complex. For example, it has been shown by Mark T. Stier, Daniel T. Jaffee and Giovanni G. Fazio of the Center for Astrophysics that massive stars are forming throughout the giant molecular-cloud complex associated with the Omega Nebula; they are not forming just in the part of the complex that might have been shocked by the nebula. Furthermore, young, low-mass stars of certain types, such as T Tauri stars, are thought to form in



acetylene chains, have never been detected on the earth even in laboratories. In this illustration the molecules are grouped according to the number of atoms in each one. The bond angles have been inferred in part from spectroscopic data. The list was compiled by Paul A. Vanden Bout and Neal J. Evans of the University of Texas at Austin and Barry E. Turner of the National Radio Astronomy Observatory.

parts of molecular complexes that have not been influenced by some external trigger. We do not yet know what fraction of stars are triggered into birth and what fraction form spontaneously, but then the formation of stars is one of the less well understood phenomena encountered in astronomy.

Cloud Lifetimes

One thing seems plain: all the currently known giant molecular-cloud complexes in the vicinity of the sun (say within 2,000 parsecs) show evidence that young stars have recently formed in them. Sometimes the complexes are associated with visible stars that can be seen to be young; sometimes stars still in the process of forming inside the complexes are detected by their infrared radiation. It therefore seems likely that the complexes begin to give rise to massive stars soon after the clouds themselves take shape. It also seems likely that if the complexes attained ages of more than 100 million years, the turbulence and viscosity within them would make them rounded and give them a clearly defined dense core. Instead the complexes are clumpy, irregular objects without an obvious central condensation. The view that the giant molecular-cloud complexes we see today are less than 30 million years old has not been accepted by all investigators. Nevertheless, the preponderance of the evidence indicates that the complexes are quite young by galactic standards, and specifically no older than a hundredth the age of the sun and the earth.

If the clouds are young, however, some mechanism must be continually forming them in the galaxy. Several possible mechanisms have been proposed. For one thing, it is suggested that molecular clouds probably stick together when they collide. (Their average collisional velocity is greater than the speed of sound in the clouds; hence the collision generates a shock wave, much like a sonic boom, that dissipates the energy of the collision. As a result the colliding clouds are not expected to rebound from each other.) Over a period of time giant clouds could emerge from the collision of many smaller clouds. The only problem is that the process is probably too slow: most calculations suggest it will work only if the clouds attain ages greater than a few hundred million years.

Another mechanism, proposed by Frank H. Shu and his colleagues at Berkeley, is that the formation of giant molecular-cloud complexes is mediated by the magnetic field permeating the interstellar medium. The field is carried by charged particles (ions) in the medium, but the ions convey the effects of the field to uncharged particles, mostly by collisions. In this way instabilities in the field can cause the medium to flow so that (Shu proposes) complexes form. A third mechanism, proposed by Elmegreen, is that the self-gravitation of the gas swept up in the spiral arms of the galaxy is sufficient in itself to make the gas coalesce into units with the size and the mass typical of a giant molecular-cloud complex. Still other mechanisms are sure to be put forward, but the solution to the problem of how the complexes form does not seem to be close at hand.

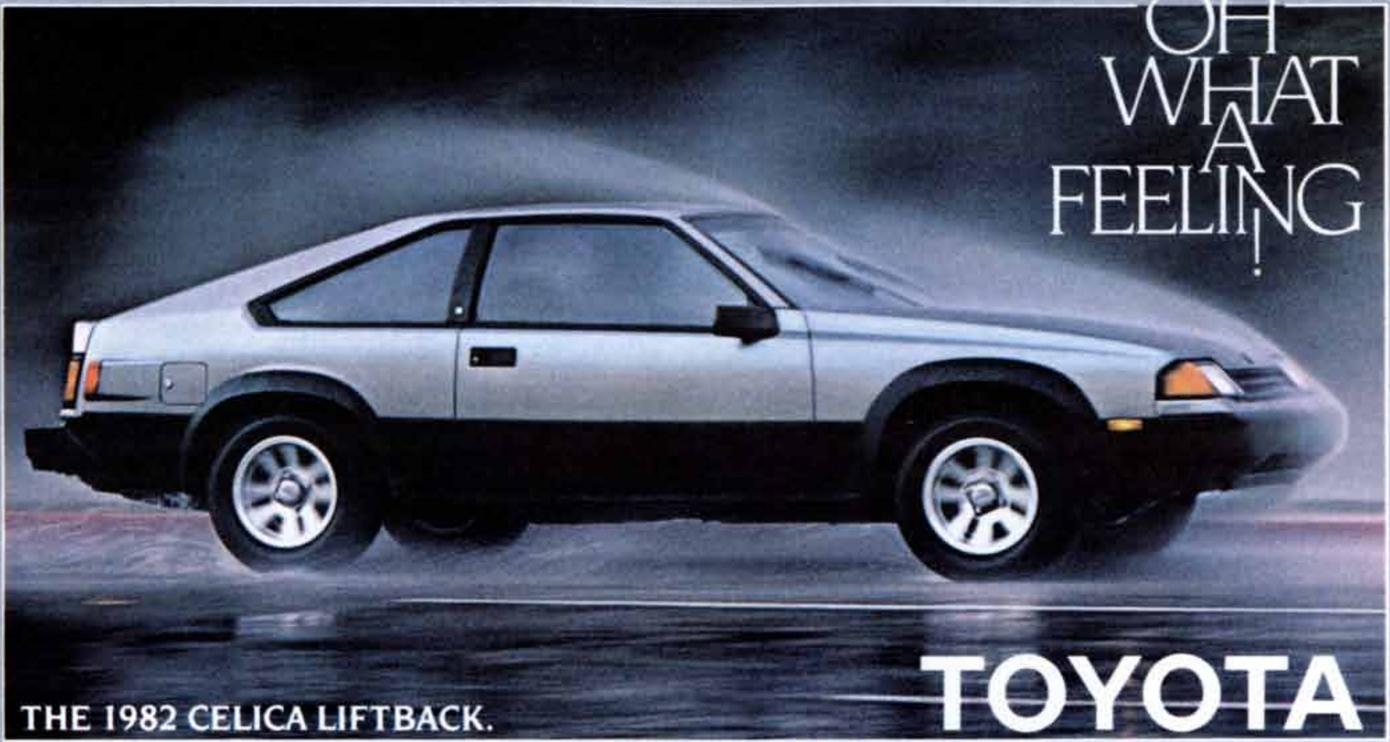
ULTRACOMPACT H II REGIONS invisible at optical wavelengths are inside a giant molecular-cloud complex in the constellation Aquila; they were detected by William J. Welch and John W. Dreher of the University of California at Berkeley with the radio telescopes in New Mexico called the Very Large Array. Each bright spot is one such region; its radio emission (*shown in false color*) is thought to result from the ultraviolet radiation of a single massive young star. Apparently the ultraviolet radiation has ionized hydrogen atoms, which in turn emit radio waves as they decelerate. At a distance from the solar system of about 15,000 parsecs each H II region can be no larger than .04 parsec in diameter; on that assumption the unseen stars inside them have only recently begun to ionize their surroundings. They may be less than 1,000 years old, which would make them the youngest cluster of massive stars known.

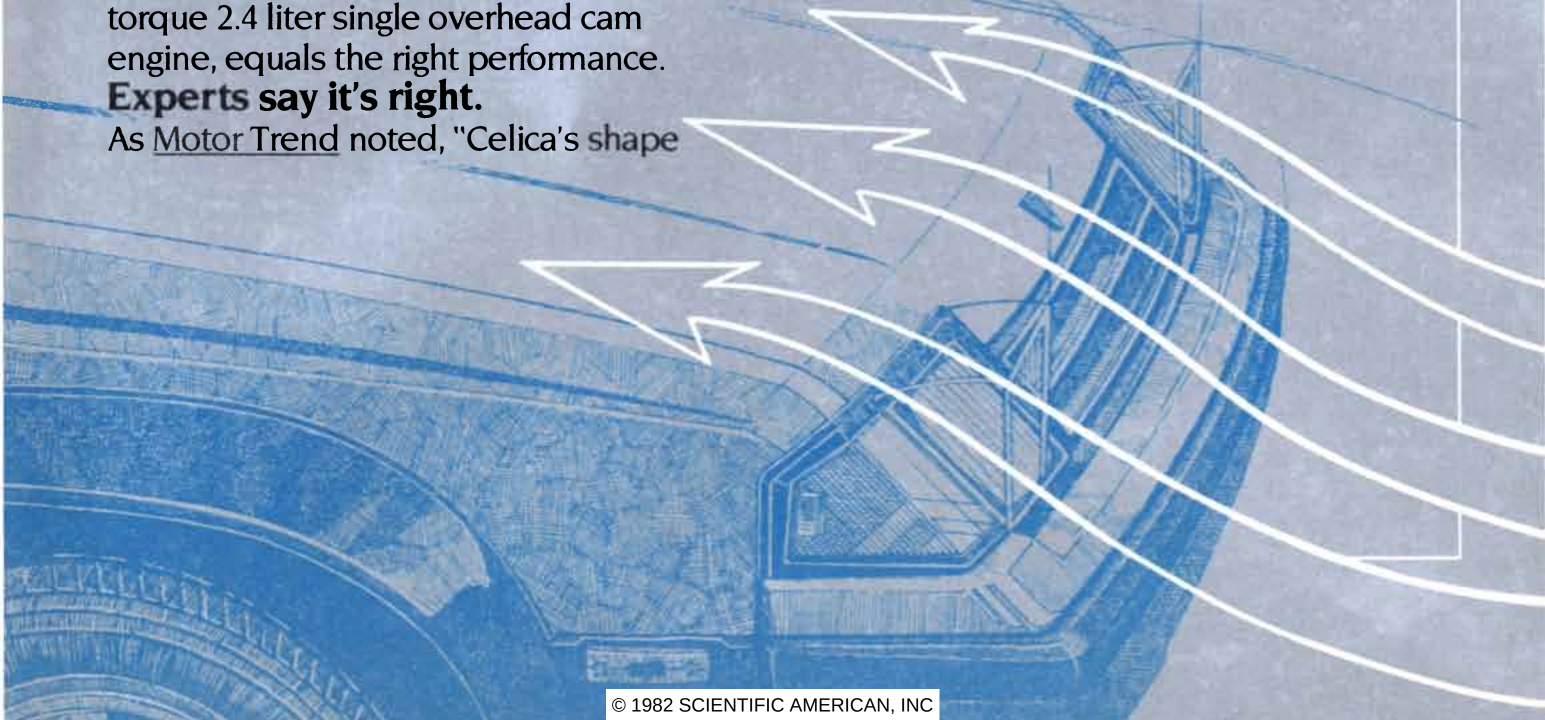
Cloud Histories

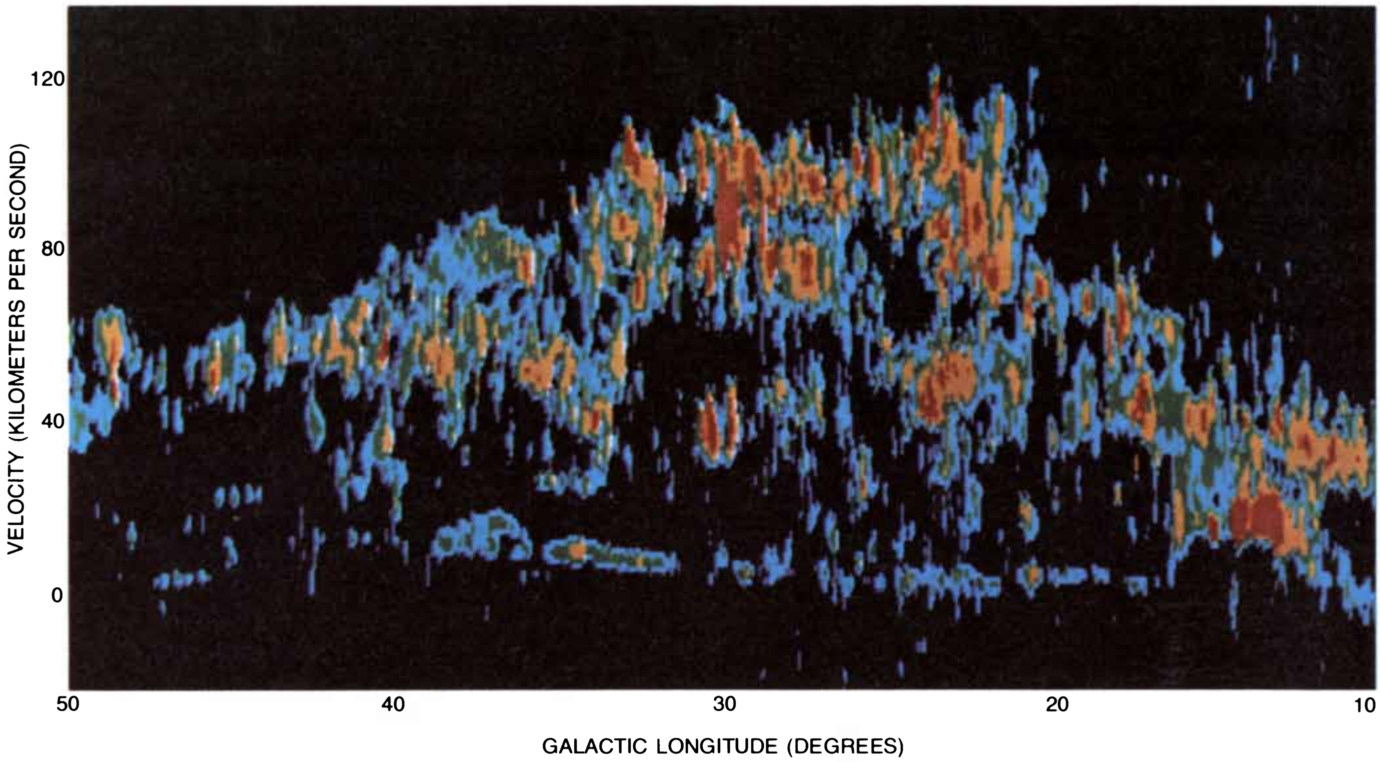
What is known about the complexes at present suggests that they go through a characteristic cycle of birth, middle age and death. Some of the details are speculative, but the following scheme is emerging.

As the gas in the galaxy moves in orbit around the galaxy's center it periodically encounters regions where the density of the interstellar medium is heightened. These regions are generated by density waves, which move through the disk of the galaxy much as sound waves move through air. The origin of the density waves, however, is gravitational. Specifically, the density waves are the response of the matter in a galaxy to a gravitational perturbation that can arise from a barlike distribution of stars in the core of the galaxy, from a companion galaxy or from both. (Our own galaxy has two massive companions, the Large and Small Clouds of Magellan. Moreover, some workers have suggested that the stars in the core of our galaxy have a barlike distribution.) The density waves can generate the galaxy's spiral-arm structure. In the regions of heightened density (which probably correspond to spiral arms or spiral-arm segments) the gas in the galaxy is somehow collected into large aggregates. They are the giant molecular-cloud complexes. If the gas is initially molecular, small clouds collect to form the giant complexes. If the gas is initially atomic, its compression causes molecules to form.

Once the complexes have arisen, stars of low mass form throughout them. Somewhere in most or all of the complexes a clump of gas is sufficiently mas-


OH
WHAT
A
FEELING!
TOYOTA
THE 1982 CELICA LIFTBACK.





SURVEY OF CARBON MONOXIDE in the inner part of the galaxy suggests that much of the molecular gas there is organized into giant cloud complexes like the ones nearer the solar system. In the chart the horizontal axis plots galactic longitude from 10 to 50 degrees. (Zero degrees signifies the direction to the center of the galaxy.) The vertical axis plots the velocity of carbon monoxide with respect to the average velocity of stars near the sun. A given velocity corresponds to a particular distance from the center. For example, the great concentration of gas (*red and yellow patches*) at velocities greater than 70 kilometers per second and longitudes from 22 to 34 degrees represents molecular clouds some 4,000 to 7,000 parsecs from the center. The survey was done by Philip M. Solomon of the State University of New York at Stony Brook and David B. Sanders and Nicholas Z. Scoville of the University of Massachusetts at Amherst.

sive to give rise to one or more *O* stars. At first these stars are invisible in optical telescopes. They can, however, be detected in the infrared. Soon they begin to ionize the gas that surrounds them. The result is a compact H II region. Even then the stars are optically invisible, but the H II region can be detected by infrared and radio telescopes.

The H II region grows until its ionized gas reaches the surface of the giant molecular complex. Now the *O* stars can be seen in the midst of the H II region they created. The radiation and the particle wind leaving the H II region dissipate some of the complex; thus they uncover some of the low-mass stars that have been forming. Meanwhile the radiation and the particle wind compress some of the remaining molecular gas, possibly triggering the formation of a new generation of massive stars. The new stars continue the processes of ionization and compression until most of the molecular gas in the complex is consumed or blown away. That takes about 30 million years. Alternatively, *O* stars form spontaneously in the complex (but generally in clusters) until the complex is consumed.

As the complex is being consumed relatively small fragments of the gas can be accelerated more or less intact into the interstellar medium by a process first described by J. H. Oort of the Leiden Observatory and Lyman Spitzer, Jr., of Princeton University. In this process one face of a fragment is ionized by stellar radiation; then the emission of radiation by the ions themselves propels the fragment out of the complex. Perhaps the process gives rise to at least some of the small clouds of molecular gas described by Bart J. Bok of the University of Arizona and known today as Bok globules. In any case the protons and electrons ejected from the complex ultimately recombine to form neutral hydrogen atoms. The gas composed of these atoms becomes part of the interstellar medium. It will resume an orbit around the center of the galaxy and will continue to follow that orbit until it once again encounters a spiral arm. The processes that form giant molecular-cloud complexes will then set in anew.

Distribution of the Complexes

The bright stars that form in giant molecular-cloud complexes clearly define the large-scale spiral structure of many galaxies other than our own. It is therefore natural to suppose the complexes in our galaxy might reveal a spiral-arm pattern. Several groups of workers have attempted to find one, but the results so far are ambiguous. For one thing, some of the surveys are incomplete. In addition the mapping of spiral arms in the inner parts of the galaxy is difficult because radio observations reveal only the velocity of interstellar carbon monoxide as a function of its direction in the sky. The velocity, which is determined by the rotation of the gas around the center of the galaxy, unambiguously gives the distance of a clump of gas from the center. It leaves unresolved, however, the choice between two possible distances from the solar system.

My own work, done in collaboration with Shrinivas Kulkarni and Carl Heiles of Berkeley, indicates that the giant molecular-cloud complexes in the outer parts of the galaxy indeed are concentrated in regular spiral arms. As for the inner parts, the surveys of interstellar carbon monoxide conducted so far show a concentration of molecular clouds at the center of the galaxy. Outside the center the greatest concentration of clouds is at radial distances between 4,000 and 8,000 parsecs. Beyond that the concentration of molecular gas falls off toward the edge of the disk of the galaxy, which lies 20,000 parsecs from the center. (The sun moves in orbit around the center of the galaxy at a distance of about 10,000 parsecs.)

The study of the giant molecular-cloud complexes and their relation to the structure of the galaxy is still in its infancy. For example, astronomers are only now beginning to realize that the

The engineer's personal computer.

Improving the engineer's productivity is another way HP computers help manufacturers sharpen their competitive edge.

The new HP 9836 is not just a general-purpose computer that's been fine-tuned for engineering. It's a full-function workstation developed expressly for the jobs you deal with most in the lab or on the bench: design, computation and testing.

Yet for all its power, the 9836 is more than just a "personal computer" —it's *personalized.* For over 12 years at HP, we've been building personal computers for engineers, scientists and other technical professionals. By working closely with them, we've been able to identify the kinds of features they need to help gather data, crunch numbers and interpret results. Quickly and easily.

That's why a large, easy-to-read CRT is part of the 9836. So are high-performance graphics and Pascal (in addition to BASIC and HPL). Even a special rotary control knob that can be used for editing programs, calibrating instruments or controlling motor speeds. Add to that a growing list of specialized engineering software and you'd think the 9836 was designed just for you!

At the heart of the 9836 is the powerful Motorola MC 68000 processor, which provides ultra-fast computation and data transfer. You can also take full advantage of the computer's eight I/O slots, choice of interface cards and additional memory boards (up to 2 megabytes) to personalize your system even further.

For example, you can plug in just the HP peripherals you need: printers, plotters, data tablets and high-speed discs (including the new 5″ Winchester). And with the built-in HP-IB standard instrumentation slot, you get access to the entire range of HP instruments.

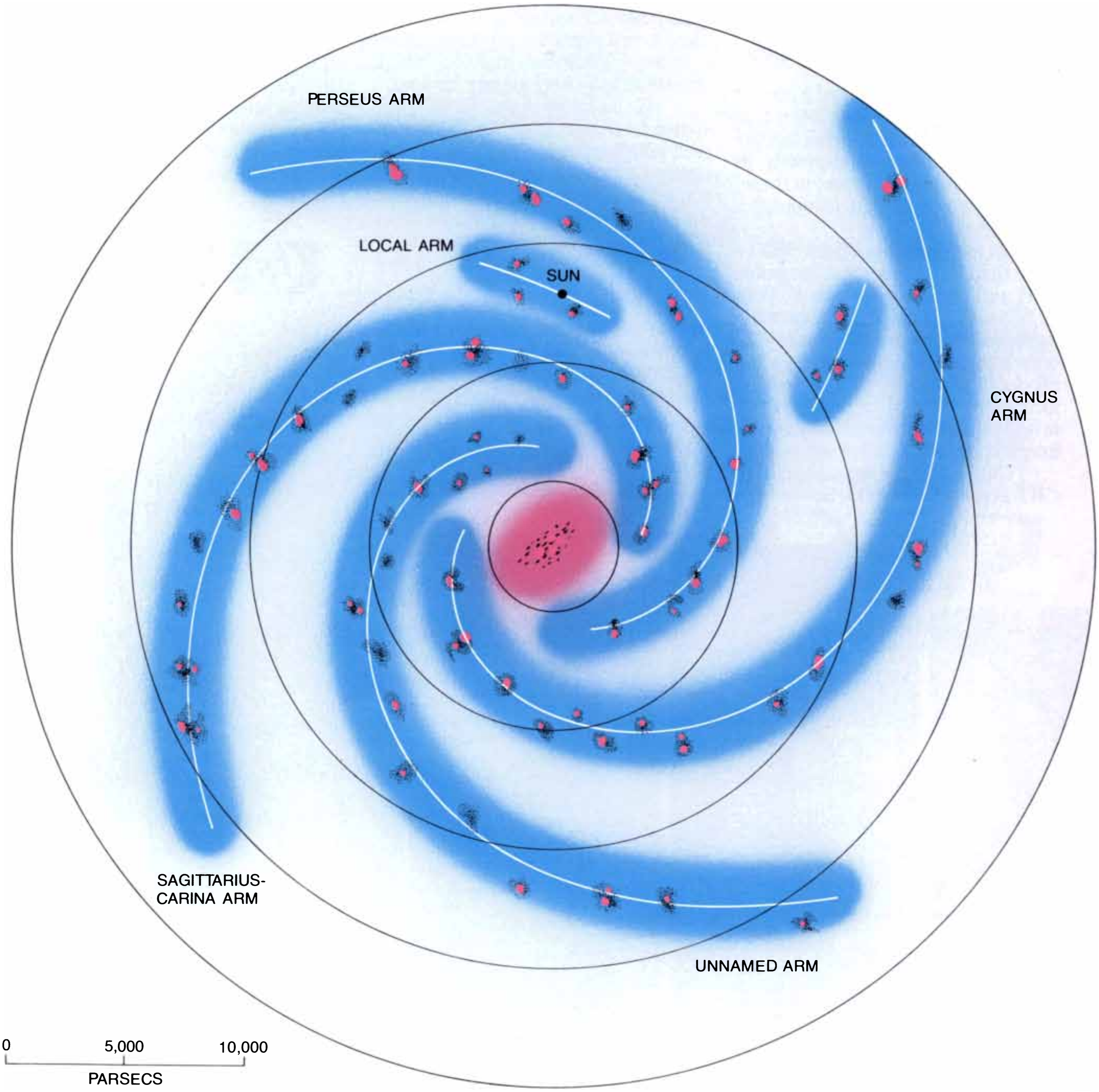
The HP 9836 is also a powerful part of our Manufacturer's Productivity Network (MPN). If you'd like a hands-on demonstration, contact your local HP sales office. Or see the 9836 in action at our Productivity '82 Seminars. For more information write Hewlett-Packard, Attn: Pat Welch, Dept. 56129, 3404 E. Harmony Rd., Ft. Collins, CO 80525.

HEWLETT PACKARD

40202

complexes may have an important effect on the distribution of stars in the galaxy not only because they give birth to stars but also because they exert gravitational and tidal forces on the stars and star clusters around them. Some 25 years ago Spitzer and Martin Schwarzschild of Princeton argued that the distribution of stars in the direction perpendicular to the disk of the galaxy might arise from the gravitational interaction of the stars and massive clouds of gas. Three years ago Stark and I showed that the giant molecular-cloud complexes in the galaxy were massive enough and large enough to account for the distribution.

The tidal forces exerted by the giant molecular-cloud complexes might be responsible for the disruption and ultimate dissipation of loosely bound clusters of stars. Furthermore, many astronomers suspect that gamma rays generated within the galaxy arise from the collision of high-energy particles and the matter in giant molecular clouds. As astronomers increasingly turn their attention to the giant molecular clouds they find that the clouds affect the galaxy in unsuspected ways. The discovery of any new major component of the universe brings with it not just the excitement of the discovery but a host of new questions as well.



PROPOSED PATTERN of the galaxy's spiral arms shows that giant molecular-cloud complexes (*black*) and H II regions (*magenta*) are confined more or less to a set of four arms. (The solar system lies in what seems to be an incomplete arm or the spur of an arm.) The densest concentration of molecular gas is shown, however, to lie within about 350 parsecs of the center of the galaxy, or a thirtieth of the distance from the center to the sun. Young, bright stars (*blue*) follow the pattern of the arms; older, fainter stars (*red*) are more evenly distributed, except that the center of the galaxy is given a barlike distribution of stars. The pattern of the arms is based on a study of interstellar atomic hydrogen in the galaxy done by the author and Shrinivas Kulkarni and Carl Heiles of the University of California at Berkeley.

The National Wildlife Federation
presents

MOUNTAIN SENTINEL

by Chapel

An exceptionally powerful
and dramatic new work
in the great tradition
of bronze animal sculpture.

Shown smaller than actual size. Overall height: 7½ inches.

Available
in a single, limited edition.
Advance ordering deadline:
April 30, 1982.

As part of its first collection of original sculptures portraying the great wildlife of North America, the National Wildlife Federation is proud to announce the creation of "Mountain Sentinel."

This dramatic new work portrays the Rocky Mountain bighorn — living symbol of America's wilderness. Superbly detailed, crafted in cold-cast bronze and meticulously hand-finished, it is both a brilliant study of the bighorn and an impressive work of art.

"Mountain Sentinel" has been created for the National Wildlife Federation by one of America's most gifted wildlife sculptors — Chapel. And he has shown the bighorn in its element — high on a snowy precipice. The proud head with its great curving horns. The powerful hindquarters. The agility of its stride over a fallen cedar. All are captured with an artistry that is completely true to nature — a quality valued throughout the long tradition of bronze animal sculpture.

To faithfully capture all the rich detail and subtle nuances of the sculptor's art, the work will be crafted in cold-cast bronze — a medium capable of holding even finer detail than traditional hot-cast bronze. Each sculpture will be individually cast from a carefully prepared sculptor's blend of powdered bronze and resins. Once hardened, the work will be meticulously finished with a fine patina — hand-rubbed to a soft luster.

"Mountain Sentinel" will be issued in a single, limited edition — available for only a comparatively brief period of time. This advance opportunity expires April 30, 1982, and the edition will be permanently closed at the end of the year. There is a further limit of one sculpture per person. A Certificate of Authenticity from the National Wildlife Federation will accompany each sculpture.

As an original sculpture that will enrich your home, "Mountain Sentinel" will be a truly rewarding acquisition. A work in the honored tradition of fine animal sculpture. One that will remain a source of enjoyment and a most treasured family possession.

To reserve this important new work, return the attached form directly to the appointed issuing agent, The Franklin Gallery. No payment need accompany your order, but please mail it by April 30th at the latest.

© 1982 FG

ADVANCE ORDER FORM

MOUNTAIN SENTINEL

by Chapel

Valid only if postmarked by April 30, 1982.
Limit: One per person.

The National Wildlife Federation
c/o The Franklin Gallery
Franklin Center, Pennsylvania 19091

Please accept my order for "Mountain Sentinel," to be crafted for me in cold-cast bronze at $195.* plus $2.50 for shipping and handling.

I need send no money now. When my sculpture is ready to be sent to me, I will be billed for my deposit of $39.50* and, after shipment, for the balance in four monthly installments of $39.50* each.

**Plus my applicable state sales tax*

Signature ____________________

ALL ORDERS ARE SUBJECT TO ACCEPTANCE.

Mr.
Mrs.
Miss ____________________

Address ____________________

City ____________________

State, Zip ____________________

20

The Giant Squid

This deep-sea cephalopod is so seldom seen, dead or alive, that it seems mythical. Numerous specimens have, however, been studied, and they have begun to reveal the animal's anatomy and its ecology

by Clyde F. E. Roper and Kenneth J. Boss

In *Moby-Dick* Herman Melville describes a sea creature of "vast pulpy mass, furlongs in length..., long arms radiating from its center and curling and twisting like a nest of anacondas." He apparently had in mind the giant squid; his description reflects the meager information that was available about the animal in his time. Indeed, until the crew of a French warship sighted one in 1861 and managed to haul in part of it the animal was quasi-mythical. Even now, when a considerable number of specimens have been reported, the giant squid (weighing up to 1,000 pounds and having an overall length of some 18 meters, or nearly 60 feet, if the tentacles are extended) remains largely mysterious. No living specimen has ever been maintained in a research institution or an aquarium. Most of what is known comes from strandings, when the squid is dead or dying; from capture by fishermen of animals that soon died, and from specimens removed from the stomach of toothed (fish- or squid-eating) whales. On the basis of this information one can state what a mature giant squid looks like and can say quite a bit about its internal anatomy. On other matters, such as its habitat and method of reproduction, one can offer only educated guesses based on what is known of related oceanic squids.

Teuthologists, the specialists who study cephalopods (the group of marine animals that includes the squid, the cuttlefish and the octopus), have placed the giant squid by itself in the genus *Architeuthis* of the family Architeuthidae. Nineteen nominal "species" have been described, some solely on the basis of individual parts of the animals. Until recently the tendency was to classify nearly every find as a new species. Although authorities differ, it has been suggested that the 19 nominal species can in fact be encompassed by only three: *Architeuthis sanctipauli* in the Southern Hemisphere, *A. japonica* in the northern Pacific and *A. dux* in the northern Atlantic. Some evidence, based mainly on the relative size of the head and the outline of the caudal fin, has been marshaled to indicate that as many as five species live in the Atlantic, two in the northern Pacific and several more in the southern Pacific.

In 1980 a specimen of medium size—about 10 meters in total length—was stranded on Plum Island in Massachusetts. The animal was preserved and was displayed for several months at the New England Aquarium in Boston. It is now in Washington at the Smithsonian Institution's National Museum of Natural History, where it is being studied in detail. The study provides much of the data for the following description, in which we adhere to the standard method for avoiding ambiguity about anatomical positions by relying on the terms dorsal for back or top, ventral for front or bottom, caudal for at or near the tail, anterior for forward, posterior for rearward, proximal for near a reference point and distal for far. The normal swimming position of the giant squid is horizontal, so that the animal's dorsal side is at the top and its ventral side is at the bottom.

The giant squid's cylindrical head may reach nearly a meter in length; it is connected to the body proper by a "neck" bearing a circumferential collar or sheath with a dorsal locking cartilage. A crown of eight thick, muscular arms and two very long and thin but muscular tentacles surrounds the buccal (mouth) apparatus and extends from the anterior end of the head. (Cephalopod means head-foot.)

Although the arms are proportionately long in young individuals, they are relatively much shorter than the tentacles in adults. An arm may attain a circumference of 50 centimeters at the base and a length of three meters. Each arm bears on its inner surface low, weak and sometimes scalloped protective membranes that border two rows of suckers. The suckers gradually decrease in size distally until at the tip of the arm they are merely tiny knobs.

The two ventral arms in males are hectocotylized, that is, specialized to facilitate the fertilization of the female's eggs. With them the male transfers spermatophores to the female in mating.

ADULT GIANT SQUID can be as much as 18 meters (60 feet) long and can weigh as much as 450 kilograms (1,000 pounds). This drawing is based on numerous specimens or parts of them that have been drawn, photographed or preserved over the past century. The squid in the draw-

These arms differ among species in their length and diameter and in the extent of their modification for the mating function. Distally the arms bear, as continuations of the two rows of suckers, two rows of rectangular pads separated by a deep furrow.

The two tentacles, with which the squid seizes its prey in a motion rather like the thrusting and closing of a pair of pliers, are about 25 centimeters in circumference at the base and can reach a length of more than 10 meters. The stalk of a tentacle is bare along the base; alternating small suckers and adhesive knobs may appear farther along its length. The suckers and knobs increase in size and frequency toward the club: the slightly expanded distal end of the tentacle. The manus, or palm, of the club has four rows of finely toothed suckers. The larger medial ones (in two rows) are about two and a half times the diameter of the smaller ones in the marginal rows. The diameter of a tentacular sucker may reach 5.2 centimeters. The distal end of the club, the dactylus (finger), is pointed and attenuated but is covered with hundreds of small suckers.

All the suckers of a giant squid are shaped like a suction cup. Each sucker is set on a muscular pedicle, or short stalk, that can be moved by the animal. The perimeter of a sucker is rimmed by a sharply toothed ring of chitin (the hard material that forms the outer covering of many crustaceans and insects) that adheres to the surface of the prey when the sucker is applied. No specialized hooks appear in place of suckers, as they do in certain other oceanic squids.

Imprints or scars from squid suckers have been found on the skin of sperm whales and even in their stomach. The diameter of these scars has been variously exaggerated, sometimes to as much as 20 centimeters (eight inches). Such reports have led by extrapolation to a distorted estimate of the maximum length of giant squids. The most reliable evidence suggests that the average diameter of the suckers on the arms is about 2.5 centimeters and that the maximum diameter is 5.2 centimeters, the size sometimes found on a tentacular club.

The mantle, or body, of the giant squid is more or less narrowly cone-shaped. It tapers to a bluntly pointed tail. In adults a short, stout, taillike projection extends beyond the fins; juveniles do not have it. The fins are flexible but not strongly muscular, suggesting that they serve as stabilizing vanes.

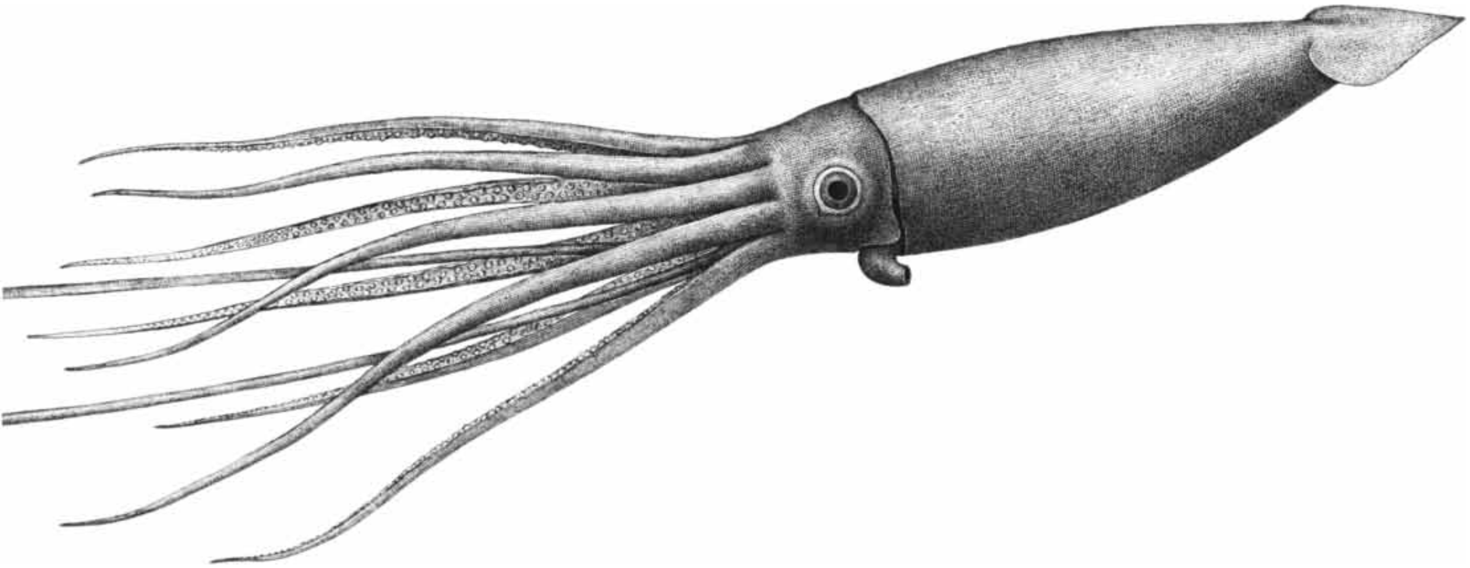
As a modification of the molluscan foot a large, muscular funnel rises ventrally behind the head at the anterior end of the mantle. With it the squid propels itself by squirting water out of its mantle cavity. The funnel is highly mobile, so that the squid can dart forward, backward, up, down or to the side. Inside the funnel is a flaplike valve that prevents the backflow of water between squirts. Ventrally at the base of the funnel on each side is a groove of cartilage that interlocks with corresponding cartilaginous ridges on the inner surface of the mantle when the squid is expelling water. This funnel-locking mechanism prevents water from escaping around the neck, forcing all the water to go through the funnel.

The giant squid also has a translucent internal supportive structure variously called the gladius or the pen. It is the remnant of an archaic internal calcareous shell that is still found in more primitive cephalopods such as the cuttlefish. The gladius lies in a sac in the musculature of the mantle, extending posteriorly from the anterior edge of the mantle to the taillike posterior extremity. It functions as a skeletal rod for the attachment of muscles and as a supporting staff for the elongated body.

The multilayered integument that envelops the body, the head and the arms is a dark purplish red to maroon dorsally and slightly lighter ventrally. The color of the dorsal and ventral surfaces of the arms is less intense than that of the lateral ones. The color arises not only from a background pigmentation but also from a layer of chromatophores: pigment-bearing cells that can change the coloration of the integument by expanding or contracting. In the squid the chromatophores probably serve to modify the color of the integument according to behavioral requirements or changes in the amount of light in the water. The internal surface of the mantle and some of the viscera also have dark reddish pigmentation, an uncommon feature in oceanic squids. Photophores, or bioluminescent organs, are not known in *Architeuthis.*

The giant squid's eyes are enormous, larger than the headlights of an automobile. With a diameter approaching 25 centimeters (10 inches), they are the largest eyes in the animal kingdom. They are positioned laterally on the head and are circular in outline. Each eye has an adjustable lens and a dark iris but no cornea.

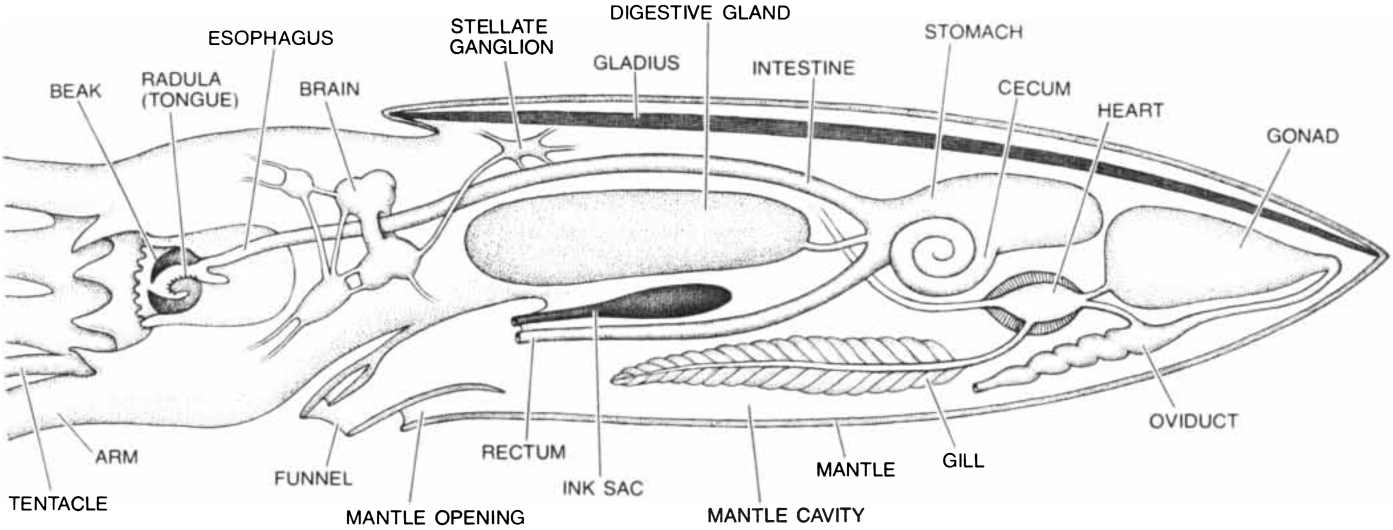
The mouth is at the center of the circular crown of arms. Powerful chitinous jaws, encased in a muscular mass and capable of rotation and protrusion, are utilized to bite prey into chunks of a size suitable for swallowing. These parrotlike beaks, which can be more than 15 centimeters in length, consist of upper and lower mandibles. The strong upper mandible bears a pointed and acute-angled rostrum that forms a cutting edge with the lower mandible, which characteristically has a short, blunt rostrum and rounded winglike extensions. By



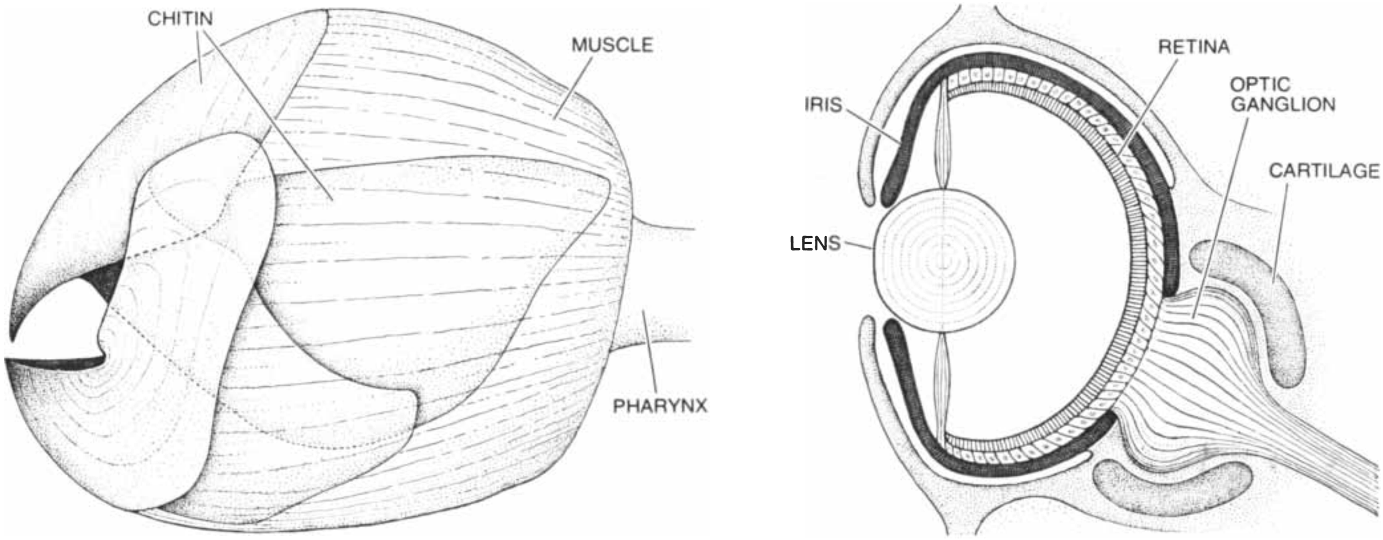
ing is depicted seizing a fish by thrusting out its long tentacles as it attacks, propelling itself forward by squirting water through the backturned funnel that is visible below the eye. By appropriately turning the funnel the squid is able to move in any direction. In the class of cephalopods the giant squid is placed in the genus *Architeuthis;* the number of species in the genus has yet to be firmly established.

careful examination the isolated jaws of individual squid species can be distinguished from those of other species. Hence it is possible to identify the otherwise unrecognizable digested contents of whale stomachs and the inclusions in ambergris, the waxy substance made by the sperm whale to enable it to purge itself of indigestible squid beaks.

The radula, or rasping tongue, is a characteristic feature of mollusks and usually consists of a long cuticular ribbon bearing transverse rows of chitinized teeth with cusps of various shapes. In *Architeuthis* the radula is small for an animal of such huge proportions. It is nonetheless impressive compared with the radulas of other mollusks, being about 100 millimeters long and a little more than 10 millimeters wide. Each row has a three-cusp central tooth and three smaller teeth on each side. Food bitten into chunks by the jaws is forced into the buccal cavity by the bolting



INTERNAL ANATOMY of a female giant squid is indicated on the basis of present knowledge. The gladius, which is also known as the pen, is a fairly stiff structure that serves as a supporting rod for the animal's long body and also as a point of attachment for muscles.



MOUTH AND EYE of the giant squid are among the animal's noteworthy features. The powerful beak consists of chitin, the hard material that forms the external covering of many crustaceans and insects. With its beak the squid cuts its prey into pieces it can swallow. The eyes of the giant squid are the largest in the animal kingdom, each one being approximately 25 centimeters (10 inches) in diameter.

CLUB END OF A TENTACLE of the giant squid is distinguished by a manus, or palm, where the tentacle becomes slightly wider. The manus has four rows of finely toothed suckers, the larger ones being about five centimeters in diameter and the smaller ones about two centimeters. Each sucker can be moved. One sucker and its toothed chitinous ring are shown above the end of the tentacle at the left.

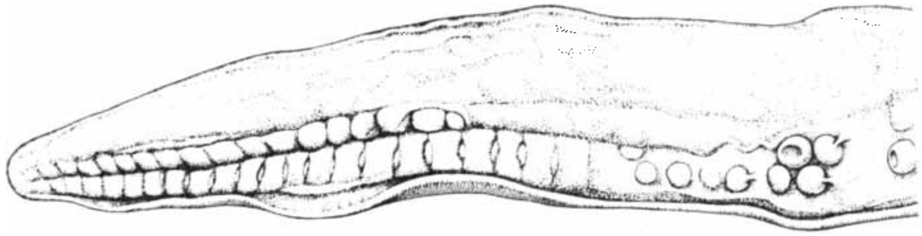
movement of the radula. In addition many minute backward-sloping denticulations (the pharyngeal teeth) in the cuticle lining the pharynx facilitate swallowing and ensure that food moves inward to the alimentary canal.

The alimentary canal continues into a muscular esophagus that forces food by peristaltic contractions into the thick-walled stomach and the associated cecum. Digestive enzymes are secreted by the massive salivary glands and by the single medial "liver," or digestive gland, and the anterior pancreas. Assimilation is further advanced in the cecum. Wastes pass through the short intestine and out through a flapped rectum, which discharges near the internal opening of the funnel. The wastes are flushed out through the funnel with the water expelled for propulsion.

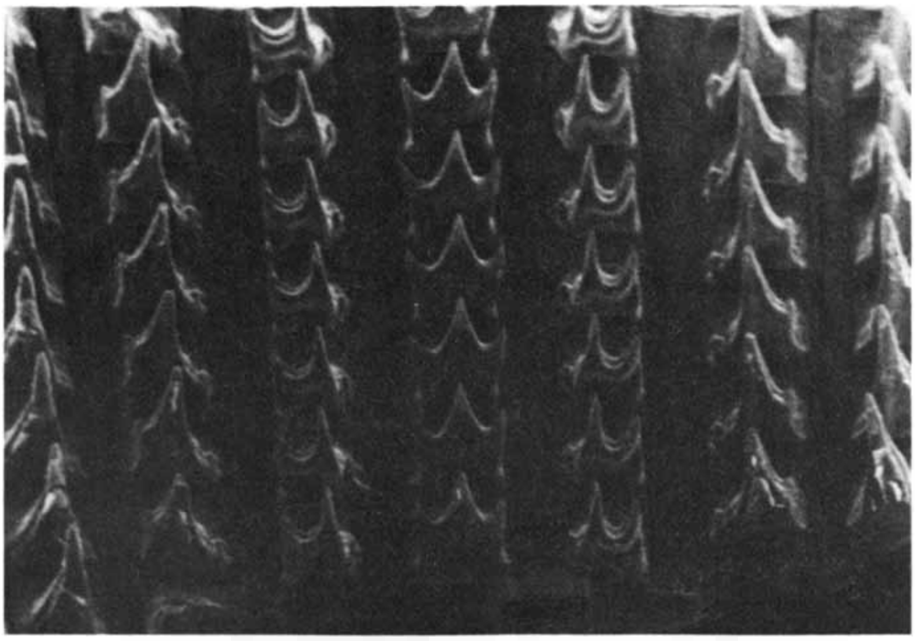
The squid obtains oxygen through a pair of long gills, which have many (sometimes more than 100) of the lamellae, or leaflets, that are the basic structures for the exchange of oxygen and carbon dioxide. Another feature is the ink sac characteristic of most squids. In the giant squid the sac is large and elongated, with a long duct that empties into the rectum. The black, mucous ink is thought to serve the giant squid as the ink of other squids does when it is extruded through the funnel in an escape reaction, that is, it maintains a cohesive shape resembling that of the squid, presumably confusing a predator as the squid jets away.

A female squid produces enormous numbers of whitish eggs. The eggs are relatively small, being from .5 millimeter to 1.4 millimeters long and .3 to .7 millimeter across. One specimen carried more than 5,000 grams (11 pounds), or perhaps a million eggs.

The female reproductive apparatus consists of (1) a single median ovary, lying posteriorly in the visceral mass; (2) paired, convoluted oviducts through which the mature eggs pass out into the mantle cavity, and (3) large nidamental glands, which manufacture a gelatinous mass that envelops the eggs in the mantle cavity. The site and method of fertilization are not known, but presumably it takes place as the mature eggs leave the oviducts. Whether the eggs float in the ocean or are attached to the ocean bottom is also not known, but it seems likely that they float in huge gelatinous masses because this mode is characteristic of all other known oceanic squids. The females of many shallow-water squid species have a receptacle for spermatophores below the mouth, but no such structure is found in *Architeuthis* females. Moreover, no spermatophores have been found implanted in female giant squids, although such implantation (in the mantle cavity or at the opening of the mantle near the neck) is common in other oceanic squids.



SPECIALIZED ARM of the male squid serves for transferring spermatophores to the female. Two of the male giant squid's eight arms are specialized in this way; they are called hectocotylized arms. They are distinguished by two rows of rectangular pads (in place of suckers).



RADULA, or tongue, of the giant squid is shown at an enlargement of 22 diameters in this scanning electron micrograph. The teeth slope toward the back of the squid's mouth and assist the movement of food into the esophagus. The entire radula is about 10 centimeters long.

SUCKER SCARS of a giant squid appear on the skin of a sperm whale, the chief predator of *Architeuthis*. Such scars, made when a squid fought to escape from the whale, have often been cited as evidence for squids of monstrous size. Since a scar grows as the whale grows, however, scars are unreliable evidence for the size of the squid unless they are known to be recent.

The males are externally differentiated from the females by their two hectocotylized arms. The testis consists of a white filamentous mass embedded posteriorly in the visceral mass. The spermatophores, long tubes filled with sperm, are manufactured in the complex spermatophoral apparatus, which is on the left side anterior to the testis.

The long, thin-walled spermatophoral sac (often called Needham's sac because it was first discovered and described by John Turberville Needham) with its basal seminal vesicle is attached to the visceral mass along the left side of the intestine in the mantle cavity. As such structures go it is enormous, reaching a length of a meter and serving as a storage chamber for hundreds (perhaps thousands) of spermatophores packed parallel to one another.

The distal extension of Needham's sac is often termed the penis, but that is probably a misnomer because the structure seems not to be an intromittent organ. Even in a small mature male it may be some 80 centimeters long and may protrude as much as 5.5 centimeters beyond the free edge of the mantle. The tip of the "penis" is mushroom-shaped and has in it a ventral slit about 15 millimeters long.

The spermatophores range between 10 and 20 centimeters in length, apparently varying in total length and in the size of the constituent parts depending on the species and the stage of development. As in other squids, the spermatophore of *Architeuthis* has a proximal thread and cap, a springlike ejaculatory cell, a cement body and a distal sperm mass enveloped by a tunic. An external coating of jelly that envelops the spermatophore is peculiar to *Architeuthis.*

It is likely that the giant squid, like other squids, has a mechanism whereby the cap and thread are pulled, much like the pin of a hand grenade, triggering the release of the tightly coiled ejaculatory apparatus and ejecting the sperm-containing capsule and the cement body. The cement body functions as an adhesive mechanism to attach the sperm mass in or near the mantle cavity of the female. Males apparently become sexually mature quite early; specimens with a mantle length of less than a meter have been found to have completely formed spermatophores.

Even though a giant squid is notably large and heavy out of the water, it is quite buoyant in the water. The buoyancy is due to the relatively high concentration of ammonium ions (NH_4^+) in the muscles of the mantle, head and arms. The concentration of ammonium ions probably explains why dead or dying squids rise to the surface and are often washed ashore. Ammonium ions have a specific gravity of 1.01, which is lower than that of seawater (1.022 at a depth of 50 meters and a temperature of 28 degrees Celsius). Without the ammonium ions the tissues of the squid are heavier than seawater, their average specific gravity being 1.046, but with them the animal can maintain its level in the water without having to expend energy by constant swimming.

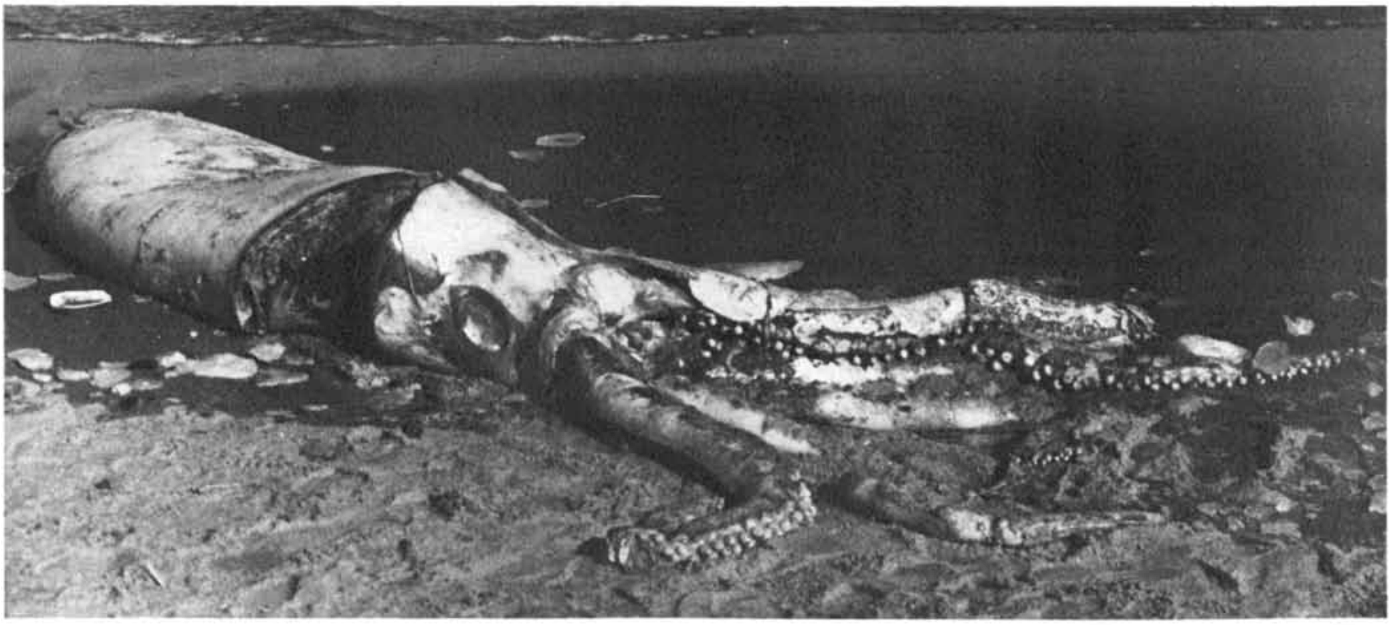
A side effect of this characteristic is that a freshly stranded giant squid is made unpalatable by the strong, bitter taste of ammonia. In fact, the first observation that *Architeuthis* concentrates ammonium ions in certain muscles was made by three teuthologists, including one of us (Roper), who cooked a piece of giant squid for a party celebrating the completion of a doctoral examination. Subsequent analysis of the tissue confirmed the observation.

Although certain scholars have described *Architeuthis* as a strong swimmer, that is probably not the case. It is true that some oceanic squids are spectacular swimmers, but compared with them *Architeuthis* has a poorly developed musculature. In addition its fins are relatively small and weak, and the locking apparatus of the funnel is considerably less complex and powerful than the one in the strong swimmers.

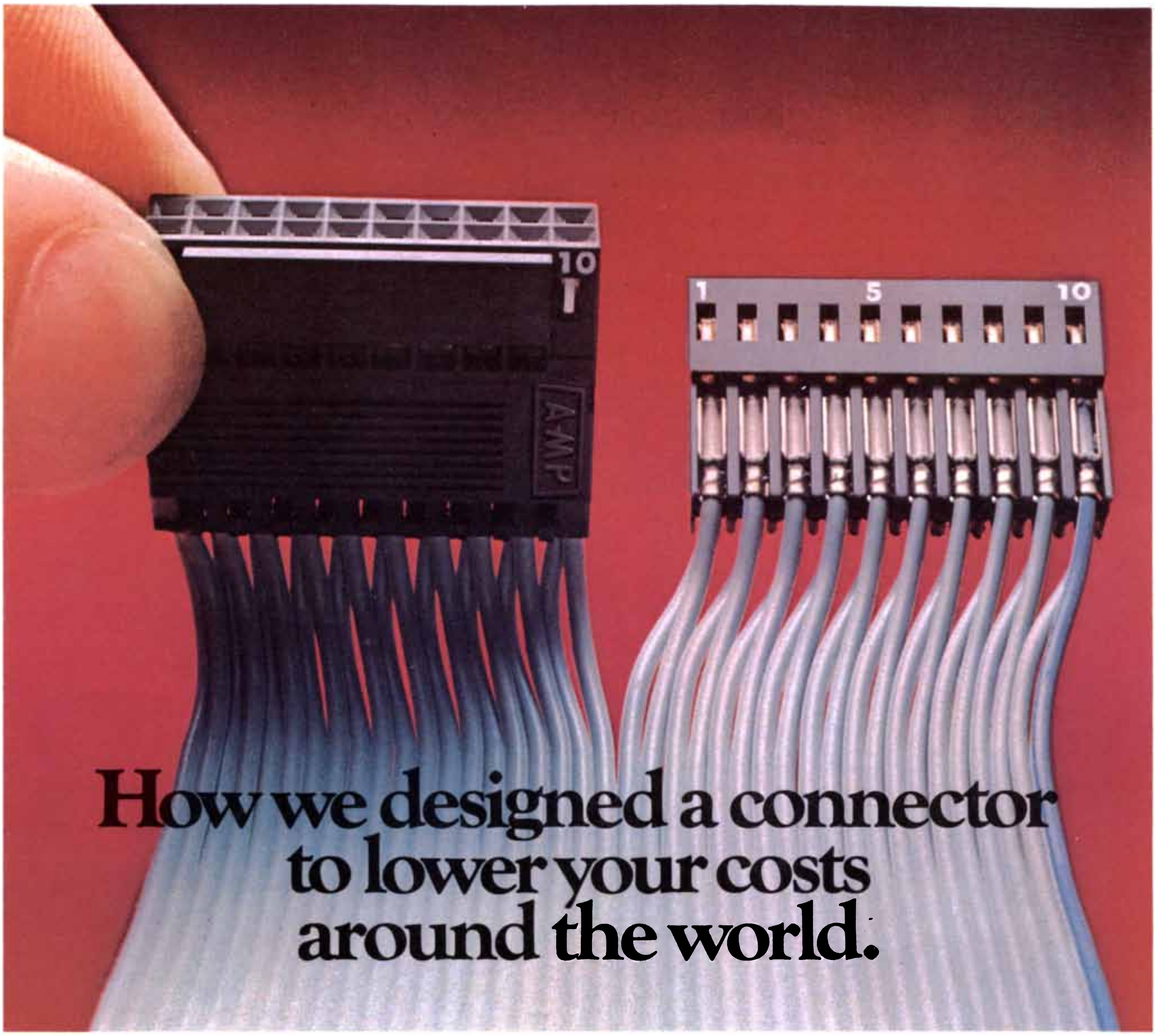
Little is known about what the giant squid eats. Most of the specimens collected have an empty stomach; if the squid were not sick and "off its feed," it would have been unlikely to come to the surface or to be stranded. Even when something is found in a squid's stomach, the chances of identifying it are slim because the beak and the radula reduce the prey to small pieces and the digestive enzymes work fast.

Nevertheless, various workers have guessed or apparently observed that *Architeuthis* feeds on small fish and large invertebrates such as other cephalopods. The few published accounts based on remnants of food found in the stomach of giant squids indicate that the animals do indeed feed on fish and other squids. Since *Architeuthis* appears to be a relatively poor swimmer, it presumably is a somewhat passive and sluggish predator, unable to chase and capture large, active prey.

As for the predators of the giant squid, the principal one is the sperm, or cachalot, whale, *Physeter catodon.* Immature giant squids are eaten by certain mid-water fishes. Notwithstanding the limitations of *Architeuthis* as a swimmer, it is formidable in size and dexterity and has arms that are heavy, suckered and



STRANDED GIANT SQUID was washed ashore on Plum Island in Massachusetts in 1980. It was an adult of medium size, about 10 meters in total length. Most of the skin was worn off by the sand, and the tentacles (along with the tips of some of the arms) were missing. An indication of the size of the animal is given by the human footprints in the foreground and the seashells on the smooth part of the beach.



How we designed a connector to lower your costs around the world.

If you're a progressive company, you need the flexibility to keep costs down yet maintain the same performance— anywhere in the world.

That's why if you use electrical connectors, it pays to specify AMP.

An example is our modular interconnection.

Designed for use with printed circuit boards, its building block simplicity has made it universally acceptable. Now it is particularly cost effective because of mass termination technology—making up to 50 connections in the time it took to make one.

As a result, this connector is helping to increase productivity for customers around the world.

Can we help you?

Whether you're looking for connectors or better ways to apply them, call us. We'll give you what the world expects from the industry's leader in specialized engineering. Products designed for performance. Products which help reduce the cost of making yours.

AMP has a better way.

For ways to save you money, for products and innovations that help improve productivity, we have the better ways.

Write for our brochure, "Why AMP Means Productivity." AMP Incorporated, Harrisburg, PA 17105.

Auckland • Barcelona • Brussels • Buenos Aires • Dublin • Frankfurt • Harrisburg • Helsinki • s-Hertogenbosch • Kowloon • London • Luzern • Mexico City
Montreal • Oslo • Paris • San Juan • Sao Paulo • Singapore • Stockholm • Sydney • Taipei • Turin • Toronto • Tokyo • Valley Forge • Viby • Vienna

AMP means productivity.


New. Ra
REN
elf REN
MAG
Fuego


Fuego
cy. Fuego.
RENAULT
American Motors

highly motile. The large eyes, representing an efficient detection system, and the ink sac are among the giant squid's protective mechanisms.

In the literature of giant squids one occasionally finds an account of a battle between a giant squid and a sperm whale at the surface of the sea. We must surmise that such a struggle is an attempt by the squid to escape from the whale rather than an attack on the whale by the squid. The scars of giant-squid suckers on the skin around the mouth and head of sperm whales attest to the reality of these battles.

Sperm whales prey on numerous kinds of fish, crustaceans, octopuses and squids, but much of their diet consists of *Architeuthis.* Although only a few giant-squid beaks may be found in a sperm whale's stomach along with hundreds of beaks from other squid species, the sheer size of a single giant squid may take up a third of the volume of a whale's stomach. The fact that a sperm whale's gut is usually found to contain large numbers of other prey but only one *Architeuthis* suggests that giant squids may be solitary animals, except possibly during the mating period.

Much has been written about the maximum size attained by *Architeuthis,* with assertions of total lengths exceeding 75 meters. There is no firm evidence for such assertions. The usual basis for them is the size of sucker scars on whales, but since a scar grows as a whale grows, it is unreliable evidence for the size of a squid unless it is demonstrably recent.

The largest specimen recorded in the scientific literature measured approximately 20 meters in total length. (It was stranded on a beach in New Zealand in 1880.) A significant part of this length, probably from 10 to 12 meters, consisted of the tentacles, which in a dead squid are notably elastic and easily stretched. In all other squid species the length of tentacles is always regarded as an imprecise component of measurement. The largest giant-squid mantle lengths known to us are in the range from five to six meters, the largest head lengths about one meter.

Even with the fairly large number of records of *Architeuthis* now available, it is still impossible to identify the precise habitat of this elusive animal. Most of the records come from strandings and from the stomach contents of sperm whales, neither of which give direct information about habitat. Sperm whales are known to feed in the depth range from 10 to 1,000 meters, and there is strong evidence that they go as deep as 2,000 meters. It is clear from their stomach contents that they feed partly along the bottom. Hence the *Architeuthis* individuals in sperm-whale stomachs could have been captured in mid-water or on the bottom. An educated guess, supported by the record of *Architeuthis* tissue recovered from the stomach of a deep-sea shark captured on the bottom at 1,246 meters near the Azores, is that as adults giant squids live near or on the bottom at a depth of about 1,000 meters. Even this evidence is indirect, however, because some species of bottom-living fishes and sharks often come off the bottom to feed.

It is remarkable that few giant squids have been caught in fishing nets. A few captures have been made in nets being trawled along the bottom at depths of from 200 to 375 meters. Since the specimens were reported to be alive at the time of capture, we must assume that they were in or near their natural habitat, probably on the bottom but possibly in mid-water when the net was being hauled in.

On the other hand, a fresh 12-foot section of a giant squid's tentacle was brought up 150 miles off the coast of California by a mid-water trawl at a depth of 600 meters over a bottom depth of 4,000 meters. Sightings of giant squids swimming at the surface have been reported from off Newfoundland near the Grand Banks, where the water is less than 100 meters deep, and from the central Pacific over depths exceed-

DATA ON THE GIANT SQUID have come mainly from strandings and from specimens found in the stomach of sperm whales. The sightings are less reliable, since at a distance an observer might mistake another kind of large squid for a giant squid. In addition to these ing 4,000 meters. In recent years specimens of *Architeuthis* have been found in the vicinity of the Hawaiian Islands, where there is virtually no continental shelf and the bottom drops off sharply to several thousand meters.

Then there are the strandings. Many of the specimens are washed ashore dead, but enough are so fresh (or even barely alive) as to indicate that they were alive not too long before or too far from where they came ashore. Still, large squids, like large whales, may be alive when they are stranded but may also be far from their normal habitat because they are sick.

The fact that so few giant squids are captured in nets is intriguing, particularly now that huge mid-water and bottom trawls are deployed by commercial fishermen and by research vessels. Do the squids detect the approach of a net and avoid capture, as so many other oceanic cephalopods do? Or do they live in habitats not normally entered by commercial fishermen or exploratory biologists because they are known to be unproductive, like the middle depths of the open ocean, or because they are too rocky and craggy and therefore dangerous to nets, like the deep canyons and edges of continental slopes? The fact that such questions still have to be asked about giant squids indicates how much remains to be learned about them.

data there are 148 other records, mostly from near Japan and the southern tip of Africa.

Superheavy Magnetic Monopoles

Isolated north and south magnetic poles are predicted to exist but have never been observed. A new theory may explain why: such particles seem to be too massive, slow-moving and rare

by Richard A. Carrigan, Jr., and W. Peter Trower

In 1269 Petrus Peregrinus, an early French investigator of the magnetic properties of materials, described the orientation of bits of iron near the surface of various lodestones. He noted that the lines of force around such a natural magnet are invariably concentrated at two points, just as the meridians of the earth come together at opposite geographic poles. The analogy led him to designate the two points the north and south poles of the magnet. Subsequent observations have confirmed that all ordinary magnetic objects have paired regions of opposite polarity, that is, all magnets are dipoles.

It is easy to conceive of an isolated north or south magnetic pole. Speculation about the possible existence of such magnetic monopoles has persisted for centuries, but there has been no evidence of them. Interest in the idea became more focused in 1931, when the British physicist P. A. M. Dirac showed that an important observed property of electrically charged particles could be explained by assuming the existence of elementary particles bearing a magnetic charge. Dirac's conjecture stimulated a flurry of theoretical papers on the expected properties of the hypothetical monopoles, and several experiments were undertaken to detect them. None of the attempts was successful.

Recent efforts to forge a unified theory of the fundamental forces of nature have again drawn attention to the absence of magnetic monopoles. Indeed, one proposed theory requires the creation of monopoles in the first instants of the great explosion in which the universe was presumably born. The theory goes on to offer an explanation of how the monopoles could exist without having been detected in any of the earlier searches: they would be extraordinarily massive and so would have properties quite different from those of ordinary particles. Several novel experiments are now being prepared to look for superheavy magnetic monopoles among the surviving particles of the big bang.

The dipolar nature of ordinary magnetic materials can readily be demonstrated. If iron filings are sprinkled on a piece of paper held over a bar magnet, the filings trace out a pattern of smooth arcs extending from one end of the magnet to the other. The arcs represent the magnetic field lines between the poles. Where the lines are close together the magnetic field is strong; where they are far apart the field is weak. Cutting the magnet in half does not isolate the poles; instead two smaller bar magnets are created. (It is possible to simulate the field surrounding an isolated magnetic pole by placing the magnet on end, with one pole directly under the paper; in this case, however, the lines of force connecting the poles are merely removed from the plane of the paper.)

An analogous electric dipole can be made by depositing electric charges of opposite sign on the ends of an insulating rod. When the electric dipole is cut, two isolated electric poles are created. The reason is that each electric pole represents a clustering of individual electrically charged particles: negatively charged electrons at one end and positive ions at the other. When the poles are separated, the aggregations of charge are not affected. The possibility of isolating electric poles but not magnetic ones is a fundamental distinction between electricity and magnetism.

The explanation of this distinction has been known for more than a century. The magnetism of an ordinary object, such as a bar magnet, arises not from clusters of magnetically charged particles but from loops of electric current. For example, the magnetic field of a solenoid (a cylindrical electromagnet) is set up by the electric current circulating in the coil. On a submicroscopic scale similar currents are generated by the circulation of electrons around atomic nuclei. In nonmagnetic materials the atoms, and hence the currents, are randomly oriented. If the atoms are somehow aligned, the material shows a net magnetization. In a permanent magnet the atoms retain their alignment even when the orienting force is removed. Cutting such a magnet in half cannot isolate the poles because each atom is effectively a dipole.

This basic distinction between electricity and magnetism is at the heart of the theory of electromagnetic phenomena formulated by James Clerk Maxwell in 1864. In Maxwell's theory the possibility of isolated magnetic charges was ignored, since none had ever been observed. Instead all magnetism was explained in terms of moving electric charges. Over the past century Maxwell's theory has been put to many experimental tests and has never been found wanting. That fact alone severely limits the contexts in which magnetic monopoles might be found.

Dirac's contribution to the subject stemmed from his effort to explain the quantization of electric charge: the observation that electric charge appears only in multiples of the charge of the electron and the proton. Dirac showed that if an isolated magnetic pole exists anywhere in the universe, electric charge must be quantized everywhere. Until recently Dirac's magnetic-monopole hypothesis was the only explanation of the observed quantization of electric charge. (The existence of the particles called quarks, with charges of one-third and two-thirds the magnitude of the electron's charge, would not alter Dirac's conclusion. Many physicists think the quarks must be permanently confined to the interior of particles that invariably have an integer value of net charge.)

The reasoning behind Dirac's proposal can be presented in nonmathematical terms as follows. An atom that has been excited to an elevated energy state tends to revert abruptly to a lower energy state, simultaneously emitting its excess energy in the form of a photon, or quantum of electromagnetic radiation. The photon not only removes energy from the atom but also carries away some of

SCIENCE/SCOPE

Solar-powered ion propulsion will be demonstrated on a spacecraft in the near future. Compared to conventional chemical propulsion, ion propulsion saves weight and is better for spacecraft control and interplanetary travel. It generates small but exact thrust for long durations. Hughes has developed a mercury ion thruster system having a specific impulse of 2500 seconds. It is being incorporated for NASA into a prototype Ion Auxiliary Propulsion System (IAPS) to be launched from the Space Shuttle on the U.S. Air Force P80-1 spacecraft.

A new all-optical logic device could make many electronics systems immune to effects of natural or man-made "noise," including lightning strikes and radio interference. Hughes scientists have fabricated a high-speed optical device that uses no electronic signals. It is made of discrete components, including four reflecting surfaces and a slab of non-linear material (gallium arsenide). The device has shown optical bistability (flip-flop behavior) with switching times of 3 nanoseconds and switch energies under 100 microjoules. Although propagation delays have kept the device's speed under the theoretical limit up to 10 gigahertz, the speed will be increased by further miniaturizing of the device on an integrated optic chip. The device could be used in fault-tolerance computers, flight control systems, and ultra high-speed signal processors.

A safely concealed gunner could guide a missile toward a battlefield target with a new fiber-optic communications system. The concept calls for a missile with an imaging seeker in its nose to be fired toward an enemy force. What the missile sees is relayed back to the gunner over a glass thread that pays out from a spool in the missile's aft end. The cable, unlike ordinary wire, can transmit broadband signals required for video. The gunner looks at a display and picks a target. Guidance commands are transmitted automatically to the missile over the fiber-optic link. Hughes and principal subcontractor ITT Electro-Optical Products Division are developing the Integrated Fiber-Optic Communications Link for the U.S. Army for possible use in a low-cost anti-armor missile.

An ultramodern facility spanning 1.75 million square feet will be the showcase where outstanding Hughes engineering will combine with advanced manufacturing techniques and production processes. Our complex is nearly completed, so we're looking for experienced and graduating engineers to work on such programs as: infrared thermal imaging systems, laser rangefinders and designators, and missile launching and guidance systems. Send your resume to Dan O'Daly, Hughes Electro-Optical and Data Systems Manufacturing, Professional Employment, P.O. Box 924, Dept. SSA, El Segundo, CA 90245. Equal opportunity employer.

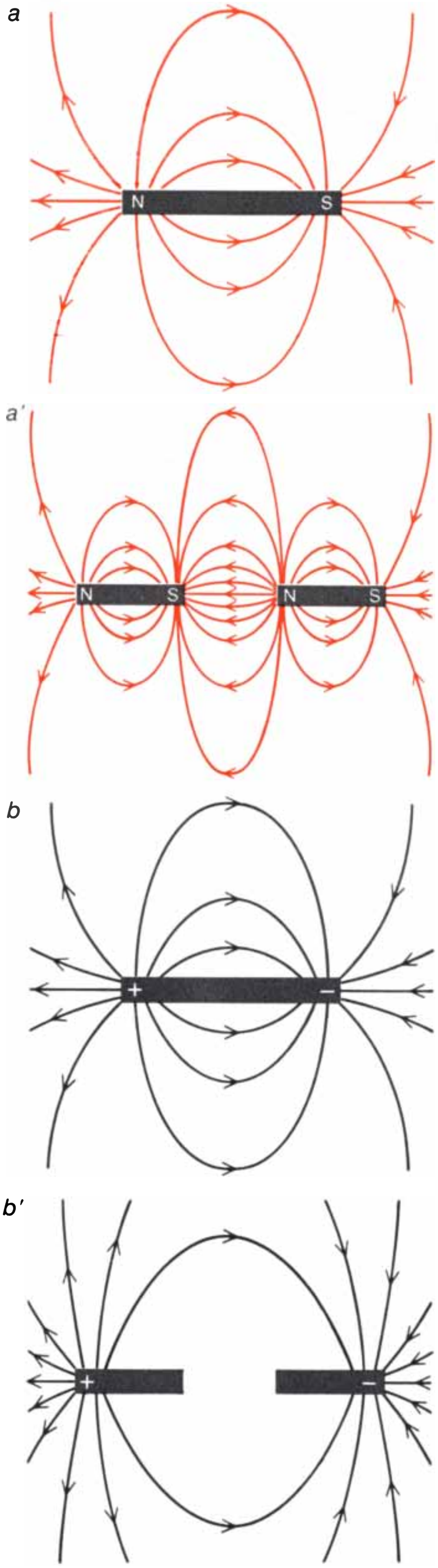
Cables and wire harness assemblies for the Space Shuttle's solid rocket booster were tested to withstand the rigors of launch and space travel by an automatic circuit analyzer. United Space Boosters Inc., a subsidiary of United Technologies Corp., tested its cables, harnesses, and ground support equipment on a high-voltage Hughes FACT® (Flexible Automatic Circuit Tester) system. The FACT unit performs up to 9,000 continuity tests and 3,000 leakage tests per minute. The system's speed and flexibility help to cut production costs.

Creating a new world with electronics

HUGHES

HUGHES AIRCRAFT COMPANY
CULVER CITY, CALIFORNIA 90230

(213) 670-1515 EXTENSION 5964



DIPOLE FIELDS are set up by a bar magnet (*a*) and by an analogous structure consisting of an insulating rod with opposite electric charges deposited at the ends (*b*). When the magnet is cut in half, two smaller dipoles are created (*a'*). When the electric analogue is cut in half, the field remains dipolar because the electric charges that generate the field remain in place (*b'*). Magnetic field lines are shown in color, electric field lines in black.

the atom's intrinsic angular momentum. Hence the electromagnetic field represented by a beam of light, say, can be considered to have a certain amount of angular momentum.

It is a fairly straightforward task to calculate the amount of angular momentum in an electromagnetic field. Indeed, more than 80 years ago J. J. Thomson, in a textbook on electricity and magnetism, suggested this exercise for students: Determine the angular momentum of an electromagnetically bound system consisting of a single electric charge and a single magnetic charge. The solution reveals that the angular momentum of the system depends on the product of the electric charge and the magnetic charge but is independent of the distance between them. In other words, the electric charge and the magnetic charge may be separated by the radius of an atom or by the radius of the universe; in both cases the angular momentum of the electromagnetic field is the same.

Now, the angular momentum of any system of this kind is known to be quantized: the smallest amount of angular momentum in nature is equivalent to the unit called Planck's constant, and all larger amounts are multiples of this unit. If it is assumed that the angular momentum of the system is equal to some number of Planck units and that the magnetic charge has a definite value, then it follows that the electric charge is also fixed. In this way Dirac was able to show mathematically that if magnetic charge exists, electric charge must be quantized.

Dirac's theory incorporated a curious mathematical construct he called a string. A Dirac string resembles an infinitely long solenoid with a magnetic monopole at one end; the rest of the string threads off into the distance. The string is an impediment to detailed calculations based on Dirac's model. Recent developments in mathematical physics, however, have made it possible to divest the monopole of its troublesome tail.

Dirac's quantization condition endows the magnetic monopole with certain properties. For example, to account for the observed quantization of the angular momentum of an electromagnetic field the minimum unit of magnetic charge must be about 70 times as large as the corresponding unit of electric charge. According to another prediction of Dirac's, every particle of matter, including the magnetic monopole, must have an antimatter counterpart. Just as the electron has its antiparticle (the positron, discovered in 1932, four years after Dirac foretold its existence), so the magnetic monopole is expected to be matched by a magnetic antimonopole. In keeping with Peregrinus' arbitrary nomenclature, one particle of such a pair is called a north monopole and the other a south monopole. Dirac's theory made no prediction about the mass or size of the magnetic monopoles or about their abundance in the universe.

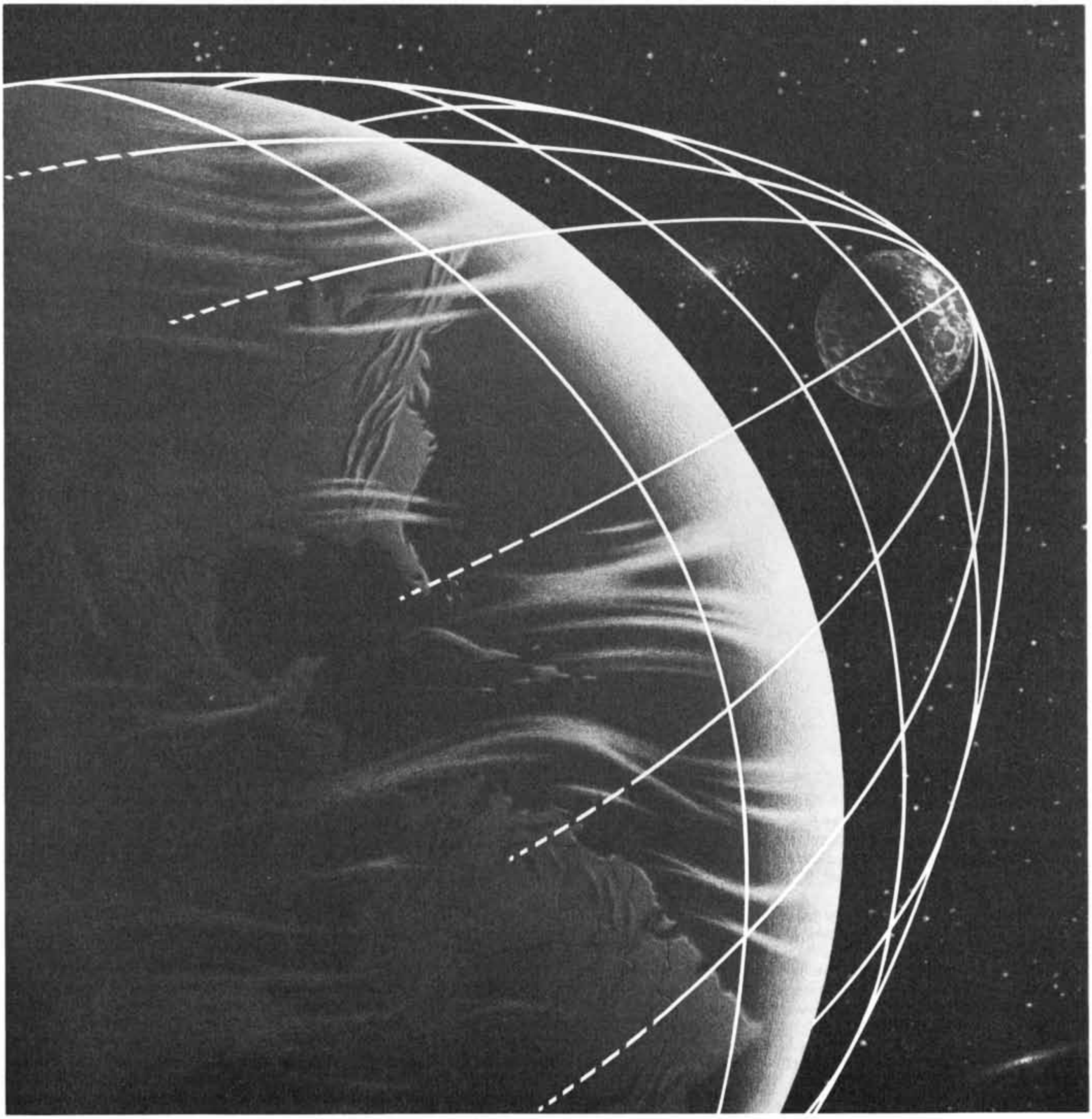
Some interesting effects arise when Maxwell's equations of electromagnetism are augmented to include magnetic charges and magnetic currents. For example, as the velocity of a moving electric charge approaches the speed of light, its properties should increasingly resemble those of a magnetic charge; similarly, a moving magnetic monopole would begin to take on the properties of an electric charge at a speed approaching the speed of light. These transformations, which follow from Einstein's special theory of relativity, have been confirmed experimentally for moving electric charges but not of course for moving magnetic charges.

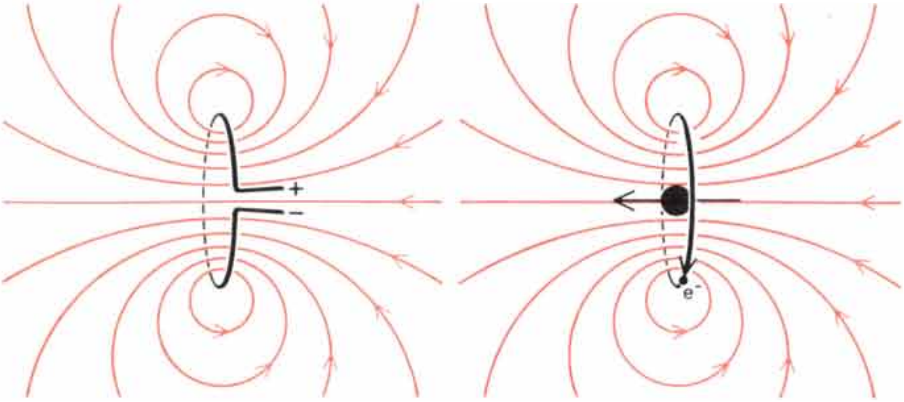
A moving electric charge can lose energy by ionizing matter (that is, by detaching electrons from their atoms). Typically, energy is lost at the rate of a few million electron volts for every centimeter traveled through a substance. The energy required to ionize an atom is generally a few tens of electron volts, and so a moving electric charge can ionize hundreds of thousands of atoms per centimeter.

Because of the much stronger charge of the magnetic monopole, it would ionize atoms some 10,000 times more effectively. Thus a magnetic monopole passing through a photographic emulsion of the type employed by physicists to detect electrically charged particles would leave a track thousands of times darker than the track left by an electric charge moving at the same speed. Because the monopole would lose energy to the ionization process so quickly, it would slow down much sooner on entering a substance than an electrically charged particle with the same kinetic energy does.

Just as an electric field can accelerate an electrically charged particle, so a magnetic field could accelerate a magnetic monopole. Because of the greater pole strength of the magnetic particle, however, it would gain energy faster in a magnetic field than an electrically charged particle does in an equivalent electric field. A magnetic monopole traversing a superconducting coil one meter long would gain more energy than a proton acquires in the largest particle accelerator yet built.

The physics of magnetic monopoles has another curious feature, which can be made apparent by imagining that the flow of time could be reversed. In a thought experiment suggested by Robert K. Adair of Yale University a proton is moving through a magnetic field, which causes it to follow a curved path. In one case the magnetic field is produced by an electric current in a coil.

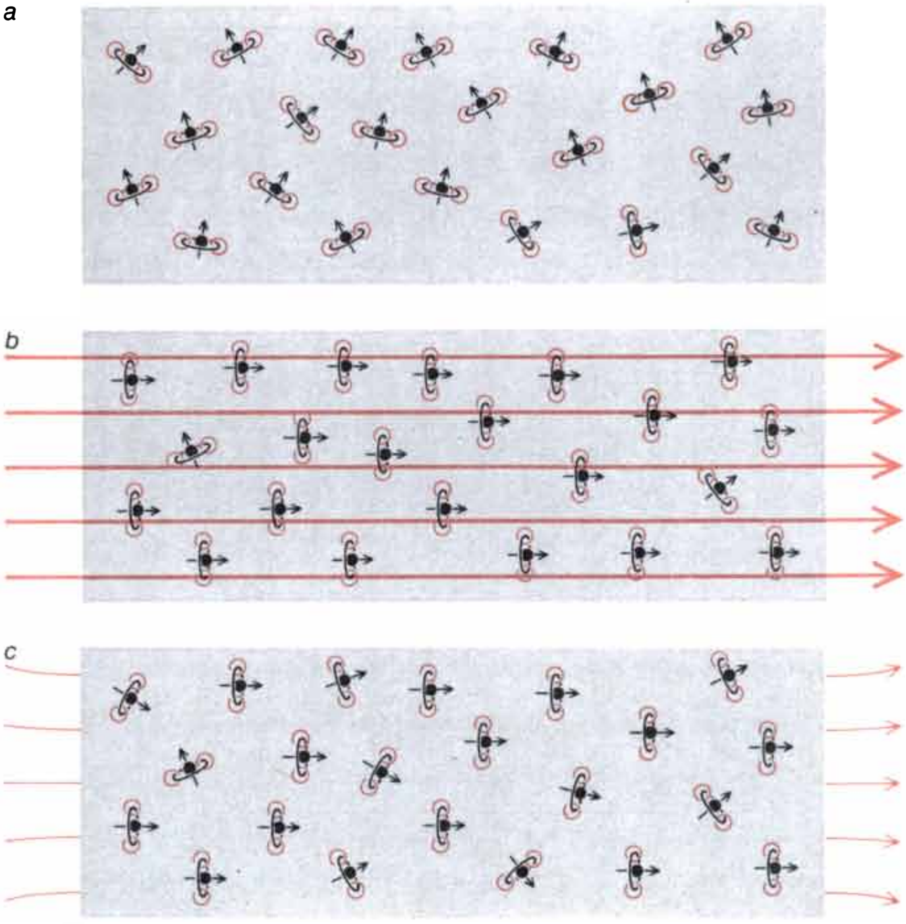




CIRCULATING ELECTRIC CHARGE in a loop of wire generates a dipole magnetic field with its axis oriented at a right angle to the plane of the loop (*left*). The movement of bound electrons around the nucleus of an atom constitutes a similar loop of current and endows the atom with a corresponding dipole field (*right*). Only one representative electron is shown.

The effect of reversing time is to reverse both the motion of the proton and that of the electrons making up the current; hence the magnetic field is also reversed. Under these circumstances the proton simply retraces the same path in the opposite direction; the path of the proton is said to be invariant with respect to time reversal.

Now suppose the magnetic field arises not from an electric current but from the presence of a magnetic monopole. Reversing time does not alter the polarity of the monopole and therefore leaves the direction of the magnetic field unchanged. As before, the proton reverses



ATOMIC MAGNETS are randomly oriented in an ordinary, nonmagnetic iron bar (*a*). The atoms can be aligned by the application of an external magnetic field (*b*). When the external field is removed (*c*), many of the atoms remain aligned, forming a permanent magnet. Cutting such a magnet in half has no effect on the atomic currents and hence does not serve to isolate the magnetic poles. For simplicity the bar is represented as having a single magnetic domain.


POPULAR
POP electronics
POP & photo

WE ARE CALCULATORS!
AUTHORIZED HEWLETT PACKARD DEALER

HP-41C	$186.95
HP-41CV	249.95
PRINTER/41C	279.95
CARD READER/41C	164.95
HP-32E SCI-STATS	39.95
HP-37E BUSI-MGMT	57.50
HP-38C BUSI-PROG	108.95
HP-33C SCI-MEM	68.50
HP-34C SCI-MEM	108.95
HP-11C SCIENTIFIC	99.50
HP-12C FINANCIAL	112.95

- HP ACCESSORIES IN STOCK
- MAIL ORDERS PROMPTLY SHIPPED

POPULAR ELECTRONICS & PHOTO
49 WEST 23RD STREET
NEW YORK, NY 10010 • 212/691 • 6045
ORDER TOLL-FREE 1 • 800 • 847 • 4204

columbia pacific university

NONRESIDENT BACHELOR, MASTER AND DOCTORAL DEGREES FOR THE ACCOMPLISHED INDIVIDUAL

Columbia Pacific University, the largest nonresidential graduate university in the U.S., has been authorized by the State of California to grant nonresident Bachelor, Master and Doctoral degrees in numerous fields, including Business, Psychology, Engineering and Health.

Degrees are earned through a combination of full academic credit for life and work experience, and completion of an independent study project in the student's area of special interest. The time involved is normally six to 12 months. The cost is about $2,000.

Columbia Pacific University is attracting accomplished individuals, members of the business and professional community desiring to design their own projects, and receive academic acknowledgement for their personal achievements. May I send you more information?

R.L. Crews, M.D., resident
COLUMBIA PACIFIC UNIVERSITY
150 Shoreline, Suite 5404
Mill Valley, CA 94941
Call Toll Free USA: 800-227-1617, ext. 480
California only: 800-772-3545, ext. 480

© 1982 SCIENTIFIC AMERICAN, INC

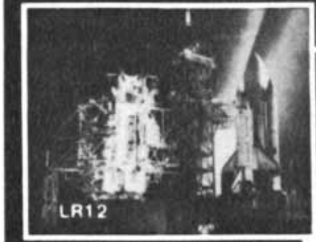
direction, but it does not retrace its path. In short, the proton's path in the field of a monopole depends on the direction of time, an effect that violates the principle of time-reversal invariance.

The predicted effects of a magnetic monopole when time is reversed were for many years viewed as a serious argument against its existence. In 1964, however, Val L. Fitch, James W. Cronin and their colleagues from Princeton University, in an experiment done at the Brookhaven National Laboratory, discovered an effect much like a violation of time-reversal invariance in the decay of the particles called neutral kaons. The full theoretical significance of this finding has only recently begun to be understood. As this understanding has increased, some of the opposition to the idea of magnetic monopoles has abated.

Given the tantalizing properties of magnetic monopoles, what is the status of the experimental search for evidence of their existence? Soon after every new particle accelerator is commissioned magnetic monopoles are looked for in the debris of the initial high-energy particle collisions; searches of this kind have become virtually a rite of passage. Monopoles have also been sought among the by-products of collisions between cosmic rays and atoms in the atmosphere. Experiments of another kind have attempted to detect monopoles among the iron atoms of terrestrial and extraterrestrial substances. So far none of these searches has succeeded.

How are the experiments done? One approach is to look for monopoles in iron that has been exposed to a beam of high-energy particles from an accelerator. If monopoles are produced by such a beam, some of them should bind to the iron by inducing an opposite magnetic charge in the material (much as a note-holding magnet, say, clings to the door of a refrigerator). A powerful electromagnet should then be able to pull the monopoles out of the iron. Monopoles liberated in this way would be detected in particle counters designed so that only very strongly ionizing particles would be recorded. Samples of iron ore collected from the rocky outcroppings of old mountains are another potential source of material for this method of extracting monopoles. Indeed, this approach was initially taken in the 1940's by Willem Malkus of the University of Chicago.

Another detection method, first discussed in the 1960's, was implemented in the 1970's by Luis W. Alvarez and his colleagues at the Lawrence Berkeley Laboratory of the University of California. In their device a sample of material suspected of harboring magnetic monopoles is passed repeatedly through a superconducting coil. On each pass of a magnetic monopole the electric current


LR12


LR31


LR38


LR44

in the coil would presumably increase by a small amount. Because the coil is superconducting the incremental induced current would persist indefinitely. The task then becomes one of measuring the extremely small signal induced by multiple passes of a single monopole. By means of this technique Alvarez and his colleagues were able to show that the density of magnetic monopoles in rock samples recovered from the surface of the moon is less than one for every 10^{28} protons. Even at this limiting abundance, however, there could still be an average of one monopole in every 20 kilograms or so of matter.

A less direct way of hunting for magnetic monopoles is to look for signs of the creation and destruction of a monopole-antimonopole pair. In theory a pair of this type could be created when a high-energy photon passes near a proton, just as an electron-positron pair is known to be produced [*see illustration on page 116*]. The oppositely charged monopoles would exist for only a moment, however. They would immediately be attracted to each other, bending their paths in a way that would cause them to spew out photons of *bremsstrahlung,* or "braking radiation." They would soon come together and annihilate each other, converting their mass into additional photons.

This hypothetical mechanism was invoked by Malvin A. Ruderman and Daniel Zwanziger of New York University in the mid-1960's to explain a few unusual cosmic-ray events recorded in the late 1950's. Each of the events consisted of a single jet of tightly bunched, very energetic, photon-induced showers of electrons in which there was no evidence of any strongly interacting nuclear particles. Two subsequent experiments done with the large proton accelerator at the Fermi National Accelerator Laboratory (Fermilab) found no evidence of showers of this type. (One of the experiments was done by a group from the Virginia Polytechnic Institute and State University and the other by a group from the University of Michigan.) If the process is real, it might take place only at energies higher than those currently available in particle accelerators. Alternatively, the process could be much rarer than the cosmic-ray evidence suggests.

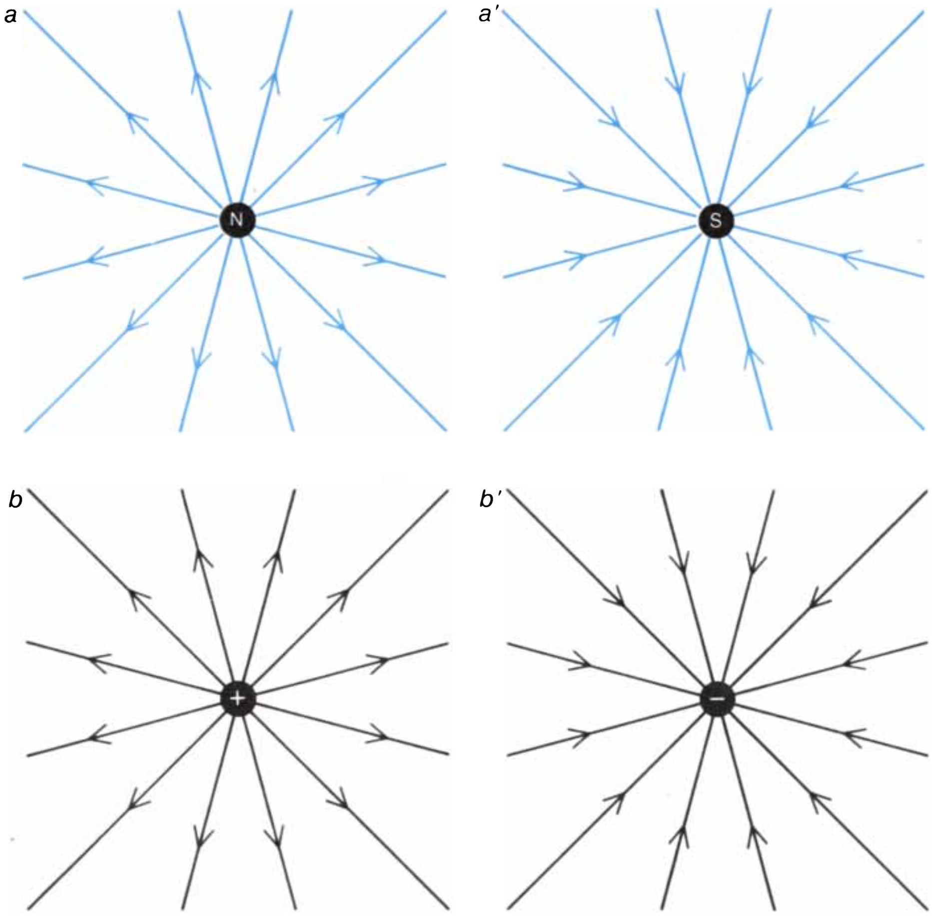
In 1975 the world of physics was jolted by the announcement that a magnetic monopole had been discovered. The claim was made by investigators at the University of California at Berkeley and the University of Houston. Their evidence was an anomalously thick, dark track, presumably of cosmic-ray origin, recorded in a stack of photographic emulsions and plastic sheets. The detector had been exposed to cosmic rays while it was suspended from a balloon flown at high altitude for two and a half days. Soon after the announcement the interpretation of the event as evidence of the passage of a magnetic monopole came under widespread criticism. A cosmic-ray experiment of this kind is characterized by its area-time factor: a measure of the area of the detector multiplied by the exposure time. The detector in which the candidate track was seen had an area-time factor roughly a million times smaller than that attained in previous searches in which no monopole had been seen. In the analysis of moon rocks by Alvarez and his colleagues, for example, the rocks had been exposed to particles of all kinds for billions of years.

Other problems with the monopole interpretation of the event subsequently led the experimenters to suggest instead that the track might have been caused by the passage of a superheavy atomic nucleus or a massive antiparticle. One benefit of the episode is that it has inspired a careful evaluation of the pivotal question of how a magnetic monopole would lose energy through ionization. Even so, the question remains unsettled.

At about this time the prospects for magnetic-monopole hunters suddenly brightened. Working independently, Gerard 't Hooft of the University of Utrecht in the Netherlands and Alexander M. Polyakov of the Landau Institute for Theoretical Physics near Moscow found that a certain class of theories of elementary-particle interactions not only allow magnetic monopoles but also demand them. Furthermore, the new theories, which are called gauge theories, indicate that the monopoles must be much more massive than any particle previously seen or even predicted. In retrospect it was hardly surprising that none had been detected up to then.

Besides having an extraordinarily large mass, the magnetic monopoles proposed by 't Hooft and Polyakov differ in several other respects from the original Dirac monopoles. For one thing, the 't Hooft-Polyakov monopoles do not require a string. For another, they are not pointlike particles, although they are expected to be too small to be directly measured.

The superheavy 't Hooft-Polyakov monopoles play an important role in attempts to construct a "grand unified theory" to describe three of the four known forces of nature. Two of these forces, the electromagnetic force and the weak nuclear force, are already linked by the highly successful electroweak theory, which treats them as different manifestations of a single underlying force. The goal of most of the current attempts to create a grand unified theory is to arrive



SYMMETRY between magnetic monopoles and electrically charged particles such as the proton extends in theory to their antimatter counterparts. A north monopole (*a*) would have as its antiparticle a south monopole (*a′*), just as the proton (*b*) has as its antiparticle the antiproton (*b′*).

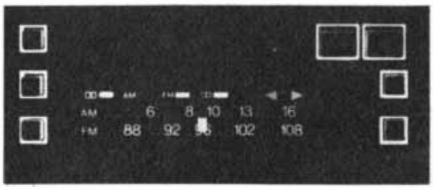
The Blaupunkt CR-2010. Richer, purer sound than you ever thought possible in a moving vehicle.

Autosound is only as good as the equipment it passes through. Which means that, in the case of the Blaupunkt CR-2010, the sound is exceptional.

The CR-2010 coddles and shapes highs and lows into a sound as full-bodied, as richly-textured, as anything you'll hear from a home stereo.

4 Channel Amplifier

Blaupunkt increased the conventional two channels to four, each with a maximum output of 7.5 watts. Even when hooked up to a front end, two-speaker system, the CR-2010's crisp reproduction will surprise you. Add two rear speakers and the home stereo effect is complete—sound that surges to new heights of clarity and richness while holding its delicate balance through the magic of a built-in front-to-rear fader.



The essential controls are fully illuminated.

Holds Signal Longer

Drive away from the signal source of your favorite stereo station and what happens? Reception breaks up into a barrage of crackles and hisses. Not with the CR-2010. Thanks to the "Soft MPX" feature your Blaupunkt automatically shifts reception from stereo to mono *before* the hissing sets in.

Higher Volume without Distortion

The CR-2010 has a pre-amp output jack that lets you bypass the built-in amp and plug directly into a high power amplifier. Yet the boost in volume doesn't come at the expense of distorted sound. And at lower volume the clarity is actually enhanced.

Of course, you get Dolby Noise Reduction—but for FM reception as well as tape. A Sendust Alloy tape head reproduces a fuller range of recorded frequencies.

Hear the Full Blaupunkt Line

The CR-2010 retails for only $396.00** and is the latest in a full line of Blaupunkt AM/FM stereo cassette radios priced from $290.00.**

CR-2010 Features

- **4 channel (4 x 7.5W)**
- **Autoreverse Cassette**
- **Sendust Alloy Tape Head**
- **Auto Hi-Cut Filter**
- **Soft Mute**
- **Soft MPX**
- **Night Illuminated Controls**
- **Dolby Noise Reduction Circuit**

*Dolby is a registered trademark of Dolby Laboratories.

Blaupunkts can be installed in virtually any car, import or domestic. For more information, write:

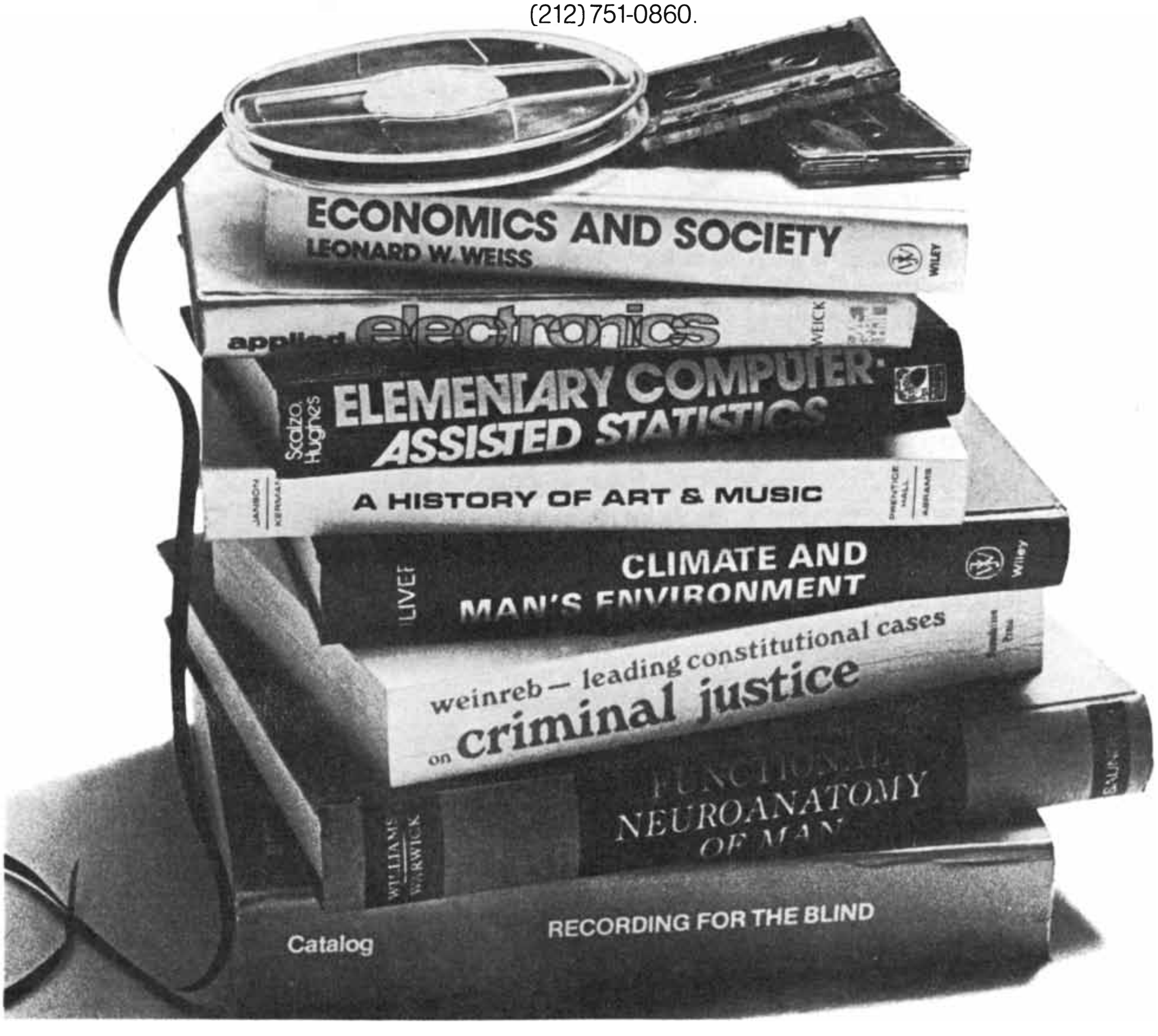
Robert Bosch Sales Corporation
2800 South 25th Avenue
Broadview, IL 60153

Robert Bosch Canada, Ltd.
6811 Century Avenue
Mississauga, Ontario L5N 1R1

**Suggested U.S. retail prices. The Blue Dot and Blaupunkt are registered trademarks of Blaupunkt Werke GmbH.

●BLAUPUNKT

©1982 Robert Bosch Sales Corporation

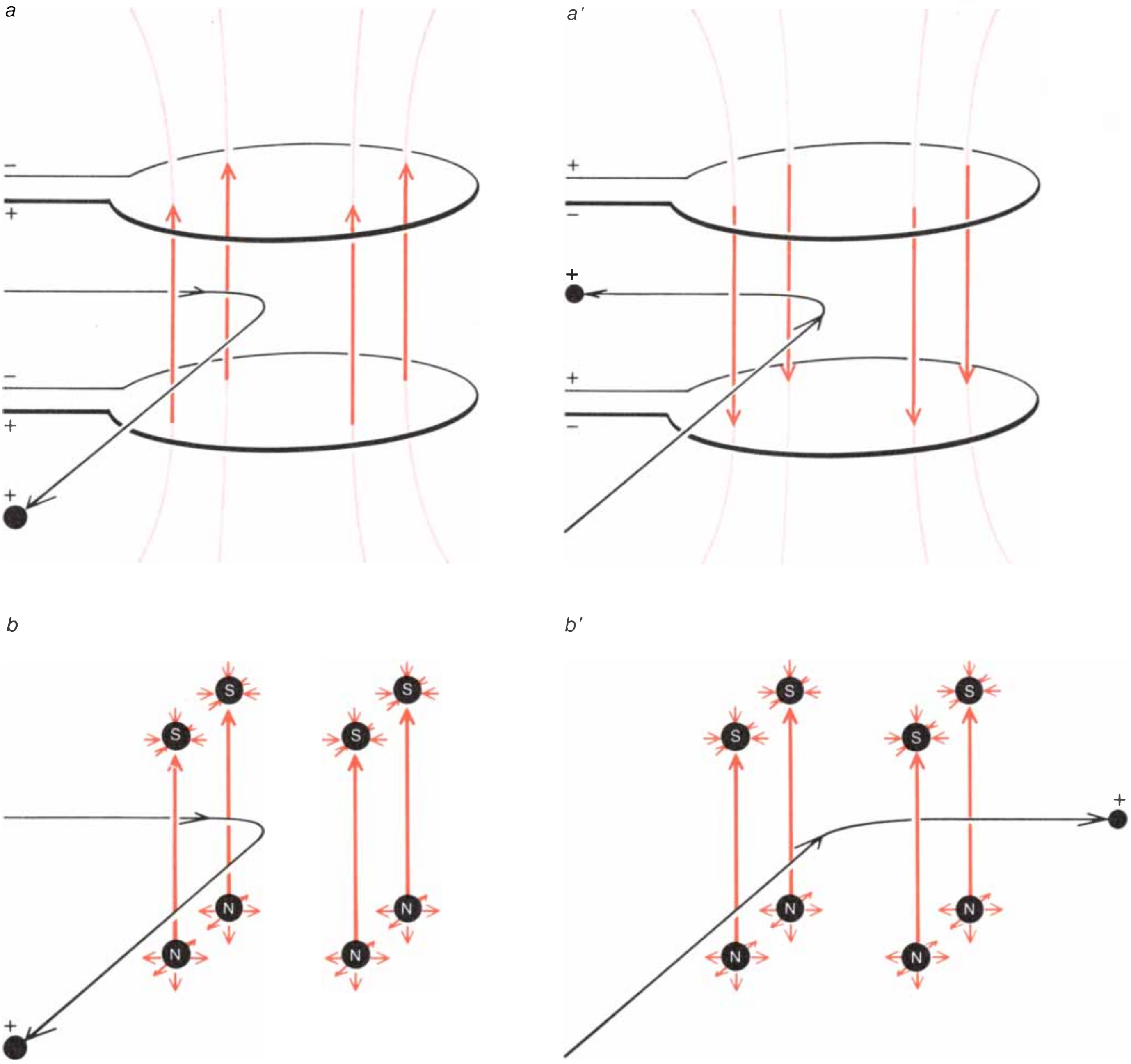

ECONOMICS AND SOCIETY
LEONARD W. WEISS
WILEY
applied electronics
WEICK
ELEMENTARY COMPUTER-ASSISTED STATISTICS
Scalzo Hughes
A HISTORY OF ART & MUSIC
JANSON KERMAN
PRENTICE HALL
ABRAMS
CLIMATE AND MAN'S ENVIRONMENT
OLIVER
Wiley
weinreb – leading constitutional cases on criminal justice
FUNCTIONAL NEUROANATOMY OF MAN
WILLIAMS WARWICK
SAUNDERS
Catalog
RECORDING FOR THE BLIND

at a more comprehensive mathematical structure that would incorporate both the electroweak force and the strong nuclear force (omitting only gravity, the fourth known force).

One of the most interesting consequences of such grand unified theories is the prediction that the proton will decay into other particles. To account for such a disintegration novel particles called leptoquarks are introduced to change quarks (the supposed constituents of particles such as the proton) into leptons (particles such as the electron that respond to the weak nuclear force but not to the strong one). The leptoquarks would be extremely heavy, perhaps 10^{14} times as heavy as the proton. If, as the 't Hooft-Polyakov hypothesis suggests, a superheavy magnetic monopole is associated with the leptoquarks, the monopole would have a mass of about 10^{16} proton masses, or roughly 20 nanograms. This is an exceedingly large mass for an elementary particle, comparable to that of a paramecium or an amoeba.

Superheavy monopoles could have been created only in the first 10^{-35} second after the birth of the universe. The big bang of creation would have been the only event hot enough (almost 10^{30} degrees Kelvin) to generate such particles. Both north and south magnetic monopoles would have been formed, and a small fraction of them would have recombined, annihilating each other. Most of the superheavy monopoles would have escaped an early death, however, and there is no reason to think they would not have survived to the present. It is unclear where the monopoles would have collected as the universe evolved, but then it is also unclear how the universe evolved from the big bang into the galactic structures we see today.

One feature of our own galaxy makes



TIME REVERSAL would have a peculiar effect on an electrically charged particle moving through the magnetic field produced by a magnetic monopole. In *a* a proton is shown moving along a curved path through a perpendicular magnetic field generated by electric currents flowing in a pair of wire loops. If the direction of time is reversed (*a'*), both the currents (and therefore the magnetic field) and the motion of the proton would be reversed. The path of the proton, however, would be invariant: the particle would simply retrace the same path in the opposite direction. In *b* a proton is shown moving along a similar path through the magnetic field produced by an idealized array of north and south monopoles. In this case reversing time would leave the magnetic field unchanged (*b'*). Although the proton would reverse direction, it would not retrace its path; such a result would be a violation of the principle of time-reversal invariance.

it possible to set a stringent limit on the number of magnetic monopoles that could be present locally. The galaxy has a magnetic field, and although it is weak (on the order of a hundred-thousandth of the earth's magnetic field), it extends over enormous distances. A monopole caught in the galactic field would be accelerated to very high energy and would eventually escape from the galaxy's gravitational attraction. Because the monopoles would be extremely massive, however, they would still move slowly (at velocities of a few thousandths the speed of light).

Eugene N. Parker of the University of Chicago has pointed out that if there were too many monopoles of this kind in the galaxy, they would destroy the galaxy's magnetic field. His argument is based on the fact that the magnetic field of the galaxy is generated by the large-scale circulation of electrically charged particles. As a magnetic monopole is accelerated it would drain energy from the galactic field by slowing the currents of moving electric charges. The existence of a galactic magnetic field therefore places an upper limit on the population of magnetic monopoles in the galaxy.

The maximum density of monopoles in the universe is related to the unresolved question of whether the universe will continue to expand forever or will eventually collapse. The outcome depends on the amount of matter in the universe. The amount of visible mass (that is, the mass of luminous objects such as stars) is not quite enough to cause the expansion to slow, stop and ultimately reverse. Unless there is some additional, unseen mass the expansion of the universe will gradually slow, but it will never stop.

Is there enough invisible matter in the universe to affect its fate? What form of matter could be present in great abundance and yet remain undetected? One possibility is that the missing mass consists of neutrinos, which rarely interact with other particles and never give off light. The neutrino was long thought to be massless, but there is now much speculation that it does have a small mass. Because neutrinos are thought to fill the universe with an average density of approximately a million per cubic centimeter, even a small neutrino mass could make a significant contribution, although probably not much more than the visible mass.

In all likelihood superheavy magnetic monopoles would radiate little light and hence would also be an invisible component of the universe. If the monopoles have a mass 10^{16} times the proton's mass, it would not take many of them to greatly augment the total mass. At a concentration of just one monopole for every 10^{16} protons there would be roughly as much mass in monopoles as there is in luminous matter. It does not seem feasible for the total mass of the universe to be much more than 10 times the visible mass, and so as an upper limit the ratio of magnetic monopoles to protons must be less than one to 10^{15}.

In 1979 John P. Preskill, then a graduate student at Harvard University, combined the grand unified model of the strong, weak and electromagnetic forces with standard cosmology to argue that there should be about one magnetic monopole for each proton in the universe. From the analysis of the expansion rate of the universe, on the other hand, it appears that there should be less than one monopole per 10^{15} protons. Preskill formulated a dilemma: Either the role of magnetic monopoles in the grand unified theories is wrong or standard cosmology is wrong. One way out of Preskill's dilemma is to adjust the cosmological model, allowing more monopole-antimonopole annihilations in the early moments of the universe. Another option is to suppress the estimated initial production rate of magnetic monopoles by some unspecified mechanism.

Recently George Lazarides, Qaisar Shafi and Thomas Walsh of the European Organization for Nuclear Research (CERN) in Geneva considered how the predicted density of superheavy magnetic monopoles in the universe could be reduced by adjusting either particle theory or cosmological theory. They concluded that the interaction of monopoles with the galactic magnetic field sets a limit on the ratio of magnetic monopoles to protons of about one to 10^{20}. Given that abundance, some 200 monopoles per year would be expected to pass through an area of one square kilometer. A more conservative estimate, based on a more uniform distribution of monopoles in the universe, would result in a flux of a few monopoles per year per square kilometer.

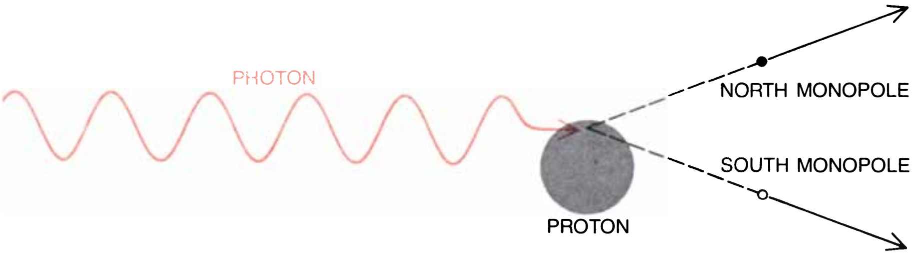
For the first time the theory of magnetic monopoles thus provides estimates of the expected mass and flux of magnetic monopoles. Armed with these estimates, however rough they are, the experimenter now has a fresh field to explore. The predicted flux of magnetic monopoles is small, but not so small that it is out of the question to look for them.

One place to look for superheavy monopoles is in large-scale natural effects. Indeed, this is the direction originally suggested by Parker a decade ago. One of us (Carrigan) has speculated on the fate of monopoles in the material that accreted to form the solar system. As the earth condensed, for example, magnetic monopoles would have sunk toward the center under the influence of the planet's gravitational and magnetic fields. North monopoles would have collected near the south geomagnetic pole and vice versa.

From the geologic record it is known that the earth's magnetic field has reversed many times. Such a field reversal would cause the two separated populations of monopoles to migrate toward and then through each other. During their journey some monopoles and antimonopoles would be annihilated, liberating the enormous energy embodied in their mass. From the measured heat flow at the surface of the earth one can set a rough limit on the number of monopoles trapped in the core; the number calculated in this way is consistent with other experimental limits on the abundance of superheavy monopoles.

A more straightforward approach is to construct a detector specifically to search for these heavy, rare particles. The design of such a detector, however, is not obvious. Indeed, the art of searching for massive monopoles is now at one of those engaging moments in science when a wealth of ideas, many of them quite bizarre, are at war on paper and over lunch tables. Massive monopoles are expected to travel slowly, at speeds far below the velocity of light. Just what would happen when a slow-moving monopole struck an atom is not clear.

The collision of a superheavy monopole and a stationary atomic nucleus would be like a steamroller hitting an ant. A cosmic-ray monopole could lose a huge amount of energy to many such encounters as it plowed its way ponderously through the earth, and it might still emerge virtually unscathed from the other side. Under these circumstances it is difficult to predict what degree of ionization would be observed in



PARTICLE-ANTIPARTICLE PAIR, consisting of a north monopole and a south monopole, could be created when a high-energy photon, or quantum of electromagnetic radiation, interacts with an electrically charged particle such as a proton. The mutual attraction between the monopoles, however, would cause them to collide, converting their mass back into photons.

a detector. One view holds that there are enough fast-moving monopoles for detectors relying on ionization to be used in the search. The other view is that ionization will be seen rarely and only weakly, so that unconventional techniques would be needed to detect monopoles. In either case, an extremely large detector is clearly needed if the experimenter is to observe a monopole event in his lifetime.

One detector that records the light generated by ionization and covers many square kilometers has been developed by Haven E. Bergeson, George L. Cassiday and Eugene C. Loh of the University of Utah. The device, called the fly's-eye detector, is an array of photomultiplier tubes directed at the night sky; it registers the light given off by secondary particles produced by rare ultrahigh-energy cosmic-ray interactions in the upper atmosphere. As the secondary particles shower down toward the earth they collide with nitrogen atoms in the atmosphere, causing them to scintillate. The fly's-eye detector can distinguish this light only if the total energy released in the event is equivalent to at least 100 million proton masses. The passage of a magnetic monopole, even with the most optimistic estimate of its ionization rate, would give rise to less than a ten-thousandth of the light needed to set off the detector.

The ability of such a detector to respond to particle-induced scintillations is limited by background illumination from stars, overflying aircraft and other sources such as the beacon lights of distant radio towers. One suggestion for alleviating the problem of background light is to mount a detector on the rim of the Grand Canyon, point it into the canyon and record data only on cloudy nights. Even this measure, however, would reduce the background light only tenfold. In addition it might be difficult to get hikers to forgo campfire meals for several years. Perhaps a fly's-eye detector could be installed in a large cave or a salt mine such as those now being used to look for proton decay.

Another large-volume detector now being planned is the Deep Underwater Muon and Neutrino Detector (DUMAND), which will be sensitive to events within a cube of ocean about a kilometer on a side [see "A Deep-Sea Neutrino Telescope," by John G. Learned and David Eichler; SCIENTIFIC AMERICAN, February, 1981]. DUMAND will respond to the Cerenkov radiation emitted when a particle moves through the seawater faster than the speed of light in water. Unfortunately superheavy magnetic monopoles would probably move too slowly to give off Cerenkov radiation.

Some of the largest existing scintillation detectors, such as the giant neutrino detectors at Fermilab and CERN, are too small by a factor of about 100 to have a good chance of observing magnetic monopoles if the flux is limited by the galactic magnetic field. Nevertheless, it may still be useful to turn these detectors to searching for monopoles during accelerator shutdowns, when otherwise they would be unemployed, since the experimental limits for slow-moving, superheavy particles are not well defined. A preliminary search just completed by Jack D. Ullman of Lehman College in New York employed a detector half a meter square operating for several months, thereby providing the only experimental limit obtained to date. Definitive experiments will have to be 10,000 times as sensitive.

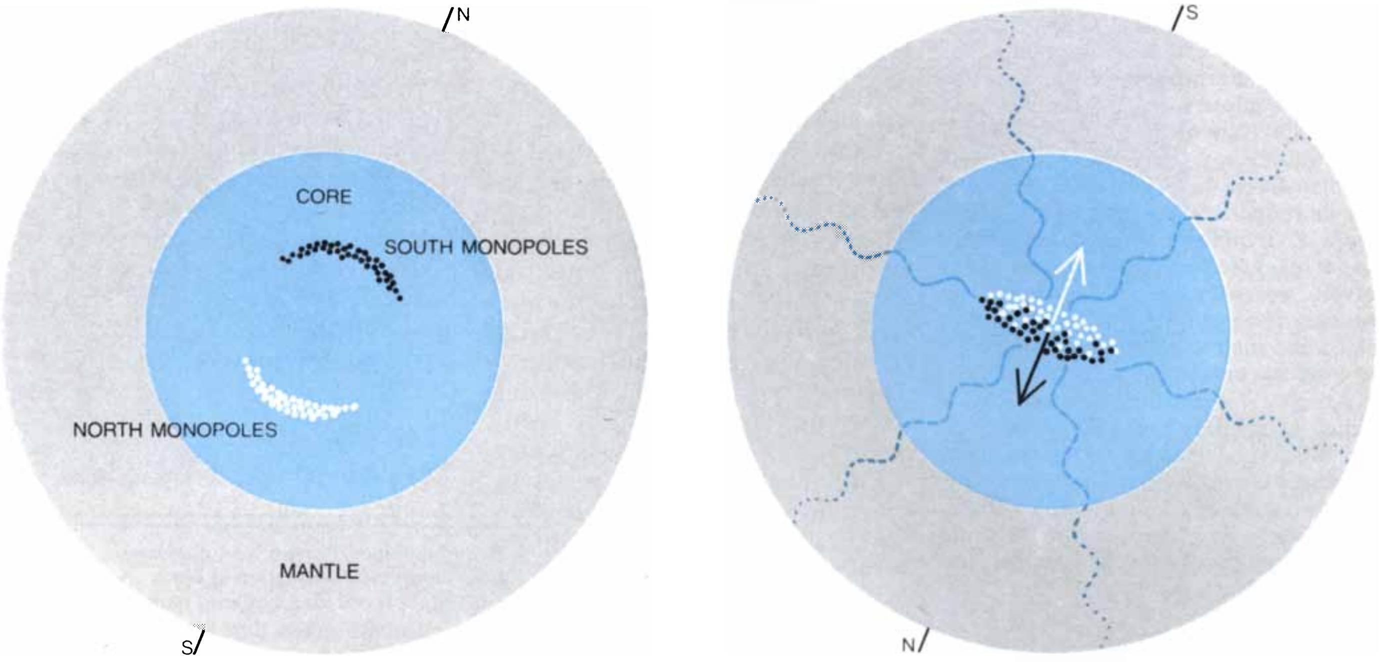
The contrary view holds that all searches with ionization detectors are doomed to failure because the slow-moving, superheavy monopoles will cause no ionization. The passage of any charged particle through metal is accompanied by eddy currents, however, regardless of the particle's speed and of whether its charge is electric or magnetic. The eddy currents can be expected to create acoustic pulses, which might be detected. Carl W. Akerlof of the University of Michigan has checked the possibility of constructing a spherical metal detector along these lines and has concluded that a signal could be detected above the background of thermal noise only if the detector is cooled to a few millidegrees above absolute zero. This difficult technical requirement is compounded by the necessity of making the detector large enough to sense the low expected flux of monopoles.

One comparatively simple strategy for detecting superheavy monopoles calls for a superconducting coil similar to the one employed by Alvarez and his colleagues. The detector is called a SQUID (for superconducting quantum interference device), and it registers changes in an electric current when a free cosmic-ray monopole passes through it. Blas Cabrera of Stanford University is currently searching for monopoles with a superconducting niobium coil five centimeters in diameter. Cabrera also has a second monopole search under way. In the latter experiment he inflates a cylindrical, superconducting lead bag one meter long and 20 centimeters in diameter, expelling most of the trapped magnetic flux. If a magnetic monopole penetrates the bag, it will leave trapped magnetic flux in the parts of the wall where it entered and left. By mapping the magnetism of the bag periodically, differences in the flux patterns can be attributed to the passage of a monopole. If a monopole is detected, rough information on its direction will be recorded.

Recently David B. Cline of Fermilab and the University of Wisconsin at Madison and Carlo Rubbia of CERN and Harvard University have initiated an ambitious monopole-detection effort. They plan to mount a superconducting detector under an iron-ore processing plant in Wisconsin. The plant heats more than a million tons of ore per year to a temperature of 1,700 degrees Celsius. At this temperature any magnetic monopoles trapped in the iron would be released, allowing them to fall through the detector.

The story of the magnetic-monopole conjecture is unlike any other in physics. Begun half a century ago by one of the giants of modern physics, the hunt for monopoles has been a fertile field for theoretical speculation but to date it has been barren of any supporting experimental evidence. The discovery of a magnetic monopole would rank as one of the finds of the century, comparable to the discovery of the positron, Dirac's other great prediction. If the monopole were found to be very massive, the case for some form of grand unified theory of elementary-particle interactions would be strengthened.

In the more likely event that no magnetic monopoles are found, the negative evidence will continue to be viewed as inconclusive. The necessary experiments will be difficult, and even if they are perfectly executed, a null result will not bring much clarity to the situation. Nevertheless, the ultimate vindication of the monopole idea rests in these searches, since physics is in the final analysis an experimental science.



MONOPOLES TRAPPED IN THE EARTH would tend to collect at two places in the earth's core. According to an estimate made by one of the authors (Carrigan), the equilibrium points where the population of monopoles would be densest would be less than 1,600 kilometers from the center of the earth. North monopoles would accumulate closer to the south geomagnetic pole and vice versa (*left*). Following a reversal of the earth's magnetic field (an event observed repeatedly in the geological record) the two segregated populations of monopoles would migrate through each other (*right*). Some pairs of north and south monopoles would meet and annihilate each other, converting their mass into energy. Even a very small density of superheavy monopoles could contribute noticeably to the outflow of heat at the earth's surface. Accordingly the observed heat flux sets a limit on the number of monopoles that could be trapped in this manner.

Variations in Medical Care among Small Areas

The amount and cost of hospital treatment in a community have more to do with the number of physicians there, their medical specialties and the procedures they prefer than with the health of the residents

by John Wennberg and Alan Gittelsohn

There is a city in Maine where the surgical procedure of hysterectomy (removal of the uterus) was done so frequently in the past decade that if the rate persists, 70 percent of the women there will have had the operation by the time they reach the age of 75. In a city less than 20 miles away the rate of hysterectomy is so much lower that if it persists, only 25 percent of the women will have lost their uterus by age 75. What could account for the disparity? It seems unlikely that there would be any large difference in the general health of the populations of the two neighboring cities, and after looking into the matter we have found none. The populations are similar in economic status. Differences in the number of physicians, the supply of hospital beds and coverage by medical-insurance plans cannot explain the difference in the rate of surgery. Instead the most important factor in determining the rate of hysterectomy seems to be the style of medical practice of the physicians in the two cities. In one city surgeons appear to be enthusiastic about hysterectomy; in the other they appear to be skeptical of its value.

We have examined the rate of surgery and other forms of medical treatment in 193 small areas in the six states of New England. The overall rate of surgery varies more than twofold among the areas. The total rate in a given area is correlated strongly with the number of surgeons there and with the number of hospital beds per capita; these are factors that themselves vary substantially. The amount spent per capita on treatment in hospitals is also quite different from one area to the next. The rates of three of the most common surgical procedures (hysterectomy, prostatectomy and tonsillectomy) vary even more dramatically: the highest rate is six times the lowest one. Even in communities with the same overall rate of surgery the rates of individual procedures can differ greatly. Hysterectomy, prostatectomy and tonsillectomy cause much controversy among physicians. In the absence of general agreement on their value for individual patients the style of practice of the individual physician appears to take precedence.

The substantial variation from area to area in the consumption of medical care and in its per capita cost is sustained by the policies of hospital boards and administrators, regulatory agencies and providers of medical insurance. The policies seldom take into account the existing level of health care in a community; a common result is an increase in medical services in areas that already have high rates of consumption. When such inequities develop, the people receiving the greater number of medical and surgical procedures do not necessarily benefit, particularly when the procedures entail substantial risk.

The 193 areas employed in our work cover the states of Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island and Vermont. Our aim in constructing the areas was to specify the population that attends one local hospital. Except in cases of disorders requiring elaborate treatment (such as cardiac surgery) people are generally treated at a nearby hospital. The attitude of the physicians at that hospital therefore has a strong influence on the rate of a given procedure in the surrounding area.

The analysis of medical care in small geographic areas has been made possible by the establishment of computer-encoded records of hospital admissions in specific regions. In Maine, Rhode Island and Vermont there are registries that include information on each patient admitted to a hospital. The registry lists the patient's age, sex, place of residence, diagnoses, surgical procedures, dates of admission and discharge and health on discharge. For Connecticut, Massachusetts and New Hampshire our data on hospital admissions come from studies in which only the hospital and the patient's place of residence are recorded. The expansion of health-insurance coverage, particularly the passage of the Federal Medicare Act in 1966, has yielded additional information about the medical care of specific populations.

To construct the geographic areas we extracted from the records each patient's residence and the community in which treatment was received. In the records we utilized, the residence of the patient is recorded in the form of his Zip Code, minor civil division (township, for example) or census tract. For each of these small units of residence we determined the community in which residents are most likely to be hospitalized. All townships, census tracts and so on whose residents were most likely to go to a particular community to be treated were combined to form a hospital area.

The 193 hospital areas defined in this way generally have populations of between 10,000 and 200,000, which is large enough for them to have stable rates of medical procedures. In almost all the areas the majority of hospital treatment is provided by facilities within the area.

By counting the surgical procedures done on the population of a hospital area in a given period, the per capita rate of surgery can be calculated. Similar methods give the rate of other kinds of medical treatment. Insurance-reimbursement rates can be calculated by totaling the reimbursements received by residents and dividing by the number of residents who are members of an insurance program. The number of hospital beds per capita is also readily determined.

Although the hospital areas can be employed in a variety of analyses, much of our work has concerned rates of surgery because the information on surgical procedures in regional record-keeping systems has been shown to be more reliable than that for other forms of treatment or for diagnosis.

After adjusting for differences in age

among populations we have calculated the rates of hospital admission for surgical procedures in the 11 most populous hospital areas in each of three states: Maine, Rhode Island and Vermont. Procedures done on residents of a hospital area are counted toward the area's total whether the operation took place within the area or outside it. Among the hospital areas in each state the overall rate of surgery varies by a factor of about two.

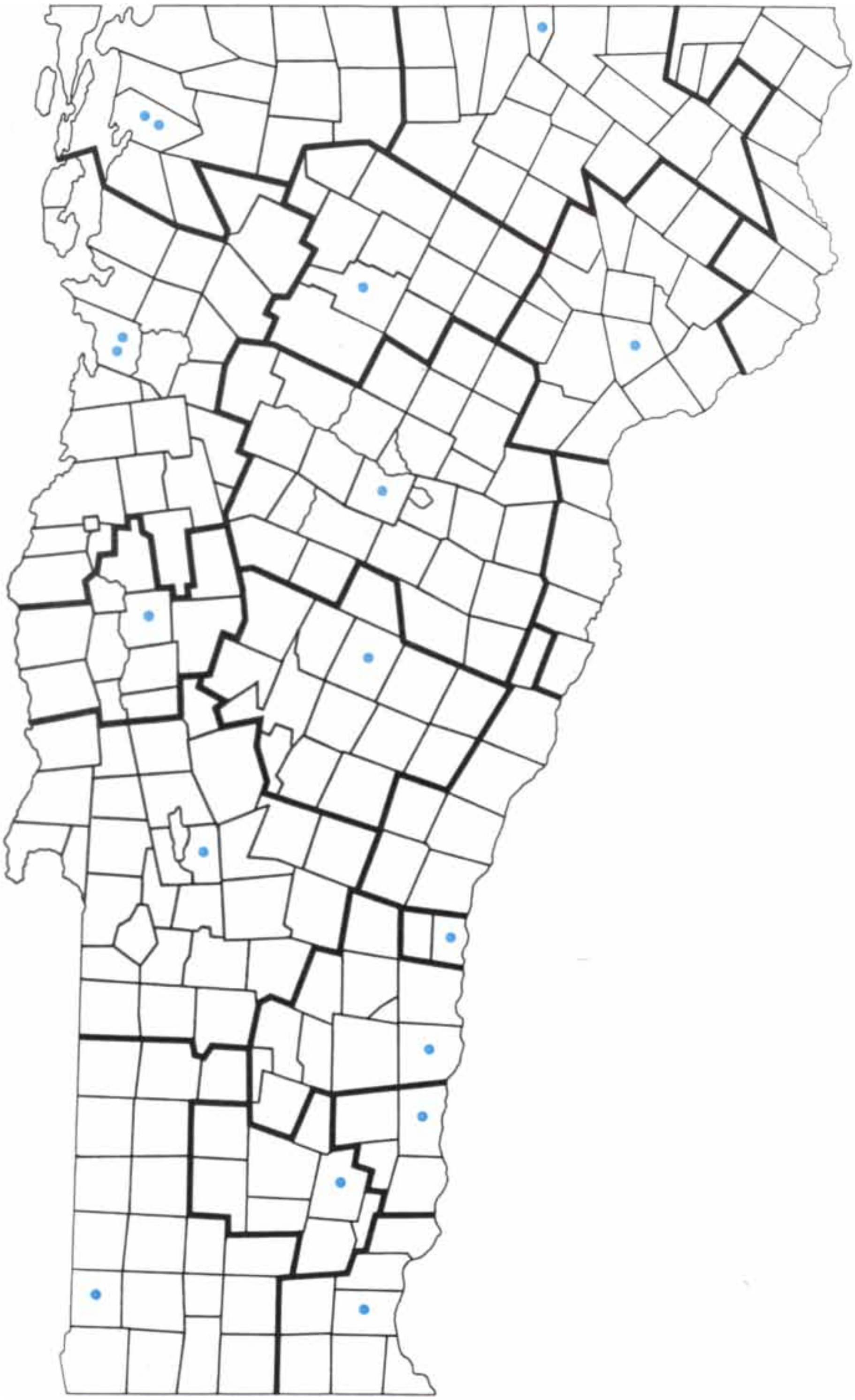
The variation in the rates of certain common procedures are more dramatic than the variation in the total. The highest hysterectomy rate in the 33 areas (some 90 procedures per 10,000 women per year in 1975) is about four times the lowest rate. The highest rate of prostatectomy is also about four times the lowest rate. For tonsillectomy the highest rate (about 60 procedures per 10,000 people per year) is six times the lowest rate. In many cases the difference between the extreme rates for a procedure and the average rate for all the areas is statistically significant, indicating that the difference is unlikely to be a result of chance variation.

Because of the wide range of rates residents of different areas face very different probabilities of having surgery. In one area of Vermont the tonsillectomy rate from 1969 through 1971 was such that if it had persisted, 60 percent of all children would have had their tonsils removed by age 20. In a second Vermont area the rate was such that only 8 percent would have had their tonsils removed by age 20. In the area with fewer tonsillectomies, however, the prostatectomy rate was such that 59 percent of all men would have had their prostate gland removed by age 80. In a neighboring area only 35 percent would have had a prostatectomy by age 80.

It is important to note that such large disparities in the rate of surgery are not observed for all common surgical procedures. The rates of cholecystectomy (surgical removal of the gallbladder) and appendectomy, for example, vary by a ratio of less than three to one. In few cases is the difference between the rate for an individual area and the average rate statistically significant. The rate of herniorrhaphy (surgical repair of hernia) varies even less, and most of the variation seems attributable to chance.

The ratio of hospital beds to population and the average cost of being hospitalized also vary greatly among New England communities. In the 11 most populous hospital areas in Vermont the highest ratio of hospital beds to population is 6.8 beds per 1,000 people; the lowest is 3.7. In Connecticut the highest ratio is more than four per 1,000; the lowest is less than two. (The ratios have been adjusted to compensate for residents treated outside their hospital areas.) The Federal Health Planning Program has specified four beds per 1,000 people as its standard for health-care planning; the ratios in New England thus range from well below the standard to well above it. Furthermore, the difference between small areas is so great that the number of beds per capita in a state or a county (a measure often used in health planning) is not a reliable indicator of conditions in each community.

The average reimbursement per area resident by agencies that provide medical insurance is another highly variable quantity. In Boston, the area of Massachusetts with the highest rate of Medi-



HOSPITAL AREAS of Vermont are small areas in which most residents go to a single hospital. The hospital areas are indicated by heavy lines; boundaries between townships within an area are indicated by light lines; hospitals are indicated by circles. The six New England states have been divided into 193 hospital areas for an analysis of geographic variations in medical care. Because most care in each hospital area is provided by physicians at the local hospital the attitudes of those physicians have a large effect on rates of treatment. Although the residents of most hospital areas appear to have similar health, rates of common surgical procedures vary as much as sixfold from one area to another. The number of hospital beds per capita and the amount spent per capita on hospital care also vary greatly. In two areas there are two hospitals. In the areas without hospitals the residents receive most of their care in New Hampshire.

care reimbursement, an average of $640 was paid to each person enrolled in the program in 1975. Across the Charles River in Cambridge the amount was $540. In Manchester, N.H., less than 50 miles away, it was $176.

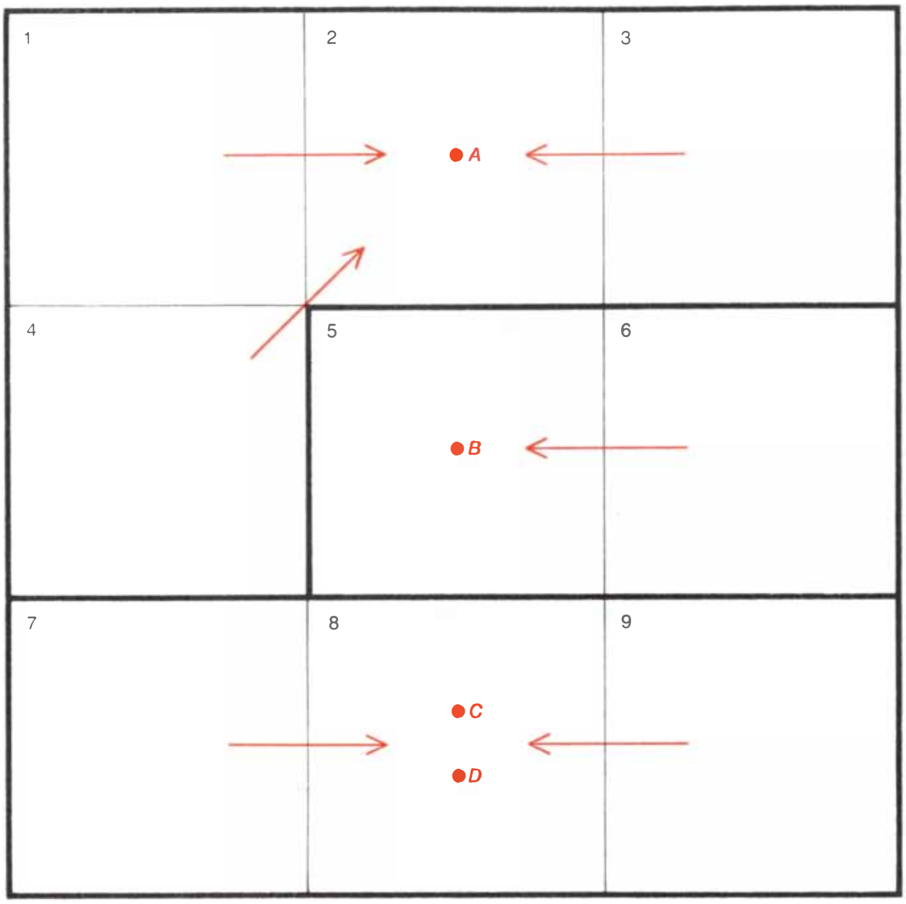
The amount spent per capita on all treatment in hospitals is also quite inconsistent. In 1975, $324 was spent per capita in Boston; in Providence, $225 was spent. In New Haven, on the other hand, per capita expenditure was $153; in Hanover, N.H., it was $120. In all these areas the majority of admissions are to major teaching hospitals. The services provided in the hospitals are probably similar; one would not expect such a disparity in the amount spent on treatment.

Our work has shown that residents of some areas receive much more medical treatment than others and spend more on that care. Why? One might assume that such differences are caused by better health in some communities, but this does not appear to be the case. Surveys we made in selected hospital areas show that differences in the health of residents and in the other factors that affect the demand for medical care account for only a small amount of the difference in the consumption of services.

Ronald M. Andersen and Lu Ann Aday of the University of Chicago have listed certain characteristics that influence whether or not an individual will seek medical treatment. The most important factors are those that affect the person's health or the perception of health. A variety of studies, however, have found that "enabling factors" such as income, health-insurance coverage and education can also have a strong effect. In addition "predisposing factors" such as skepticism or faith in medicine appear to play a role. In the course of our work we have tried to determine how much of the observed difference in medical care could be explained by such characteristics. With Floyd J. Fowler, Jr., of the University of Massachusetts at Boston we interviewed residents of six Vermont hospital areas. The residents were chosen to provide a representative sample of the local populations.

The total surgery rate and the amount spent per capita on hospital treatment differed as much as twofold in the six areas; rates of some surgical procedures varied even more. The interviews, however, showed that the residents differed little in the factors affecting the consumption of medical services. The average numbers of episodes of acute and chronic illness in each area were similar, as were the proportion of people with an income below the poverty level, the proportion with various kinds of health insurance and the proportion with access to a physician. Indeed, approximately equal proportions of the people in the areas visit a physician each year, as would be expected in populations of similar wealth and health. The large differences in surgical rates and the amount spent on hospital care must therefore be traced to factors that come into play after patients have contact with physicians.

What is it that takes effect after the patient sees a physician to increase the patient's chance of being hospitalized or of having surgery or a diagnostic procedure? The crucial factor appears to be the system of medical care in the community. Although health and other demographic factors do not differ much among the six areas, the number of hospital beds and the number of physicians in proportion to the population vary widely. Moreover, the supply of hospital facilities and the types of physicians who practice in the area are closely correlated with overall consumption rates. Where there are many hospital beds per capita and many physicians whose specialty or style of practice requires frequent hospitalization, there is more treatment in hospitals and greater expenditure per capita for hospital care. In hospital areas where there are many general surgeons the surgery rate is high.



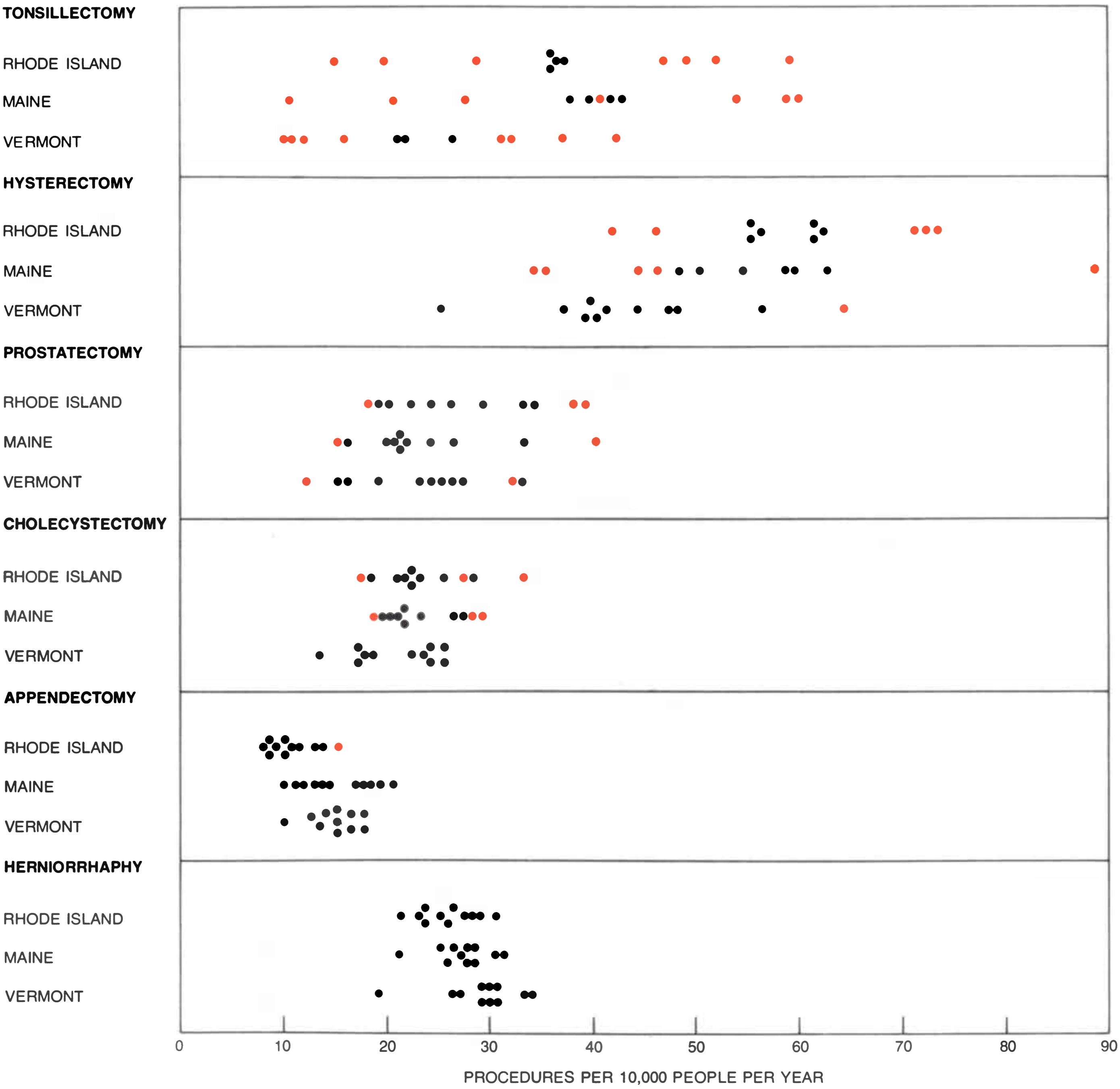
	TOWNSHIP								
HOSPITAL	1	2	3	4	5	6	7	8	9
A	55	1,236	200	19	76	4	8	18	49
B	0	82	0	16	85	826	10	7	18
C AND *D*	3	37	4	18	0	82	1,421	831	2,332
OTHER HOSPITALS	6	104	129	10	32	92	386	72	230

GEOGRAPHIC DIVISIONS from which the 193 hospital areas were assembled included townships, census tracts and Zip Code units. Patients' addresses and the hospitals where they were treated were taken from regional registries of hospital treatment. (Whether the town, the Zip Code or the census tract was recorded depended on the state.) For each small geographic division it was determined which community the residents were most likely to go to for hospital treatment. The illustration shows nine townships whose boundaries are indicated by light lines. There are four hospitals in the townships, indicated by circles: one hospital each in townships 2 and 5 and two hospitals in township 8. The table indicates where the residents of each township went for hospital treatment in a single year. On the basis of this information the townships were combined to form the three hospital areas whose boundaries are indicated by heavy lines. In each hospital area most of the residents are treated at hospitals in a single community.

The surgery rate and the rate of hospitalization are also high in communities where a large proportion of the general practitioners do surgery. In areas where there are many internists many diagnostic tests are given.

That the overall rate of surgery is influenced by the ratio of surgeons to population has been known for some time. In 1970 John P. Bunker of the Stanford University School of Medicine found that the total surgery rate in the U.S. was about twice that in Britain, where there are fewer surgeons per capita. In 1973 a study supported by the American College of Surgeons showed an analogous relation in regions of the U.S.

The total rate of surgery and the likelihood of being admitted to a hospital for treatment thus depend on the supply of physicians and hospital beds in the area. The wide variations in the rates of individual procedures, however, are not caused by differences in the supply of resources alone. Our work suggests that such variations are due to differences in the style of medical practice of local physicians. We examined the total surgery rate and the rate of individual procedures in the five most populous hospital areas in Maine. In three of the areas the total rate is close to the average for the state, in one area the rate is above average and in one it is below average. In each area, however, a different surgical procedure is the commonest one; all the commonest procedures are among those whose rates vary widely. For example, hysterectomy is the commonest procedure in one of the areas but the least common in another,



RATES OF SURGICAL PROCEDURES vary greatly among hospital areas. The rates shown are for the six commonest surgical procedures for the repair or removal of an organ in the 11 most populous hospital areas of Maine, Rhode Island and Vermont. The rate of tonsillectomy varies about sixfold among the 33 areas; the rates of hysterectomy and prostatectomy vary about fourfold. Moreover, many of the extreme rates for these procedures differ from the average rate for the state by an amount that is statistically significant (*color*). There is much disagreement among physicians on the value of these procedures and the conditions for which they should be done. The rates of cholecystectomy (gallbladder removal) and appendectomy vary less, and few of the extreme rates for these procedures differ from the state average by an amount that is statistically significant. The rate of herniorrhaphy (surgical repair of hernia) varies least.

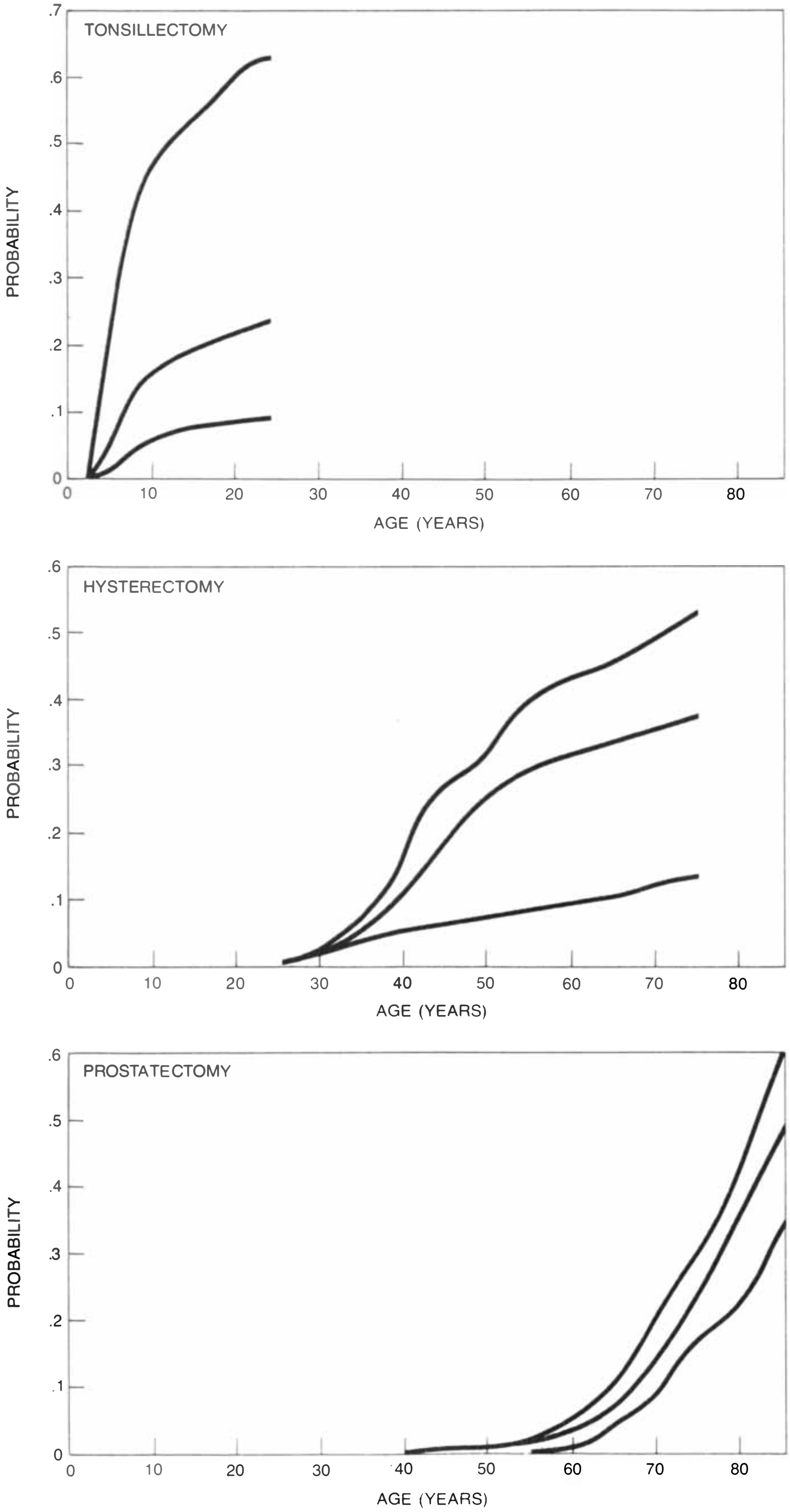
although the two areas have the same overall rate of surgery.

In each of the five areas of Maine the rates of common surgical procedures constitute a "surgical signature" that tends to be consistent over many years, unless physicians leave the area or enter it. In each signature the rates of some procedures exceed the state average; those of other procedures fall below the average. Nora Lou Roos, Leslie L. Roos and their colleagues of the University of Manitoba Faculty of Medicine reached similar conclusions after analyzing variations in small areas in Manitoba.

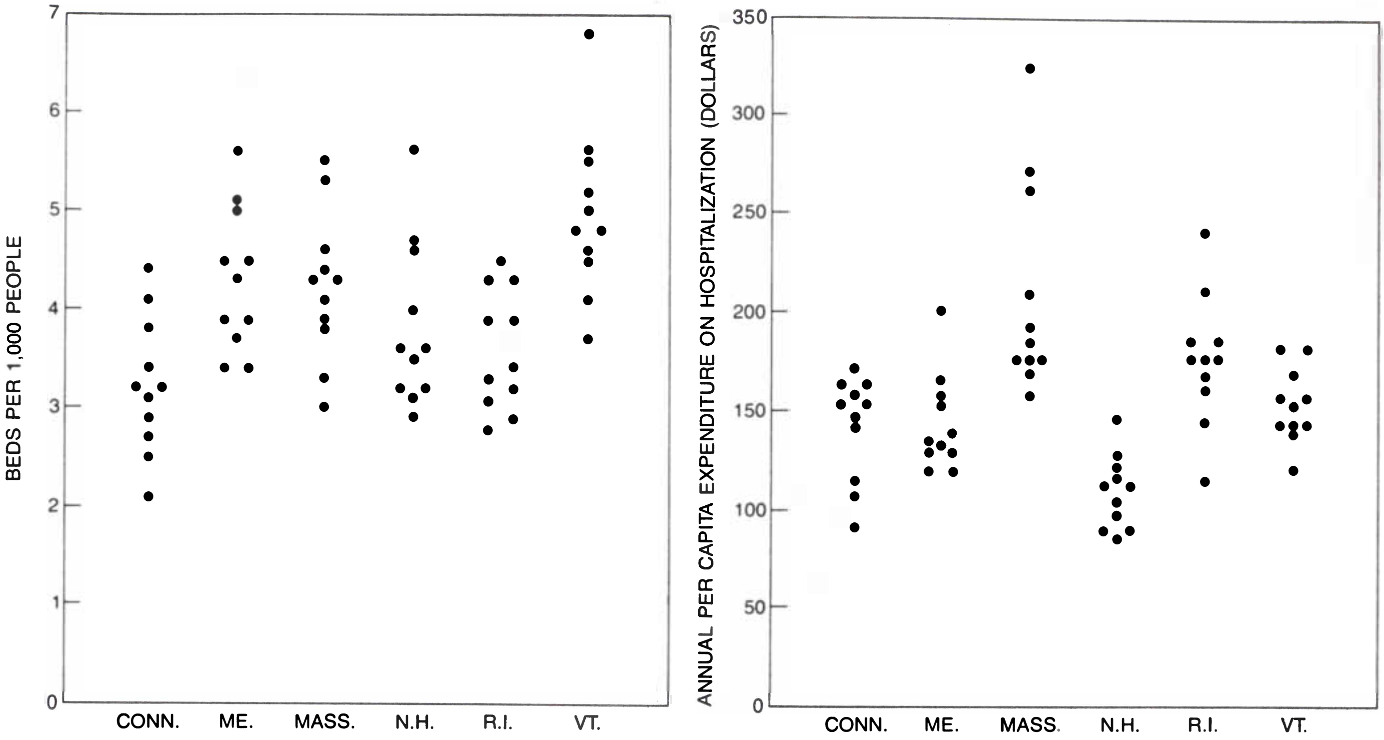
We have accounted for the factors that might influence the rates of surgical procedures, including the health of residents, the supply of hospital beds and the number of physicians. Even taken together, these factors cannot explain all the variation in rates of individual procedures. The strongest remaining hypothesis is that the judgments and preferences of physicians give rise to the surgical signature.

Some of the most persuasive evidence that the style of practice adopted by physicians has a strong influence on surgery rates comes from studies in which physicians are told of geographic variations in the rates. The studies also show that physicians' attitudes can be changed. In the 1950's Paul Lembke of the University of Rochester employed information similar to ours to calculate per capita rates of surgery in communities near Rochester, N.Y. He also persuaded physicians there to undertake an audit of surgical procedures. Soon afterward the rates were reduced in some areas where they had been high. We followed a similar course in Vermont, except that no formal audit was made. Information on the rate of tonsillectomy in each hospital area was given to the Vermont Medical Society. In the area with the highest rate physicians established the requirement that a second opinion be obtained before a tonsillectomy was done. As a result the probability that a child living in the area would have a tonsillectomy before age 20 declined from 60 percent to less than 10 percent. It had been suggested that if tonsillectomy became much less common at the local hospital, the people of the area would go to other nearby hospitals to obtain the surgery for their children. This did not happen, implying that demand by residents for the procedure had not been a major factor in maintaining the high tonsillectomy rate.

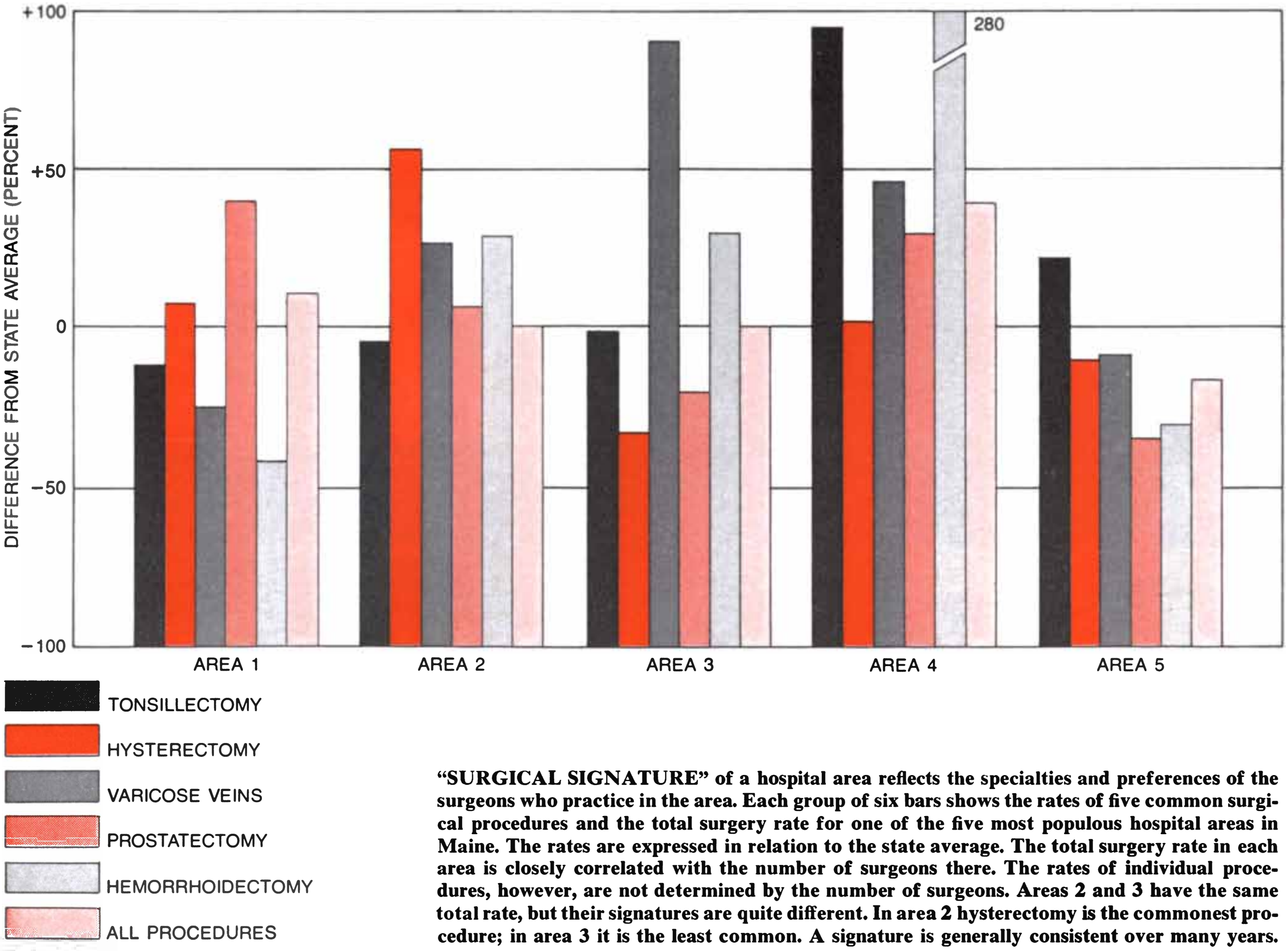
How can the decisions made by physicians vary so widely from one community to another a few miles away? It seems that the procedures whose rates vary the most are the ones whose risks and benefits are least well established in the medical profession. In some instances the value of the procedure itself has been questioned; in others the crite-



PROBABILITY OF HAVING SURGERY in 11 Vermont hospital areas has a wide range of variation as a result of the differences in surgical rates. In each graph the lines indicate the fraction of the population that would have had the procedure if the rates of surgery prevailing in about 1970 had persisted. The top line is for the area with the highest rate of the 11. The middle line is based on the average rate for the state. The bottom line is for the area with the lowest rate. In one hospital area 60 percent of all children will have had their tonsils removed by age 20, whereas in another area fewer than 10 percent will have a tonsillectomy by that age. The fractions that will have had hysterectomy and prostatectomy also show substantial differences.



NUMBER OF HOSPITAL BEDS per capita and the annual amount spent per capita on hospital treatment also show the influence of geographic variations in medical care. The data are for the 11 most populous hospital areas in each of the six New England states. The number of hospital beds per 1,000 people (adjusted for the number of people who leave their hospital area for treatment) ranges from about two to more than six. The ratios thus range from well below to well above the four beds per 1,000 established by the Federal Health Planning Program as a standard. Furthermore, the variation in each state is so great that the number of beds per capita in the state or county as a whole (a measure often employed by health-planning agencies) bears little relation to the conditions prevailing in each community. The average amount spent on treatment in hospitals in 1975 ranged from less than $100 per capita to more than $300 in the 66 areas.



"SURGICAL SIGNATURE" of a hospital area reflects the specialties and preferences of the surgeons who practice in the area. Each group of six bars shows the rates of five common surgical procedures and the total surgery rate for one of the five most populous hospital areas in Maine. The rates are expressed in relation to the state average. The total surgery rate in each area is closely correlated with the number of surgeons there. The rates of individual procedures, however, are not determined by the number of surgeons. Areas 2 and 3 have the same total rate, but their signatures are quite different. In area 2 hysterectomy is the commonest procedure; in area 3 it is the least common. A signature is generally consistent over many years.

ria for selecting patients for the operation are not definitive. Tonsillectomy, for example, was once done almost routinely for minor inflammation but is now usually reserved for more serious cases. Some practitioners, however, retain the older attitude. In the case of hysterectomy there is general agreement on its necessity in uterine cancer. The operation is most often done, however, for a variety of less threatening conditions; the appropriateness of the procedure in these circumstances has been widely questioned.

In contrast, the procedures whose rates vary little are those that provoke little disagreement. Inguinal hernia, for example, is easily recognized; the treatment of choice is surgical repair. Only where the simultaneous presence of other conditions makes surgery dangerous is any other therapy employed, at least in the U.S. Consequently the rate of surgical repair of hernia is relatively constant in the hospital areas.

The uncertainty of the medical profession about the controversial procedures can be great indeed. Ira M. Rutkow and George D. Zuidema of the Johns Hopkins University School of Hygiene and Public Health and one of us (Gittelsohn) recently surveyed a group of randomly selected surgical specialists. Each surgeon was given a set of fabricated case histories and asked whether he would recommend a particular surgical procedure for them. There was a marked divergence of opinion. For the three fictitious cases related to hysterectomy 25 percent of the surgeons thought none of the cases warranted surgery; 5 percent thought all three did. The remaining 70 percent recommended surgery in one case or two cases. Similar inconsistencies appeared for breast surgery, varicose-vein surgery, tonsillectomy, gallbladder removal, cataract surgery and prostatectomy.

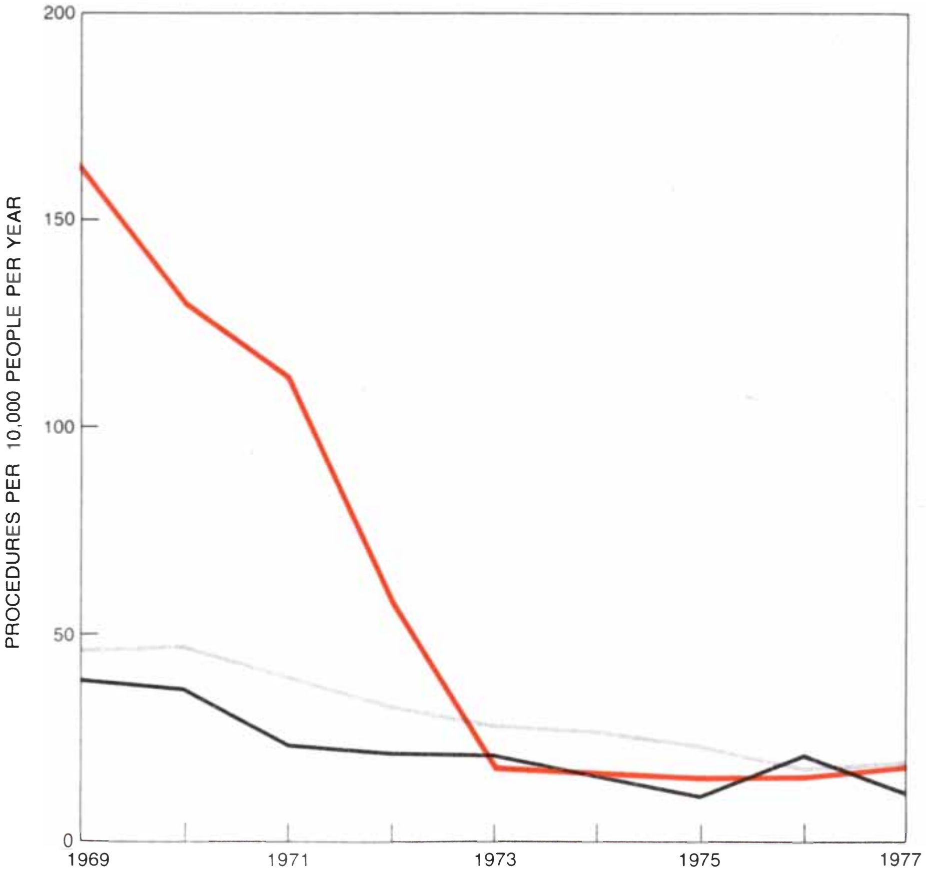
Earlier studies had also demonstrated extreme conflicts of opinion. In 1934 workers from the American Child Health Association chose 1,000 schoolchildren to be examined by physicians, who were to determine whether or not they should have their tonsils removed. Six hundred children had already had the procedure and were removed from the sample. The remaining 400 were examined by school physicians, who recommended that 45 percent have a tonsillectomy. Those that remained after the first round of examinations were examined by another group of physicians who recommended that 46 percent of them have their tonsils out. A third examination by still another group of physicians led to 44 percent of the remainder having tonsillectomy recommended. After three successive rounds only 65 of the original 1,000 children had not had tonsillectomy recommended for them.

For many common illnesses well-designed clinical studies to test alternative forms of therapy have not been done. For this reason there is conflicting information on whether a particular procedure will improve a patient's health or the quality of his life. Many diagnostic and therapeutic techniques are adopted or discarded on the basis of fashion or a physician's personal experience rather than on more reliable grounds.

In the absence of authoritative standards differences among physicians in perceptions of illness and preferences for treatment appear to be the cause of much variation in rates of surgery and other kinds of treatment. The variation is perpetuated by regulatory agencies and providers of medical insurance. In determining whether new hospital facilities should be constructed, for example, regulatory agencies often rely on the occupancy rate of local hospitals and the average cost of treating a particular kind of case there. Such measures of economic efficiency are not strongly correlated with the per capita consumption of medical services in the community or with the amount spent per capita on medical services. Utilization of narrow economic criteria may lead to an increase in medical care in areas where rates are already well above the average.

Our colleagues Richard J. Greene, Harvey M. Sapolsky and Drew Altman of the Massachusetts Institute of Technology have examined all decisions by regulatory agencies concerning the building of new hospital facilities in Maine, Rhode Island and Vermont from 1975 through 1978. Hospitals in areas where the ratio of hospital beds to population was above the standard of four per 1,000 were just as likely to seek permission to add beds as those in areas where the ratio was below the standard. Moreover, the likelihood that health-planning officials would agree that new beds were needed had no relation to the ratio.

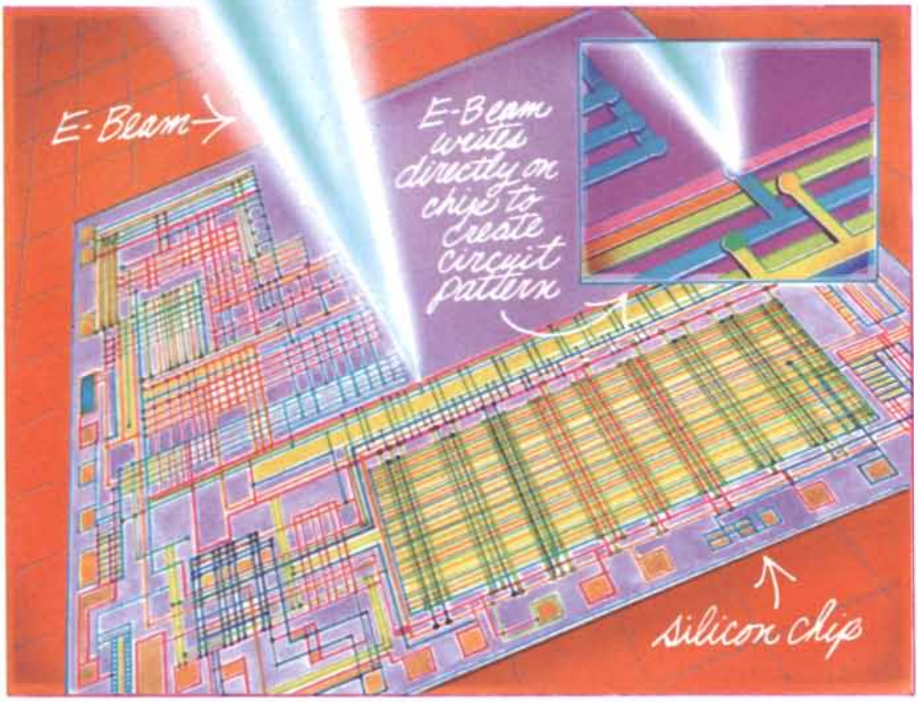
The cause of such irrational decisions is the tendency to consider a hospital in isolation from the level of medical care in the surrounding community. An extreme example is a medium-size facility in Maine that cares for acutely ill patients. In 1975 hospital officials applied for permission to add 11 beds. The officials stated that more physicians were being recruited to the area and that their patients would need the additional facilities. Because of the recruitment of



TONSILLECTOMY RATE in Vermont demonstrates the strong influence of physicians' preferences on how often a surgical procedure is done. The colored line is for the hospital area with the highest tonsillectomy rate in the state. The gray line shows the state average. The solid black line is for the area with the lowest rate in the state. In 1971 physicians in the area with the highest rate were told of that fact. They established the practice of obtaining a second opinion before a tonsillectomy was done, and the rate fell to a level roughly equal to the lowest in the state. (The decline before 1971 followed a single physician's leaving the hospital area.)


You deserve National attention.

OUR RESEARCH IN E-BEAM LITHOGRAPHY WILL MAKE OUR 1.25 MICRON VLSI CHIP LOOK LIKE HORSE AND BUGGY TECHNOLOGY.



WE'RE USING E-BEAMS TO DEVELOP VERY LARGE SCALE INTEGRATED CIRCUITS WITH SUB-MICRON FEATURE SIZE.

At Honeywell, solid state electronics is a key technology. We've already built one of the largest in-house semiconductor operations in the United States. But we're fast expanding our current capabilities to meet the growing demand for integrated circuits with extremely small feature size. Recently, our research in very large scale integrated circuits helped us win a contract for Phase 1 of the Department of Defense's VHSIC program.

In Phase O of VHSIC, we produced a bipolar microcircuit with 1.25 micron feature size and more than 7000 gates on a 200 X 250 mil chip. But that's only the first step. We're now pushing optical lithography to its limits, fabricating high-density, near-micron microcircuits. With the help of our advanced design automation system, we are developing a bipolar chip with 30,000 gates and a CMOS chip with 300,000 transistors.

At the same time, we're developing CMOS and bipolar microcircuits with 0.5 micron feature size. But, the micron barrier will not be broken with conventional optical lithography. Sub-micron technology will require significant advances in four major fields: electron-beam lithography, dry etching, ion implantation and low temperature processing.

For the last six years, Honeywell has been conducting e-beam research. Currently, computer controlled e-beam equipment is used to print the masks that are used in near-micron circuit printing. In August, 1980, we began using an e-beam system for direct writing of very large scale integrated circuits in the development laboratory at our Solid State Electronics Division.

One of the problems in e-beam direct writing is the proximity effects from high-energy electrons colliding with the atoms of the resist. At our Corporate Technology Center, we're working on an improved polymer resist that will not degrade during direct write.

Dry etching techniques, like plasma etching and reactive ion milling, are another focus of Honeywell research. Because dry etching techniques work vertically, they are preferred to wet etching, which tends to undercut the layers of the resist. At Honeywell, we're now in the process of designing our own reactive ion milling equipment.

We've taken tremendous strides in our VLSI capabilities. But there are still many possibilities in sub-micron chip technology yet to be explored.

If you are interested in learning more about Honeywell's research and development of VLSI, you are invited to correspond with Dr. J. M. Daughton, Vice President, Solid State Development Center. If you have an advanced degree and are interested in a career in systems analysis, solid state electronics, sensors, design automation, or material sciences, please write to Dr. K. C. Nomura, Vice President and General Manager, Solid State Electronics Division. Both may be reached at this address: Honeywell, 12001 State Highway 55, Minneapolis, MN 55441.

Honeywell

Although Honeywell engineering is world-wide, the bulk of corporately-funded research and applied research is done in Minneapolis. The most recent Quality of Life Study conducted by Midwest Research Institute shows Minnesota to be one of the best places in the country to live and work considering cultural, social, economic, educational and political factors.

This ideal environment is further enhanced by Honeywell's affiliations with universities across the country. We have an ongoing program of seminars with Berkeley, MIT, Stanford, Carnegie Mellon, the University of Illinois, Cornell, Purdue, Oregon Graduate Center, and the University of Minnesota.

physicians the number of admissions to the hospital had already begun to rise. The occupancy rate was then 75 percent (close to the recommended standard). It was estimated that the recruitment of more physicians would raise the occupancy rate above the acceptable level.

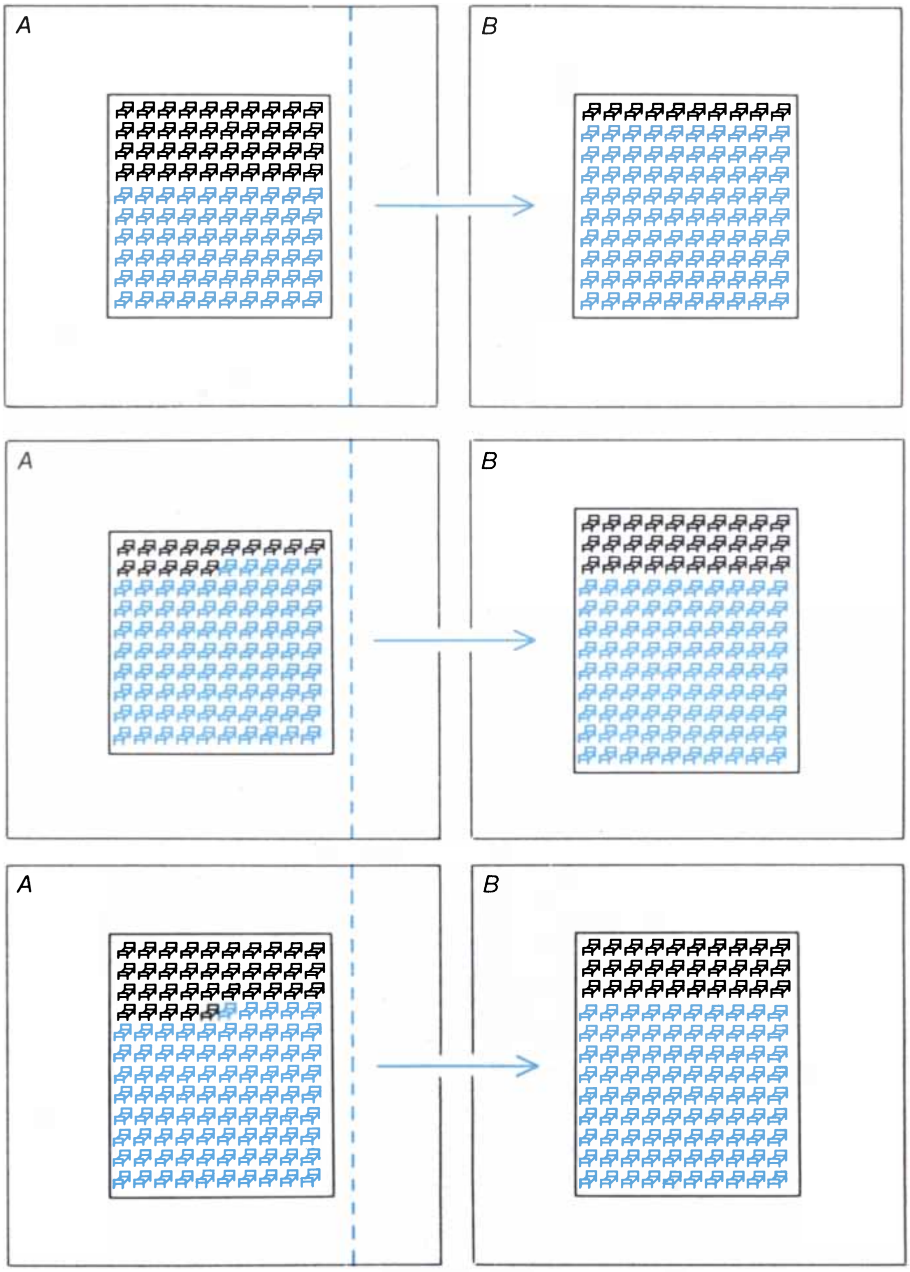
At the time the application for new beds was made the area from which most of the hospital's patients came had 6.3 beds per 1,000 people; 80 percent of the beds were in the hospital whose management had made the request. With the addition of 11 beds the ratio would be 7.0 per 1,000, which is 75 percent higher than the Federal standard. Furthermore, the rate at which people in the area were admitted to hospitals was 246 per 1,000 per year, the second-highest rate in the state and 50 percent higher than the state average. The total surgical rate was one-third higher than the state average. The per capita rates of four of the six common surgical procedures were more than double the state average and were increasing as the supply of physicians grew. These facts were not considered by the state health-planning officials, who approved the application for new beds. This case is one of several we have seen in which regulatory decisions contributed to disparities among small areas without evidence of a need for more medical care.

Paradoxically, a low hospital-occupancy rate may also contribute to increases in medical care in areas with hospitalization rates that are well above average. Low occupancy rates are taken to be a sign of inefficiency. A hospital with a rate of less than 75 percent often comes to the attention of regulatory agencies and health planners. Hospital managers in such a situation fear they will be compelled to reduce the number of beds in the hospital; a common response is to recruit additional physicians. Such decisions are generally made without considering the need for the physicians or the need for the services the physicians are likely to provide.

The measures that provide the only direct estimate of the level of health care in a community are population-based: the number of hospital admissions per capita, the number of occupied beds per capita and the amount spent per capita on hospital care. The last two measures give the best estimate of the consumption of medical services in a community. They depend much more strongly on the per capita admission rate than on the average length of stay in local hospitals or the average cost of a stay. The narrow measures of efficiency, however, are the ones more commonly employed by health planners. The population-based rates can be derived from studies of small areas. Their employment would make analysis of variation an explicit consideration in regulatory decisions, and it could help to reduce some of the wide variation in services among hospital areas.

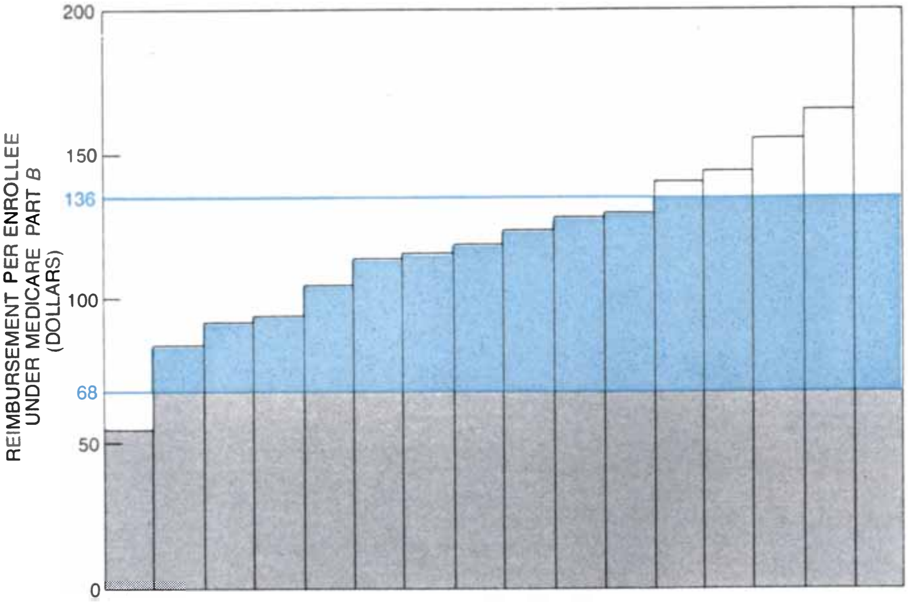
The policies of agencies that provide medical insurance also contribute to disparities in medical care. As a result of their policies residents of areas where rates of consumption are low subsidize residents of areas where rates are high. Again, the reason for the imbalance is that the local level of medical care is overlooked. The contributions to insurance pools (in the form of private insurance premiums or taxes) are determined by rates of utilization of medical care averaged over large regions. In areas where per capita expenditures for medical care are lower than the average, part of the premium goes to pay for services received by people in the areas with high rates of consumption.

We examined Medicare contributions and reimbursements in 16 Vermont hospital areas. The contributions of those enrolled in the Medicare Part *B* pro-

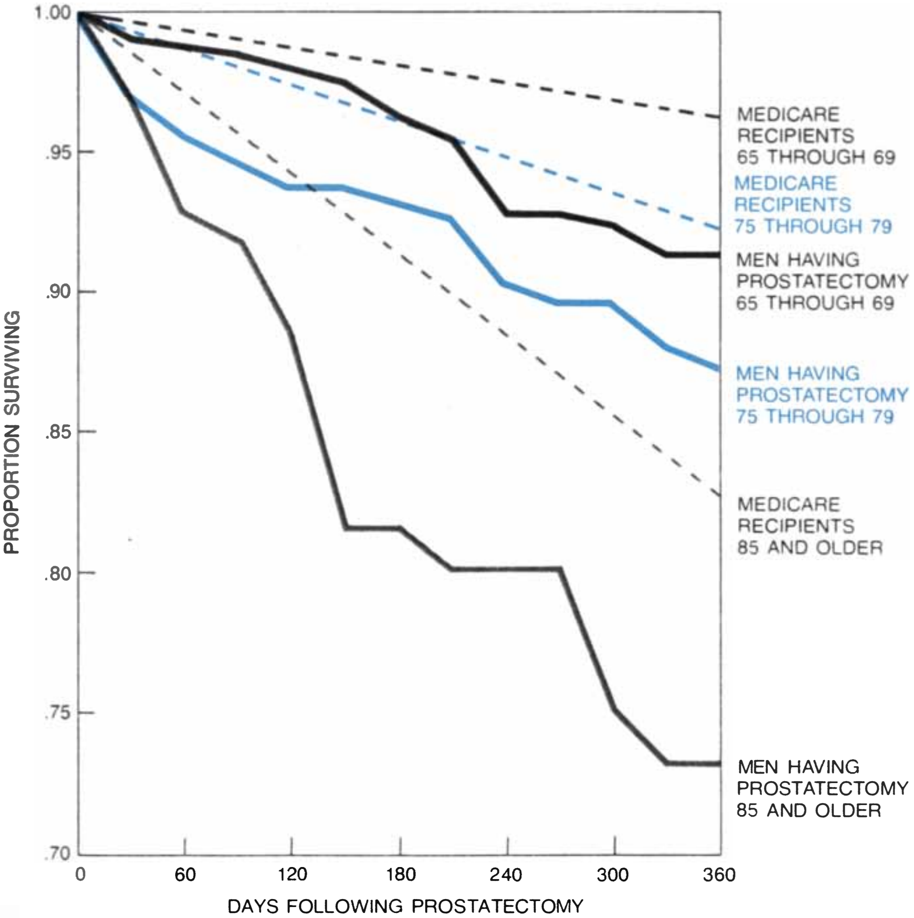


OCCUPANCY RATE of local hospitals has little relation to the per capita level of hospital treatment in a community. Regulatory agencies and hospital administrations that utilize the occupancy rate as a measure of the efficiency of the hospital contribute unwittingly to excessive medical care. The illustration shows two hospital areas, each of which has 25,000 residents and a 100-bed hospital. Each area thus has the ideal ratio of beds to people (four per 1,000) set by Federal agencies. One-fifth of the hospitalized patients from area *A* go to the hospital in area *B*, which has an occupancy rate of 90 percent, above the recommended level. The management of that hospital asks permission to add 20 beds, a request granted because of the high occupancy rate. At the hospital in area *A* the occupancy rate is 60 percent, often taken as a sign of inefficiency. Hospital officials recruit physicians, whose patients raise the rate to 85 percent. Application is then made to add 20 beds and is approved. As a result the two areas now have a total of 240 hospital beds, or almost five beds per 1,000 people, well above the standard.

Phil and Steve Mahre, World Cup Race at Aspen



POLICIES OF INSURERS that do not take into account local variations in medical practice effectively subsidize treatment in areas with a high rate of care. The illustration shows the average reimbursement per enrollee in Medicare Part *B* in 16 Vermont hospital areas in 1972. Enrollees paid $68 each; the amount was matched by the Federal Government. In one area the average reimbursement was less than $68; residents of that area received less than they had paid out to the Medicare program. In 10 areas reimbursement was between $68 and $136. In five areas an average of more than $136 was paid to each enrollee. The medical care of enrollees in those five areas was thus subsidized by Medicare recipients and taxpayers in other areas.



COST IN LIVES of high rates of surgery may be considerable. Of the procedures for which the authors found substantial geographic variation prostatectomy is the riskiest. The graph compares the proportion of men who have had a prostatectomy (for a reason other than cancer of the bladder or the prostate gland) who survive for a year after the procedure with the proportion of all male Medicare recipients of the same age who survive for a year. In the three age groups studied the risk of death is much higher among those who have had a prostatectomy; moreover, the added risk of death persists throughout the year after the procedure.

gram are uniform: in 1972 they were $68 per capita. The average reimbursement per capita in the U.S. that year was $136; Federal taxes paid the $68 difference. In the area where reimbursement was lowest, the average amount received was less than $68 per person enrolled; Medicare recipients in the area were thus receiving less in services than they had paid in premiums. In 10 areas the average recipient got more than $68 but less than $136. In five areas recipients were getting more than $136. In those communities the physicians who treated Medicare recipients and the hospitals in which the care was provided were in effect receiving contributions from enrollees and taxpayers in regions where less medical treatment was given.

When all forms of insurance are taken into account, the value of such subsidies can be large even in small communities. We examined two neighboring hospital areas in Vermont. Both areas have roughly 11,000 residents, distributed about equally by age and sex. Surveys of households in the two areas showed that residents had similar rates of illness and amounts of medical insurance, and they visited their physicians about as often. They had, however, quite different rates of admission to hospitals. We calculated how much would have been spent on hospital care by residents of the two areas if their rates of hospitalization had been the average for the state. In one area $3 million more than the average had been spent over a 15-year period. In the other area $4 million less than the average had been spent. Most of the excess $3 million in the area where the rate was high was spent at the local hospital, benefitting the local economy. In the area where the rate was low, on the other hand, the savings were not realized in the local economy. Much of the difference was transferred to other communities (including the neighboring one) in the form of insurance premiums and taxes.

Such subsidies might be more troubling to residents of areas with low rates of hospitalization if they realized that the need for services is no greater in areas where rates are high. Moreover, the imbalance may not be beneficial to the recipients of the greater medical care. Without controlled clinical trials it is not possible to determine conclusively which populations are getting better medical care. Nevertheless, it is clear that more medical treatment is not necessarily better treatment.

Some of the procedures whose rates vary widely are quite risky. Of those we considered the riskiest is prostatectomy. Some 1.3 percent of patients who have the procedure die in the hospital. To get an idea of the risks associated with variations in medical care we calculated how many people would have died in the U.S. in 1975 from prostatec-

MERRILL LYNCH INTRODUCES FLEXPLAN III

If you're concerned over the long range financial security of yourself and your family, it's something you should consider immediately.

Flexplan III, the new tax-deferred annuity developed by the University Life Insurance Company of America, a subsidiary of the Baldwin-United Corporation. It is available exclusively through Merrill Lynch Life Agency.

FLEXPLAN III FEATURES CURRENT YIELDS OF 14.25%* TO HELP FIGHT INFLATION, AND TAX DEFERMENT TO NULLIFY TODAY'S TAX BITE.

Your money has a golden opportunity to earn today's high interest rates, and multiply faster because of tax-deferred compounding.

That means the money you would normally pay in income taxes remains in your account earning more interest every year. Your principal earns interest. And your interest earns interest!

For the first year of your annuity, your interest rate is 14.25% guaranteed. For the next 9 years, you are guaranteed it will never be lower than 7.5%, 4% thereafter. In addition, if you fail to receive at least 13.50% annually, you may withdraw your principal and interest without surrender charge.

Suppose you need money for an emergency? You can make one withdrawal of up to 10% of the accumulated value per year without charge

A very important feature of Flexplan III: the 5% surrender charge disappears after six years

You also gain tax benefits when you start receiving your annuity income, because you will presumably be in a lower tax bracket. And, part of this income is tax free, as a return of principal. Annuity contracts can also simplify probate.

You decide when you want to start receiving income, and how you'd like the checks to come — all at once, in monthly payments for the rest of your life, over a short period of time, or a variety of other ways.

The choice is yours.

For complete information, mail the coupon, or call toll free Monday through Friday, 8:30 a.m. to 12 midnight. Eastern Time, 1-800-526-3030, ext. 504. In New Jersey, call 1-800-742-2900, ext 504.

14.25%
(TAX-DEFERRED)

A GUARANTEED SAFE WAY TO BUILD THE RETIREMENT INCOME YOU NEED.

Mail to: Merrill Lynch, P.O. 2021, Jersey City, NJ 07303
Yes, I would like complete information, without charge or obligation, on how I can build a bigger nest egg with the new tax-deferred annuity, Flexplan III.

Name______________________________

Address____________________________

City____________ State____________ Zip________

Business Phone____________ Home Phone____________
Merrill Lynch customers, please give name and office address of Account Executive:
____

Merrill Lynch
Merrill Lynch Life Agency Inc.
A breed apart.

504

© 1981 Merrill Lynch Life Agency Inc.

*Based on rates effective as of Jan. 27, 1982. Current rates will be provided upon request.

© 1982 SCIENTIFIC AMERICAN, INC

tomy if the lowest rate in Maine had been the national average and if the highest rate had been the national average. Under the low rate there would have been 1,900 deaths from prostatectomy; under the high rate there would have been 6,800 deaths. The difference of almost 5,000 deaths suggests the possible cost of geographic variations in health care. Under the high Maine rate one man in 100 who had reached the age of 65 would die in the hospital following a prostatectomy.

To the person who must choose whether to have a prostatectomy the most important consideration is the balance of risks and benefits from the procedure. In evaluating the balance claims information from Medicare is useful; it enables us to follow for long periods patients who have had a prostatectomy. Their chances of surviving can then be compared with the chances of those who have not had the operation. We examined records of about 1,500 men who had had a prostatectomy in Maine in 1976 and 1977. On the basis of this sample we estimate that a 67-year-old man who has a prostatectomy for a condition other than cancer of the prostate or bladder has about a 9 percent chance of dying in the year after the procedure. For men in this age group who have not had the procedure the risk of dying within a year is about 3.5 percent.

The savings in lives and money that would result from making rates of medical care correspond to the health needs and preferences of informed consumers might be considerable. How could this be achieved? If physicians in an area where rates are high are made aware of that fact, rates may fall, eliminating some unnecessary surgery. Reliable studies of the effects of various kinds of treatment might lead to a consensus on their value and could provide consumers with more information on which to base their decisions. If regulatory agencies consider the level of health care in the community, some of the tendencies toward excess may be restrained. If premiums for medical insurance reflect the amount of treatment per capita in a given area, subsidization could be reduced.

The most important factor, however, may be the emergence of an informed consumer of medical services. When patients are aware that different forms of treatment are available, they can demand information on risks and benefits and make their own preferences known. If they know that rates of surgery are high at the local hospital, they may choose another. If they realize that a particular operation is a controversial one, they may seek the opinion of a second and even a third physician. Informed patients may therefore be the most important factor in making rates of treatment reflect health needs and eliminating unnecessary medicine.

Choose one of these six sets and save up to 282^{50}.

You simply agree to buy 4 books within the next two years.

The Compact Edition of The Oxford English Dictionary

The "most complete, most scholarly dictionary of the English language"—*The Christian Science Monitor.* Through photo reduction, the original 13-volume set has been reproduced in this two-volume *Compact Edition.* Magnifying glass included.

$125.00	Publisher's price
-$105.05	Your savings
$ 19.95	Book-of-the-Month Club price

The Random House Library of Painting & Sculpture

Your own guided tour of the world's greatest painting and sculpture in four lavishly illustrated, slipcased volumes: Vol. I, *Understanding Art,* Vols. II and III, *The History of Art,* Vol. IV, *Dictionary of Artists and Art Terms.*

$100.00	Publisher's prices total
-$ 85.00	Your savings
$ 15.00	Book-of-the-Month Club price

Jefferson and His Time by Dumas Malone

This landmark six-volume work is the definitive portrait of the most complex, elusive and fascinating of our presidents. "A masterly achievement of scholarship, the finest biography of Jefferson we have or are likely to have and a monumental triumph of the senior American historian"—*The New York Times Book Review.*

$109.90	Publisher's prices total
-$ 94.90	Your savings
$ 15.00	Book-of-the-Month Club price

The Story of Civilization by Will and Ariel Durant

For almost half a century Will and Ariel Durant traced the continuity of world history—the religions and philosophies, the political and economic tides, the arts and sciences, the customs and conquests—to show the foundations of society today. The Durants' 11-volume illustrated masterwork is history come alive in all its dimensions. The enormous scope of the work covers ancient and recent civilization, including Oriental as well as Western history. A magnificent centerpiece for any home library.

$307.45	Publisher's prices total
-$282.50	Your savings
$ 24.95	Book-of-the-Month Club price

The Encyclopedia of Philosophy

The most comprehensive encyclopedia of philosophy ever published, this set traces ancient, medieval, modern, Eastern and Western thought. An essential and rewarding reference source for home libraries.

$150.00	Publisher's price
-$130.05	Your savings
$ 19.95	Book-of-the-Month Club price

Benefits of Membership

Membership in the Book-of-the-Month Club begins with your choice of the extraordinary works offered here. Because our prices are generally lower than the publishers', you will save on works such as these throughout your membership as well as on the finest new fiction and nonfiction titles. In fact, the longer you remain a member, the greater your savings can be. Our Book-Dividend® plan, for which you become eligible after a brief trial enrollment, offers at least 70% off the publisher's price on art books, reference works, classics, books on cooking and crafts, literary sets and other contemporary works of enduring value. Nevertheless, all Book-of-the-Month Club books are equal in quality to the publishers' originals; they are not condensed versions or cheaply made reprints.

BOOK-OF-THE-MONTH CLUB®
America's Bookstore® since 1926.

Available in 13 records or 8 cassettes.

The Nine Beethoven Symphonies & The Five Piano Concertos/Ashkenazy

A special offering of two of the most acclaimed recordings of Beethoven's works. Sir Georg Solti conducts the Chicago Symphony in the nine-record set of Beethoven's complete symphonies, hailed by *Time* as a "glorious musical combustion." In collaboration with soloist Vladimir Ashkenazy, Solti and the Chicago Symphony perform Beethoven's five concertos. *High Fidelity* calls it "a major addition to the Beethoven discography."

$115.74	List prices total
-$ 99.79	Your savings
$ 15.95	Book-of-the-Month Club price

Book-of-the-Month Club, Inc., Camp Hill, Pennsylvania 17012

The Short Life of Évariste Galois

Legend has it that the young mathematician wrote down group theory the night before he was fatally shot in a duel. More careful investigation shows that Galois's remarkable ideas took somewhat longer to mature

by Tony Rothman

In the early morning hours of May 30, 1832, the French mathematical prodigy Évariste Galois, who was then 20 years old, wrote to his friends Napoleon Lebon and V. Delauney:

"I have been provoked by two patriots.... It is impossible for me to refuse. I beg for your forgiveness for not having told you. But my adversaries have put me on my word of honor not to inform any patriot. Your task is simple: prove that I am fighting against my will, having exhausted all possible means of reconciliation; say whether I am capable of lying even in the most trivial matters. Please remember me since fate did not give me enough of a life to be remembered by my country.

I die your friend,
É. Galois"

On the same night Galois also wrote to his friend Auguste Chevalier:

"I have made some new discoveries in analysis. The first concerns the theory of equations, the others integral functions.

"In the theory of equations I have investigated the conditions for the solvability of equations by radicals; this has given me the occasion to deepen this theory and describe all the transformations possible on an equation even though it is not solvable by radicals. All of this will be found here in three memoirs....

"Make a public request of [Carl Gustav Jacob] Jacobi or [Carl Friedrich] Gauss to give their opinions not as to the truth but as to the importance of these theorems. After that, I hope some men will find it profitable to sort out this mess."

Galois's desperate state during the writing of these letters was fully warranted in view of the subsequent events. Shortly after sunrise on the morning he completed the letters he left his room at the pension Sieur Faultrier in Paris and confronted a political activist named Pescheux d'Herbinville in a duel of honor on the banks of a nearby pond. There Galois was shot in the abdomen and abandoned. A passerby found him and he was taken to the Hôpital Cochin, where he died the next day. Fourteen years later the manuscripts he had left behind for Chevalier were published by the French mathematician Joseph Liouville, and the extraordinarily fecund branch of mathematics called group theory was born.

Few tales in the history of science can equal the high romance of the known facts about the life and death of Galois. Yet because the facts of the story are so compelling it is easy to read too much into Galois's letters, and it is tempting to sift through the events that led up to the duel for an explanatory thread that can match the melodrama apparent in his writings.

It is known, for example, that at age 17 Galois was instrumental in creating a branch of mathematics that now provides insights into such diverse areas as arithmetic, crystallography, particle physics and the attainable positions of Rubik's cube. It is also a matter of record that at the same age Galois failed for a second time the mathematics examination for admission to the École Polytechnique. He studied instead at the École Normale in Paris, but by the time he was 19 he had been expelled from the school and twice arrested and imprisoned for his political activities. Shortly before the duel he had become involved in an unhappy love affair, which in one of his last letters he seemed to link with the duel itself. "I die," he wrote, "the victim of an infamous coquette and her two dupes."

Unfortunately several of Galois's 20th-century biographers have not resisted the temptation to arrange, interpret and embellish such facts. The story of Galois known to most people today is derived from popular accounts, such as those by the physicist Leopold Infeld and the astronomer Fred Hoyle. The most influential version of the story has been that of Eric Temple Bell, the mathematician whose 1937 book *Men of Mathematics* is probably the best-known introduction to the lives of great mathematicians.

In the popular retellings of the tale Galois is presented as a misunderstood genius, oppressed by the stupidity of his teachers, ignored by the mathematical establishment and goaded by the events of the times into political activities that squandered his energies and eventually cost him his life. Most remarkable of all, according to these accounts, is that throughout the political turmoil and even during his imprisonment Galois continued to develop his mathematical ideas in his head and finally wrote them down the night before the duel. Bell's description of the final night is worth quoting because it has probably given the greatest impetus to the Galois myth:

"All night long he had spent the fleeting hours feverishly dashing off his scientific last will and testament, writing against time to glean a few of the great things in his teeming mind before the death which he saw could overtake him. Time after time he broke off to scribble in the margin 'I have not time; I have not time,' and passed on to the next frantically scrawled outline. What he wrote in those last desperate hours before the dawn will keep generations of mathematicians busy for hundreds of years."

Recently, with the help of Marc Henneaux and Cecile DeWitt-Morrette of the University of Texas at Austin, I have studied some of Galois's writings and the later scholarly work on his life. Although it is clear from these materials that the major events in Galois's life have been known for some time, the reconstructions by Bell and others reveal more about the stereotypes of scientific genius that appeal to the popular imagination than they do about Galois. The true romance of Évariste Galois is a fascinating story in its own right, and it bears telling on the 150th anniversary of his death.

Apart from letters, official records and other contemporary documents the principal source on the life of Galois is an 1896 biography by Paul Dupuy, a historian and the general superintendent of the École Normale, the college Galois had attended 66 years earlier. According to Dupuy, Galois was born on October 25, 1811, in Bourg-la-Reine, a suburb of Paris. His father, Nicholas-Gabriel Galois, supported Napoleon and headed the town's liberal party; he was elected mayor of Bourg-la-Reine in 1815 during the Hundred Days, Napoleon's first return from exile.

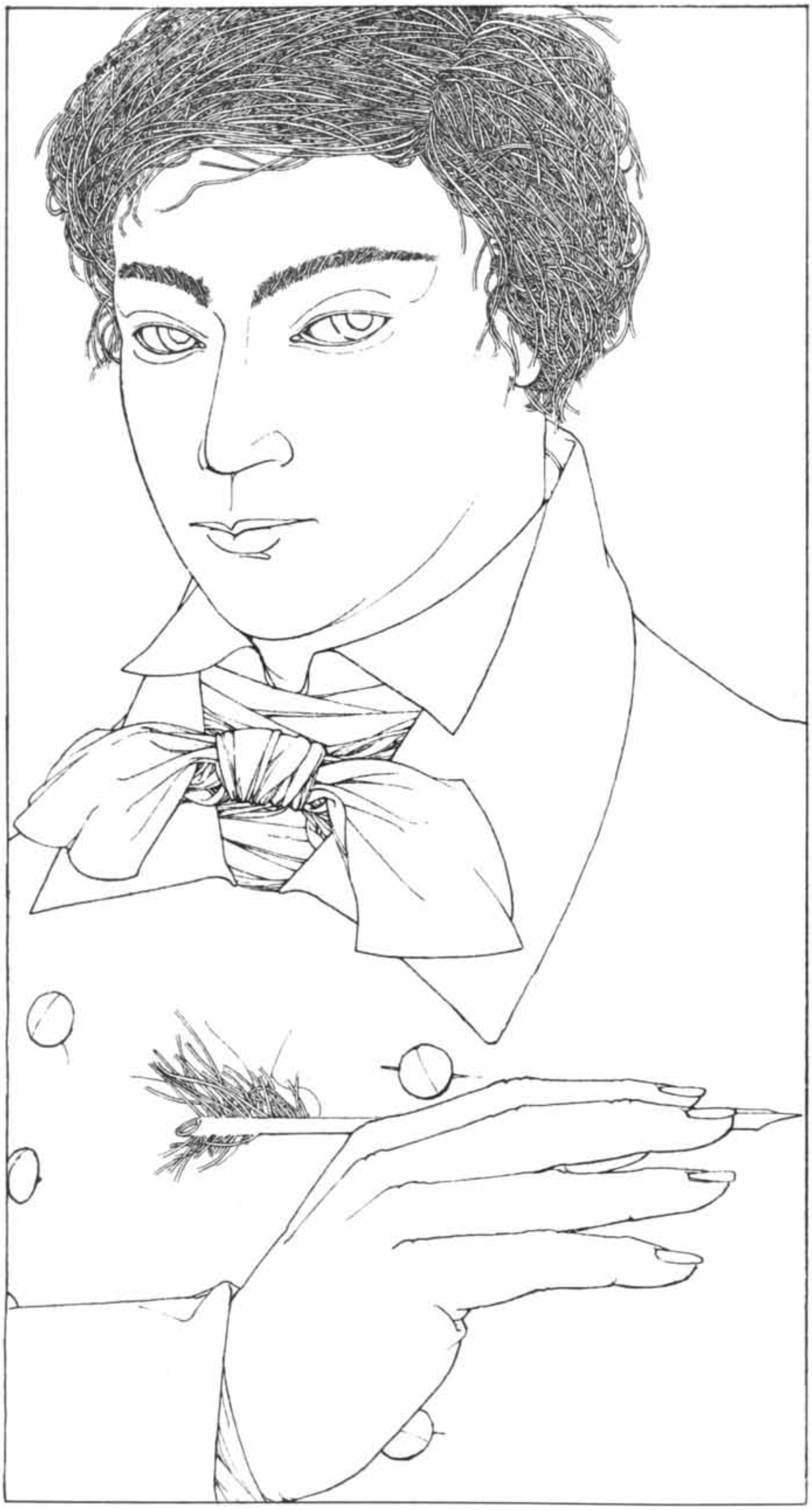
For the first 12 years of his life Évariste was educated by his mother, Adeläide-Marie Demante Galois. She gave her son a solid background in Greek and Latin, and she passed on to him her skepticism toward established religion. It is unlikely, however, that the young Galois was exposed to mathematics in any more detail than the usual lessons in arithmetic; a mathematics education was not considered particularly important at the time. There is no record of previous mathematical talent on either side of the family.

Galois's formal education began in 1823, when he was enrolled in the Collège Royal de Louis-le-Grand, the Paris preparatory school that was the alma mater of Robespierre and Victor Hugo (and is still operating today). At Louis-le-Grand, Galois immediately began to develop his political sensibilities. His liberal, or antiroyalist, sympathies, acquired from his parents, were in accord with the political opinions of most of the other students.

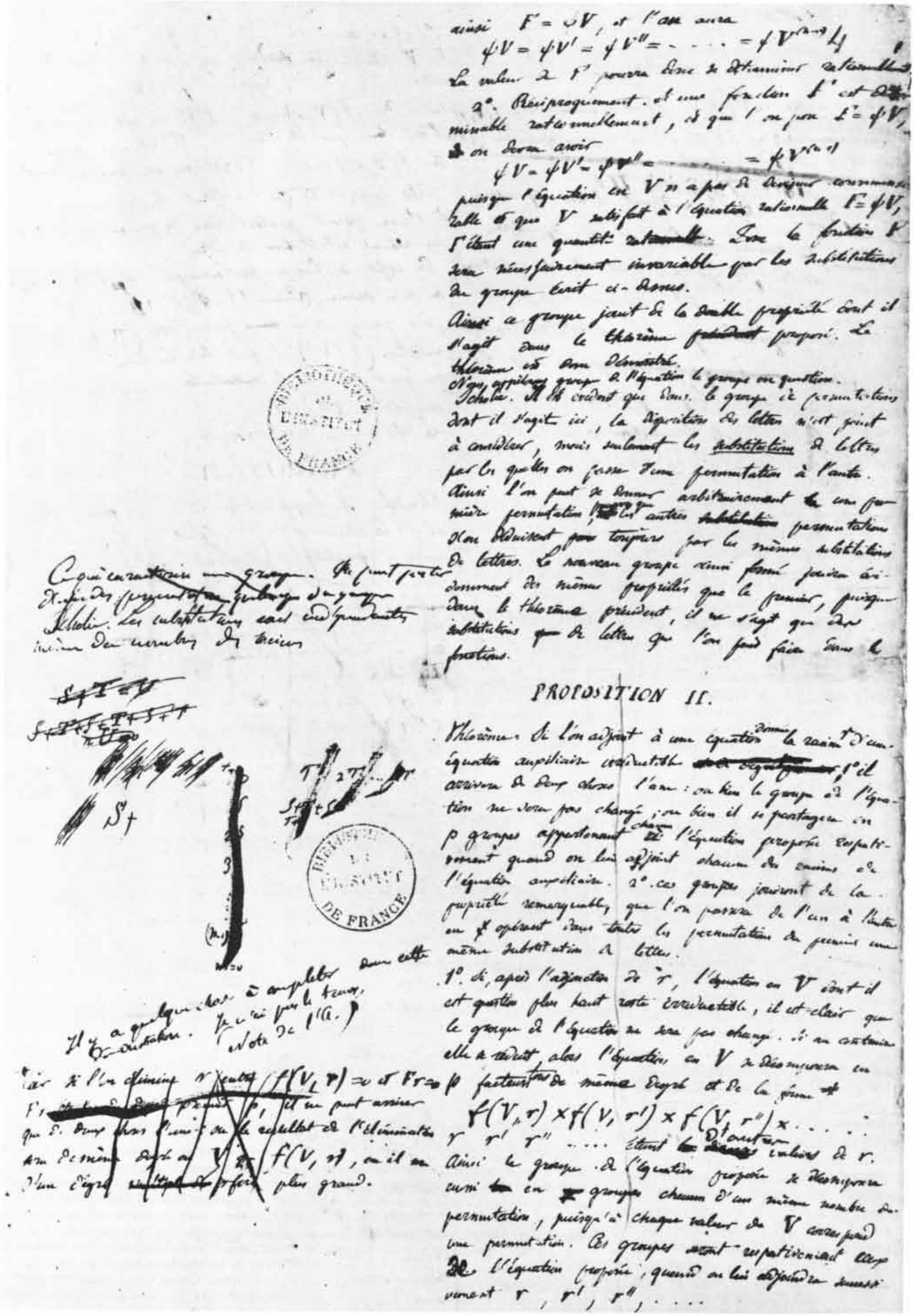
During Galois's first term, however, relations between the students and a newly appointed headmaster of the school were badly strained. The students suspected the headmaster of planning to return the school to Jesuit administration; the Jesuits were leaders of the right-wing backlash that followed the Napoleonic era. The students staged a minor rebellion: they refused to sing at chapel, to recite in class or to toast Louis XVIII at a school banquet. The headmaster summarily expelled 40 students he suspected of leading the insurrection. Although Galois was not expelled (and it is not known whether he participated in the uprising), the arbitrary action of the headmaster undoubtedly helped to foster Galois's distrust of authority.

There is little evidence that Galois was a poor student or that his intellectual growth was stunted by inferior teachers at Louis-le-Grand, as the popular accounts would have it. In his first few years he won several prizes in Greek and Latin and half a dozen honorable mentions. The historian of science René Taton calls his progress brilliant. Nevertheless, during Galois's third year his work in rhetoric was inadequate and he had to repeat the year. Contrary to Bell's statement that Galois's poor work in rhetoric was a result of his preoccupation with algebra, it was only after this setback that Galois enrolled in his first course in mathematics. He was then 15.

The course, taught by Hippolyte Jean Vernier, awakened Galois's genius for mathematics. He raced through the usu-



DRAWING OF GALOIS by David A. Johnson depicts the mathematician at the age of 17, while he was a student at the Collège Royal de Louis-le-Grand. At the time Galois had studied mathematics for only two years, yet he had already published a paper on continued fractions and had begun the investigations into the theory of equations that led him to consider an abstract algebraic theory of sets of objects he called groups. Credit for the development of group theory must also be given to several other mathematicians of the late 18th and early 19th centuries, notably Paulo Ruffini, Neils Henrik Abel and Joseph Louis Lagrange. The distinction of founding group theory, however, is usually accorded to Galois. Johnson's drawing is based on the two known renderings of Galois. One was done when Galois was 15 and the other was completed from memory by Galois's brother Alfred in 1848, 16 years after Évariste's death.



ainsi $F = \varphi V$, et l'on aura

$$\varphi V = \varphi V' = \varphi V'' = \ldots\ldots = \varphi V^{(n-1)}$$

La valeur de F pourra donc se déterminer rationnellement.

2°. Réciproquement. Si une fonction F est déterminable rationnellement, et que l'on pose $F = \varphi V$, on devra avoir

$$\varphi V = \varphi V' = \varphi V'' = \ldots\ldots = \varphi V^{(n-1)}$$

puisque l'équation en V n'a pas de diviseur commensurable, et que V satisfait à l'équation rationnelle $F = \varphi V$, F étant une quantité ~~rationnelle~~. Donc la fonction F sera nécessairement invariable par les substitutions du groupe écrit ci-dessus.

Ainsi ce groupe jouit de la double propriété dont il s'agit dans le théorème ~~précédent~~ proposé. Le théorème est donc démontré.

Nous appellerons groupe de l'équation le groupe en question.

Scholie. Il est évident que dans le groupe de permutations dont il s'agit ici, la disposition des lettres n'est point à considérer, mais seulement les substitutions de lettres par lesquelles on passe d'une permutation à l'autre. Ainsi l'on peut se donner arbitrairement une première permutation, pourvu que les autres ~~substitution~~ permutations s'en déduisent toujours par les mêmes substitutions de lettres. Le nouveau groupe ainsi formé jouira évidemment des mêmes propriétés que le premier, puisque dans le théorème précédent, il ne s'agit que des substitutions que l'on peut faire dans les fonctions.

PROPOSITION II.

Théorème. Si l'on adjoint à une équation donnée la racine r d'une équation auxiliaire irréductible ~~[illegible]~~, 1° il arrivera de deux choses l'une : ou bien le groupe de l'équation ne sera pas changé ; ou bien il se partagera en p groupes appartenant chacun à l'équation proposée respectivement quand on lui adjoint chacune des racines de l'équation auxiliaire. 2°. ces groupes jouiront de la propriété remarquable, que l'on passera de l'un à l'autre en opérant dans toutes les permutations du premier une même substitution de lettres.

1°. Si, après l'adjonction de r, l'équation en V dont il est question plus haut reste irréductible, il est clair que le groupe de l'équation ne sera pas changé. Si au contraire elle se réduit, alors l'équation en V se décomposera en p facteurs tous de même degré et de la forme

$$f(V, r) \times f(V, r') \times f(V, r'') \times \ldots\ldots$$

$r\ r'\ r'' \ldots\ldots$ étant ~~les diverses~~ d'autres valeurs de r. Ainsi le groupe de l'équation proposée se décomposera aussi en groupes chacun d'un même nombre de permutations, puisqu'à chaque valeur de V correspond une permutation. Ces groupes seront respectivement ceux de l'équation proposée, quand on lui adjoindra successivement r, r', r'', ...

~~Ce qui caractérise un groupe. On peut partir d'une des permutations quelconques du groupe~~

Scholie. Les substitutions sont indépendantes même du nombre des racines

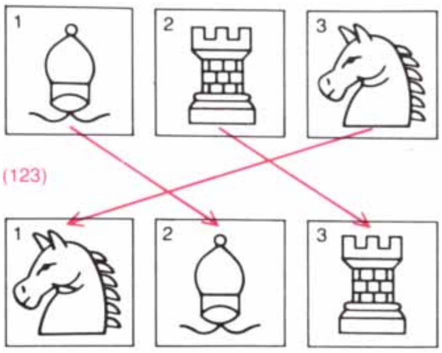
~~$S + T = U$~~

~~$S + T + S + T + S + T$~~

S_T

Il y a quelque chose à compléter dans cette démonstration. Je n'ai pas le temps. (Note de l'A.)

~~Car si l'on élimine r entre $f(V, r) = 0$ et $Fr = 0$, Fr étant de degré premier p, il ne peut arriver que de deux choses l'une : ou le résultat de l'élimination sera de même degré en V que $f(V, r)$, ou il sera d'un degré ~~multiple de~~ p fois plus grand.~~



IDEA OF A GROUP can be illustrated by considering the group $S(3)$, which is the group of permutations of three objects. An element of $S(3)$ operates on the objects by rearranging them. The permutation (123) moves the object in the first square into the second, the object in the second square into the third and the object in the third square into the first. Because there are six possible arrangements of three objects there are six elements in $S(3)$.

al texts and went straight for the masters of the day, devouring Adrien Marie Legendre's work on geometry and the memoirs of Joseph Louis Lagrange: *The Resolution of Algebraic Equations, The Theory of Analytic Functions* and *Lessons on the Calculus of Functions.* It was undoubtedly from Lagrange that Galois first learned the theory of equations, to which he was to make fundamental contributions over the next four years. Vernier seems to have appreciated his student's talents: his remarks in Galois's trimester reports carry high praise such as "zeal and success" and "zeal and progress very marked."

With his discovery of mathematics Galois's personality underwent a striking change. He began to neglect his other courses and aroused the hostility of his teachers in the humanities. His rhetoric teachers called him "dissipated" on the trimester reports, and the words "withdrawn," "bizarre" and "original" appear on his evaluations. Even Vernier, although not seeking to cool Galois's passion for mathematics, urged him to work more systematically. Galois did not follow the advice: he decided to take the entrance examination for the École Polytechnique a year early and without the usual preparatory courses in mathematics. Evidently lacking some of the basics, he failed.

Galois considered the failure an injustice, and it hardened his attitude toward authority. Nevertheless, he continued to progress rapidly in mathematics and enrolled in the more advanced course at Louis-le-Grand taught by Louis-Paul-Émile Richard, a distinguished instructor. Richard immediately recognized Galois's abilities and called for his admission to the École Polytechnique without examination. Although the recommendation was not followed, Richard's encouragement produced spectacular results. In March, 1829, while Galois was still a student, his first paper was published. It was titled "Proof of a Theorem on Periodic Continued Fractions" and appeared in *Annales de mathématiques pures et appliquées* of Joseph Diaz Gergonne.

The paper, however, was a minor aside. Galois had already turned to the theory of equations, the topic he had first explored in the works of Lagrange. At age 17 he was taking on one of the most difficult problems in mathematics, a problem that had confounded mathematicians for more than a century.

In 1829 the central question for the theory of equations was: Under what conditions can an equation be solved? More precisely, what was sought was a method of solving an equation having a single variable x whose coefficients are all rational numbers and whose highest-power term is x^n. The method was to be a general one that could be applied to all such equations, and it was to rely on only the four elementary operations of arithmetic (addition, subtraction, multiplication and division) and the extraction of roots. If the solutions or roots of an equation can be obtained from the coefficients solely by these operations, the equation is said to be solvable by radicals.

From a historical perspective it was natural to expect that solving an equation of the nth degree would call for no operations more elaborate than the extraction of nth roots. The solution to the general quadratic, or second-degree, equation $ax^2 + bx + c = 0$, known to the Babylonians, requires the extraction of the square root of a function of the coefficients, namely $b^2 - 4ac$. Hence the general quadratic equation is solvable by radicals. Similarly, the general solution to the cubic equation, devised by the Italian mathematicians Scipione dal Ferro and Niccolò Fontana, or Tartaglia, in the early 1500's, requires taking cube roots of functions of the coefficients. The solution to the general fourth-degree equation, first achieved by the Italian mathematician Lodovico Ferrari at about the same time, requires the extraction of fourth roots.

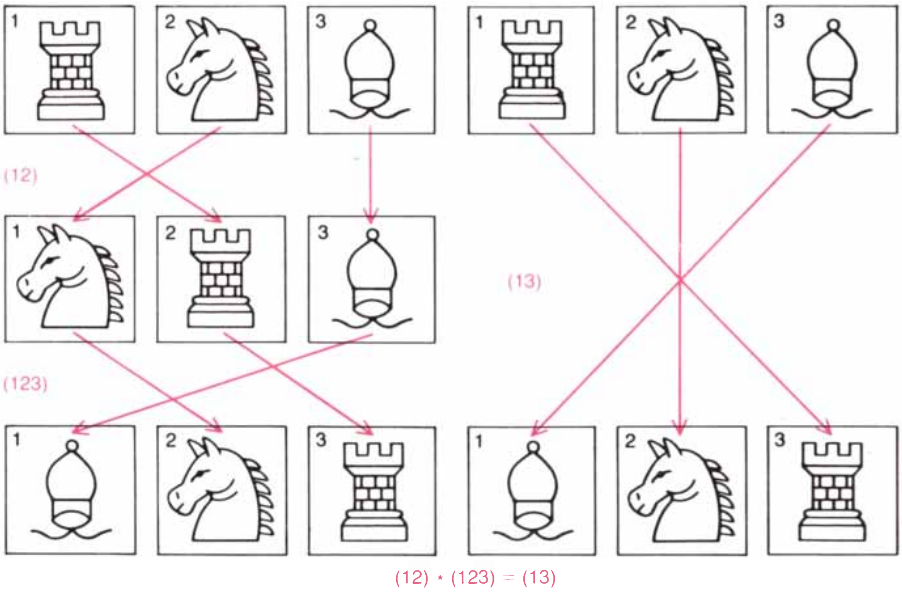
By the time of Galois, however, nearly 300 years of effort had not yielded a solution by radicals to general equations of the fifth degree or higher. A number of mathematicians had come to suspect that such general solutions are not possible, even though in special cases, such as the equation $x^7 - 2 = 0$, the solution can be found by radicals. (In this instance one solution is $\sqrt[7]{2}$.) Galois provided definitive criteria for determining whether or not the solutions to a given equation can be found by radicals. Perhaps even more remarkable than Galois's findings in the theory of equations were the methods he developed to study the problem. His investigations led to a theory with applications far outside the theory of equations, now called the theory of groups.

Galois submitted his first papers on what was to become group theory to the French Academy of Sciences on May 25 and June 1, 1829, near the end of his final year at Louis-le-Grand. Less than two months later he was to take the entrance examination to the École Polytechnique for the second time, but meanwhile events in his life took an unfortunate turn. On July 2, a few weeks before the examination, Évariste's father suffocated himself in his Paris apartment. The Jesuit priest of Bourg-la-Reine had forged Mayor Galois's name to a number of malicious epigrams directed at Galois's own relatives. The senior Galois could not bear the scandal. The entrance examination therefore took place under the worst-possible circumstances. Furthermore, Évariste apparently declined to follow the examiner's suggestions for exposition and was failed for a second and final time. The two disasters crystallized his hatred for the conservative hierarchy then ruling France.

Forced to consider the less prestigious École Normale (then called the École Préparatoire), Galois took the baccalaureate examinations required for entry in November, 1829. This time he passed on the basis of an exceptional score in mathematics, and he was given university status at about the time his first papers on group theory were to be presented to the Academy of Sciences. The papers, however, did not receive a hearing.

When the papers were received by the Academy, Augustin Louis Cauchy, then the most eminent mathematician in France and a staunch supporter of the conservative restoration, was appointed referee. Cauchy had already investigat-

MARGINAL NOTE on one of the papers left behind by Galois on the morning of his duel is the most famous of the documents cited in support of the legend that Galois wrote down his ideas on group theory in a single night. The note reads: "There is something to complete in this demonstration. I do not have the time (author's note)." ("Il y a quelque chose à completer dans cette démonstration. Je n'ai pas le temps (Note de l'A.).") According to the familiar account of the life of Galois by Eric Temple Bell, the phrase "I do not have the time" is written frequently in the manuscripts. Actually the page reproduced here is the only place the phrase appears. The rapid handwriting of the note contrasts sharply with the careful and deliberate hand of the body of the text, which suggests that Galois did not write the paper the night before the duel but only edited it. Indeed, the paper had already been submitted to the Academy of Sciences and returned to Galois by Siméon Denis Poisson with suggestions for reworking it.



"MULTIPLICATION" of one element in $S(3)$ by another element is carried out by determining the arrangement of objects that results from the operation of the first permutation and then applying the second permutation to this arrangement. The single permutation that brings about the same rearrangement is called the product of the two permutations. In general the multiplication of groups is not commutative: the product of two elements depends on the sequence they are applied. Thus (12) * (123) is equal to (13) but (123) * (12) is equal to (23).

SECOND ELEMENT

FIRST ELEMENT

	(1)	(123)	(132)	(12)	(13)	(23)
(1)	(1)	(123)	(132)	(12)	(13)	(23)
(123)	(123)	(132)	(1)	(23)	(12)	(13)
(132)	(132)	(1)	(123)	(13)	(23)	(12)
(12)	(12)	(13)	(23)	(1)	(123)	(132)
(13)	(13)	(23)	(12)	(132)	(1)	(123)
(23)	(23)	(12)	(13)	(123)	(132)	(1)

MULTIPLICATION TABLE for the six permutations of three objects provides a verification that the permutations satisfy the properties of a group. For every pair of permutations a and b the table shows that their product $a * b$ is itself a permutation. There is an identity element, namely the element (1), with the property that $a * (1)$ is equal to a. For every element a there exists an element called the inverse of a, or a^{-1}, with the property that $a * a^{-1}$ is equal to (1). The inverse of (123), for example, is (132). Finally, the associative law, which states that for any permutations a, b and c the product $a * (b * c)$ is equal to $(a * b) * c$, can be checked in the table. The permutations in color form a subset of the six permutations. Their multiplication table, also in color, shows that they too form a group. Such a group within a group is called a proper subgroup.

ed permutation theory, a forerunner of group theory, and he later wrote extensively on group theory itself. Although legend has it that Cauchy lost, forgot or discarded Galois's manuscripts, it is far more credible that Cauchy recognized their importance and handled them with care. Indeed, a letter discovered in 1971 by Taton in the archives of the Academy makes it clear that on January 18, 1830, Cauchy had planned to give Galois's results a full hearing before the Academy. Cauchy wrote: "I was supposed to present today to the Academy...a report on the work of the young Galois. ... Am indisposed at home. I regret not being able to attend today's session and I should like you to schedule me for the following session for the...indicated subjects."

The following week, however, when Cauchy read a paper of his own to the Academy, he did not present Galois's work. Why this happened remains a subject of speculation. Taton conjectures that Cauchy urged Galois to expand his work and submit it for the Academy's Grand Prize in mathematics. Although Taton's conjecture cannot yet be documented, Galois did submit an entry for the prize in February, the month before the deadline. The entry was sent to Jean Baptiste Joseph Fourier, the mathematician who invented what is now called Fourier analysis, in his capacity as Perpetual Secretary to the Academy. Fourier died in May, however, and Galois's manuscript could not be found among Fourier's effects. Galois later attributed his bad luck to malicious intent on the part of the Academy, accusing the prize committee of rejecting his paper out-of-hand because his name was Galois and because he was still a student. The Galois legend passes on such accusations at face value, but there can be little doubt that Galois's attitude toward authority was becoming somewhat paranoid.

In spite of the setbacks Galois remained a productive mathematician and began to publish in Baron de Férussac's *Bulletin des sciences mathématiques, astronomiques, physiques et chimiques,* a far less conspicuous forum than the meetings of the Academy. His articles make it clear that in 1830 he had progressed beyond all others in the search for the conditions that determine the solvability of equations, although he did not yet have the complete answer in hand. By January, 1831, however, he had reached a conclusion, which he submitted to the Academy in a new memoir, written at the request of the mathematician Siméon Denis Poisson. The paper is the most important of Galois's works, and its existence more than a year before the duel makes nonsense of the story that all Galois's work on the theory of groups was written down in a single night.

In order to understand Galois's work

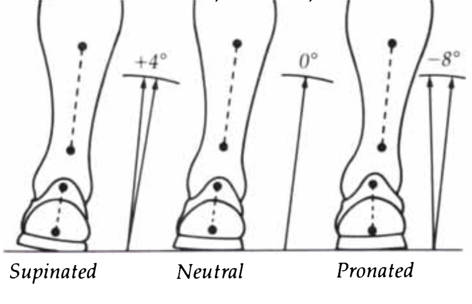
HELPING PRONATORS GET BACK ON THEIR FEET.

It's beginning to sound like the only reason runners get hurt nowadays is because they have one unspeakable fault. They pronate.

Depending on whom you talk to, the wages of pronation include everything from stress fractures and Achilles tendinitis to crippling knee injuries.

Not much fun. And not surprising that a lot of people are out there trying to put a stop to it.

But is pronation really such a bad habit? What, in fact, is it?



Computer digitized film data show that most runners' feet first contact the ground in a supinated position then roll through a neutral into a pronated position.

Well, unseen by the naked eye, as your feet are running forward, they are also rolling from one side to the other. Landing first on the outer edge, moving to the neutral plane, then continuing inward to what is called a pronated position.

Frankly, they have good reason to behave this way. If they didn't pronate, the shock of hitting the ground would be considerably worse.

So contrary to anything else you may have heard, pronation is not an unnatural act. Everybody does it. Some more than others.

From our own survey and those of others, it appears that about 12 percent of the population are likely to suffer because they pronate too much. About 8 percent, because they pronate too little.

All very interesting. But if high-speed cameras are required to see what's really going on, how can you tell if you even have a problem? Much less what it is?

You can't for sure. But you can get a good idea the next time you take a shower. Make two sets of footprints. One while standing, one sitting. A certain change in shape is normal, but too little change means your feet are rigid. Worry less about pronation and more about adequate cushioning. The kind you'll find in our Columbia, Aurora and Terra T/C.

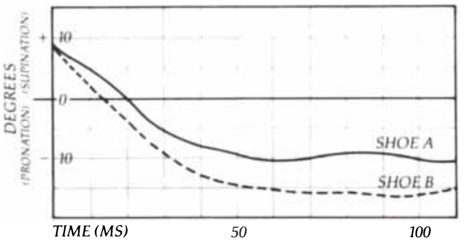
If, on the other hand, the second set is much fatter than the first, your feet are flexible. And likely to find pronation irresistible.

If that's the case, there are things to look for in a shoe. Like a stiff heel counter. Make sure it fits snug, otherwise it won't be effective. Normally, you're also better off with a firm midsole. And good arch support.

Those are just a few of the properties we designed into the Nike Centurion and Internationalist.

But when it came to the Equator, we went even further. Recognizing that the base of the heel bone is like a small billiard ball that loves to roll, we decided to approach the problem from the ground up. Right under the Spenco® sockliner we put a soft orthotic to cradle the heel and bring rock and roll under control.

After a few more innovative touches, we wound up with a shoe that proved effective in reducing pronation an average of 22 percent.



Two different shoe models tested by nine runners. After 54 trials, shoe A exhibits an average of 4.9° less maximum pronation than shoe B.

If you need it, we've got it. But make sure. Because as bad as pronation can be, we suspect it's getting a bum rap. Many of the problems laid at its feet may actually be the result of overtraining. Especially when the difficulty doesn't recur in the same part of the body every time.

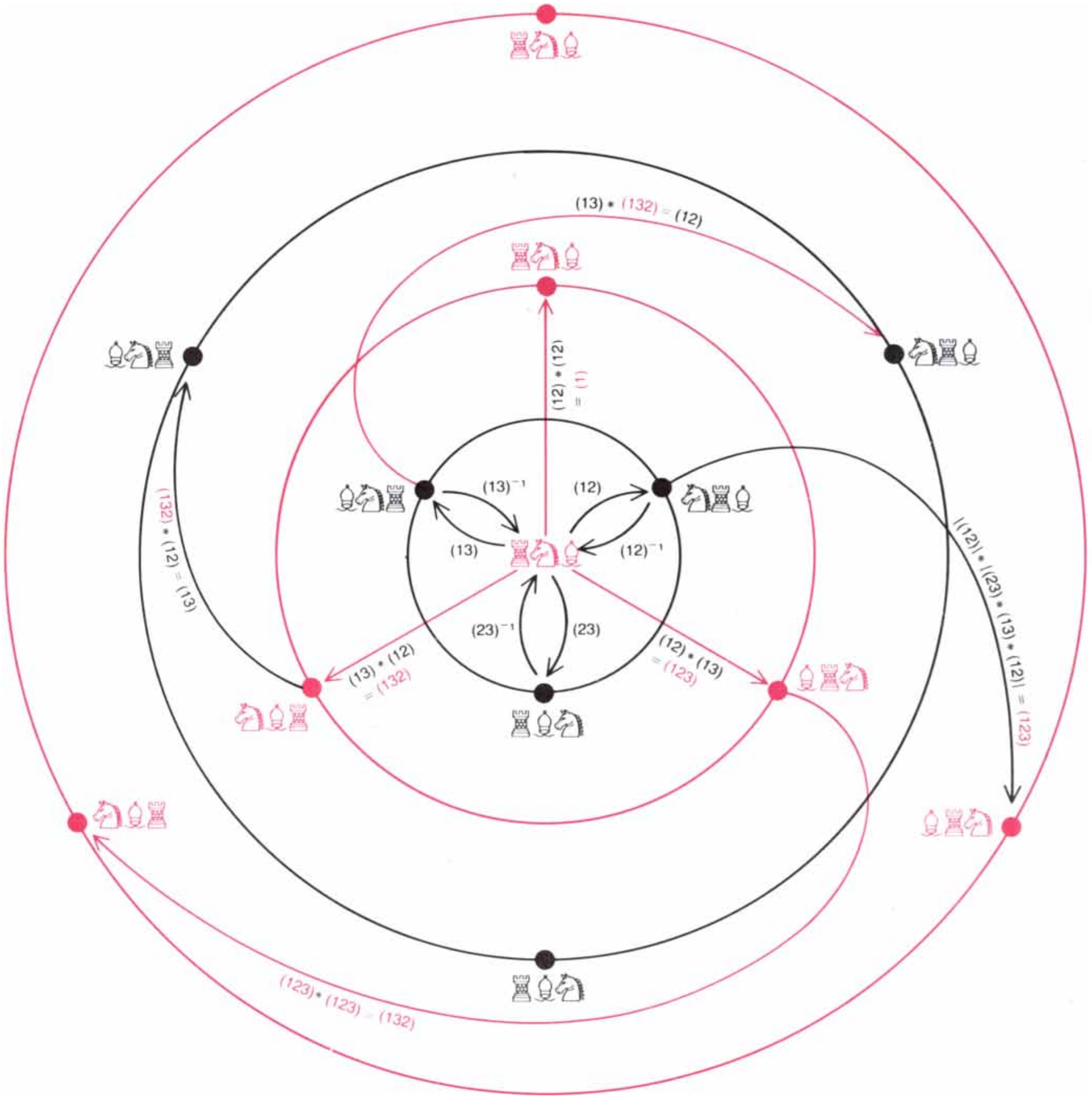
We're not the only ones with an opinion, however. There are several excellent works on the subject, including Harry F. Hlavac's *The Foot Book.*

So listen to the experts. Listen to your doctor. But more important, listen to your body. It has a vested interest in your well being. And if something isn't working, your body is sure to let you know.

But please. Don't make it cry to be heard.

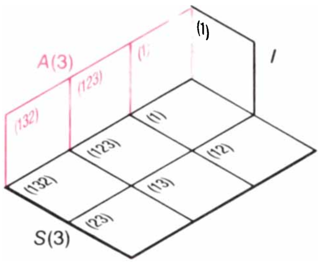
NIKE
Beaverton, Oregon

it is unprofitable to study the original papers. Poisson did his best to understand the 1831 manuscript, but he finally recommended that the Academy reject it, encouraging Galois to expand and clarify the exposition. Poisson also criticized one of Galois's proofs as inadequate, although the statement being proved could be shown to be true through a result proved by Lagrange. According to Peter Neumann of the University of Oxford, the criticism is completely accurate. Galois's arguments are presented in a concise form that makes them extremely difficult to follow, and they are not free from error. With the benefit of a century and a half of clarification, however, it is now possible to set forth the essentials of the theory in an accessible form. To this end I have had the assistance of the astro-



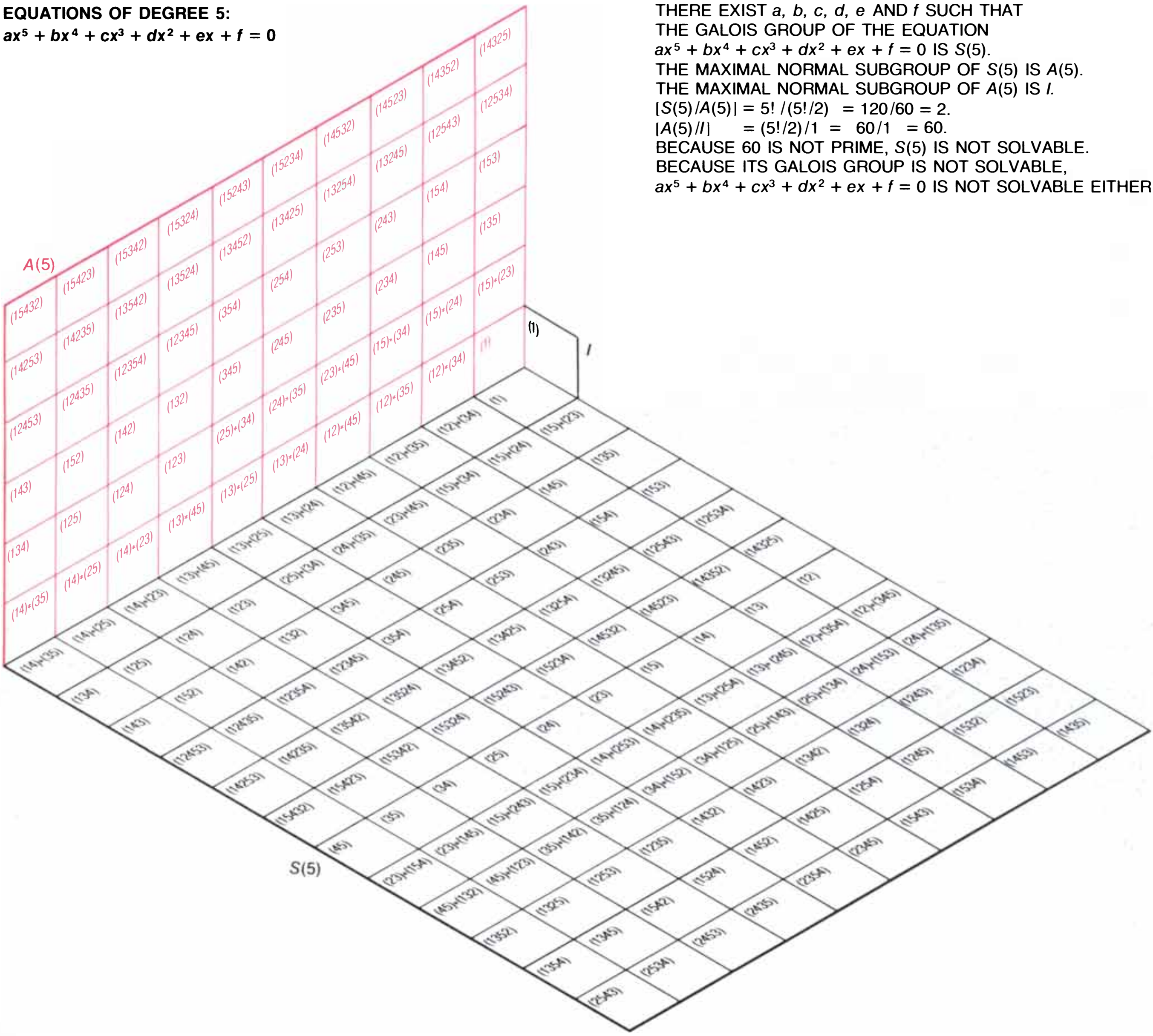
PERMUTATIONS IN $S(3)$ can invariably be written as the product of permutations that interchange only two objects. If a permutation can be written as a product of an even number of such interchanges, it is called an even permutation; otherwise it is an odd permutation. If an even permutation (*colored circles*) is multiplied by an even permutation (*colored arrows*), the product is an even permutation; if an even permutation is multiplied by an odd permutation (*black arrows*), the product is odd. Similarly, if an odd permutation (*black circles*) is multiplied by an even permutation, the product is odd, whereas if an odd permutation is multiplied by an odd permutation, the product is even. The even permutations form a subgroup, namely the subgroup printed in color in the bottom illustration on page 140. This subgroup is called the alternating group, or $A(3)$. A subgroup such as $A(3)$ is called a normal subgroup of $S(3)$ if for any element h in $A(3)$ and any element g in $S(3)$, the element $g * h * g^{-1}$ is also an element of $A(3)$. To prove that $A(3)$ is a normal subgroup of $S(3)$ suppose g is an even permutation. Then $g * h * g^{-1}$ is the product of three even permutations, which is also an even permutation and so is a member of $A(3)$. If g is an odd permutation, $g * h * g^{-1}$ is the product of an odd by an even by an odd permutation, which is again an even permutation. Hence $A(3)$ is a normal subgroup. By a similar argument it can be shown that for any number n, $A(n)$ is a normal subgroup of $S(n)$. The number of elements in a subgroup must divide the number of elements in the parent group without remainder. Because $A(n)$ has half as many elements as $S(n)$, $A(n)$ includes the maximum number of elements that a proper subgroup of $S(n)$ can have. $A(n)$ is the maximal normal subgroup.

EQUATIONS OF DEGREE 3:
$ax^3 + bx^2 + cx + d = 0$



THERE EXIST a, b, c AND d SUCH THAT
THE GALOIS GROUP OF THE EQUATION
$ax^3 + bx^2 + cx + d = 0$ IS $S(3)$.
THE MAXIMAL NORMAL SUBGROUP OF $S(3)$ IS $A(3)$.
THE MAXIMAL NORMAL SUBGROUP OF $A(3)$ IS I.
$|S(3)/A(3)| = 3!/3 = 6/3 = 2.$
$|A(3)/I| \quad = 3/1 = \quad 3.$
BECAUSE 2 AND 3 ARE PRIME, $S(3)$ IS SOLVABLE.
BECAUSE ITS GALOIS GROUP IS SOLVABLE,
$ax^3 + bx^2 + cx + d = 0$ IS SOLVABLE ALSO.

EQUATIONS OF DEGREE 5:
$ax^5 + bx^4 + cx^3 + dx^2 + ex + f = 0$



THERE EXIST a, b, c, d, e AND f SUCH THAT
THE GALOIS GROUP OF THE EQUATION
$ax^5 + bx^4 + cx^3 + dx^2 + ex + f = 0$ IS $S(5)$.
THE MAXIMAL NORMAL SUBGROUP OF $S(5)$ IS $A(5)$.
THE MAXIMAL NORMAL SUBGROUP OF $A(5)$ IS I.
$|S(5)/A(5)| = 5!/(5!/2) = 120/60 = 2.$
$|A(5)/I| \quad = (5!/2)/1 = \quad 60/1 \quad = 60.$
BECAUSE 60 IS NOT PRIME, $S(5)$ IS NOT SOLVABLE.
BECAUSE ITS GALOIS GROUP IS NOT SOLVABLE,
$ax^5 + bx^4 + cx^3 + dx^2 + ex + f = 0$ IS NOT SOLVABLE EITHER.

SOLUTION OF EQUATIONS was the problem for which Galois developed group theory. A general method of solution would employ only addition, subtraction, multiplication, division and the extraction of roots, and it could be applied to any equation of degree *n*, where *n* is the highest power to which a variable is raised. Galois proved that no such method exists when *n* is equal to 5 or more. Every equation of degree *n* can be associated with the group $S(n)$ or with some subgroup of $S(n)$; the group associated with an equation is now called the Galois group of the equation. Galois showed that an equation can be solved by the methods of arithmetic and root extraction only if its Galois group is what he defined as a solvable group. A group is solvable if it generates a series of maximal normal subgroups whose composition factors (determined from the number of elements in the parent group and the subgroups) are prime numbers. The composition factors generated by $S(3)$ and its series of maximal normal subgroups are all prime numbers. It follows that all third-degree equations are solvable. When *n* is equal to 5 or more, however, it can be proved that the maximal normal subgroup of $A(n)$ is the identity group *I*, which includes only the identity element. Since $A(n)$ is the maximal normal subgroup of $S(n)$, the composition factors of $S(n)$ when *n* is equal to 5 or more are not all prime numbers. Hence there are equations of degree 5 or greater that cannot be solved by the allowed methods.

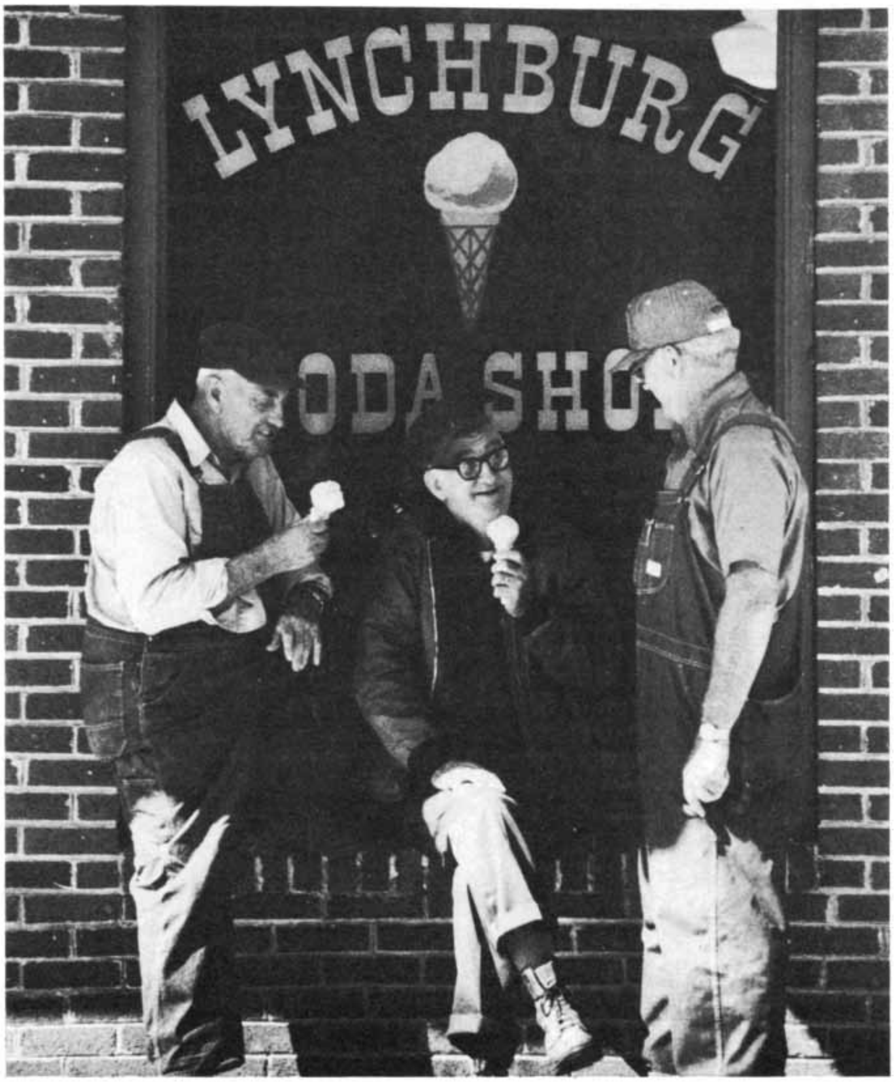
physicist Adrian C. Ottewill of Oxford.

What is a group? On its deepest level group theory concerns the symmetries inherent in any system. Imagine a snowflake whose points or vertexes are equally spaced at angles of 60 degrees. If the snowflake is rotated about an axis through its center by 60 degrees or any integral multiple of 60 degrees, its basic pattern remains unchanged, even though any particular vertex may change its position. An operation that leaves a pattern invariant in this sense is called a symmetry operation.

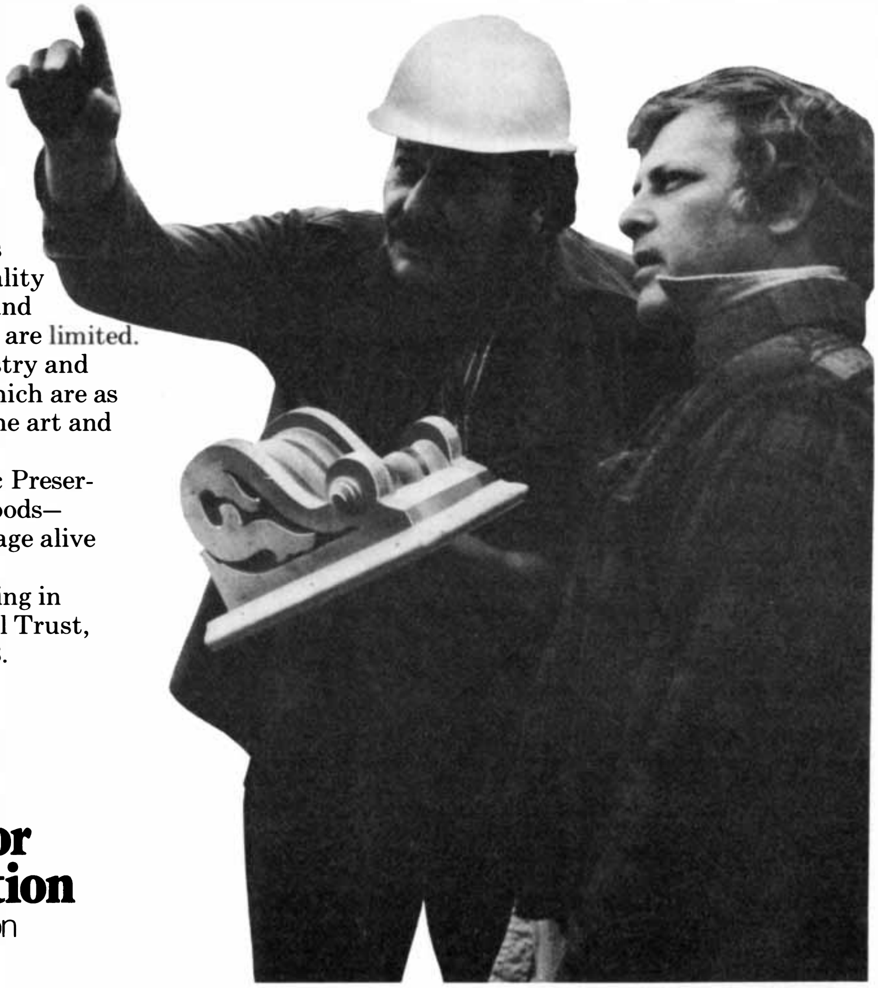
If two rotations by integral multiples of 60 degrees are carried out in sequence, the snowflake remains invariant, and the position assumed by the vertexes is one that could have been reached by a single operation. For example, a 60-degree counterclockwise rotation followed by a 240-degree clockwise rotation is equivalent to a 180-degree clockwise rotation. In general, if $R(n)$ denotes a rotation through $60n$ degrees, and if the result of performing first one such operation and then another is written $R(n) * R(m)$, then for all integers n and m the expression $R(n) * R(m)$ is equal to $R(n + m)$. Mathematically the equivalence states that the "product" of two symmetry operations is also a symmetry operation.

There are three other important properties of the rotations of a snowflake. First, a rotation through zero degrees, or $R(0)$, always leaves the pattern invariant, since it does nothing. The product of any rotation $R(n)$ and $R(0)$ is $R(n)$, so that $R(0)$ plays much the same role in rotations as the number 1 plays in ordinary multiplication. $R(0)$ is therefore called the identity rotation. Second, a rotation $R(n)$ followed by a rotation in the opposite direction by the same amount, which can be denoted $R(-n)$, returns the pattern to its starting point. Thus the product $R(n) * R(-n)$ is equivalent to $R(0)$. The rotation $R(-n)$ is called the inverse of the rotation $R(n)$. Third, the expression $R(m) * R(n) * R(p)$ is unambiguous, because $[R(m) * R(n)] * R(p)$ is equivalent to $R(m) * [R(n) * R(p)]$. This is a formal property of the operation *, by means of which two rotations are combined, called the associative property.

The four properties that hold for combinations of snowflake rotations are characteristic of any set of symmetry operations on any system; they are called the group properties. The system need not be a geometric pattern such as a snowflake. For example, an equation is also a system whose symmetries can be described by the group properties. In abstract terms a group consists of elements or symmetry operations a, b, c and so forth, and a rule denoted by * for combining any two elements. The elements of the group and the rule * are assumed to satisfy the closure criterion, which



STEINHARDT-HANSON
OFFICE PRODUCTS
BUSINESS FURNITURE
OFFICE MACHINES
GIFTS • GREETING CARDS
DOUGLAS



states that for any elements *a* and *b* in the group, $a * b$ is also an element of the group. The group must include an identity element 1, which is defined so that for any element *a* in the group, $a * 1$ is equal to *a*. Moreover, for every element *a* there must be some inverse element a^{-1} with the property that $a * a^{-1}$ is equal to 1. Finally, the elements of the group and the operation are assumed to satisfy the associative property, which states that $(a * b) * c$ is equal to $a * (b * c)$.

The theory of groups is one of the most fruitful areas of mathematical research; Bell is correct when he writes that it will keep mathematicians busy for hundreds of years. One of the most important recent achievements in group theory was a proof announced at a meeting of the American Mathematical Society in January, 1981, by Daniel Gorenstein of Rutgers University. Gorenstein showed that a list of 26 groups called sporadic finite simple groups is a complete list. In a sense this finding implies that the components or building blocks of any group with a finite number of elements have now been exhaustively classified.

Another set of non-numerical elements that satisfies the properties of a group is the group of permutations on a fixed number of objects. The permuted objects might be chess pieces, for example, or letters of the alphabet. It is essential to recognize, however, that the elements of the group are neither the chess pieces nor the letters but rather the functions that generate the various permutations. To find the "product" of two elements *a* and *b* of the group (that is, to find $a * b$) one finds the result of the first permutation on the set of objects and applies the second permutation to this result.

Suppose three chess pieces are arranged so that a rook is on a square labeled 1, a knight is on a square labeled 2 and a bishop is on a square labeled 3. One element of the permutation group for these objects can be written (12); it takes the object on square 1 and moves it to square 2 and takes the object on square 2 and moves it to square 1. The effect of the element (12) on the arrangement rook-knight-bishop is to exchange the rook and the knight, generating the arrangement knight-rook-bishop. If the operation is then done again, it exchanges the pieces on the same squares, recreating the arrangement rook-knight-bishop. Thus the group element (12) is its own inverse.

Another group element, designated (123), moves the object on square 1 to square 2, the object on square 2 to square 3 and the object on square 3 to square 1. Suppose the initial arrangement rook-knight-bishop is again transformed by the element (12), giving rise to the arrangement knight-rook-bishop. Now the element (123) is applied and generates the arrangement bishop-knight-rook. This final arrangement could have been reached in one step from the initial arrangement by applying the permutation (13), which interchanges the object on square 1 with the object on square 3. Thus the result of the permutation (12) followed by (123) generates the same arrangement of objects as the permutation (13) does. Symbolically, then, $(12) * (123) = (13)$.

The number of permutations or arrangements of *n* objects is equal to *n* factorial, written *n*!. The factorial of a number *n* is the product of all the whole numbers from 1 to *n* inclusive; 5!, for example, is equal to $1 \times 2 \times 3 \times 4 \times 5$, or 120. Hence the number of elements in $S(n)$, the permutation group for *n* objects, is *n*!. The number of elements in a group is called the "order" of the group. $S(3)$, the permutation group for three objects, includes the 3! (or six) permutations (1), (12), (13), (23), (123) and (132). Here (1) is the identity permutation that leaves any arrangement of objects unchanged.

It turns out that certain subsets of the set of elements in a group can by themselves satisfy all the properties of a group, in which case they are said to form a subgroup. If the number of elements in the subgroup is less than the number of elements in the parent group, the subgroup is called a proper subgroup. For example, it is easy to verify that [(1), (12)] is a group, and so it is a proper subgroup of $S(3)$.

For any proper subgroup *H* of a group *G* a number called the composition factor can be defined: it is the order of the parent group divided by the order of the subgroup and is generally written $[G/H]$. The composition factor of the subgroup [(1), (12)] with respect to the group $S(3)$ is 6/2 or 3. According to an elementary theorem of group theory, which I shall not prove here, the order of any subgroup must exactly divide the order of its parent group, so that the composition factor is invariably a whole number.

Galois introduced three critical concepts whose interrelations enabled him to prove that there is no general method for solving an equation of the fifth degree or higher when all solutions must be found by radicals. First Galois noted that every equation can be associated with a group of permutations. Such a group is a representation of the symmetry properties of the equation; it is now called the Galois group.

To appreciate the properties of the Galois group, consider any third-degree equation whose coefficients are rational numbers. It can be proved that such an equation has three roots, although the proof does not reveal whether the roots can be found by radicals. If the roots are designated *u*, *v* and *w*, one can form polynomial functions of them, such as $u - v$

"INFAMOUS COQUETTE" whom Galois blames for his troubles in a letter written the night before the duel was probably the same woman whose name appears frequently in the margins of Galois's papers. In the manuscript reproduced above the name "Stéphanie" can be read under the name "Évariste," and Galois has also combined the initials "S" and "E" into a single cursive diagram. From letters and other manuscripts it is clear that Galois's angry epithet is his reaction to an unhappy love affair with a woman he had met only a few months before the duel. She has been identified as Stéphanie-Félicie Poterin du Motel, the daughter of a Parisian physician.

or $uv + w - 1$. Any such function can be converted into a related function by permuting the roots u, v and w. For example, the permutation (12) interchanges u and v and so converts the function $u - v$ into the function $v - u$. Many functions of the roots are changed in value by such a permutation, but some of them are not. For example, the function $u + v + w$ is not changed in value by any permutation of u, v and w. Since the group $S(3)$ includes all possible permutations of u, v and w, $u + v + w$ is said to be invariant under $S(3)$.

It is possible to show that the value of $u + v + w$ is a rational number for any third-degree equation with rational coefficients. Other polynomial functions of the roots may be rational for some equations and irrational for others, depending on the coefficients of the equation. If the value of such a function is rational, there exists some group of permutations of u, v and w that do not change the value of the function. The Galois group of an equation is the largest group of permutations that meet this requirement for every rational-valued polynomial function of the roots. In other words, for any polynomial function of the roots that has a rational value, every permutation in the Galois group leaves the value of the function unchanged. When a permutation of the roots does not change the value of any rational-valued polynomial function of the roots, the roots are indistinguishable for that permutation. Hence the larger the number of elements in the Galois group is, the more permutations there are for which the roots are indistinguishable. For this reason the Galois group is a powerful way of representing the symmetry properties of an equation.

Calculating the Galois group for a given equation is generally difficult, although it can always be done in principle without knowing the values of the roots of the equation. For Galois's purposes, however, the calculation was not necessary. All he needed to show was that there are invariably equations of degree n whose Galois group is the largest-possible group of permutations of the roots, namely $S(n)$.

The second concept introduced by Galois was the concept of a normal subgroup. A subgroup H of a group G is normal in G if and only if the following condition is satisfied: When one "multiplies" any element h of the subgroup H on the left by any element g of the parent group G, and then "multiplies" the product on the right by g^{-1} (the inverse element of g), the result is an element of the subgroup H. Symbolically, if H is normal in G, there is an element h' in H such that $h' = g * h * g^{-1}$. For example, one can verify that [(1), (123), (132)] is a normal subgroup of $S(3)$ [*see illustration on page 142*].

If a finite group G has any normal subgroups at all, there must be one subgroup whose order is the largest of all the normal subgroups of G; it is called the maximal normal subgroup of G. Similarly, a maximal normal subgroup may in turn have a maximal normal subgroup of its own, and the sequence of maximal normal subgroups continues until the smallest normal subgroup possible is reached. Any group G therefore generates a sequence of maximal normal subgroups. If the sequence is labeled $G, H, I, J, \dots$, then a series of maximal normal composition factors can be defined: $[G/H]$, $[H/I]$, $[I/J]$ and so on.

The third important concept of Galois's theory is the concept of a solvable group. Galois calls a group solvable if every one of the maximal normal composition factors generated by the group is a prime number. The maximal normal subgroup of $S(3)$, for instance, is [(1), (123), (132)]. In turn the maximal normal subgroup of [(1), (123), (132)] is [(1)]. The composition factor defined for $S(3)$ and its subgroup [(1), (123), (132)] is 6/3 or 2, and the composition factor for the group [(1), (123), (132)] and its subgroup [(1)] is 3/1 or 3. Since both 2 and 3 are prime numbers, $S(3)$ is a solvable group.

The term solvable group is well justified by Galois's theory: he was able to show that an equation is solvable by radicals if and only if the Galois group of the equation is a solvable group. In order to prove that equations of the fifth degree or higher cannot in general be solved by radicals, Galois had to show that there are equations of this kind for which the Galois group is not a solvable one. As it happens the group $S(n)$ is not a solvable group when n is equal to 5 or more [*see illustrations on pages 142 and 144*]. Since for all such values of n there are equations of degree n for which $S(n)$ is the Galois group, the general equation of the fifth degree or higher is not solvable.

By the time Galois's work on group theory was nearly completed, the events of his life had become markedly political. In July, 1830, the republican opponents of the restored monarchy took to the streets and the revolution forced the Bourbon king Charles X into exile. While the students of the left-wing École Polytechnique played an active role in the fighting, Galois and his fellows at the École Préparatoire were locked up in the school by its director. Incensed, Galois tried to scale the walls; he failed and thereby missed the brief revolution.

Although the Bourbon abdication seemed a great victory for the republicans, it proved a short-lived one. Louis

Philippe was placed on the throne, to the disappointment of Galois and like-minded liberals. In the months following the revolution Galois joined republican societies, met republican leaders (notably François Vincent Raspail) and probably took part in the riots and demonstrations that were racking Paris. He joined the Artillery of the National Guard, a branch of the militia made up almost entirely of republicans. In December his break with the École Préparatoire became official. He wrote a letter that called the director of the school a traitor for his actions during the July revolution; not surprisingly, Galois was expelled.

The impression of Galois that one gains from the events of this period is not that of a victim of circumstances, as legend would have it. Instead he appears to have been a hothead whose extreme actions consistently got him into trouble. A letter by the mathematician Sophie Germain implies that Galois regularly attended sessions of the Academy of Sciences and habitually insulted the speakers. After his expulsion from the École Préparatoire he moved to his mother's house in Paris but proved to be so difficult to live with that she fled.

The climactic event of the turbulent spring of 1831 took place on May 9 during a republican banquet celebrating the acquittal of 19 artillery officers who had been accused of plotting to overthrow the government. According to the memoirs of Alexandre Dumas (père), Galois stood to propose a toast. "To Louis Philippe!" he said, raising a glass and a dagger at the same time. For this provocative act he was arrested the next day and held for more than a month in the prison at Sainte-Pélagie.

At the ensuing trial Galois's defense claimed that the toast had been "To Louis Philippe, if he betrays" but that "if he betrays" had been drowned out in the uproar. Whether the jurors believed the defense or were moved by Galois's youth (he was then 19) is not known, but they acquitted him in minutes. Nevertheless, on Bastille Day, July 14, 1831, less than a month after his acquittal, Galois was arrested again, this time for illegally wearing the uniform of the Artillery Guard. The guard had been disbanded as a threat to the crown and Galois's gesture was therefore an act of defiance. This time he spent eight months in Sainte-Pélagie.

The prison term was devastating: Galois was alternately despondent and raging. Raspail, who was serving a sentence at the same time, later recalled that Galois once, while drunk, had to be restrained from an attempt at suicide. Later, according to Raspail, Galois confided a chilling vision of his demise: "I shall die in a duel for the sake of some worthless girl [quelque coquette de bas étage]. Why? Because she will invite me to avenge her honor, which another will have compromised." After a fellow prisoner was shot, it seems Galois accused the prison superintendent of arranging the shooting. Galois was subsequently confined to the dungeon, perhaps as a result of the accusation.

In all this turmoil, however, the worst blow was the rejection of Galois's 1831 paper. In the scathing preface to his memoirs, which he wrote while he was in prison, he stated: "I tell no one that I owe anything of value in my work to his advice or encouragement. I do not say so because it would be a lie."

The end of Galois's life has always had a particular fascination for theorists. Biographers have been unwilling to accept at face value the implication of his own words, namely that the duel was a result of a personal quarrel. Instead his biographers have looked for prostitutes, agents provocateurs and political opponents to account for his death. There is no evidence to support any of these conjectures.

In the middle of March, 1832, Galois was transferred from Sainte-Pélagie to the nursing home Sieur Faultrier because of a cholera epidemic in Paris. There he apparently met the "infamous coquette." The involvement was brief, but it is absurd to suggest that the girl was a prostitute or a conspirator who helped to arrange his assassination. The epithet "infamous coquette" has been associated with the words "quelque coquette de bas étage" and so taken as confirming the prostitute story. According to the account by Raspail, however, the latter phrase was spoken by Galois a year before the duel; it may even have been Raspail's own invention. Moreover, on May 25, six days before his death, Galois alludes to a broken love affair in a letter to his friend, Auguste Chevalier: "How can I console myself when in one month I have exhausted the greatest source of happiness a man can have, when I have exhausted it without happiness, without hope, when I am certain it is drained for life?"

Who was the woman? Two fragmentary letters were written to Galois in the weeks before the duel, suggesting a personal quarrel in which Galois was more of a participant than he admitted. The first letter begins: "Please let us break up this affair. I do not have the wit to follow a correspondence of this nature but I will try to have enough to converse with you as I did before anything happened...." The second letter is similar in tone, and both of them bear the signature "Stéphanie D." In the Galois manuscripts Carlos Alberto Infantozzi of the University of the Republic in Uruguay has managed to read a name that Galois had erased: Stéphanie Dumotel. Further detective work by Infantozzi shows she was Stéphanie-Félicie Poterin du Motel, the daughter of a resident physician at the Sieur Faultrier. She later married a language professor.

It is also unlikely that the man who killed Galois was in the pay of an antirepublican plot, in spite of the assertion of Galois's brother Alfred that Évariste was murdered. According to Dumas, Galois's adversary was Pescheux d'Herbinville, not a political enemy but an ardent republican. Indeed, d'Herbinville was one of the 19 officers of the Artillery Guard whose acquittal was the occasion of Galois's defiant toast to the king. Moreover, when agents of the crown were exposed during the revolution of 1848, d'Herbinville was not among them. A summary of an article recently sent to me by Taton indicates that the duel was between friends and unfolded as a kind of Russian roulette in which only one pistol was charged.

Galois's mathematical writings the night before the duel were actually confined to making editorial corrections on two manuscripts and to summarizing the contents of these and one other paper in a long letter to Chevalier. The first paper was the one rejected by Poisson; the second was a fragmentary version of an article that had already been published in Férussac's *Bulletin.* The third has not been found and its content is known only from the summary in the letter; it apparently concerned integrals of general algebraic functions.

What of the famous words "I do not have the time" that Galois is supposed to have written repeatedly in frustration at being unable to complete his work? The phrase does appear, in the margin of the first memoir, but only once. Appended to it in parentheses is the comment "Author's note."

I do not believe that the facts about the life of Évariste Galois as I have presented them diminish his stature as a mathematician in the slightest degree. Many manuscript fragments indicate he carried on his mathematical investigations not only while in prison but also up to the time of his death. That he could work productively through such turbulent times is testimony to the extraordinary fertility of his imagination. Quite apart from the circumstances under which the work was done, there is no question that Galois developed one of the most original ideas in the history of mathematics.

His reputation is not served, however, nor is the history of science, by a legend that insists a scientific genius must be above reproach in his personal life, or that any contemporary who does not appreciate his genius is either a fool, an assassin or a prostitute. The notion that genius is not tolerated by mediocrity is too old a platitude to be adopted uncritically as accurate history. From this point of view a genius would have to be recognized as such even when standing at a banquet table with a dagger in his hand.

THE AMATEUR SCIENTIST

"Floaters": visual artifacts that result from blood cells in front of the retina

by Jearl Walker

The "floaters" one sees while looking at a featureless background are classified as entoptic phenomena, meaning that they arise within the visual system. The floaters are interesting to watch, and they reveal some interesting things about the workings of the visual system.

My experience, which is probably similar to yours, is that when I stare at a featureless background such as a blank wall or a clear sky, my field of view is spotted with small circles and larger hairlike forms. I can never quite bring them into focus. In fact, when I shift my gaze in an effort to see one of them better, the object flits away.

The small circles are easy to ignore. The larger objects are more disturbing. In the field of view of my right eye I see a structure that resembles a stiff rope tied into several large knots with loose ends. This structure is so large that it often interferes with my reading. I can even see it against a complex background. When the background is a page in a book, the object is dark. Against a computer screen, where the background is dark and the symbols are light, the object appears as a fuzzy white region. I can also see it when I close the eye while facing a source of strong illumination such as a bright sky.

One of the best early discussions of floaters is in *Treatise on Physiological Optics,* by Hermann von Helmholtz. An interpretation of the objects in terms of modern physics was offered in these pages 20 years ago by Harvey E. White and Paul Levatin [see "'Floaters' in the Eye," by Harvey E. White and Paul Levatin; SCIENTIFIC AMERICAN, June, 1962]. Much of what follows builds on the ideas presented in that article.

Although floaters seem to lie on the surface of the cornea, they are actually near the retina. They float in the fluid just in front of the fovea: the small area of acutest vision where the visual receptors are densest. The fovea (or its place in one's field of view) is sometimes called the point of fixation.

Floaters originate with red blood cells that have leaked out of the retina. Once the cells are released they swell into spheres, losing their hemoglobin and thus their red color. The spheres drift through the liquid region in front of the fovea individually or in strings. The individual cells are responsible for the small-circle floaters; the strings give rise to the hairlike structures. The object I see in my right eye seems to be an unusually long string of some 70 blood cells that is wound around itself several times. I could be wrong; several shorter strings tangled together would give the same appearance.

Each swollen cell is some eight micrometers in diameter. A floater, however, is not the direct image of a blood cell and usually is not even the simple shadow of one. Most of the time it is a diffraction pattern projected by the blood cell onto the retina. The pattern consists of bright and dark bands arising from the constructive and destructive interference of the light passing the cell on its way to the photoreceptors of the retina. When the cell is an isolated one, the pattern consists of concentric bright and dark rings. When several cells are linked, the pattern is elongated.

The liquid layer in which the cells float results from partial liquefaction of the vitreous humor, the transparent material filling most of the eye. Cells relatively distant from the retina cast large patterns that are probably too fuzzy to be distinguished. Cells closer to the retina create smaller, distincter patterns. Many people find that the number of floaters increases with age. Perhaps more blood cells are released by an older retina. Moreover, the liquid layer near the retina often expands with age.

One of the peculiarities of a diffraction pattern from a small object is that the center of the pattern can be bright. In many optical phenomena light can be modeled as either a wave or a ray. With floaters the limits are stricter. If light passed a cell as a collection of rays, the cell would cast a dark shadow on the retina. Since the cell is very small, however, the light behaves not as rays but as waves. They diffract around the edge of the cell, sending light into what should be the shadow region. At the retina the light interferes, creating the bright and dark bands constituting the pattern of a floater. Once past the cell the light can be regarded as being in ray form; at the retina only the wave model serves.

The process is depicted in the illustration on page 156 for two shafts of light passing a cell on opposite sides. The shafts are shown both as rays and as waves that are in phase, that is, in step with one another. The light diffracts around the cell and spreads into the shadow region. The shafts arrive at the retina after traveling equal distances to the center of the shadow region. Since they begin in phase and traverse equal distances, they arrive in phase and therefore interfere constructively, making the center of the shadow region bright.

Slightly off the center the shafts arrive out of phase because they travel different distances to reach the spot. When the path difference amounts to half a wavelength of light, the waves interfere destructively at the retina, leaving darkness at that spot. Symmetry creates a ring of darkness around the center of the pattern.

Slightly farther from the center is a bright band. Here shafts of light arrive from opposite sides of the cell approximately in phase and so interfere constructively. Additional bright and dark bands lie farther from the center. A floater consists primarily of the central bright spot and the first of the bands.

Y. P. Hwu of Wytheville Community College in Virginia achieved a similar diffraction pattern from blood cells by putting diluted blood on a transparent slide, covering it with a cover slide and illuminating the preparation with laser light. The laser beam was expanded by the objective lens of a microscope (of 10-diameter magnifying power) positioned just in front of the slide. Diffraction patterns of the blood cells appeared on a screen placed 50 centimeters from the slide.

The floaters in my field of view seem to drift when I tilt my head. The drift can be deceiving because the perceptive apparatus reinverts an image that is upside down at the retina. If a floater drifts downward past the retina because of gravity, the brain applies the same interpretation, making the floater seem to move upward.

I can make floaters move if I tilt my head or suddenly glance down and then up. The motion of the eye shifts the vitreous humor and the liquid layer near the retina. The resulting flow of liquid carries the floaters past the retina. If the flow is truly upward, the floaters appear to drift downward.

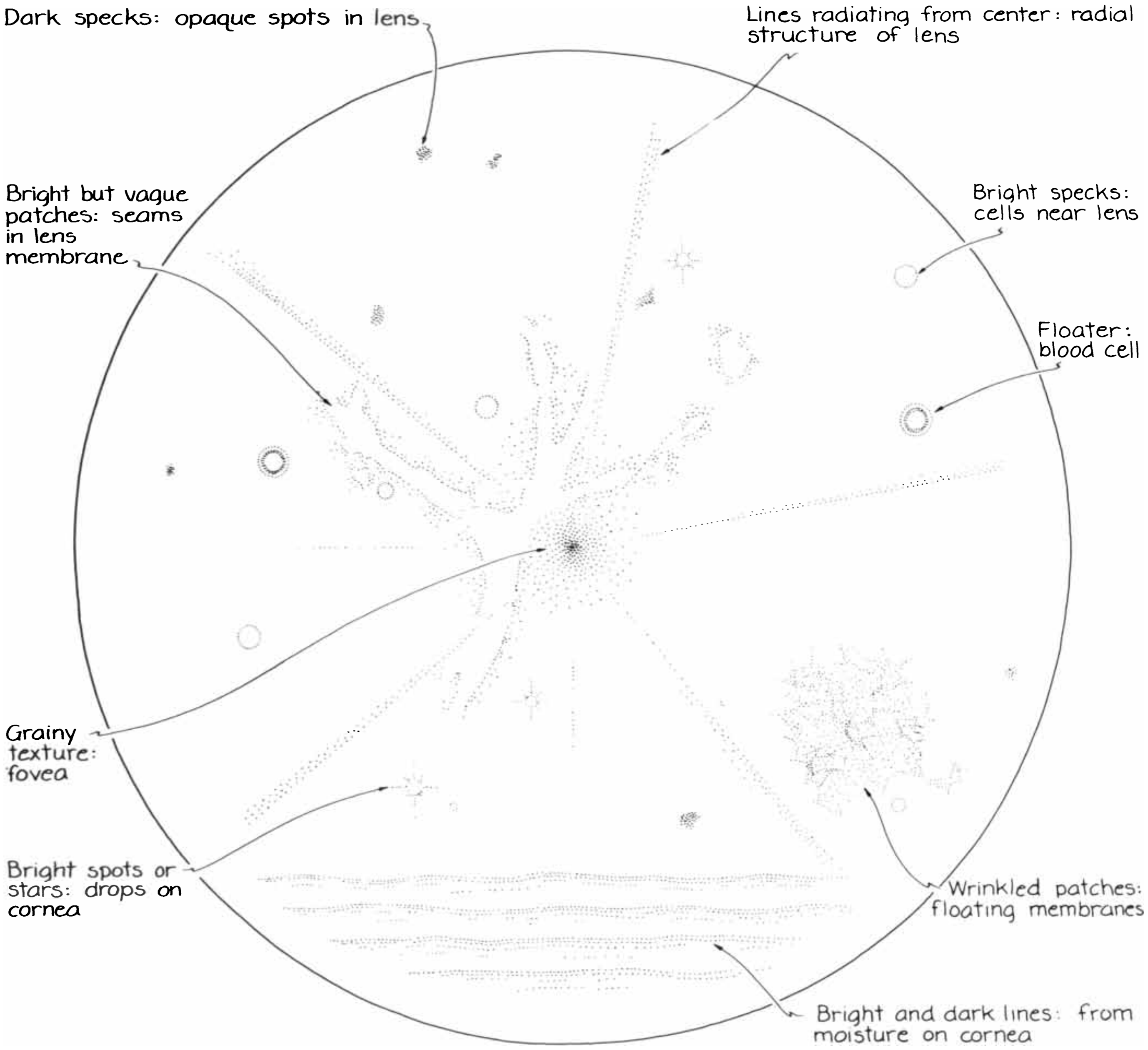
White and Levatin investigated the appearance of floaters when the viewer lay on his back and stared upward. The position makes the floating cells drift closer to the retina, where they settle into the foveal depression. The motion decreases the distance between a cell and the surface on which its diffraction pattern forms. As a result the diffraction pattern shrinks because it depends on the difference in path lengths traveled by the light rays.

For example, the first dark band arises when rays from opposite sides of the cell travel distances that differ by half a wavelength of light. When the cell lies closer to the retina, this condition is met by a spot closer to the center of the pattern. The dark band appears smaller in the field of view.

Because the patterns are smaller the reclining observer can more easily detect the ones from individual cells. They become sharper, and more bands are apparent. Eventually a cell moves so close to the fovea that the pattern loses resolution because several bands fall across adjacent photoreceptors.

Floaters can be examined with the aid of a pinhole held up to the eye. The pinhole not only serves to keep ambient light from entering the eye but also provides conditions in which the diffraction patterns from a cell are more distinct. So far this discussion has been proceeding on the basis that the light passing a cell is entirely in phase, but that is not the case when the light is from many sources. If the light comes through a pinhole, however, the shafts passing a cell are more nearly in phase. The light therefore creates a sharper diffraction pattern.

A pinhole can be made by putting a sheet of aluminum foil over the hole in the end of a spool of thread and punching a small hole in the foil with a pin. More uniform and permanent pinholes can be made by following the instructions I gave in this department last November for the construction of pinhole cameras. You can also buy a pinhole mounted on a small flashlight from the Edmund Scientific Co. (101 East Gloucester Pike, Barrington, N.J. 08007). If you make your own pinhole, you will need a diffuse light source on the opposite side of it. A hand-held slide viewer works well, as does any small flashlight. *Do not use a laser.* Although laser light is ideally suited for experiments involving optical interference, it can easily damage your retina. The light is so coherent (the waves are almost exactly in phase) that the beam can dissipate enough energy in a small region of your retina to destroy the photorecep-



A few of the objects that may be seen when a featureless background is viewed through a pinhole

Physician, did you miss any of these significant developments in medical science?

- ***Campylobacter fetus* subsp. *jejuni* is associated with a colitis that can clinically and sigmoidoscopically resemble acute idiopathic ulcerative colitis. Stool cultures are in order for *C. fetus* before beginning nonspecific anti-inflammatory therapy.**

- **Coumarin derivatives cross the placenta. A recent study shows that the consequences for the fetus can be severe. These include embryopathy, stillbirth, and premature delivery.**

- **Nonsteroidal anti-inflammatory drugs may produce a marked reduction in glomerular filtration rate; with termination of the drug, GFR returns to normal.**

- **Pittsburgh pneumonia agent (PPA) and atypical *Legionella*-like organisms (ALLO) are particularly dangerous in immunosuppressed patients.**

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.

SCIENTIFIC AMERICAN *Medicine* is the busy clinician's answer to this problem.

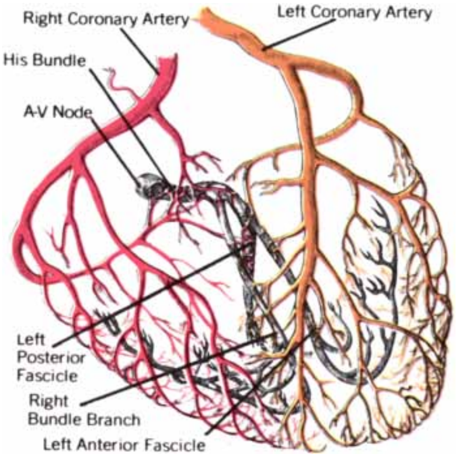
Because its authors update SCIENTIFIC AMERICAN *Medicine* every month, it is always current. Because the new information appears in a single source, it is there when you need it.

This 2,000-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 is approved by the American Academy of Family Physicians for 32 Elective Hours of CME credit.



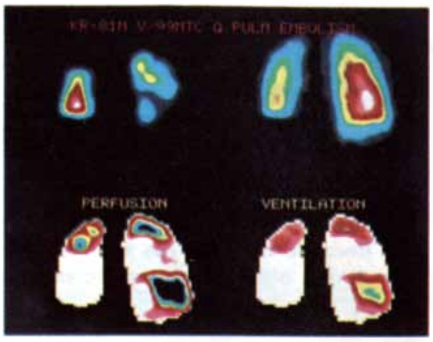
Branches of the right and left coronary arteries supply blood to the AV node and intraventricular conduction system.

SCIENTIFIC AMERICAN *Medicine* is lucidly illustrated with drawings and photographs. Some examples are seen here and on the facing page.

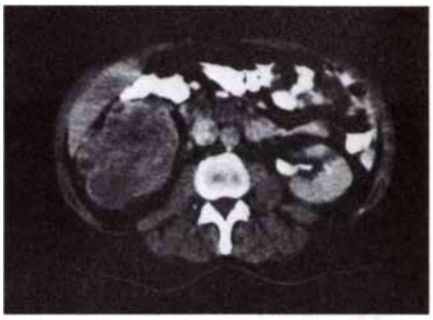
Trial Offer

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 $220 for the full 12 months (renewal is currently $170); 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

5. Hematology

Stanley L. Schrier, M.D., Stanford University School of Medicine

6. Immunology

John David, M.D., Harvard Medical School and Robert B. Brigham Hospital

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

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 Peter Bent Brigham Hospital

10. Nephrology

Roy H. Maffly, M.D., Stanford University School of Medicine and Palo Alto Veterans Administration Medical Center

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

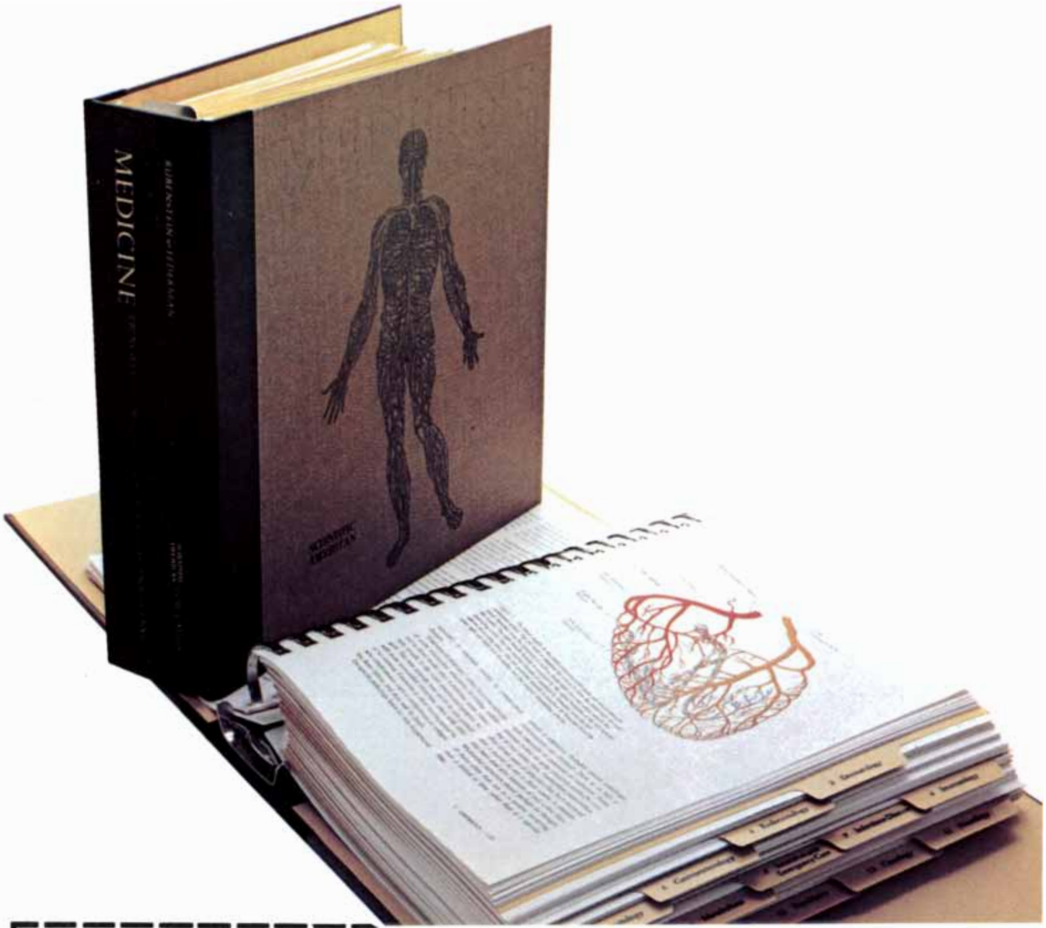
Stephen M. Krane, M.D., Harvard Medical School and Massachusetts General Hospital

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 $220 for the first year of service is tax deductible, as is current renewal at $170. 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 $220*
☐ Please bill me.

* Please add sales tax for California, Illinois, Massachusetts, Michigan, Ohio and New York

Name ______________________________

MD Specialty ______________________________

Address ______________________________

City ______________ State ______________ Zip ______

Signature ______________________________

Subscribers outside of the U.S. and possessions will be charged extra for shipping and handling; updates will be sent by surface routes unless airmail delivery is requested. Please allow 6-8 weeks for delivery. 01


Spektrum
DER WISSENSCHAFT
SCIENTIFIC AMERICAN
POUR LA
SCIENCE
édition française de
SCIENTIFIC AMERICAN
INVESTIGACION Y
CIENCIA
Edición en español de
SCIENTIFIC AMERICAN
サイエンス
SCIENTIFIC AMERICAN
1
1981
日本版
LE SCIENZE
edizione italiana di
SCIENTIFIC AMERICAN
科学
SCIENTIFIC AMERICAN
中译本
SCIENTIFIC AMERICAN
$2.00
March 1981
INTERNATIONAL

tors, even though you may feel no pain.

The floaters derive ultimately from the blood vessels of the retina, an intricate network that actually lies in front of the photoreceptors. One might suppose the vessels would cast shadows on the visual field. They do, but you rarely perceive them. Any constant visual pattern is ignored by the visual system after a few seconds. A classical explanation is that the system becomes fatigued by a constant pattern, which therefore fades from perception. A stationary external object remains perceptible only because the eye constantly moves in small jerks that prevent the formation of a constant image on the retina. Since the circulatory system is fixed to the retina, its shadows remain constant and are therefore imperceptible.

You can see the network under special circumstances. Some morning when you wake up in a sunlit room, look with one eye at a featureless wall or the ceiling. You will see a faint colorless outline of the vessels on your retina, but only for a few seconds. Then the visual system begins to ignore the pattern and it fades from perception. You might be able to see it again if you close your eye briefly and then open it. The network will resemble a tree that sprouts from the "blind spot," the point on the retina where the larger blood vessels come through it from the rear.

The network is often called Purkinje's tree because it was first described by Johannes Purkinje in 1823. Purkinje, a pioneer in physiological optics, was able to outline the circulatory system of his eyes with the aid of a small light source. You can do the same with a pinhole. You might have to experiment with the size of the pinhole, however, to enhance your perception of the network.

Slowly move the pinhole around in front of one eye while keeping the other eye closed. The light goes past the circulatory network in the illuminated eye at constantly shifting angles. The shadow of the network therefore moves across different photoreceptors, with the result that the pattern remains perceptible.

The light source on the other side of the pinhole can be any relatively bright, diffuse source. An illuminated screen of the kind normally employed for viewing photographic slides works well. Even the clear sky will do if you can somehow shield your eye so that it receives light only through the pinhole.

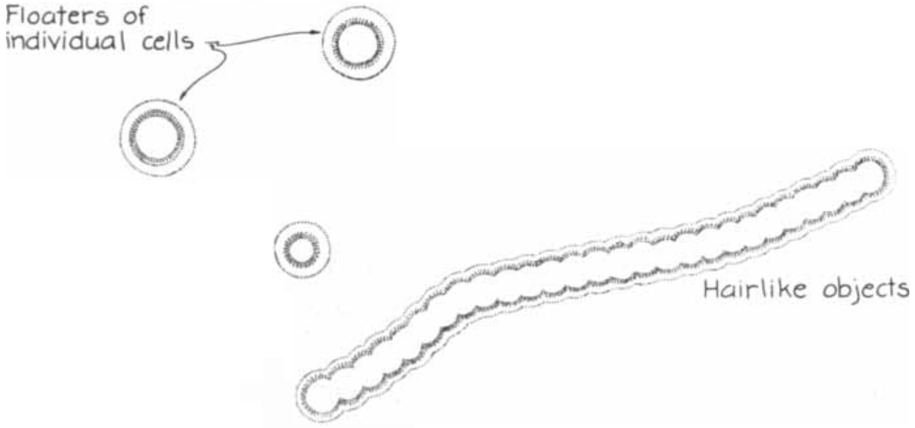
William Oldendorf suggests using a penlight lightly touching a closed eyelid. The light should be positioned on the closed upper eyelid at the outside corner of the eye, just above the junction of the lids. Do not apply pressure to the eye and definitely do not touch the eyeball itself. The light enters through the eyelid and casts a shadow of the circulatory system onto the retina. To keep the image from standing still and fading out move the penlight around in a small circle. Vary the frequency of the movement until you get the clearest impression of the system.

Tom M. Cornsweet states in his book *Visual Perception* that a sharp image of the circulatory system can be seen when the light is provided by optical fibers. Such fibers can be bought from the Edmund Scientific Co., along with a small light source to attach to one end of a bundle of them. The fiber light source works better than a penlight because the fibers are very thin and therefore cast sharper shadows.

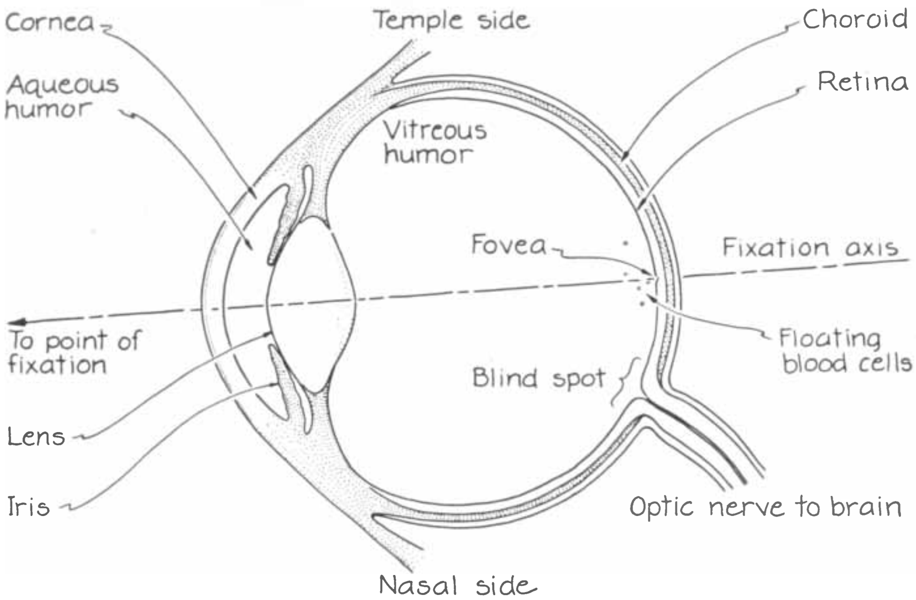
Cornsweet also points out that the best view of the retinal blood vessels can be got if the light is predominantly blue (with wavelengths approximating 415 nanometers). Light at those wavelengths, obtained with a blue filter, is absorbed by the hemoglobin in the red blood cells. If the light consists of only those predominantly blue wavelengths, the blood vessels cast darker shadows. Edmund Scientific sells color filters that serve the purpose.

The standard explanation of why the retinal circulatory system is not normally visible does not really explain the cause of the perceptual fading of a constant signal. Recent work by A. E. Drysdale of the University of Reading suggests that several factors may be involved. The general lack of contrast in the shadows cast by the network may be partly responsible. The fading might also be due to a mechanism in the visual cortex that specifically inhibits perception of the network.

When I stare at a featureless background, I occasionally see another entoptic phenomenon associated with the retinal circulatory system. Tiny specks dart across most of my field of view in a seemingly random manner and then disappear. As with the floaters I cannot bring these objects into focus. Indeed, none of them pass through the fove-



Characteristic shapes of floaters



The location of the blood cells that give rise to the floaters al region in my field of view, which is marked by a slightly grainy texture.

If I hold my gaze as steady as possible, however, by fixing it on some small, distant object, the specks pass repeatedly along the same paths. I do not always see the specks when I look at the sky. I am not sure what it is I do that succeeds in bringing them out, but it probably entails not only fixing my gaze but also relaxing the lenses of my eyes so that they are focused for distance.

The specks are due to the circulation of blood through the retina. If you carefully monitor your pulse as you view them, you will find they appear in correlation with the pulse. The specks are luminous and have relatively dark tails. They always move in single file and none ever overtakes another. Their speed varies; they move faster during the systolic (contraction) phase of the heartbeat and slower during the diastolic (dilation) phase. The specks are thought to be due to the flow of white blood cells through the retinal blood vessels.

You might find the specks more pronounced after physical exercise, since the pulse rate and blood pressure are then increased. The specks might be more visible if you suddenly bend over, thereby changing the blood pressure in your eyes. You can also vary the visibility of the specks by pressing lightly on an eyelid. This action increases the pressure on the retina and therefore increases the pressure in the retinal circulation. A slightly stronger pressing of the eye exerts enough pressure on the retina to make the blood flow decrease, almost eliminating the specks. (If you try this experiment, you may also see the luminous figures called phosphenes that I described last May. They are an entirely different entoptic phenomenon.)

I see the specks best against a clear blue sky. The sky is free of interfering features, and the blue light enhances the speck effect. You can create similar conditions by looking toward a bright featureless background through a blue filter. A filter that passes only light with wavelengths of about 415 nanometers is ideal because this light is absorbed by the hemoglobin in the red blood cells. The light is not absorbed by the white blood cells, so that the motion of the white cells is in sharper contrast. Since the specks arise from the circulation of blood through the retina, they follow the circulation pathways (rather than darting randomly as was my first impression). No specks pass through the foveal region because there are no blood vessels there.

Another entoptic effect can be observed if you view a brightly lighted sheet of white paper while looking alternately through blue and yellow filters. After a switch from yellow to blue a small dark ring surrounds the foveal point in your field of view. Within the ring is a darker spot. Not everyone can perceive these structures and no one sees them for longer than a few seconds. The phenomenon is called Maxwell's spot after James Clerk Maxwell, who is better known for his theories on the nature and propagation of electromagnetic waves. Filters of other colors will work, but one of the filters should pass more blue light than the other. If purple light is passed by the filter replacing the blue one, the spot is pink.

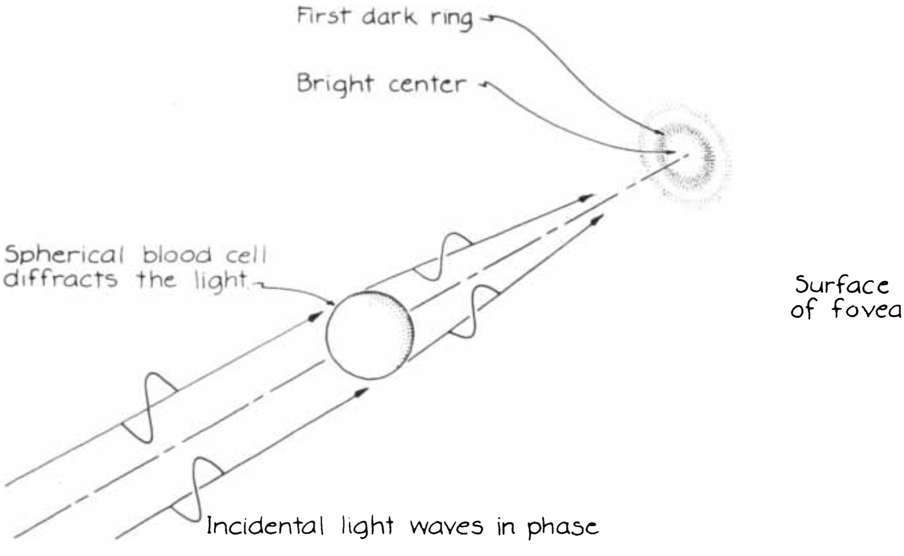
Although the origin of Maxwell's spot is a subject of debate, it appears to be due to the pigmentation of the macula lutea, or yellow spot, that lies across the foveal depression. The location of the macula lutea is not precise, but the spot is most pigmented in that depression. The molecules of the pigment absorb in the blue range; therefore when blue light traverses the macula on its way to the underlying photoreceptors, the light is absorbed. For the few seconds necessary for the eye to adapt to the blue light, Maxwell's spot is apparent. Then it fades. It can be restored briefly if the eye is adapted to light of a different color before it is presented with blue again.

If you examine entoptic phenomena with a pinhole, you are likely to see a variety of other structures. Just after I blink I see bright spots with an ill-defined gray surround. In addition bright and dark horizontal lines gently descend across the entire view. Both effects are caused by the transparent fluid left on the cornea by a blink. The fluid (droplets or a film) refracts the light passing through it. In a droplet the refraction concentrates light toward the center of what should be the geometric shadow of the droplet. Hence the center of the image is bright and the perimeter is relatively dull. The uneven liquid film retreating across the cornea after a blink creates lines of concentrated light along with dimmer ones.

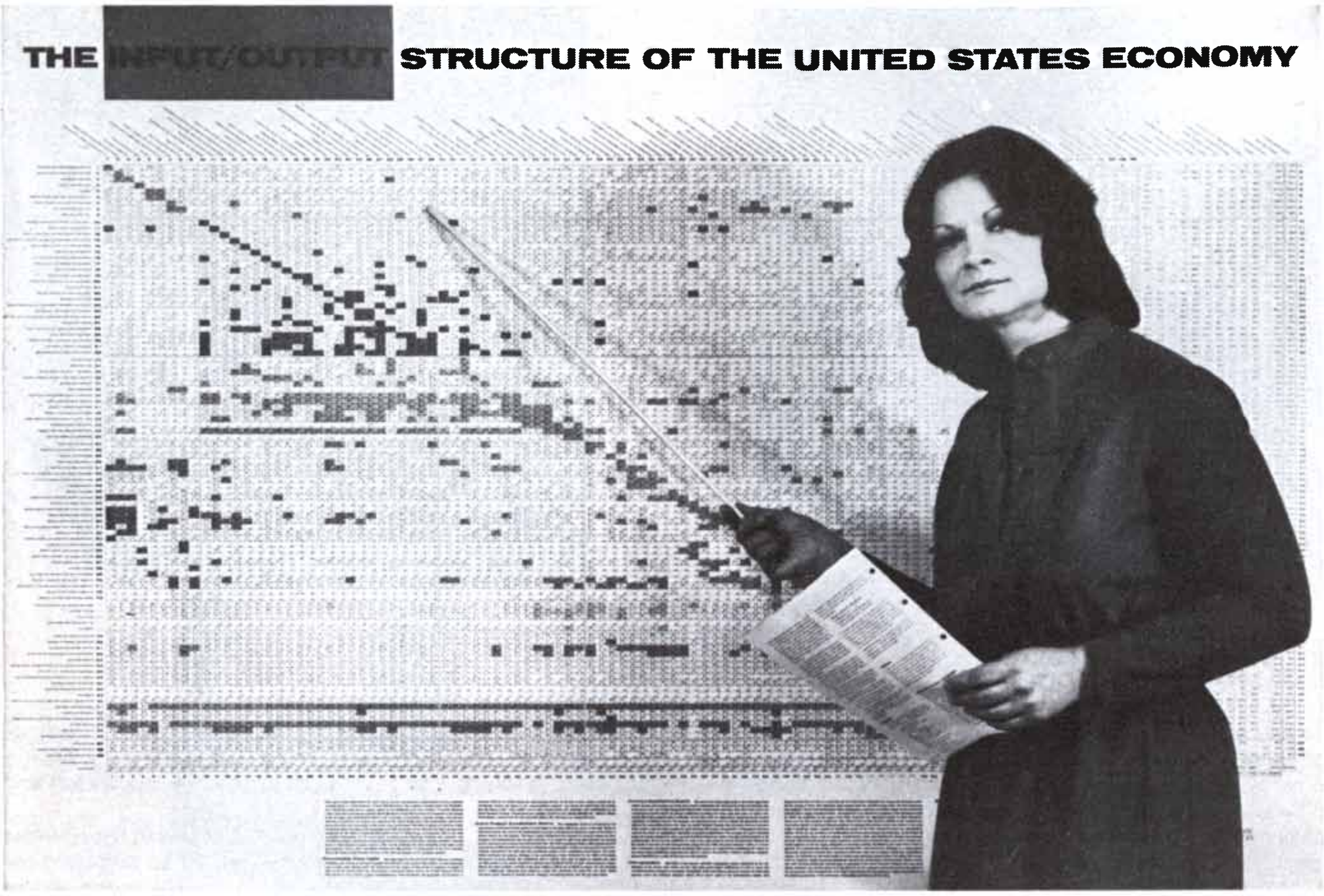
Several other entoptic features were described by Helmholtz, but their interpretation is not complete. You might see small, relatively bright specks that resemble air bubbles. They probably are not the normal floaters; they seem to lack the floater's characteristic pattern of concentric circles. They might be objects in the medium between the cornea and the lens.

You might also see stationary dim and dark lines radiating from the center of your field of view. Helmholtz suggested that these lines result from the radial structure of the lens of the eye. Dark specks, due to small opaque regions in the lens, may also be present. Stationary bright patches with small arms apparently result from seams within the membrane on the front of the lens. The radial structure may originate when the membrane separates from the cornea at an early stage of fetal development. Occasionally I see bright, wrinkled patches floating across my field of view. They appear to be patches of membrane floating in the eye just as the floaters do, but I do not know their source.

A new entoptic phenomenon was recently described by Christopher W. Tyler. Before you open your eyes in the morning you may notice a peculiar spot in your field of view that is either much darker than the rest of the field or (apparently more rarely) brighter than the field. The spot subtends no more than about a degree in the field of view. It is more likely to be noticed near the center of the field but not necessarily in the foveal region. It may suddenly fade and then reappear later, seemingly without cause. Sometimes spots appear as pairs, sometimes as circles and sometimes as



How a blood cell casts a diffraction pattern on the fovea



WHAT MAKES THE U.S. ECONOMY TICK?

The editors of SCIENTIFIC AMERICAN have prepared a wall chart displaying for the 1980's the Input/Output Structure of the U.S. Economy based on the latest interindustry study from the U.S. Department of Commerce.

The SCIENTIFIC AMERICAN Input/Output wall chart does for economics what the table of elements does for chemistry. It answers at a glance questions about the linkage between the microeconomics of the firm and the macroeconomics of the system; about the web of technological interdependencies that tie industry to industry; about the industry-by-industry direct and indirect consequences of swings in public and private spending; about the impact of change in technology, and about any other topic you can think of. You are rewarded by surprise as well as by confirmation of your hunches. For teaching and practical and theoretical studies, here is a powerful, graphic tool.

In the familiar format of the SCIENTIFIC AMERICAN Input/Output wall charts for the 1960's and 1970's, the wall chart for the 1980's measures 65″ × 52″ and is printed in eight colors. Each of the nearly 10,000 cells in the 97-sector interindustry matrix shows (1) the interindustry commodity flow, (2) the direct input/output coefficient and (3) the "inverse" coefficient. Where the direct input/output coefficient exceeds .01, the cell is tinted in the color code of the industrial bloc from which the input comes. This device, combined with triangulation of the matrix, brings the structure of interindustry transactions into graphic visibility.

A supplementary table displays, industry by industry, the capital stock employed; the employment of managerial, technical-professional, white-collar and blue-collar personnel; the energy consumption by major categories of fuel, and environmental stress measured by tons of pollutants.

The editors of SCIENTIFIC AMERICAN are happy to acknowledge the collaboration, in the preparation of this wall chart, of Wassily Leontief, originator of input/output analysis—for which contribution to the intellectual apparatus of economics he received the 1973 Nobel prize—and director of the Institute for Economic Analysis at New York University.

Packaged with the chart is an index showing the BEA and SIC code industries aggregated in each of the 97 sectors.

SCIENTIFIC AMERICAN

415 Madison Avenue, Dept. I/O, New York, N.Y. 10017

YES, send me the new SCIENTIFIC AMERICAN 1980's INPUT/OUTPUT wall chart—printed in eight colors on heavy paper measuring 65″ × 52″.

I understand my chart(s) will be shipped prepaid, uncreased*, flat-rolled and in a sturdy mailing tube. (Allow 4-6 weeks for delivery)

I enclose a check for $______for______chart(s) at $35 each, saving me $5 each in handling charges.

Bill me for $______for______chart(s) at $40 each.

Name____________________________________

Company__________________________________

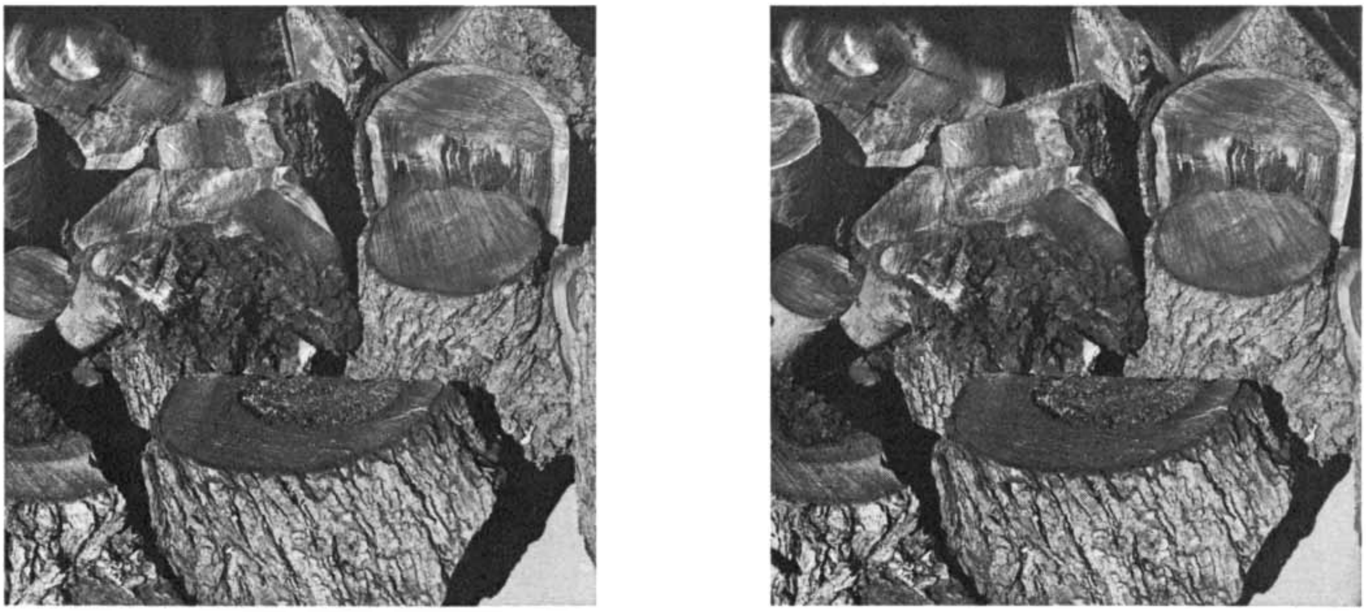
Street Address______________________________

City, State, Zip Code_________________________

4/82

*Charts shipped international require one fold.

Make check payable to SCIENTIFIC AMERICAN. Please add sales tax for California, Illinois, Massachusetts, Michigan, Ohio and New York.



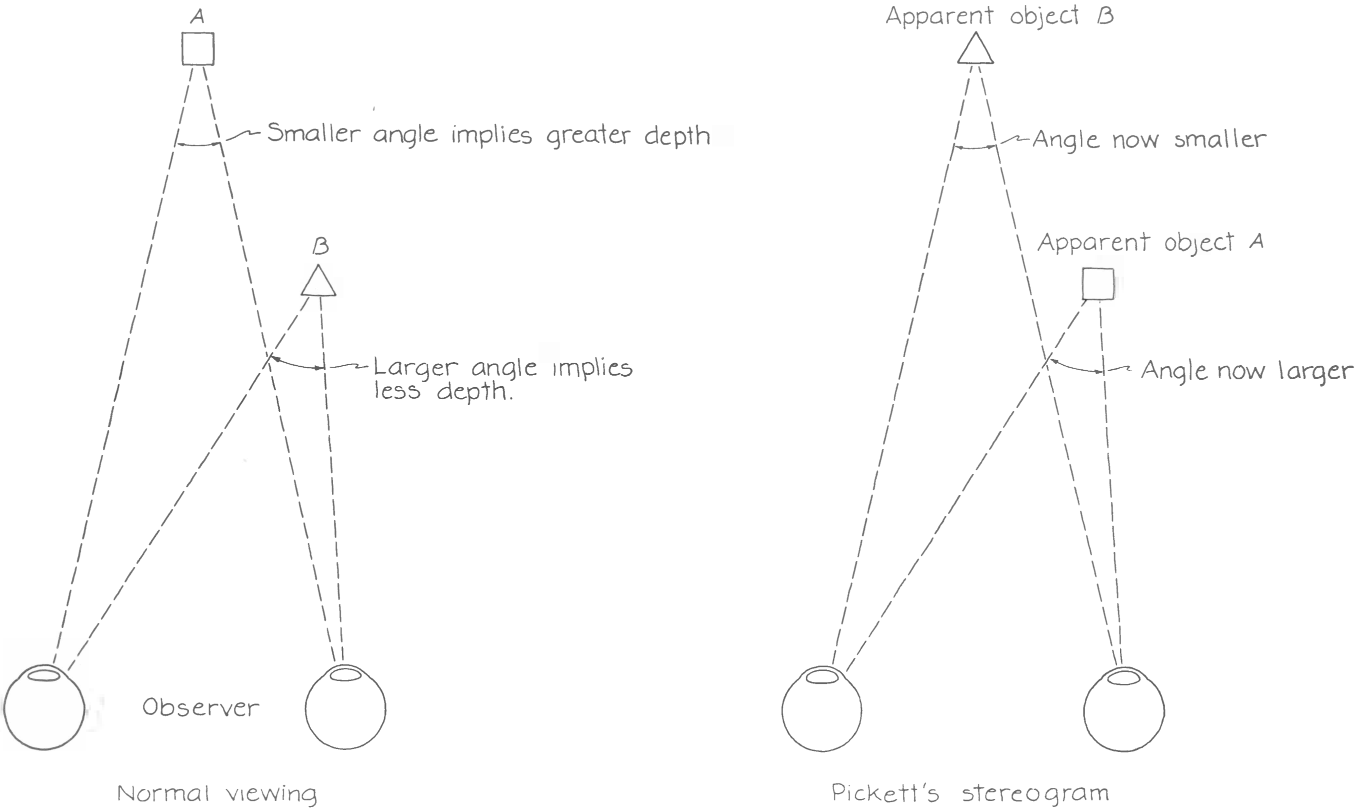
Theodore C. Pickett's "flopped" stereoscopic pair

elongated and irregular shapes. Generally they are distincter if you do not move your eyes.

Tyler calls such a spot a perceptual microaneurysm. It has much the same appearance that a physiological microaneurysm in the retina has to someone who has suffered from one. A microaneurysm is a bulging outward of a blood vessel. It is seen by a physician examining the retina as a tiny black spot. The cause of the perceptual type of black spot is not known.

A novel kind of stereogram has been sent to me by Theodore C. Pickett of Sante Fe, N.M. He photographed a woodpile in a stereographic pair in the conventional manner. Then he mounted the transparencies backward. He did not exchange the transparencies left and right.

Pickett's result appears in the illustration above. You might be able to view the setup without a stereoscope if you are able to stare at the picture until you fuse the separate images. I can do this



Why a flopped stereoscopic pair presents such a strange scene

Introducing High Technology.

The magazine for expanding your technical awareness.

HIGH TECHNOLOGY is the new magazine designed with your needs in mind . . . the needs of the technical professional. You *need* to stay on top of changes in your field but you're on the lookout for the advances in other areas that will help you do your job better. You're an individual who'd like to be on the frontier of *all* of today's changing technologies, but you realize you don't have the time to read everything that interests you.

Now there's HIGH TECHNOLOGY. The magazine that will expand your technical knowledge, with up-to-date, in-depth information on all significant technological breakthroughs. Whether you're striving for specific insights or a broad overview, HIGH TECHNOLOGY will bring you a concise, expert review of advances in computers, space, industry, electronics, energy, medicine, transportation, communication, construction, the military, genetic engineering, robotics and more.

Now's the time to get acquainted with HIGH TECHNOLOGY. Send for your Charter Subscription today and we'll send you a complimentary issue of this timely, new publication. Review your free issue and see if it isn't everything you want it to be. Then, either pay the $18.00 subscription price *or*, write "cancel" on the bill and return it to us. Even if you don't become a HIGH TECHNOLOGY subscriber, you can keep your free issue with no obligation.

Order Your Charter Subscription To HIGH TECHNOLOGY Today.

If the subscription order card and coupon are missing, simply write your name, company name, address on a separate piece of paper. Mail to: HIGH TECHNOLOGY, P.O. Box 2808, Boulder, Colorado 80322.

Special Introductory Offer

12 issues for only **$18.**

5SAB6

FREE ISSUE

If your first issue isn't all you expect, write "cancel" on our bill and return it. Keep the first issue FREE.

NAME

COMPANY MAIL STOP/BLDG.#

ADDRESS

CITY STATE ZIP

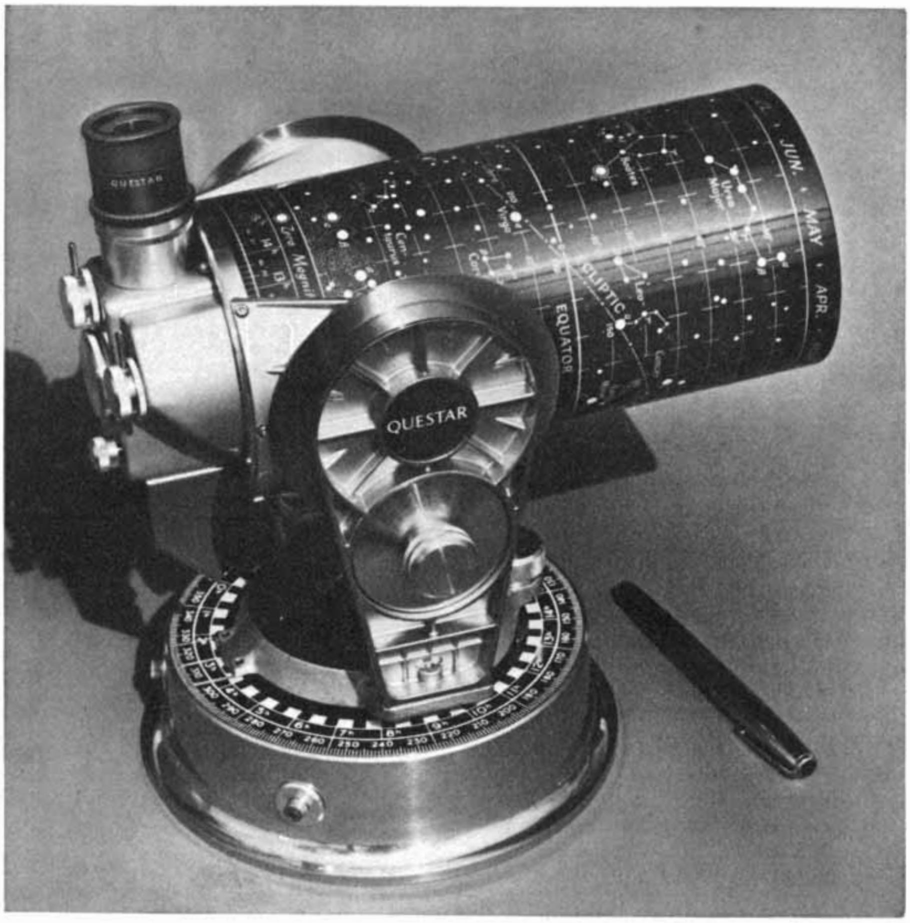
Send no money. We'll bill you later.

Rates good in U.S. only. Please allow 4-6 weeks for delivery. **Your subscription may be tax deductible.**

MAIL TO:

HIGH TECHNOLOGY
P.O. Box 2808,
Boulder, CO 80322

© 1982 SCIENTIFIC AMERICAN, INC



with some difficulty. I try to adjust the convergence of my eyes as if I were looking at something farther away than the photographs. Then I readjust my focus (essentially without thought) until they overlap. (Extra images lie on each side of the overlapping pair, but I ignore them.) As the overlapping pictures slide one over the other in my field of view they suddenly snap together to provide a three-dimensional image.

When photographs are mounted in a stereoscope in the usual way, the observer assigns depth to objects in them by examining their relative positions and sizes. Presumably he makes the same analysis of depth that he would make in looking at the original scene. Although his experience and various clues to depth may help, the critical clue seems to be the angular position of an object as seen by each eye.

Suppose two objects lie in the field of view. One object (call it *A*) is farther away than the other (*B*), but neither is notably distant. When the eyes view *A*, the lines of sight converge with a certain angle between them at *A*. When *B* is viewed, the lines of sight converge on *B* with a larger angle because *B* is closer. From the difference in the two angles the observer attributes depth to the two objects.

When the scene is photographed stereoscopically, the position of the camera is shifted between exposures to match the separation between the eyes. The photographs are then placed in a stereoscope, which facilitates the fusion of the two images. The scenes mimic the original scene. Although the photographs are flat representations of the scene, the slightly different views of the same objects cause the observer to assign depth as before. As before, the lines from the eyes to the apparent object *A* subtend a smaller angle than the lines from the eyes to the apparent object *B*. The observer therefore attributes greater depth to *A*.

Many investigators have experimented with a left-and-right exchange of photographs, which results in an inversion of depth so that truly distant objects seem to be nearby and vice versa. Pickett's stratagem plays a similar trick on the observer. The objects are reversed left and right in each photograph. The lines to the apparent objects *A* and *B* now converge with a larger angle for *A* than for *B*. As a result the truly distant object is interpreted as being relatively nearby and the truly nearby object seems to be distant.

Photographs made and mounted following Pickett's directions can create a perplexing view of the world. Some of the strangeness results from a conflict between the observer's experience and what he sees. Depth is assigned not only according to the apparent angles of objects but also according to the relative size of the images on the retina. If two identical objects are viewed when one is closer to the observer than the other, depth is easily assigned because the more distant object creates a smaller image on the retina. When a stereographic pair is viewed in Pickett's setup, this clue conflicts with the strong clue from the analysis of angles. The angular analysis usually prevails, but the conflict in clues creates an eerie sight for the observer.

The assignment of depth according to angles dominates in normal vision until the objects are quite distant (perhaps 450 meters away). Then either experience or the relative size of the retinal images determines the assigned depths. Usually objects farther away than about 450 meters are perceived by the observer as lying on a flat plane at approximately that distance.

The most startling of Pickett's photographs include both distant and nearby objects. When a stereographic pair containing such objects is viewed in Pickett's reverse mounting, the distant objects no longer appear to lie on a relatively flat and distant plane. They seem to be closer and to have greater disparity of depth. The truly nearby objects are relegated to the distant flat plane and appear to have much less disparity in depth.

If you would like to experiment with Pickett's procedure, you might try photographic scenes with objects that have a considerable range of depth. A view over a valley with both nearby and distant trees might be ideal. To make such a picture, however, you will have to follow certain special procedures. Increase the separation between the two positions of the camera by about 17 feet for each mile of range to the most distant of the objects you will be photographing. Exclude from your prospective picture anything that is nearer than 25 percent of that maximum distance, otherwise the stereographic pair will yield a double image of the nearer object. Finally, arrange for some distant object to be in the center of each photograph of the stereographic pair.

In my November column I described a method of "pinspeck" photography devised by Adam Lloyd Cohen. It turns out that the principle underlying pinspeck photography was first, and quite independently, demonstrated by Ronald Cowart of the University of Texas and described in his doctoral dissertation as "the inverse pinhole camera." After further development by Alfonso Zermeno, James M. Hevesi and Lee M. Marsh the device was patented in 1978. The patent is held by the University of Texas and all patent rights are currently licensed to Texas Medical Instruments, Inc., of San Antonio.

EXTRAORDINARY OFFER
50% OFF

LEARN SPANISH, FRENCH, ITALIAN OR GERMAN LIKE A MEMBER OF THE U.S. FOREIGN SERVICE CORPS.

This course (*specifically designed by the U.S. Department of State for Diplomats with little or absolutely no knowledge of the language to be studied*) is **complete, unedited** and includes the official Foreign Service Institute book with text for basic and advanced instruction. It begins in English and will gradually build your skills until you can comfortably handle the entire lesson in your new language. In a matter of weeks, you could be speaking your new language like a member of the U.S. Foreign Service Corps.

ORDER ONE COURSE AT $110.00 AND GET ONE OTHER AT HALF PRICE!

What an incredible value! For the single purchase price of one course your whole family can learn a new language and keep the course for years of fun and future reference. **And, now, for a limited time, you can get the second course for 50% off.** (The advanced course, or even another language!) The complete full length Foreign Service Language Course is now being offered in two volumes:

VOL. 1, BASIC SPANISH: 15 CASSETTES (19 hours) and 464 pg. Foreign Service Institute text **$110.00**

VOL. 2, ADVANCED SPANISH: 21 CASSETTES (27 hours) and 464 pg. Foreign Service Institute text **$110.00**

(The number of cassettes in FRENCH, GERMAN and ITALIAN Courses vary slightly from the SPANISH Course.)

FREE TRIAL OFFER: Try this course for three weeks. If you aren't convinced it's the fastest, easiest, most economical way to learn a new language, return the course for a full refund with no questions asked. Order yours right now! For immediate delivery, fill in the information on the bottom of this ad and mail it with your check, Money Order or charge to your AmEx, Visa or Master Card by enclosing card number, expiration date and signature.
Order both BASIC and ADVANCED course in any language and get 50% off the second course!

Mail to: AMPRO PRODUCTIONS
150 Fifth Avenue, New York, NY 10011.
NOW you can order by phone 24 hours a day!
Call collect (9AM–6PM EST) (212) 243-7726
New York state residents add proper sales tax.
N.J., N.Y., CT. residents can also buy direct at our Fifth Avenue office (at 20th Street)
Fill in the line below with the language you wish to learn . . . then check basic, advanced or both.

SATISFACTION GUARANTEED MONEY BACK OFFER.

(Language) ____________________
☐ Basic $110.00 ☐ Advanced $110.00 (Anyone ordering 2 courses may order additional gift courses at $75.00 each.)

SC3

Through a revolutionary recording system beyond direct disk

Roger Williams and Peter Nero can play *your* piano. . .like magic.

Without those awkward piano rolls! In fact, you can surprise your guests with performances by many of the world's great pianists, in your living room, on *your* piano. That's the magic of the PIANOCORDER® reproducing system, a musical breakthrough that makes the *rinky-tink* sound of the "old fashioned player piano" disappear.

Hidden away inside your piano, the PIANOCORDER uses computerized cassette tapes to bring the keys and pedals to life, recreating the original performer's style with unbelievable accuracy.

You can even record your own songs and play them back instantly, note for note, a feature that makes the PIANOCORDER invaluable for learning *and* entertaining.

What's more, our extensive tape library contains over 3,000 superb performances, from classical to jazz, ragtime to pop. And we're constantly adding new music by today's top pianists—Roger Williams, Peter Nero, George Shearing and many others.

The PIANOCORDER system can be installed in almost any existing piano. If you're considering a new piano, you can also purchase our fine MARANTZ reproducing piano with the system already built in.

Write or dial toll-free 1-800-447-4700 (in Illinois, 1-800-322-4400) for information on the remarkable PIANOCORDER system from MARANTZ.

ROGER WILLIAMS & PETER NERO CONCERT TAPES NOW AVAILABLE

Pianocorder®
Reproducing System

marantz®
reproducing piano

Marantz Piano Company, Inc. • Post Office Box 460 • Highway 64-70 • Morganton, North Carolina 28655 • (704) 437-7135

© 1982 SCIENTIFIC AMERICAN, INC

BIBLIOGRAPHY

Readers interested in further explanation of the subjects covered by the articles in this issue may find the following lists of publications helpful.

METAMAGICAL THEMAS

Chopin in His Own Land (Chopin w kraju). Edited by Krystyna Kobylańska. Polish Music Publishers, 1956.

The Chopin Companion. Edited by Alan Walker. The Norton Library, 1966.

Chopin: The Man and His Music. James Huneker. Dover Publications, Inc., 1966.

Chopin. George R. Marek and Maria Gordon-Smith. Harper & Row Publishers, Inc., 1978.

ADVANCED OFFSHORE OIL PLATFORMS

Outer Continental Shelf Frontier Technology: Proceedings of a Symposium. Marine Board, Assembly of Engineers. National Research Council, National Academy of Sciences, 1980.

Tethered Buoyant Platform Production System. Graham Perrett and Robert M. Webb in *Proceedings of the Offshore Technology Conference,* Vol. 4, pages 261–274; 1980.

Recent Advances in Design and Construction of Deepwater Capability. Griff C. Lee in *Ocean Industry,* pages 71–80; November, 1980.

Introduction to Offshore Structures. W. J. Graff. Gulf Publishing Company, 1981.

NUTRIENTS THAT MODIFY BRAIN FUNCTION

Nutrition and the Brain. Edited by Richard J. Wurtman and Judith J. Wurtman. Raven Press, 1977–82.

Release of Acetylcholine from the Vascular Perfused Rat Phrenic Nerve–Hemidiaphragm. George C. Bierkamper and Alan M. Goldberg in *Brain Research,* Vol. 202, No. 1, pages 234–237; November 24, 1980.

Precursor Control of Neurotransmitter Synthesis. R. J. Wurtman, F. Hefti and E. Melamed in *Pharmacological Reviews,* Vol. 32, No. 4, pages 315–335; December, 1980.

Alzheimer's Disease: A Report of Progress in Research. Edited by S. Corkin, K. J. Davis, J. H. Growdon, E. Usdin and R. J. Wurtman. Raven Press, 1981.

Carbohydrate Craving in Obese People: Suppression by Treatments Affecting Serotoninergic Transmission. Judith J. Wurtman, Richard J. Wurtman, John H. Growdon, Peter Henry, Anne Lipscomb and Steven H. Zeisel in *International Journal of Eating Disorders,* Vol. 1, No. 1, pages 2–15; Autumn, 1981.

A MAJOR EARTHQUAKE ZONE ON THE MISSISSIPPI

Geological Investigation of the Alluvial Valley of the Lower Mississippi River. Harold N. Fisk. Mississippi River Commission, Vicksburg, Mississippi, U.S. Army Corps of Engineers, 1944.

The Mississippi Valley Earthquakes of 1811 and 1812: Intensities, Ground Motion and Magnitudes. Otto W. Nuttli in *Bulletin of the Seismological Society of America,* Vol. 63, No. 1, pages 227–248; February, 1973.

The New Madrid Earthquakes of 1811–1812. James Penick, Jr. University of Missouri Press, 1976.

Recurrent Intraplate Tectonism in the New Madrid Seismic Zone. M. D. Zoback, R. M. Hamilton, A. J. Crone, D. P. Russ, F. A. McKeown and S. R. Brockman in *Science,* Vol. 209, No. 4460, pages 971–976; August 29, 1980.

Investigations of the New Madrid, Missouri, Earthquake Region. Edited by F. A. McKeown and L. C. Pakiser. *U.S. Geological Survey Professional Paper 1236,* 1981.

GIANT MOLECULAR-CLOUD COMPLEXES IN THE GALAXY

Giant Molecular Clouds in the Galaxy. Edited by P. M. Solomon and M. G. Edmunds. Pergamon Press, Inc., 1980.

Interstellar Molecules. Edited by B. H. Andrew. D. Reidel Publishing Company, 1980.

The Milky Way. Bart J. Bok and Priscilla Bok. Harvard University Press, 1981.

THE GIANT SQUID

Pirates of the Deep—Stories of the Squid and Octopus. Paul Bartsch in *Smithsonian Institution Annual Report,* pages 347–375; 1917.

A Review of the Systematics and Ecology of Oceanic Squids. M. R. Clarke in *Advances in Marine Biology,* Vol. 4, pages 93–300; 1966.

A Specimen of the Giant Squid *Architeuthis* from New Zealand. R. K. Dell in *Records of the Dominion Museum in Zoology,* Vol. 7, No. 4, pages 25–36; 1970.

First Record of Juvenile Giant Squid, *Architeuthis* (Cephalopoda, Oegopsida). C. F. E. Roper and R. E. Young in *Proceedings of the Biological Society of Washington,* Vol. 85, No. 16, pages 205–222; 1972.

Giant Squids. K. N. Nesis in *Gidrobiologiia,* Vol. 6, No. 706, pages 55–60; 1974.

SUPERHEAVY MAGNETIC MONOPOLES

Magnetic Monopoles. Kenneth W. Ford in *Scientific American,* Vol. 209, No. 6, pages 122–131; December, 1963.

Constraints on Controversy: The Case of the Magnetic Monopole. Andrew Pickering in *Social Studies of Science,* Vol. 11, No. 1, pages 63–93; February, 1981.

VARIATIONS IN MEDICAL CARE AMONG SMALL AREAS

Costs, Risks and Benefits of Surgery. Edited by John P. Bunker, Benjamin A. Barnes and Frederick Mosteller. Oxford University Press, 1977.

A Small Area Approach to the Analysis of Health System Performance. John E. Wennberg and Alan M. Gittelsohn. Department of Health and Human Services. U.S. Government Printing Office, 1980.

The Need for Assessing the Outcome of Common Medical Practices. John E. Wennberg, John P. Bunker and Benjamin Barnes in *Annual Review of Public Health,* Vol. 1, pages 277–295; 1980.

THE SHORT LIFE OF ÉVARISTE GALOIS

La vie d'Évariste Galois. Paul Dupuy in *Annales scientifiques de l'École normale supérieure,* Vol. 13, pages 197–266; 1896.

Écrits et mémoires mathématiques d'Évariste Galois. Edited by Robert Bourgne and J.-P. Azra. Gauthier-Villars, 1962.

Évariste Galois. René Taton in *Dictionary of Scientific Biography,* edited by Charles Coulston Gillispie. Charles Scribner's Sons, 1972.

Genius and Biographers: The Fictionalization of Évariste Galois. Tony Rothman in *The American Mathematical Monthly,* Vol. 89, No. 2, pages 84–106; February, 1982.

THE AMATEUR SCIENTIST

Entoptic Images and Related Phenomena. G. S. Brindley in *Physiology of the Retina and Visual Pathway.* The Williams & Wilkins Company, 1970.

The Visibility of Retinal Blood Vessels. A. E. Drysdale in *Vision Research,* Vol. 15, No. 7, pages 813–818; July, 1975.

BMW. ONE OF THE FEW CAR COMPANIES NOT CURRENTLY INTRODUCING IMITATION BMW'S.

The annual new-car introductions have ushered in cars so strongly resembling BMW's as to suggest new meaning for the term 'coincidence.'

And a good many of these newcomers claim to perform 'like' a BMW.

This sudden enthusiasm for BMW-like performance is rather amusing.

Because while it's possible to develop enthusiasm virtually overnight, the development of engineering takes considerably longer—a fact that's amply demonstrated by the BMW 320i.

THE BMW 320i VS. UNREASONABLE FACSIMILES.

The 320i is engineered according to a belief that hasn't changed for decades:

Namely, that performance is, and will always be, the ultimate measure of a car's worth—a belief that permits a car to grow by gradual refinement, instead of hasty responses to changing fashions.

The current preference for leaner, more efficient engines, for example, has caused no radical reshuffling of priorities at BMW. The 320i is powered by a 1.8-liter engine equipped with K-Jetronic fuel injection—and has been for years.

Nor has the 320i needed any instant re-engineering to deliver better mileage. It delivers an EPA-estimated $\boxed{25}$ mpg, 36 mpg highway, figures that would be respectable even in an economy car.

(Fuel efficiency figures are for comparison purposes only. Your actual mileage may vary, depending on speed, weather and trip length. Your actual highway mileage will most likely be lower.)

In fact, despite all the flurry of emulation, the 320i—with its fully independent suspension, precise steering, and decades of technological refinements bred on racecourses—remains unique.

So much so as to awaken an almost unseemly ardor in automotive critics who are implacable for a living.

One nominates the 320i as "the sort of car enthusiasts turn into legend." He might have added that there are no reasonable facsimiles for legends.

INVEST IN AN ORIGINAL INSTEAD OF A REPRODUCTION.

Originals generally bring their owners higher prices than reproductions. Which could help account for the following:

According to the January 1982 NADA Used-Car Guide, while the value of other cars was dwindling considerably, the average 320i sold over the past 3 years retained a phenomenal 96.9% of its original price.

All things considered, the appearance of so many BMW surrogates is perfectly understandable.

If you suspect you might not be content to spend the next several years with a surrogate, we suggest you contact your nearest BMW dealer for a test drive.

THE ULTIMATE DRIVING MACHINE.

BMW, MUNICH, GERMANY.

Making sure everything

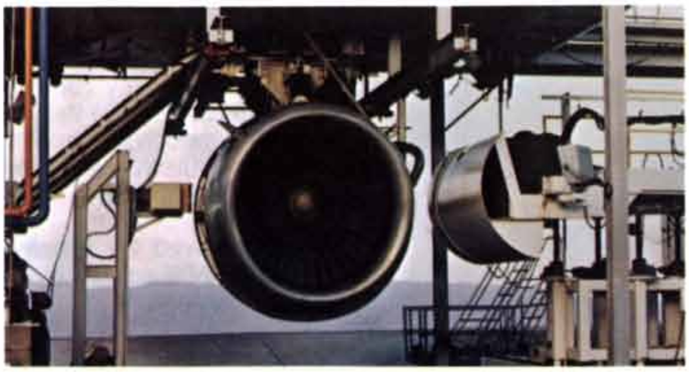
Lockheed knows how.

Across the U.S., there is a rising, insistent cry for more quality in products. At Lockheed, advances in quality are a way of life. When a fleet ballistic missile is launched, as shown above, there can be no compromise in quality. And that applies to aircraft, spacecraft, and a host of other Lockheed products large and small. The examples below cover some of the advances in quality that Lockheed has made recently in keeping pace with scientific progress.

Giving flaws no place to hide.

The cost and time-consumption of inspecting products like rocket motors and gas turbine engines can be staggering.

With rocket motors, it's a search for cracks, voids, flawed bonding, or other critical defects. With gas turbine engines, the vital subject is gas-seal clearances, because these affect fuel efficiency. The engine testing is done at the development stages and the gas seals redesigned if necessary.



Gas turbine engine under Lockheed X-ray inspection.

Until recently, such testing combined X-ray with photographic techniques. With rocket motors, however, inspecting by this photographic imagery method required several hundred separate pictures, each needing incremental rotations of the product, exposure times, individual film processing, and painstaking review of every picture.

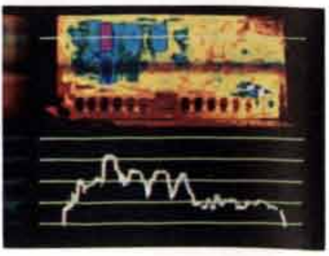
So Lockheed has developed a real-time X-ray imaging system that permits uniform, 100% inspection (rather than sampling) at far faster speed and lower cost than conventional X-ray methods. Users of the Lockheed system report savings of up to 50% in time and 66% in X-ray materials costs.

Spotting trouble before it becomes a problem.

You're looking at potential trouble in the photograph below. Actually, it's an infrared photograph of a printed circuit board. The magenta coloring identifies an integrated circuit on the board that may be too hot.

Until recently, it was almost impossible to verify the thermal integrity of the hundreds or even thousands of PC boards that go into a modern aircraft. Now Lockheed has developed a unique method of evaluating PC boards completely. This new approach involves thermal imaging. It is a non-contact, non-destructive way of testing electronic equipment.

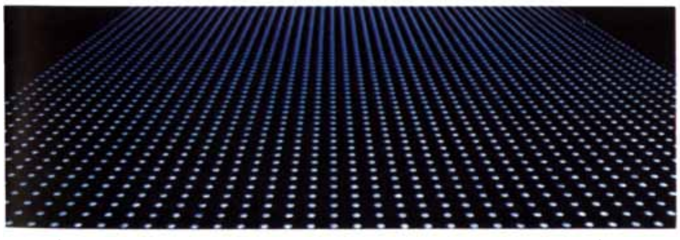
The actual test of a PC board, using infrared photography, can be made in a few minutes. Once a "hot" spot is detected, early steps can be taken to correct it.



Lockheed's thermal survey also has a long-term benefit. The infrared photograph provides a permanent temperature ecord that may be compared with future design changes in he PC board. The result: far higher reliability over the life of hese assemblies.

Found: a way to test 200,000 holes.

A big airlifter or jetliner may have as many as 200,000 noles in its airframe. Mainly rivet and fastener holes, they all

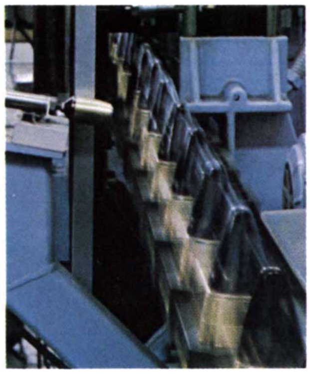


nust be checked. And rechecked after the aircraft has been n service. It's been a costly, time-consuming process.

Now Lockheed has developed two unique systems that ut costs and time enormously while increasing reliability of esting. One, a Capacitance Hole Probe tests newly drilled noles. A probe with 48 sensors is inserted into the hole. The ensors check the hole on six different layers, and a computer makes a written record of the test—all in nine seconds. ormerly, it took several minutes to test a hole with less eliability.

Lockheed has not neglected old holes either. In maintenance inspections, a computer-based Eddy Current Scanner ests an old hole for cracks in two minutes. The recent nanual method took 10 minutes with less reliability. The scanner also uses a sensor probe and makes a computer record of the findings.

50,000 fuzes a month.

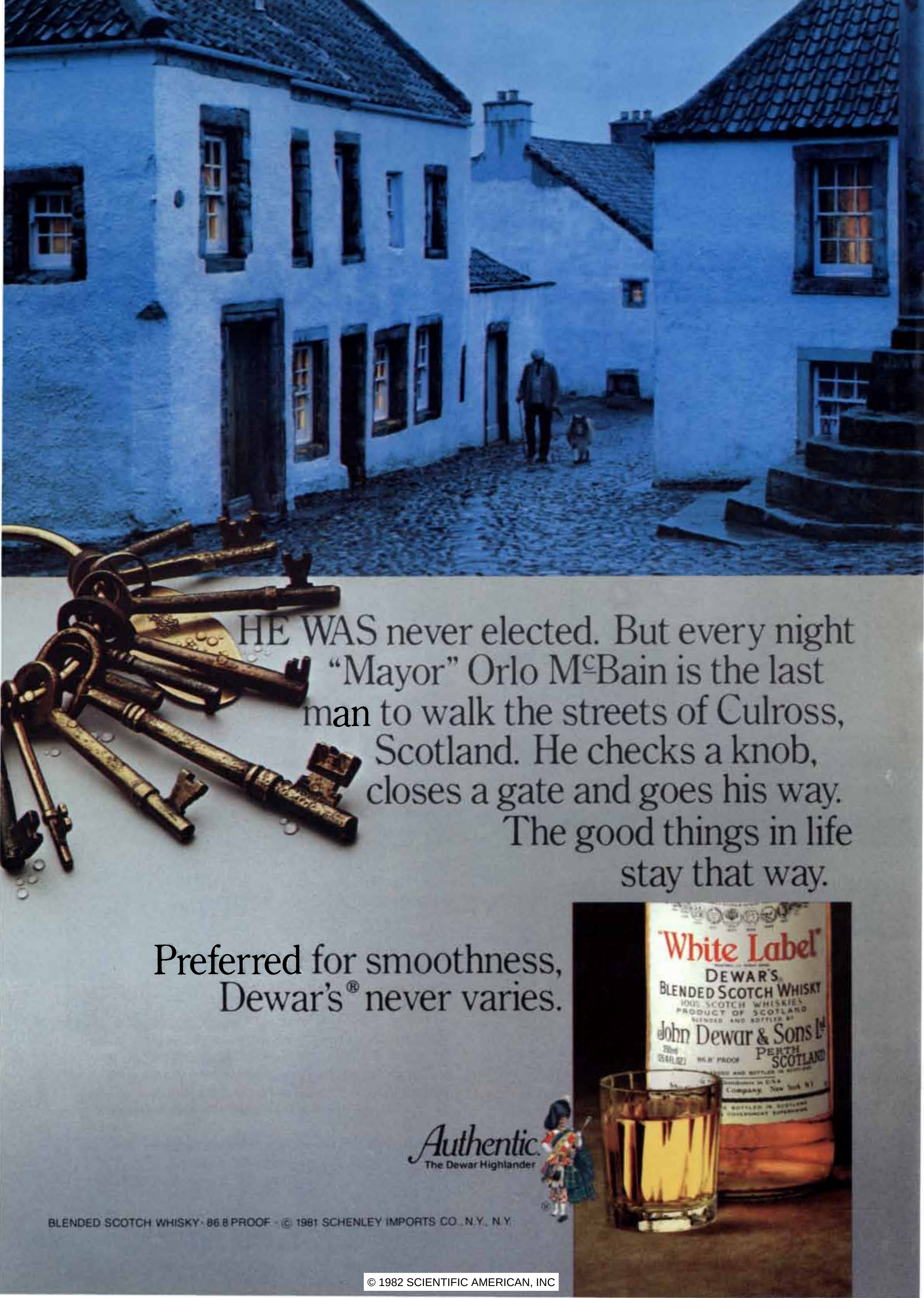


Aircraft and spacecraft roll down assembly lines in relatively small numbers. But Lockheed also is involved in mass production at a unique facility in Denville, New Jersey. There Lockheed produces 50,000 fuzes a month for U.S. Army shells. It already has manufactured more than 1,000,000 fuzes.

Designed by Lockheed under U.S. Army direction, the facility could produce well over 100,000 fuzes monthly. It is highly mechanized, and its extensive testing program is automated. Computerized stations test 100% of all electrical components as they arrive at the facility. Then, all subassemblies and final assemblies are tested. Lockheed believes this may be the most rigorous testing program of this type in the country. Result: highly reliable fuzes, produced in great quantity with enhanced productivity.

For more information about Lockheed, write for the 1981 Annual Report Highlights to A.F. Melrose, Lockheed Corporation, P.O. Box 551, Burbank, CA 91520.

Lockheed


"White Label"
DEWAR'S
BLENDED SCOTCH WHISKY
100% SCOTCH WHISKIES
PRODUCT OF SCOTLAND
BLENDED AND BOTTLED BY
John Dewar & Sons Ltd
PERTH
SCOTLAND
86.8° PROOF