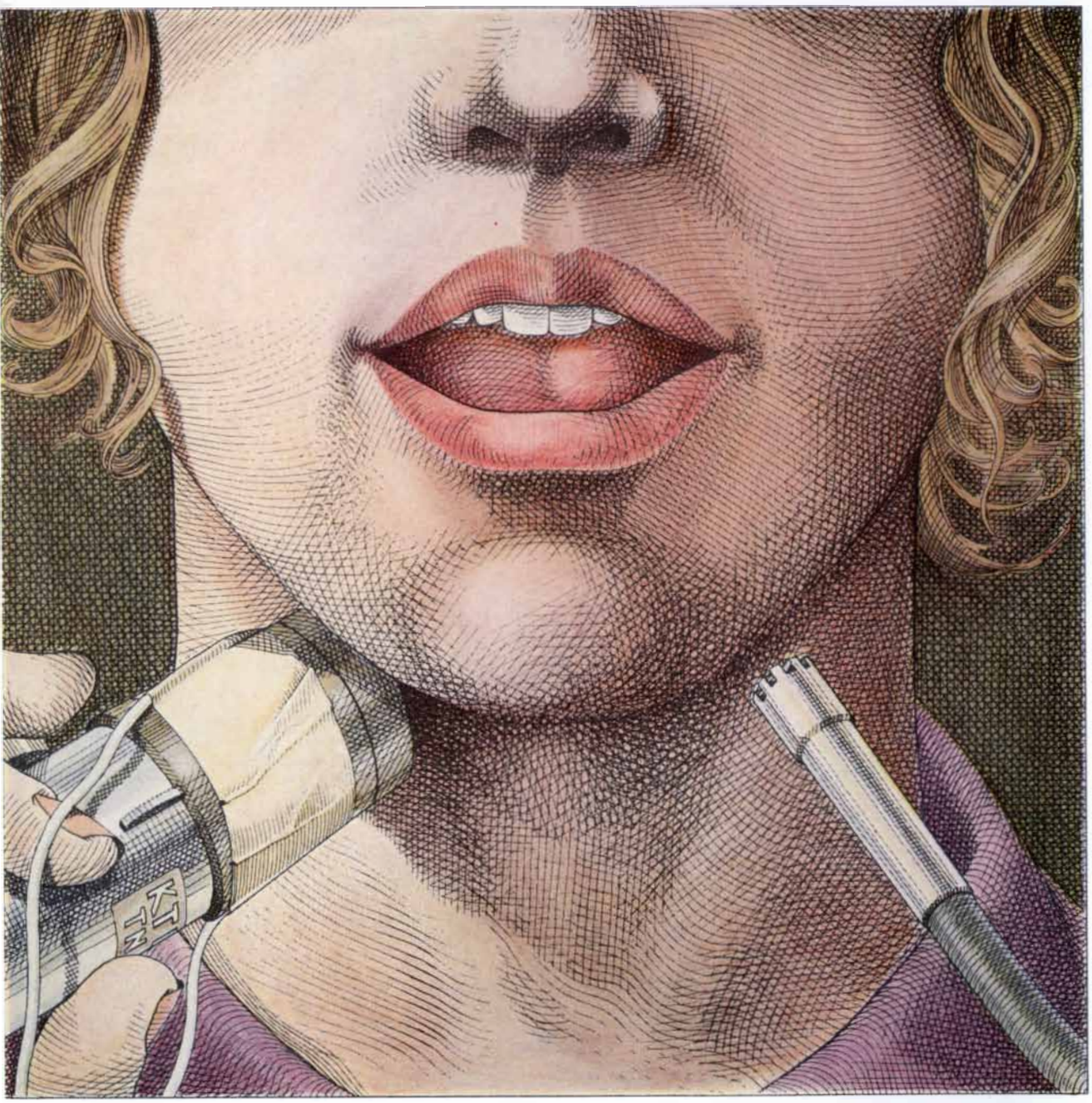
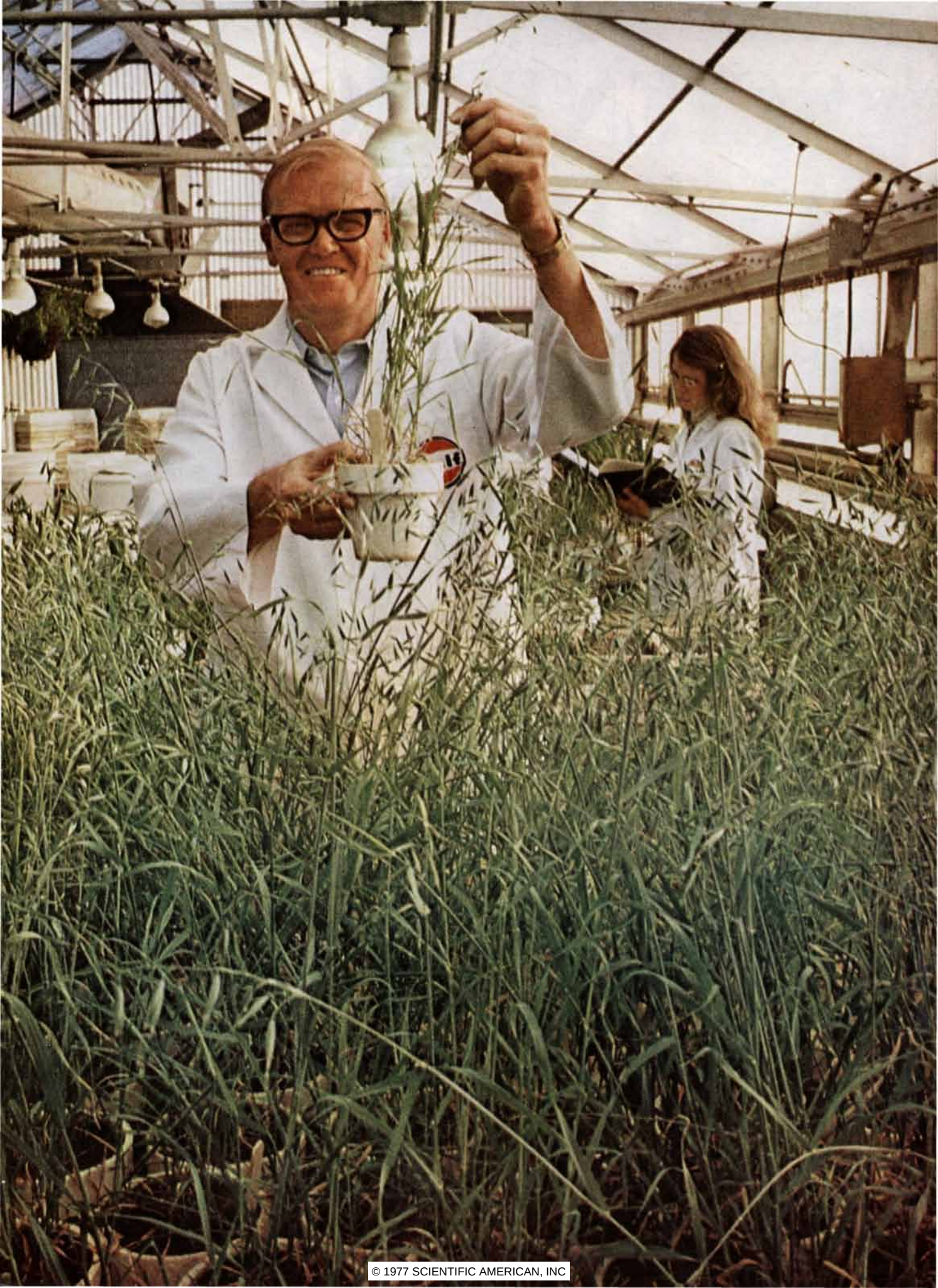
SCIENTIFIC AMERICAN



ACOUSTICS OF THE SINGING VOICE

$1.50

March 1977



"Last week I sowed my wild oats. The real fun starts in two weeks, when I try to kill them all."

"Wild oats always make people think of second-rate Victorian novels," says Gulf agricultural chemist Jim Ahle.

"Actually they're big trouble in North and South Dakota. They crowd out the wheat, and cost the farmers there over two hundred million dollars a year.

Single-minded

"A lot of things will kill wild oats. What we wanted was a single-minded herbicide that won't harm any other plants, or the people or animals who eat the grain.

"To solve that problem, we developed a product called Carbyne®. It proved to be one of the most successful wild oat controls ever developed.

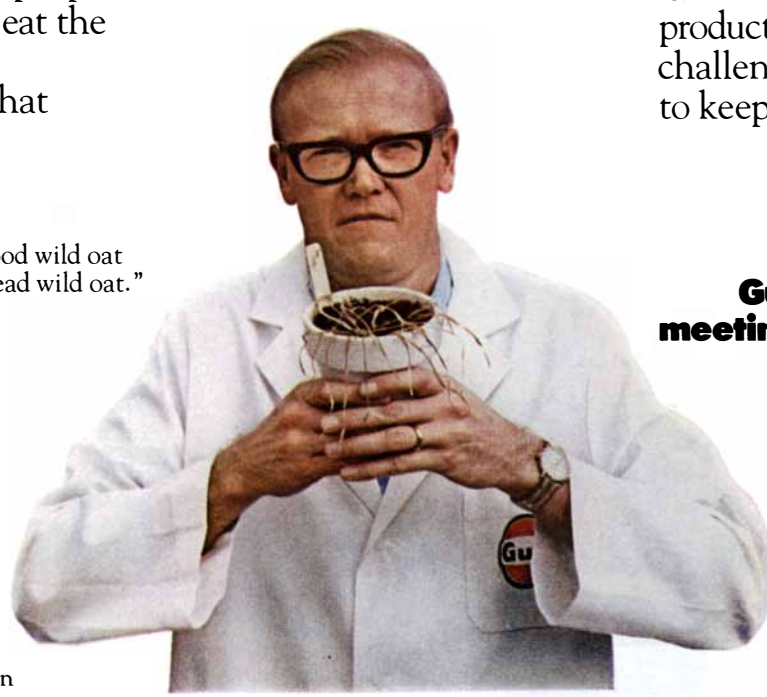
"That was several years ago, but we're still running experiments with Carbyne, either new formulations or in combination with other herbicides, so you can control a variety of weeds with one application.

"For each new formulation, we sow thousands of wheat plants and wild oats in seed beds, and in a few weeks spray them with the experimental herbicide. Then we check to make sure the weeds are dead and the grain uncontaminated.

One in 7,000

"For every herbicide that comes onto the market, at least seven thousand have been tested and rejected for one reason or another. Carbyne is a winner, but it's just one of many herbicides and pesticides we're working on here at Gulf.

"The American farmer is the most productive on earth. The challenge, as I see it, is to keep him that way."



"A good wild oat is a dead wild oat."

Gulf

Gulf people: meeting the challenge.

Gulf Oil Corporation

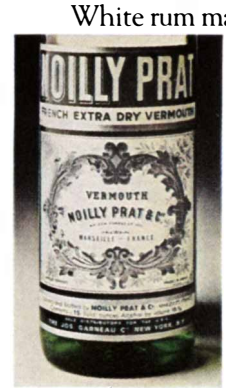
"Defy tradition. Make your next martini with white Puerto Rican rum. I recommend it emphatically."

FRANÇOIS BONNARDEL, PRESIDENT, NOILLY PRAT & CIE.

As the long-standing partner of gin and vodka, Noilly Prat is uniquely qualified to make a value judgement about martinis.

The fact that they recommend a martini made with white rum will come as a surprise only if you've never tasted one.

White rum makes a smoother martini.



Smoother than a gin martini. Smoother than a vodka martini. The Government of Puerto Rico has seen to that by requiring all white rum to age in white oak casks for at least a year.

Noilly Prat heartily approves. They know what a difference aging makes to the taste of a drink, including vermouth. After all, dry vermouth was invented by a Frenchman named Noilly.

Noilly Prat also knows that in a large-scale taste test against gin and vodka, white rum emerged the clear favorite in taste and smoothness.

To the Noilly Prat people, who care passionately about the taste of the martini, the handwriting is on the wall—just as it was when vodka surpassed gin not too long ago.

Tonight defy tradition. Combine five parts white rum from Puerto Rico with one part Noilly Prat vermouth.

We think you'll find it hard to go back.

PUERTO RICAN RUMS

For free "White Rum Classics" recipes, write: Puerto Rican Rums, Dept. SC-3, 1290 Ave. of the Americas, N.Y., N.Y. 10019

Established 1845 **SCIENTIFIC AMERICAN** March 1977 Volume 236 Number 3

ARTICLES

DEPARTMENTS

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: $18 PER YEAR, U.S., ITS POSSESSIONS AND CANADA; $22 PER YEAR, ALL OTHER COUNTRIES.

THE LONGER W
THE LONGER

In 1966, the average life expectancy of a Volvo in Sweden was 12 years. Latest figures show it's up to 16 years.

If how long a car lasts is any indication of how well it's built, you'll be interested to know that Volvos have been lasting longer.

Latest figures from the Swedish Motor Vehicle Inspection Company show that in the last ten years, the average life expectancy of a Volvo in Sweden has increased by 37%.

We can't guarantee a Volvo will last the same length of time in this country. But it's nice to know the potential for survival is there.

Of course, reducing the speed with which a Volvo travels to the junkyard isn't the only improvement we've made in ten years. We've also found a lot of ways to make that trip safer and more enjoyable.

We've introduced rack and pinion steering because it's surer and more precise.

We've improved our entire suspension system. So you get a smooth ride...even over unimproved roads.

We've redesigned our engines to make them faster and more efficient.

In high-speed passing, the Volvo 264—with its overhead cam V-6—surpasses a BMW 320i. The Volvo 240—with its fuel-injected, overhead cam 4-cylinder engine—out-accelerates a Cutlass with a V-8.

We've also made the inside of a Volvo a more comfortable place to spend the years.

VE BUILD THEM
R THEY LAST.

We've added a 12-outlet heating and ventilation system that keeps the air continuously fresh.

In 1964, we introduced orthopedically-designed bucket seats. Which we've been improving ever since.

Volvo seats now adjust in eight different directions. They have lumbar supports that you can firm or soften against the small of your back. And on the Volvo 264, the driver's seat is heated.

When it comes to safety, no one demands more of Volvo than Volvo.

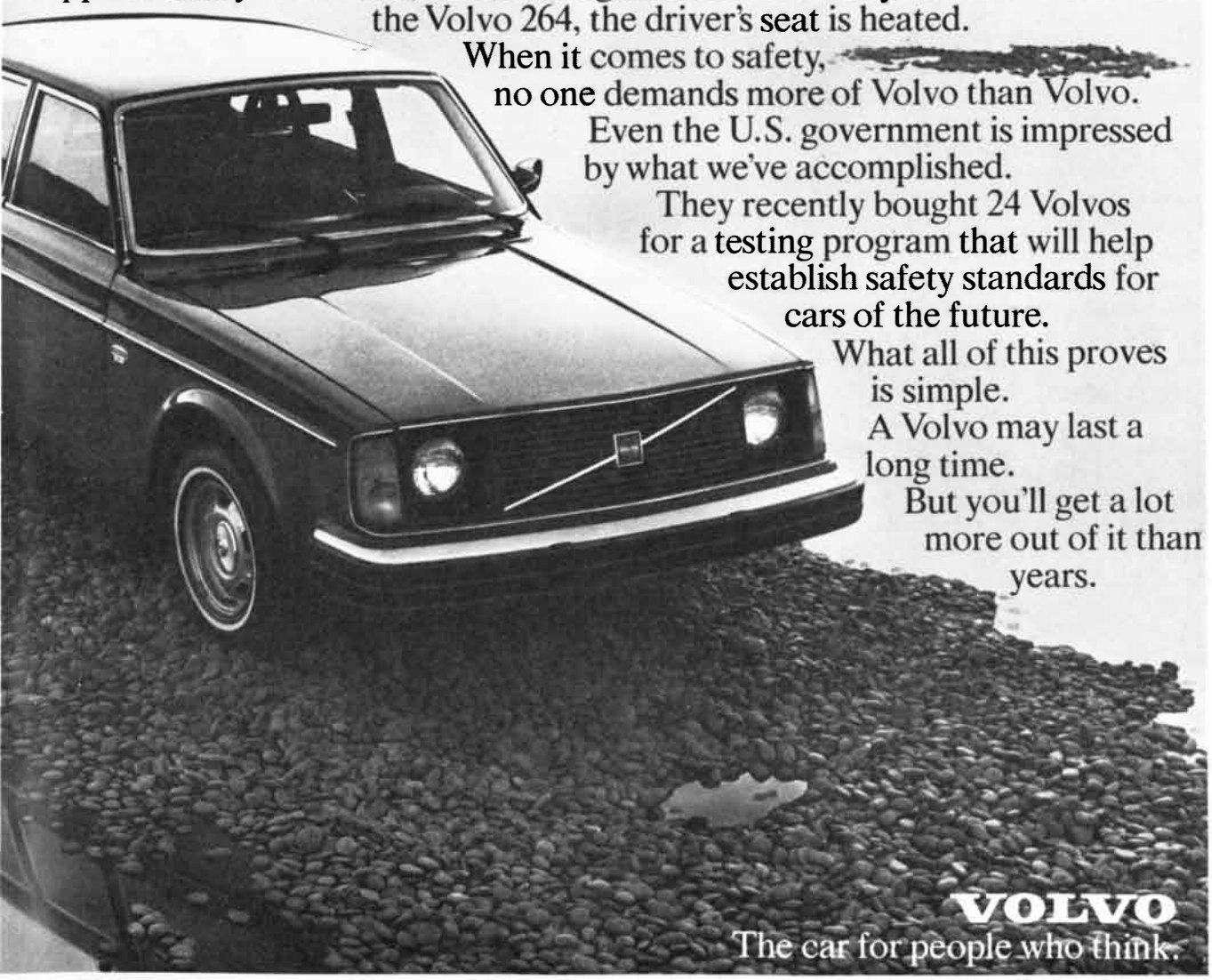
Even the U.S. government is impressed by what we've accomplished.

They recently bought 24 Volvos for a testing program that will help establish safety standards for cars of the future.

What all of this proves is simple.

A Volvo may last a long time.

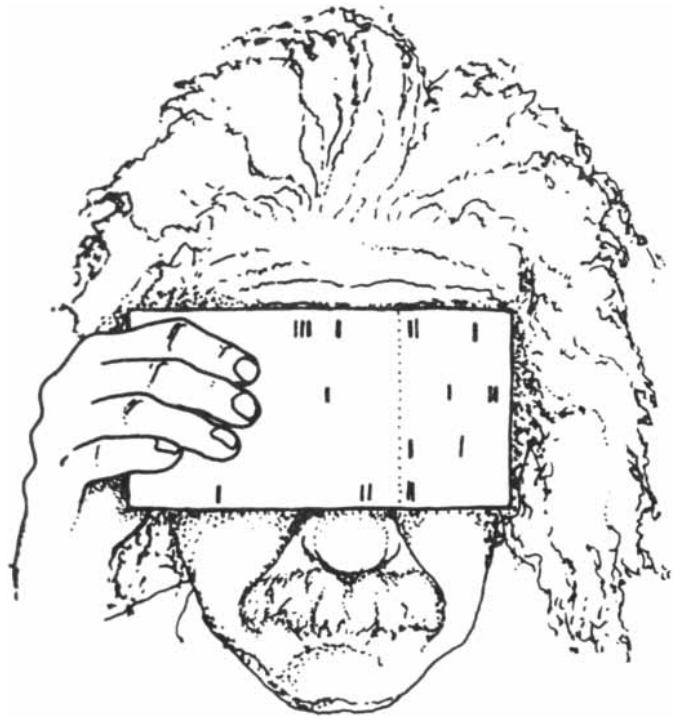
But you'll get a lot more out of it than years.



VOLVO

The car for people who think.

You don't have to be a nuclear physicist to understand an electric bill.



The numbers next to the $ sign are a lot higher today than a couple of years ago. And we can expect even bigger bills as prices increase for oil, coal, natural gas—the main sources for the energy we use.

That's why the world is turning to nuclear power as a practical alternative. Nuclear power can cut our future electricity costs and reduce dependence on those few countries that control oil production and prices. Thousands of American communities now enjoy power, light and heat from nuclear energy. By 1985, there will be at least 300 nuclear power plants in the U.S. and overseas.

Nuclear power is produced by reactors that run on uranium. After 3 or 4 years the fuel rods in a reactor are spent and taken out. They can be stored away. But they contain plutonium, as well as unburned uranium, and discarding them means losing forever a potential source of energy. However, there is a way to reprocess this material, to generate at least 50% additional electricity instead of simply throwing away this valuable resource.

Right now, there's only one facility in the entire U.S.—at Barnwell, South Carolina, built by our company in partnership with General Atomic—that is able to evaluate the benefits of reprocessing. At full operation, Barnwell will reprocess spent nuclear fuel from 50-60 reactors every year. This reusable fuel can provide the energy equivalent of almost 70 million tons of coal or nearly 300 million barrels of crude oil a year. That's a lot of energy that would be wasted without reprocessing.

Reprocessing provides another plus. By significantly reducing the amount of highly radioactive nuclear waste, the need for storage locations is lessened—an environmental as well as economic benefit.

But Barnwell is idle. It can't get started until the government clarifies its position on nuclear fuel recycling.

Barnwell can establish that effective safeguards can be maintained over the plutonium contained in spent fuel. In doing so, it will go a long way toward providing an answer to the problem of international nuclear arms proliferation.

Recent surveys show that a majority of Americans favor more nuclear power plants. Barnwell is the key to proving that reprocessing can help make nuclear power the choice of the future—an economical, safe and environmentally acceptable answer to America's need for electricity.

It doesn't take a degree in nuclear physics to make that choice. Barnwell and the American people are waiting.

Allied Chemical

For copies of this ad and information about how nuclear fuel reprocessing can help solve our energy problems, write to: Energy, Dept. SA- Allied Chemical Corporation, P.O. Box 2245R, Morristown, N. J. 07960

THE COVER

The painting on the cover depicts an experiment demonstrating a soprano's tendency to open her mouth wider as she sings higher tones. The reason, it appears, is to bring the frequency of the first "formant," a particular resonance of the vocal tract, up to match the fundamental frequency of the tone she is singing (see "The Acoustics of the Singing Voice," by Johan Sundberg, page 82). The singer holds a vibrator (*left*) tightly against her neck; a small microphone (*right*) is near her lips. She begins to sing a designated vowel sound at a specified pitch and then stops singing but maintains the position of her lips, tongue and other articulatory organs that determine the vowel sound. The vibrator supplies to the singer's vocal tract a low-pitched sound that is influenced by the articulators just as the singer's own voice would be, but whose formant frequencies are more readily analyzed than those of a high voice tone could be.

THE ILLUSTRATIONS

Cover illustration by Laszlo Kubinyi

Page	Source
27–31	George V. Kelvin
32	Georges A. Vendryes
33–35	Lorelle M. Raboni
37–43	Ilil Arbel
45	Carol Donner
46–50	Allen Beechel
51	Carol Donner
52–56	Allen Beechel
68	Bunji Tagawa
69	Martin B. Garment and Winston J. Brill
70–71	Bunji Tagawa
72	Winston J. Brill (*top*), Frank B. Dazzo (*bottom*)
73	Frank B. Dazzo (*top*), Bunji Tagawa (*bottom*)
74–81	Bunji Tagawa
83	Laszlo Kubinyi
84	Gabor Kiss
86	Laszlo Kubinyi
88	Laszlo Kubinyi and Gabor Kiss
89	Gabor Kiss
90	Johan Sundberg (*top*), Gabor Kiss (*bottom*)
91	Laszlo Kubinyi and Gabor Kiss
93	Stephen Moorbath
94–103	Andrew Tomko
107	Tom Prentiss
108	David C. Powell (*top*), Ken Lucas (*middle and bottom*)
109	James G. Morin
110	David C. Powell
111	Michael D. Lagios
112	Tom Prentiss
116–128	Alan D. Iselin
130	Corozel Project
134–139	Ilil Arbel
142	*Berggasse 19: Sigmund Freud's Home and Offices, Vienna 1938, the Photographs of Edmund Engelman.* © 1976, Basic Books, Inc.
148	*The Pantheon.* © 1976, Harvard University Press

LETTERS

Sirs:

In his article "The Quantum Mechanics of Black Holes" [SCIENTIFIC AMERICAN, January] S. W. Hawking twice makes reference to Einstein's famous dictum about God's not playing dice with the world. It might be of interest to note just how Einstein put it.

Ronald W. Clark, in *Einstein: The Life and Times,* quotes two letters from Einstein that mention the subject. The first was addressed to Max Born in 1926, and it says:

"Quantum mechanics is certainly imposing. But an inner voice tells me that it is not yet the real thing. The theory says a lot, but does not bring us any closer to the secret of the Old One. I, at any rate, am convinced that He does not throw dice."

The second letter, much later, was written to James Franck:

"I can, if the worst comes to the worst, still realize that the Good Lord may have created a world in which there are no natural laws. In short, a chaos. But that there should be statistical laws with definite solutions, i.e., laws which compel the Good Lord to throw dice in each individual case, I find highly disagreeable."

ARTHUR J. MORGAN

New York

Sirs:

Jean Mayer's article on the dimensions of human hunger [SCIENTIFIC AMERICAN, September, 1976] is worthwhile and illuminating, but it contains errors of fact and interpretation with regard to fisheries as a source of protein. In 1970–71 the world catch was steady at about 70 million tons, but it dropped in 1972 to some 66 million tons, not to "less than 55 million tons."

This decrease was caused not by general "overfishing and pollution" but by specific oceanographic conditions along the northwest coast of South America (El Niño), combined with overfishing of a single species, the Peruvian anchovy. In 1970 the catch of Peruvian anchovy was more than 13 million tons, nearly 20 percent of the total world catch and more than 20 percent of the marine catch. By 1972 it had dropped to less than five million tons, about 7 percent of the 1972 world catch. The catch of Peruvian anchovy bottomed out at two million tons in 1973 and has been improving since, under strict management measures.

World catches, including those from fresh water, have been increasing at a rate of a shade over 5 percent per year since 1955. In 1974, the latest year for which complete statistics are available, they were about 70 million tons, comparable to the 1970–71 level. Marine catches have been increasing at a shade under 5 percent per year since 1955, and in 1974 they were about 61 million tons.

The United Nations World Food Conference in late 1974 pointed out that "the annual potential yield of conventional marine species of fish, crustaceans and mollusks [has been estimated as] of the order of 118 million tons, and subsequent assessments have confirmed this estimate of over 100 million tons." A projection of the 1955–74 trend for marine fisheries shows that a 118-million-ton level would be reached by 1985, but there are at least four factors that make the picture look brighter:

1. The projection is for conventional species, that is, those now acceptable. Projections for nonconventional species, notably the Antarctic krill, start at about 100 million tons additional and range upward.

2. Production of food from the sea could be increased without increasing catches. One way would be to divert those species now used for fish meal, and converted in a relatively inefficient way to chicken and other animal sources of protein, to direct use as human food.

3. Although serious technological and economic problems make it difficult to predict the future for aquaculture, the possibility of significant increases in production in 10 or 20 years cannot be discounted.

Scientific American, March, 1977; Vol. 236, No. 3. Published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Gerard Piel, president; Dennis Flanagan, vice-president; Donald H. Miller, Jr., vice-president and secretary; George S. Conn, treasurer; Arlene Wright, assistant treasurer.

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 35 cents.

Subscription correspondence should be addressed to Subscription Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, 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

NOW:

Name any company. We'll send you a cost comparison with Phoenix Mutual.

If you ever plan to buy life insurance again—be sure you have a skilled agent to help you plan.

Be sure the company itself is solid.

And don't pay a lot more for it than you have to.

The big buyers of life insurance, such as major corporations and wealthy individuals, nearly always demand competitive quotations. They're comparison shoppers.

They and their financial advisors know that costs vary widely for identical coverage.

The best advice: keep the insurance you already own, but compare costs when considering more.

Because at the ages most life insurance is purchased...and based on the most widely used method of comparing long-range costs of various policies bought today...time and time again Phoenix Mutual is shown to have lower costs than those of competing companies.

And we have the independent, third party figures to support that statement.

Of course, no company ranks first for low cost in every situation.

Minor variations shouldn't sway you. It's those whopping variations we think you should know about.

So send us the name of any life insurance company.

We'll send you a cost comparison with Phoenix Mutual. Mail the coupon now.

Phoenix Mutual Life Insurance Company
American Row, Hartford, Conn. 06115

Yes, I'm a comparison shopper, too. Show me how Phoenix Mutual stacks up against _____ according to the guidelines described below.

Please circle number nearest your age:

30 35 40 45 50 55

Male ☐ Female ☐ Phoenix Mutual Policyholder? ☐ YES ☐ NO

NAME

ADDRESS

CITY STATE ZIP

Note: All figures used will be drawn from the 1977 edition of "Interest-Adjusted Index, Life Insurance Payment and Cost Comparisons" published by The National Underwriter. Some 300 prominent companies are included and we must restrict this offer to these companies. The most competitive policy (including preferred risk), for your age, sex and state of residence, will be used. Comparisons will be based on 20-year surrender cost for a policy with $100,000 face amount.

Phoenix Mutual Since 1851

We're saving a lot of people a lot of money.

4. The present wave of acceptance of extended-fisheries jurisdictional limits means that stocks that were overfished largely because they belonged to everyone and hence to no one may be returned to their former productivity within the next 10 to 20 years.

ROBERT W. SCHONING

Director
National Marine Fisheries Service
Washington

Sirs:

In "Urban Trees," by Thomas S. Elias and Howard S. Irwin [SCIENTIFIC AMERICAN, November, 1976], the authors remark that the American chestnut tree "is now remembered almost solely because of Longfellow's line: 'Under the spreading chestnut tree....'" Even this scanty memory is the result of a misunderstanding. I learned long ago in English class that "chestnut" here refers to the horse chestnut (*Aesculus hippocastanum*). In confirmation I quote from *The Tree Book* by Julia Ellen Rogers (Doubleday, Page, 1916), page 385: "Longfellow's 'spreading chestnut tree' was a horse chestnut. He called the tree by the name popular in England, where the word 'horse' is ordinarily left off."

RALPH P. BOAS

Northwestern University
Evanston, Ill.

Sirs:

The identity of Longfellow's spreading chestnut tree might make an interesting case for a latter-day Ellery Queen. According to several independent sources, Longfellow's tree *was* the American chestnut, *Castanea dentata*. It is cited as such in Alice Lounsberry's *A Guide to the Trees*. Furthermore, Longfellow's poem was published in the early 1840's, only four years after the introduction of *Aesculus hippocastanum* 'Baumanni,' a cultivar of the European horse chestnut, and approximately 16 years before the development of the hybrid *Aesculus* X *carnea* (the red horse chestnut). Although this is not conclusive proof, it does lend support to the notion that the village blacksmith in Cambridge, Mass., labored in the shade of a spreading (American) chestnut tree.

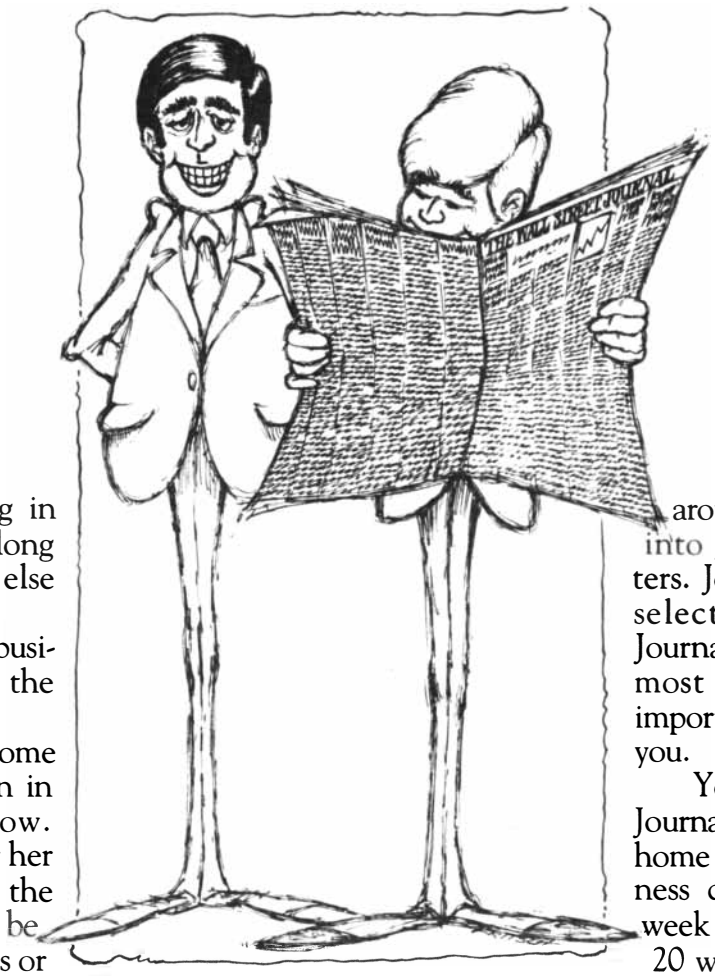
THOMAS S. ELIAS

Cary Arboretum
New York

HOWARD S. IRWIN

President
New York Botanical Garden
New York

Which would you rather be: a yes-man or a know-man?



The easiest thing in the world is to go along with what everybody else thinks is right.

But if you're in business, it may not be the smartest thing.

Because finally, some day, a man or woman in business has to know. Because finally, his or her opinion—if it is to be the right opinion—has to be based on facts. And his or her decision—if it is to be listened to and respected and acted on—has to be made on honest, unbiased, in-depth information.

How do you get that information? How do you know what's really happening in business today, what's important, what's reliable, what's factual?

May we submit that the way to know is to read The Wall Street Journal?

The Wall Street Journal has the largest staff of business-news experts in the world. Every day, reports from that staff, stationed around the globe, pour into Journal headquarters. Journal editors then select and present to Journal readers what is most significant, most important to people like you.

You can get The Journal delivered to your home or office every business day for just 86¢* a week for any period from 20 weeks to a year. Just fill out and mail in the attached reply card. Or, if the card has already been taken, write to The Wall Street Journal, 200 Burnett Road, Chicopee, Mass. 01021. We'll bill you later.

Should you read The Journal? There's no answer but "Yes."

The Wall Street Journal

*Price good in United States and possessions.

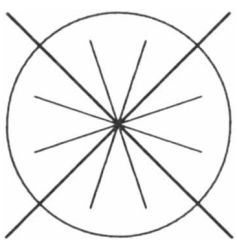
New microprocessor-based X/Y plotter: high plot quality, speed, and resolution, plus a large repertoire of useful tricks.

In complex plots, multiple traces are often difficult to distinguish and interpret. The HP 9872's four pens of different colors, seven distinct dashed and/or dotted line fonts, five built-in character fonts, plus user-defined character and symbol mode plotting capability, combine to produce easy-to-read, high-quality plots suitable for reports and presentations as well as your more mundane graphic data needs.

Hewlett-Packard's new 9872 plotter makes available a new order of X/Y graphics capability. Through microprocessor control, and interconnection via HP Interface Bus, it can provide precisely tailored solutions to the graphic display of data processed on the HP 9825 and new HP 9831 desk-top computers.

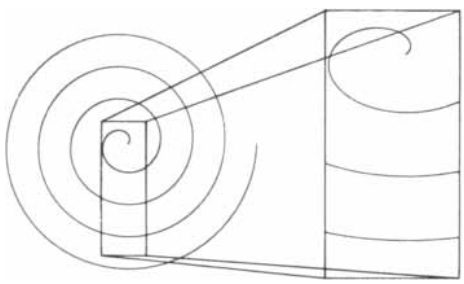
Built into the 9872's microprocessor are 38 different instructions, making possible such graphic accomplishments as:

- Automatic selection of any of four pens, either by program control or by the push of a button on the front panel.
- Stepper motor drive with addressable moves as small as 25 microns (0.025mm). The "stair-step" movements that produce curved or oblique lines can be reduced to the point that the motors will step six times to move the pen the width of a human hair. The result is exceptional line quality and precision. By the same token, the plotter can *retrace* a line without visibly widening it: the shorter radiating lines in this example have been retraced.

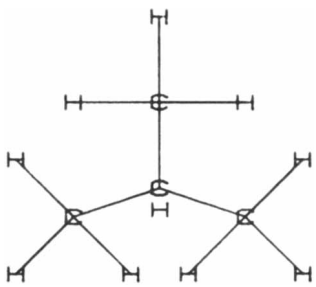


The plot thickens only a little: 0.04mm maximum when retraced by the same pen, 0.20mm maximum when another pen is used.

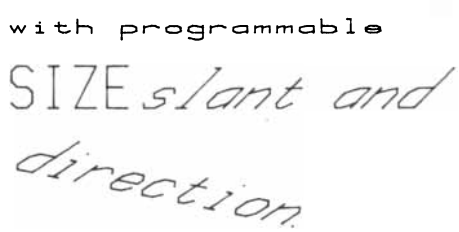
- Off-scale and window plotting. In older instruments, data that exceeded the plot area either generated error messages or were ignored. When off-scale data are encountered by the 9872, the plotter will calculate the mechanical limit intersection of the off-scale data and plot to that point, then resume plotting when the data come back on scale. Although this limit is normally set at the mechanical limits of the plotter, a data "window" of any rectangular configuration can be programmed, and positioned anywhere on the platen.



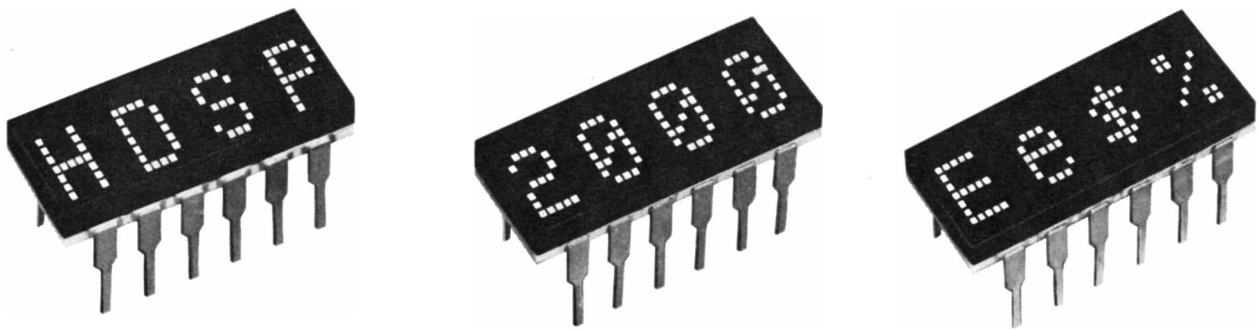
- Symbol mode plotting. The 9872 can label points continuously with characters in the symbol mode. For instance, points in a molecular diagram can be automatically labeled with element symbols—as in this representation of an isobutane molecule.



Furthermore, the plotter has five resident character set: ANSI ASCII; 9825 ASCII; and three European sets—Spanish, Scandinavian, and French/German,



Adaptable characters, a fast-moving plot, and a colorful performance, all for a base price of $4,200*. We're only sorry we can't show you one of its four-color plots in this black and white advertisement.



New high-performance alphanumeric display increases reliability through reduced parts, circuitry, and connections.

To get the kind of performance and reliability we demand in HP products that use LED (light-emitting diode) displays, we develop and manufacture the displays ourselves—along with other performance-critical semiconductor components. This new four-character LED alphanumeric display is a case in point. Usually the larger the range of characters and symbols a display generates, the greater the number of parts, circuits, and connections needed to operate it. The new HDSP-2000 display actually reverses this process, reducing system parts by a factor of 36 to 1 and the number of connections by an order of magnitude or better. It is available to help you simplify your display systems.

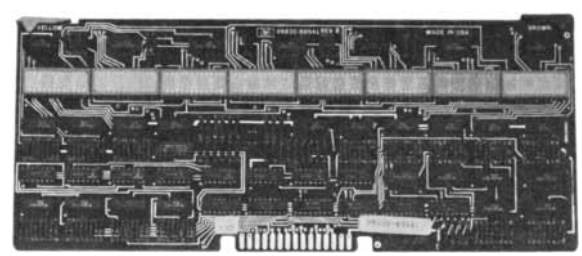
Because a significant portion of the operating electronics are incorporated in the display unit itself, the HDSP-2000 is considerably easier and less expensive to integrate into a system than other current alphanumeric units. And because it requires only 12 pin connections, the display is easier to design into an instrument, easier to troubleshoot, quicker to assemble, and can often contribute to a reduction in the size of the instrument in which it is used. By comparison, existing alphanumeric displays require as many as 35 additional circuit components, and a total of 176 connections to do the same job.

Each 5- by 7-dot character matrix is capable of generating upper and lower case letters, numerals, punctuation marks, and mathematical symbols. The 3.8mm character height permits accurate viewing at up to 3 meters.

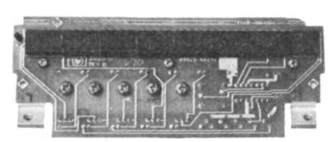
The new display is directly compatible with most common integrated circuit logic families, such as TTL, and can stack end to end to any desired line length. It is physically and environmentally rugged—well suited to portable and mobile applications—and requires only 5 volts for operation.

All together, this represents a new performance for alphanumeric displays.

HDSP-2000 displays cost $47* per four-character unit in quantities of 125, and are available from any HP franchised distributor.



Before: a row of eight 4-character displays and the circuitry needed to drive them—state of the art previous to the HDSP-2000.



After: a similar row of eight HDSP-2000 displays. Because their shift registers and LED drivers are self-contained, the associated circuitry for the equivalent function is greatly simplified.

HEWLETT hp PACKARD

1502 Page Mill Road, Palo Alto, California 94304

For assistance call: Washington (301) 948-6370, Chicago (312) 255-9800, Atlanta (404) 955-1500, Los Angeles (213) 877-1282.

Mail to: Hewlett-Packard, 1502 Page Mill Road, Palo Alto, CA 94304.
Please send me further information on

( ) HP 9872 microprocessor-based plotter
( ) HDSP-2000 LED alphanumeric display

Name _____

Company _____

Address _____

City _____ State _____ Zip _____

*Domestic U.S. prices only.

00742

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

MARCH, 1927: "So far as naval affairs are concerned the people of this country and, with a few exceptions, their congressmen are entirely at the mercy of the naval propagandists. Only the man who has made a close study of the naval problem is able to determine for himself what is the relative strength and efficiency of his own compared with other navies. The journalist is guided on his course by what the naval specialist may happen to tell him. If the truth is put before him, well and good; if not, he may prove to be the first sheep that jumps the fence and leads the entire flock into error. Take the case of the so-called armament race in the building of light cruisers, which has been so persistently dinned in the ears of the public that probably 99 out of 100 good American citizens believe that since the war such a race has been going on. As a matter of fact there never has been any such competition, there is not any such competition today and there are no indications that there ever will be."

"Anti-knock fuels will form an item of great importance to motorists during the approaching summer season. It is well known that the knock encountered in an automobile engine working under load or badly carbonized reduces its efficiency. Because of this the use of anti-knock fuels in an ordinary motor increases its efficiency as measured in the number of miles per gallon obtained from the fuel. The two types of anti-knock fuels that now enjoy the widest use are those based upon 'cracked' gasoline and those containing tetra-ethyl lead as a knock preventive. In the cracking of heavy petroleum oil to yield more gasoline there is formed in the product a considerable amount of what are called unsaturated compounds. These appear to operate to reduce the tendency of the fuel to cause knock in use. The general use of anti-knock fuels will be of unquestionable value in helping to solve the problem of a future fuel supply. If the efficiency of our automobiles can be increased by an average of even 1 per cent, the result will be equivalent to the production of nearly 100 millions of gallons of gasoline per year."

"When the early settlers occupied the virgin land of the United States, the soil was rich in such substances as are necessary to the production of crops. Wheat

THE UNBELIEVABLE ASPEN. MEET THE BELIEVERS.

In almost no time, Aspen has become one of the most popular cars around. And for good reason. It's one of the best cars around. So, with so many happy and pleased Aspen owners across the country, Aspen has become less "unbelievable." And more believable. Just ask any believer who owns a Dodge Aspen.

"I love the overall car. The size, handling, and visibility are excellent. And the price was very good. My family and friends just love it."

"I like everything about Aspen. It's what I wanted. A smart-looking car. And I got a good deal."

"I like the way Aspen handles on the road. It's very comfortable for a small car. And you get a lot of car for the price."

"I test drove an Aspen, and that was it. I like the ride, the handling, and the looks. I'm in it all day, and the comfort is good. All in all, there's nothing I don't like about my Aspen."

Buy or lease. Dodge Dealers everywhere are making Aspen easy to drive…one or a fleet…through buying or leasing arrangements. Visit your Dodge Dealer and test drive an Aspen. And be a believer, too.

Dodge Aspen.
See it, price it, drive it... and believe.

Dodge
A PRODUCT OF CHRYSLER CORPORATION

crops of 50 or 60 bushels or more to the acre were frequent, and the pioneer farmer gave little thought to fertilizer. As the years passed the yield began to decrease, until the sons of the first comers found that they were able to raise but 30 or 35 bushels to the acre. The process of decline went on until it reached a point where the farmer was able to garner from these once rich soils a harvest that ran only 10 to 15 bushels to the acre. Although it has been stated that the lack of nitrogen more than the lack of any other element is responsible for low crop yields in the United States, two other elements—phosphorus and potash—are also removed from the soil by crops and must be restored. The phosphorus for fertilizers is obtained in the United States from the beds of phosphate rock in Florida, Tennessee and Idaho. Treated with its own weight of sulphuric acid, it appears on the market as acid-phosphate fertilizer, and by weight it is the largest single fertilizing material used in this country. Potash is received mainly from French and German mines, and our annual consumption is about 250,000 tons. The principal source of nitrogen has been from the natural crude nitrate deposits of Chile. A comparatively recent development has been the extraction of nitrogen from the air, an industry that is growing very rapidly."

SCIENTIFIC AMERICAN

MARCH, 1877: "We have already laid before our readers accounts of the wonderful performances of Professor Bell's telephone, an instrument remarkable not merely for its phenomenal capabilities but also because of its having been brought to its present state of development within the period that has elapsed since June last. During that month Professor Bell exhibited the apparatus at the Centennial, working it over short distances only and causing it to transmit sound that reached the opposite terminus very much diminished in intensity. The telephone has recently carried the human voice over a distance of 143 miles (from Boston to North Conway, N.H.), so that ordinary conversational tones uttered at one end were distinctly audible at the other. Further even than this, the inventor and his assistant have talked through a wire arranged to give an artificial resistance equal to 40,000 ohms, which is more resistance than the entire length of the Atlantic cable would offer. There are, however, obstacles other than the resistance that checks the transmission of the voice over such immense distances. These the inventor is now endeavoring to overcome, and at the first favorable opportunity a practical test of the pow-

Every day, we help America save more energy than it takes to light New York.

More than enough for every bulb, neon sign, streetlamp and fluorescent light New York City burns every day.

And that doesn't include *all* the energy Combustion Engineering helps save.

Only the fuel saved by the Ljungstrom® heat exchanger, just *one* of the energy-recovery products from C-E Air Preheater.

Installed on a steam generator or process furnace, the Ljungstrom captures heat energy that would be lost through the smokestack and sends it back to work.

With the Ljungstrom, a utility, refinery or process plant uses up to 25% less fuel. In fact, over the past five years, the energy saved by Ljungstrom units around the world equals 150 million tons of coal, 190 million barrels of oil and 1.4 trillion cubic feet of natural gas.

We do more than save fuel.

For over 60 years, we've helped provide the equipment, technology and research to produce energy, conserve it, squeeze more use out of it, even create new forms of it.

C-E Power Systems, for example, designed the steam generators that produce 40% of all the thermal-electric power in the Free World.

C-E Natco helps process 25% of all the oil and gas the world uses every day.

C-E Lummus has created a clean liquid fuel by removing the ash and sulfur from coal.

For more information about our activities around the world, write Combustion Engineering, Inc., Dept. 7008-27, 902 Long Ridge Road, Stamford, Connecticut 06902.

COMBUSTION ENGINEERING

The Energy Systems Company

ers of the instrument over one of the trans-Atlantic cables will be attempted."

"The coal tar has been investigated chemically, and its constituents have been taken apart. All together at least 50 different things have been extracted from that black material. From some of these the now well-known coal-tar dyes are made. The foremost for this purpose are benzol, phenylic alcohol, anthracene and alizarine. The most important of these is perhaps anthracene. Although it was the last to be discovered, in the past three years it has effected a complete revolution in the English dyeing trade. It is obtained in the distillation of coal tar as a solid substance, which is filtered out by means of cocoa-nut matting or some such material. It yields a dye that is identical with that yielded by the madder root, so that now England can manufacture from her own resources a sufficiency of this dye material, whereas before she was dependent on imported madder from Turkey, Holland and France."

"Langendorff states that stimulation of certain regions of the cerebral hemispheres in the frog is invariably followed by muscular movements. These movements are bilateral when both hemispheres are excited simultaneously. When only one hemisphere is excited, the movements are limited to the opposite sides of the body. The 'irritable zone' is situated in the parietal region of each cerebral hemisphere."

"An interesting trial was lately had at Portsmouth, Eng., of the new British double-turreted war steamer, *Dreadnought*. This vessel represents the latest advances in marine construction for war purposes. Her load displacement is 10,950 tons; length, 320 feet; beam, 63 feet 10 inches; depth of hold, 19 feet 2 inches; thickness of side armor, 8 to 14 inches; external diameter of turrets, 32 feet 3 inches; armor, 14 inches; armament, four 38-ton guns."

"Dr. T. Sterry Hunt lately gave a very interesting address before the Chestnut Street Club, Boston, in the parlors of Mrs. J. T. Sargent. In introducing his lecture Dr. Hunt spoke of the growing belief in the unity and harmony of the forces throughout the universe. The chemical, physical and vital forces of our earth, he said, seem at first to have no influence outside of it, but the spectroscope and telescope tell us of like forces in far-away worlds. Great masses of vapor and nebulous matter in space condense to form suns and planets. The latest speculation of chemistry leads us to conclude, he said, that the chemical elements of the earth are themselves the product of growth or evolution from simpler elements."

HOW

to keep up with what's new in computer techniques and management easily—and inexpensively

The Library of Computer and Information Sciences

and start
a whale of a value

Take any 3 books for only $1 each

(values to $60.00)

if you will join now for a trial period and accept only 3 more books—at member discount prices—within the next 12 months.

(Publishers' Prices shown)

62745. MINICOMPUTER SYSTEMS—Organization and Programming (PDP-11). *Richard H. Eckhouse, Jr.* Spells out general computer organization and assembly language programming, and offers a complete guide to all important minicomputer applications. **$15.95**

41640. DATA COMMUNICATIONS DICTIONARY. *Charles J. Sippl.* Complete, up to date, cross-reference guide of over 14,500 data processing and communications terms. **$19.95**

66010. OPERATING SYSTEMS. A Systematic View. *William S. Davis.* Straightforward, useful presentation of how operating systems are designed, implemented, the kind of problems they are created to solve, the software they use to achieve these solutions. **$15.95**

82180. SYSTEMS PROGRAMMING: Concepts of Operating and Data Base Systems. *David K. Hsiao.* A broad introduction covering assembly systems, input output operations, batch processing operating systems, data base managers, more. **$16.95**

56255. INTRODUCTION TO COMPUTABILITY. *Frederick Hennie.* What are the limits to the computer's problem-solving ability? A fundamental study of the limits of digital computation and, indirectly, the expressive powers of programming languages. **$20.95**

82560. TECHNIQUES OF PROGRAM STRUCTURE AND DESIGN. *Edward Yourdon.* Over 300 illustrative problems, 90 valuable references, and four demonstration projects. **$16.95**

44900. ENCYCLOPEDIA OF COMPUTER SCIENCE. *Edited by Anthony Ralston and Chester L. Meek.* 470 articles, 1550 pages of the latest accurate information on hardware, software, systems, theory, economic, social, and professional issues, by 208 leading experts. *Counts as 3 of your 3 books.* **$60.00**

52215. HANDBOOK OF DATA PROCESSING ADMINISTRATION, OPERATIONS, AND PROCEDURES. *S. R. Mixon.* Offers a future- and user-oriented system for designing, documenting, and controlling computer center projects. Plus a wealth of practical information on systems and programming techniques, and staff training. *Counts as 2 of your 3 books.* **$24.95**

60405. MANAGING A PROGRAMMING PROJECT. *Philip W. Metzger.* Sound advice on how to manage a project from the definition and design stage through testing to installation and operation. **$15.00**

81880. SYSTEM ANALYSIS TECHNIQUES. *Edited by J. Daniel Couger and Robert W Knapp.* How to properly evaluate an existing system, determine managerial information requirements, and assess the benefits. **$15.95**

70190. PRINCIPLES OF OPERATIONS RESEARCH. Second Edition. *Harvey M. Wagner.* A complete overview of all the fundamental concepts involved in operations research. 1039 pages. **$19.95**

70250. PRINCIPLES OF SYSTEMS PROGRAMMING. *Robert M. Graham.* Examines basic problems in the design, implementation, and maintenance of major software components—assemblers, loaders, macros, time sharing systems, and more. **$16.95**

MEMBERSHIP BENEFITS • In addition to getting three books for $1 each when you join, you keep saving substantially on the books you buy. • If you continue your membership past the trial period, you will be eligible for our Bonus Book Plan—an important way to save even more, at least 70% off publishers' prices. • Your Book Club News, describing the coming Main Selection and Alternate Selections, will be sent to you 15 times a year at three to four week intervals. • If you wish to purchase the Main Selection, do nothing and it will be sent to you automatically. • If you prefer one of the Alternates or no book at all, simply indicate your decision on the reply form always enclosed with the News and mail it so we receive it by the date specified. • The News is mailed in time to allow you at least 10 days to decide if you want the coming Main Selection. If because of late mail delivery of the News, you should ever receive a Main Selection without having had the 10 day consideration period, that Selection may be returned at Club expense.

55331. INFORMATION SYSTEMS. *John G. Burch, Jr. and Felix R. Strater, Jr.* All the background needed to successfully develop information systems for just about any organization. **$12.95**

59940. MANAGEMENT INFORMATION SYSTEMS. *Paul Siegel.* Planning, forecasting, modeling, decision-making, establishing MIS roles, and implementation of MIS—are all thoroughly covered. **$16.95**

86540. THE WAY THINGS WORK BOOK OF THE COMPUTER. Encyclopedia of information science, cybernetics, and data processing. **$8.95**

76920. SCIENTIFIC ANALYSIS ON THE POCKET CALCULATOR. *Jon M. Smith.* Shows how any inexpensive pocket calculator can easily be used to handle complex computations. **$12.95**

53795. HOW TO MANAGE STRUCTURED PROGRAMMING. *Edward Yourdon.* Down-to-earth presentation of the strengths of *programmer productivity techniques* (PPT). Covers top-down design, structured walkthroughs, and management considerations for these new methods. **$19.95**

39855. COMPUTER-AIDED DESIGN OF DIGITAL SYSTEMS. *Douglas Lewin.* Shows how *computer-aided design* (CAD) can help improve the engineering process by providing interactive access to computer information bases. *Counts as 2 of your 3 books.* **$22.50**

81850. A SYSTEMATIC APPROACH TO DIGITAL LOGIC DESIGN. *Frederic J. Mowle.* Detailed presentation of the principles on which hardware functions are based. Number systems, sequential circuit design, register design techniques, flip-flop circuits, are among the topics expertly covered. **$17.95**

The Library of Computer and Information Sciences 7-7AH
Riverside, New Jersey 08075

Please accept my application for membership and send me the three volumes indicated, billing me only $1 each. I agree to purchase at least three additional books during the first 12 months I am a member, under the club plan described in this ad. Savings range up to 30% and occasionally even more. My membership is cancelable any time after I buy these three books. A shipping and handling charge is added to all shipments.

3 books for $1.00 each
Indicate by number the 3 books you want.

A few expensive books (noted in book descriptions) count as 2 choices.

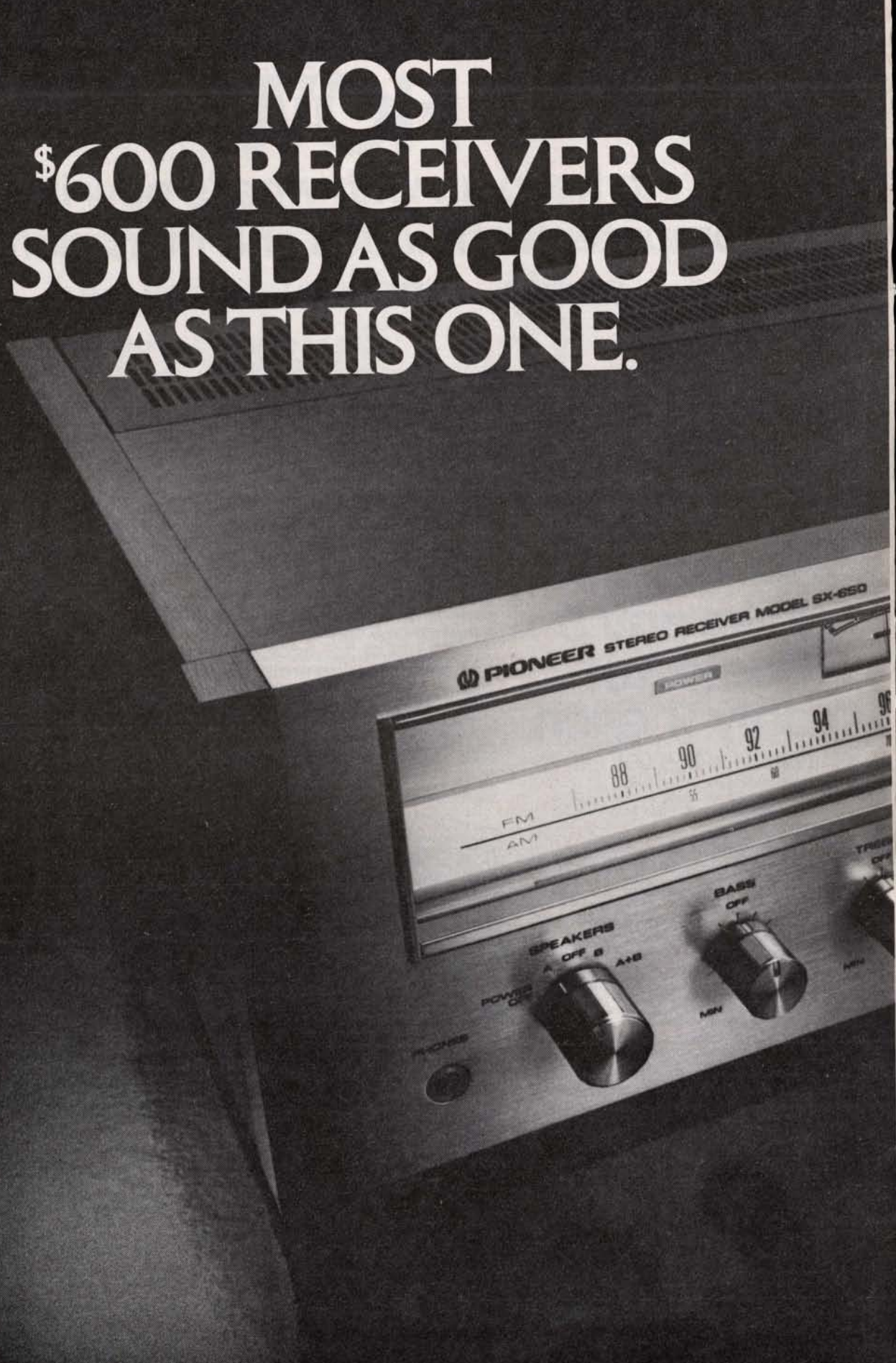
Name _____

Address _____

City _____ State _____ Zip _____

(Offer good in Continental U.S. and Canada only. Prices slightly higher in Canada.) Books purchased for professional purposes may be a tax-deductible expense.

MOST $600 RECEIVERS SOUND AS GOOD AS THIS ONE.



UNFORTUNATELY FOR THEM, THIS ONE SELLS FOR UNDER $300.*

The average $600 receiver sounds as good as the new Pioneer SX-650 until you start listening to prices.

If $600 is your kind of price, an SX-650 should qualify as your kind of receiver. Not only will it give you the kind of features and sound quality you'd expect for that kind of money; it'll also leave you with roughly half your receiver budget unexpectedly unspent.

But suppose your idea of a receiver price is somewhere under $300.* The SX-650 will sound better to you than anything you thought you could afford. Because it has more power, a wider frequency range, less distortion, and far greater versatility than most other receivers in that category.

It's a fact that the SX-650 provides a continuous power output of 35 watts per channel, min. RMS into 8 ohms, from 20 to 20,000 Hz, with no more than 0.3% total harmonic distortion. It also delivers each instrument and voice at its intended level, balanced within ±0.3dB of the RIAA curve.

The facts of its stereo separation, selectivity and sensitivity must really be experienced: sometimes only hearing is believing.

You'll also be impressed by what you don't hear from the SX-650. You won't hear the thousand miscellaneous acoustic devils that live in the limbo between FM stations on lesser receivers.

On your next visit to a high fidelity dealer, listen to a Pioneer SX-650 with any reasonably accurate speakers.

You'll find either its price or its performance amazing. Depending on which you hear first.

PIONEER®

U.S. Pioneer Electronics Corp., 75 Oxford Drive, Moonachie, New Jersey 07074.
*For informational purposes only, the SX-650 is priced under $300. The actual resale price will be set by the individual Pioneer dealer at his option.

THE AUTHORS

GEORGES A. VENDRYES ("Superphénix: A Full-Scale Breeder Reactor") has been involved in various aspects of the French effort to develop a liquid-metal-cooled fast-breeder reactor since the beginning of that program in 1957. At present he serves as director of industrial nuclear applications for the Commissariat à l'Énergie Atomique. A graduate of the École Polytechnique in Paris, he worked for a time after World War II as a civil engineer on the reconstruction of bridges and roads in southern France before he switched to scientific research. He did experimental work in nuclear physics in the Laboratoire de Synthèse Atomique under Frédéric Joliot-Curie, and in 1951 he received his doctorate in physical sciences from the University of Paris. Since joining the staff of the Commissariat à l'Énergie Atomique in 1952 he has, in addition to his work on fast-neutron reactors, been active in launching the French program on controlled nuclear fusion and has been involved in studies of the natural uranium reactor discovered in Gabon in 1972 (see "A Natural Fission Reactor," by George A. Cowan; SCIENTIFIC AMERICAN, July, 1976).

J. T. GOSLING and A. J. HUNDHAUSEN ("Waves in the Solar Wind") are respectively solar physicists at the Los Alamos Scientific Laboratory and at the High Altitude Observatory in Boulder, Colo. Gosling received his bachelor's degree in physics at Ohio University in 1960 and his doctorate in physics from the University of California at Berkeley in 1965. Immediately thereafter he went to Los Alamos and spent two years doing postdoctoral research on the solar wind. In 1967 he joined the scientific staff of the High Altitude Observatory, continuing his work on the solar wind and the physics of the solar corona. In 1975 he returned to Los Alamos. Hundhausen was a member of the staff of the theoretical division of the Los Alamos Scientific Laboratory from 1964 to 1971. He received his undergraduate and graduate training at the University of Wisconsin, obtaining his B.S. in 1958 and his Ph.D. in physics in 1965. Since 1971 he has been a member of the senior scientific staff at the High Altitude Observatory and concurrently a lecturer at the University of Colorado. His research has been generally concerned with the physical processes in tenuous plasmas, specifically with observations and theoretical models of the interplanetary plasma and its interaction with the geomagnetic field.

SOLOMON H. SNYDER ("Opiate Receptors and Internal Opiates") is professor of pharmacology and experimental therapeutics and professor of psychiatry at the Johns Hopkins University School of Medicine. He received his M.D. degree from the Georgetown University School of Medicine. After his internship at the Kaiser Foundation Hospital in San Francisco he became a research associate at the National Institute of Mental Health. In 1965 he became assistant resident in the department of psychiatry at Johns Hopkins Hospital, and a year later he joined the faculty of the School of Medicine.

WINSTON J. BRILL ("Biological Nitrogen Fixation") is professor of bacteriology at the University of Wisconsin at Madison. Born in London, Brill immigrated to the U.S. in 1950 at the age of 11. After receiving his B.A. in biology at Rutgers University in 1961, he went on to obtain his Ph.D. in microbiology from the University of Illinois at Urbana-Champaign. From 1965 through 1967 he did postdoctoral work in the biology department of the Massachusetts Institute of Technology, investigating the genetics and regulation of the degradation of amino acids in bacteria. He joined the bacteriology department of the University of Wisconsin in 1967 and was made a full professor in 1974. "Students and postdocs in my lab come from a variety of backgrounds," Brill states, "including soil science, genetics, molecular biology, chemistry, agronomy and bacteriology. It is the interaction among these people that makes the laboratory environment particularly exciting nowadays. We are always on the lookout for potential applications of concepts derived from basic research."

JOHAN SUNDBERG ("The Acoustics of the Singing Voice") is project scientist in the department of speech communication at the Royal Institute of Technology in Stockholm. He received his entire university training at the University of Uppsala, obtaining his Ph.D. in musicology and acoustics in 1966. He has been at the Royal Institute of Technology since 1963; from 1965 through 1975 he was also instructor of organ theory at the Stockholm Conservatorium of Music. "My entry into the art of music was going to an organ concert in 1952," Sundberg writes. The next year, at the age of 17, he and a friend began to build an organ. "For economic reasons the pipes were replaced by empty bottles. A vacuum cleaner supplied the air, the bellows was under the bed and there were only two wooden stops that I had built on my own and that sounded most ugly. The fifth version of that organ today has its place in our sitting room standing in a corner like a big, quiet ele-

HOW TO BE WELL READ EVEN IF YOU AREN'T WELL HEELED.

If you don't judge a book by its cover, why pay for the cover? Join QPB. Get great books by great writers in full-size, softcover editions at far less than hardcover prices.

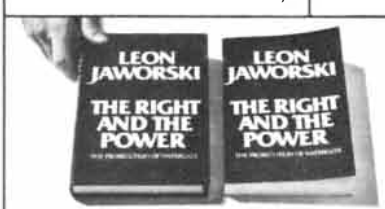
QPB has a very simple philosophy.

It's a book club for people who love to read and collect important works of fiction and nonfiction, but are stunned by current bookstore prices.

The QPB alternative gives you the titles you want in full-size, softcover editions. These are books printed on fine quality paper, books as permanent and durably bound as most hardcover editions.

You'll get at least one bonus point for each book or set you buy and when you've accumulated six bonus points, you will be able to pick another book or set free.

And with QPB's carefully selected list, numbering in the hundreds (many available in softcover *only* through QPB), no serious reader will have a problem finding books he or she has wanted to own.



106. The Right And The Power: The Prosecution of Watergate. Leon Jaworski. Hardcover Edition: $9.95. QPB Edition: $4.95. Same fine paper, same page size and same text. And only $1 if it's part of your 3-book membership choice.

198. The Best Of Life. (Photos) Hardcover: $19.95 **QPB: $8.95**

151. The Lord Of The Rings J.R.R. Tolkien. (3 Vols., Boxed) Hardcover: $25. **QPB: $8.95**

228. The Life And Work Of Sigmund Freud. Ernest Jones. **The Letters Of Sigmund Freud.** Edited by Ernst L. Freud. (2 Vols.) Hardcover: $27.95 **QPB: $11.90**

132. Cat's Cradle, Slaughterhouse-Five and Breakfast Of Champions. Kurt Vonnegut. (3 Vols.) Hardcover: $23.85 **QPB: $8.35**

222. The Americans. Daniel J. Boorstin. (3 Vols.) Hardcover: $45 **QPB: $14.85**

179. All The President's Men. Carl Bernstein & Bob Woodward. (Photos) Hardcover: $10.95 **QPB: $4.95**

Join now and pick any 3 books or sets for $1 each—with no obligation to buy another book.

381. The Photography Catalog Edited by Norman Snyder with Carole Kismaric and Don Myrus. (Illus.) **QPB: $7.95**

115. Finnegan's Wake and Ulysses James Joyce. (2 Vols.) Hardcover $23.95 **QPB: $8.90**

210. The World Of The Past Edited by Jacquetta Hawkes. (2 Vols.) Maps and Photos. Hardcover: $20 **QPB: $11.90**

376. Cat and Never Eat Anything Bigger Than Your Head & Other Drawings. B. Kliban. (2 Vols.) **QPB: $5.90**

375. The Mediterranean And The Mediterranean World In The Age Of Philip II. Volumes I and II Fernand Braudel. Translated by Siân Reynolds. (Illus.) Hardcover: $35 **QPB: $13.90**

378. Other Homes And Garbage Designs for Self-Sufficient Living. Jim Leckie, Gil Masters, Harry Whitehouse and Lily Young. **QPB: $9.95**

392. The Shattered Mind: The Person After Brain Damage. Howard Gardner. Hardcover: $10. **QPB: $3.95**

111. The Foundation Trilogy Three Classics of Science Fiction. Isaac Asimov. Hardcover: $17.85 **QPB: $4.95**

125. Fire From Heaven. Mary Renault. Hardcover: $7.95 **QPB Ed: $3.95**

134. Jung And The Story Of Our Time. Laurens van der Post Hardcover: $10. **QPB: $3.95**

379. Edith Wharton: A Biography R.W.B. Lewis. (Photos) Hardcover $15. **QPB: $6.95**

137. The Supermarket Handbook Access to Whole Foods. Nikki and David Goldbeck. Revised and Expanded Edition. (Illus.) Hardcover $7.95 **QPB: $4.95**

139. The White Goddess: A historical grammar of poetic myth Robert Graves. Amended and Enlarged Edition. Hardcover: $13.50 **QPB: $4.95**

141. Fruits Of The Shaker Tree Of Life: Memoirs of Fifty Years of Collecting and Research. Edward Deming Andrews and Faith Andrews (Photos) Hardcover: $15.95 **QPB: $8.95**

142. The Filmgoer's Companion Leslie Halliwell. The New Completely Revised & Enlarged Fourth Edition (Illus.) Hardcover: $25. **QPB: $6.95**

274. X-Raying The Pharaohs James E. Harris and Kent R. Weeks (Photos) Hardcover: $10. **QPB: $4.95**

260. The New Catalogue Of Catalogues: The Complete Guide to World-Wide Shopping by Mail. Maria De La Iglesia. (Photos) **QPB: $7.95**

253. The Adventures Of Sherlock Holmes and The Memoirs Of Sherlock Holmes. A. Conan Doyle (2 Vols.) Illus. **QPB: $7.90**

Let's try each other for 6 months...

Quality Paperback Book Club, Inc., Middletown, Pa. 17057.
Please enroll me in QPB and send the 3 choices I've listed below. Bill me $3 plus shipping charges. I understand that I am not required to buy another book, but if, after receiving *QPB Review* for 6 months, I have not bought and paid for at least 1 book in every six-month period, you may cancel my membership. 7-QB170-3

Indicate by number the 3 books or sets you want ☐ ☐ ☐

Name _____
(Please print plainly) 16

Address _____ Apt. _____

City _____ State _____ Zip _____

How membership works

1. You receive **QPB Review** 15 times each year (about every 3½ weeks). Each issue reviews a *new* Main Selection, plus scores of Alternates. All Main Selections with established publisher's list prices are offered at at least 20% discount off that list price.
2. If you want the **Main Selection** do nothing. It will be shipped to you automatically. If you want one or more Alternate books—or no book at all—indicate your decision on the reply form always enclosed and return it by the date specified.
3. **Free books.** For each book or set you take (except the first 3 you get for $1 each) you earn Bonus Points which entitle you to free books. You pay only shipping charges.
4. **Return privilege.** If *QPB Review* is delayed and you receive the Main Selection without having had 10 days to notify us, you may return it at our expense.
5. **Cancellations.** You may cancel membership at any time by notifying QPB. We may cancel your membership if you elect not to buy and pay for at least one book in every six-month period.

QPB

The first book club for smart people who aren't rich.

phant." He became interested in singing while at Uppsala, where he joined the choir at the cathedral. After completing his doctoral dissertation in 1966, he began his studies of the singing voice.

STEPHEN MOORBATH ("The Oldest Rocks and the Growth of Continents") is a fellow of Linacre College and senior research officer at the University of Oxford. He is also the Principal Investigator of the Geological Age and Isotope Research Group in Oxford's Department of Geology. He received his entire geological education at Oxford, obtaining his B.A. in 1954, his Ph.D. in 1959 and his D.Sc. in 1970.

JOHN E. McCOSKER ("Flashlight Fishes") is superintendent of the Steinhart Aquarium of the California Academy of Sciences and adjunct professor of marine biology at San Francisco State University. After receiving his B.A. from Occidental College in 1967, he obtained his Ph.D. from the Scripps Institution of Oceanography in 1973. Before receiving his Ph.D. he was a research fellow at the Smithsonian Tropical Research Institute in Panama, where he studied the marine fish fauna of the Caribbean and the eastern Pacific and their relation to the biological barrier of the Panama Canal. In 1975 McCosker led the American/French Coelacanth Expedition to Grande Comore Island. His research activities have included fieldwork in Australia, the Indian Ocean, the Caribbean, the Galápagos Islands, Central America and the California coast.

NORMAN HAMMOND ("The Earliest Maya") is director of the Corozal Project, a venture of the University of Cambridge and the British Museum set up to investigate the origins of Maya civilization. Currently he is the only Mesoamericanist working out of England, although this year he is visiting professor of anthropology at the University of California at Berkeley. He studied at Cambridge, receiving all three of his degrees there; his Ph.D. degree in 1972 was the first English Ph.D. awarded in Maya archaeology. From 1967 to 1975 he was a member of the faculty at Cambridge, in the Centre of Latin American studies. From 1975 to this year he was senior lecturer (associate professor) in the School of Archaeological Sciences at Bradford University. In addition to his main work in the Maya area he has done fieldwork in North Africa tracing Roman roads, in Afghanistan seeking Neolithic settlements and in Ecuador constructing a carbon-14 chronology for the southern-highlands zone. Since 1967 Hammond has been the archaeological correspondent of *The Times*. He is also the editor of *Afghan Studies* and coeditor of *The Archaeology of Afghanistan*, which is being published this year by Academic Press.

Before Sound Guard®, the only way to prevent your records from wearing out was not to play them.

Magnified, you can see record vinyl wearing away.

With same magnification, record vinyl shows no wear.

If you've played any record often enough, you've heard the inevitable occur. It wore out.

While "pops," "hisses," and other surface noises began making their appearance on your favorite records, high frequency sounds—like violins and flutes—began disappearing.

The villain behind this destruction is friction. (If a diamond cuts through steel, you can imagine what a diamond stylus does to vinyl records.) Fortunately, from outer space has come a solution to record degradation. It's called Sound Guard.*

A by-product of research into dry lubricants for aerospace applications, Sound Guard record preservative puts a microscopically-thin (less than 0.000003") dry film on records to protect the grooves from damage. Yet, remarkably, it does not degrade fidelity.

Independent tests show that Sound Guard preservative maintains full amplitude at all audible frequencies, while at the same time significantly retarding increases in surface noise and harmonic distortion.**

In other words, when applied according to instructions, a new record treated with Sound Guard preservative and played 100 times sounds the same as one in "mint" condition played the first time!

Sound Guard preservative comes in a kit (complete with non-aerosol pump sprayer and velvet buffing pad). It is completely safe and effective for all discs, from precious old 78's to the newest LP's including CD-4's.

Recently introduced to audiophiles, Sound Guard preservative is now available in audio and record outlets.

**For complete test results write: Sound Guard, Box 5001, Muncie, Indiana 43702.

Sound Guard® keeps your good sounds sounding good.

*Sound Guard is the registered trademark of Ball Corporation for its record preservative. © 1976 by Ball Corporation.

Established 1845 SCIENTIFIC AMERICAN March 1977 Volume 236 Number 3

Superphénix: A Full-Scale Breeder Reactor

The decision has been made to begin the construction in France of a 1,200-megawatt breeder-reactor power station. The joint European project will be the prototype of future nuclear plants

by Georges A. Vendryes

The need to resort to nuclear fission to help meet the anticipated world demand for energy over the next few decades is widely, if not universally, recognized. What is often not appreciated sufficiently, however, is the fact that if the construction of new nuclear power plants is limited to the same basic types of reactor generally in service today, the respite gained will be only a brief one. Most experts agree that at current prices the world's economically recoverable uranium reserves are inadequate to ensure a lifetime supply of fuel for light-water nuclear reactors built after the year 2000. This means that unless uranium is used in a more efficient way than it is in such reactors, it will turn out to be an energy resource not very different in scale from oil.

Only breeder reactors—nuclear power plants that produce more fuel than they consume—are capable in principle of extracting the maximum amount of fission energy contained in uranium ore, thus offering a practical long-term solution to the uranium-supply problem. Breeder reactors would make it possible to obtain some 50 times more energy from a given amount of natural uranium than can be obtained with present-day light-water reactors. Hence the minimum uranium content of economically recoverable ore could be significantly lowered. For these two reasons (of which the second is by far the more important) the useful supply of natural uranium could be greatly enlarged. Uranium would then constitute a virtually inexhaustible fuel reserve for the world's future energy needs.

Recognizing the importance of these considerations, a number of nations have undertaken intensive research programs aimed at developing an economically competitive breeder reactor before the uranium-supply situation becomes critical. Last fall a consortium of major European electric-utility companies, acting through a joint subsidiary, decided to start the construction of a 1,200-megawatt breeder-reactor power plant at Creys-Malville in France. The new full-scale breeder reactor, named Superphénix, will be described here. First, however, it is necessary to explain just what is meant by the term breeding, which serves to characterize the operation of such plants.

Two types of heavy isotope are present in the active core of every nuclear reactor. One type, called the fissile (or fissionable) isotope, undergoes most of the fission reactions and is the source both of the heat energy released by the reactor and of the neutrons that sustain the chain reaction in the core. The only fissile isotope that exists in nature is uranium 235, which constitutes .7 percent of natural uranium; the nonfissionable isotope uranium 238 accounts for the remaining 99.3 percent. Two other fissile isotopes, plutonium 239 and uranium 233, are expected to play an increasingly important role in the future as substitutes for uranium 235.

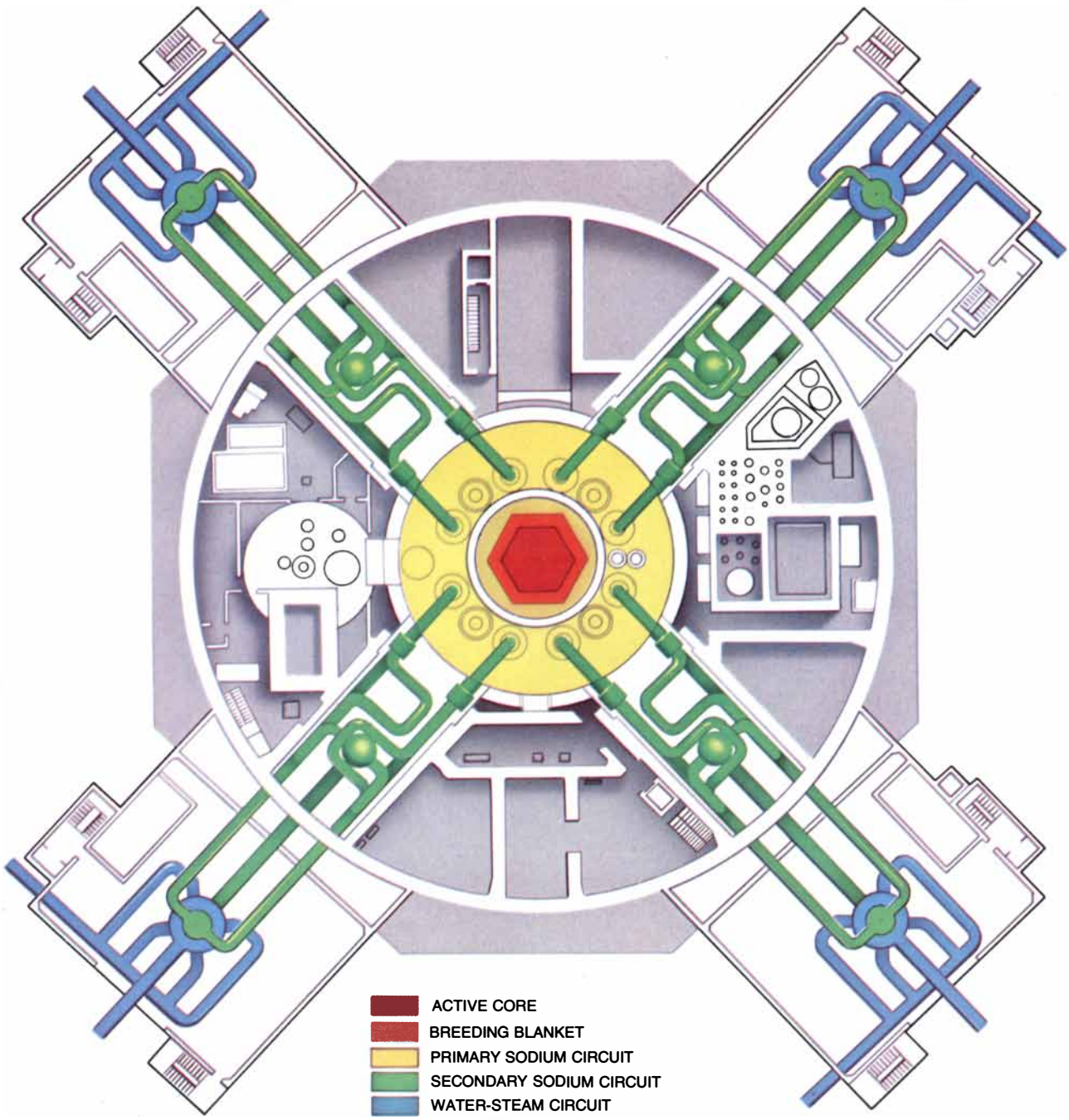
The second type of heavy isotope in the core of every reactor is said to be fertile; it undergoes practically no fission reactions, but by capturing a stray neutron a fertile nucleus can be transmuted into a fissile nucleus at the end of a series of radioactive disintegrations. A typical example of a fertile nucleus is uranium 238, which is transmuted by neutron capture into fissile plutonium 239. Similarly, fertile thorium 232, the only form of thorium extracted from the ground, can be transmuted into fissile uranium 233.

In every nuclear reactor, as the fissile nuclei are being consumed new fissile nuclei are being created by the transmutation of fertile nuclei. Most reactors in operation today, however, use either ordinary (light) water or deuterated (heavy) water to moderate, or slow, the neutron flux in the active core. In such a slow-neutron reactor it is impossible to produce as many fissile nuclei by neutron capture as are consumed. As a result the proportion of fissile nuclei in the fuel quickly falls below a certain minimum level, and the depleted fuel must be removed from service with most of the fertile nuclei still not transformed. A set of special conditions must be satisfied to raise the breeding ratio (the ratio of the amount of fissile material produced from fertile material to the amount of fissile material consumed during the same period) to a value greater than 1. The most favorable conditions for breeding are obtained when fissile plutonium 239 and fertile uranium 238 are used together in a fast-neutron reactor, in which the neutrons from the fission reactions are not slowed down by a moderating substance such as water between the time they are emitted by one fission reaction and the time they cause the next reaction. Only under these conditions can the breeding ratio be raised to a value significantly higher than 1.

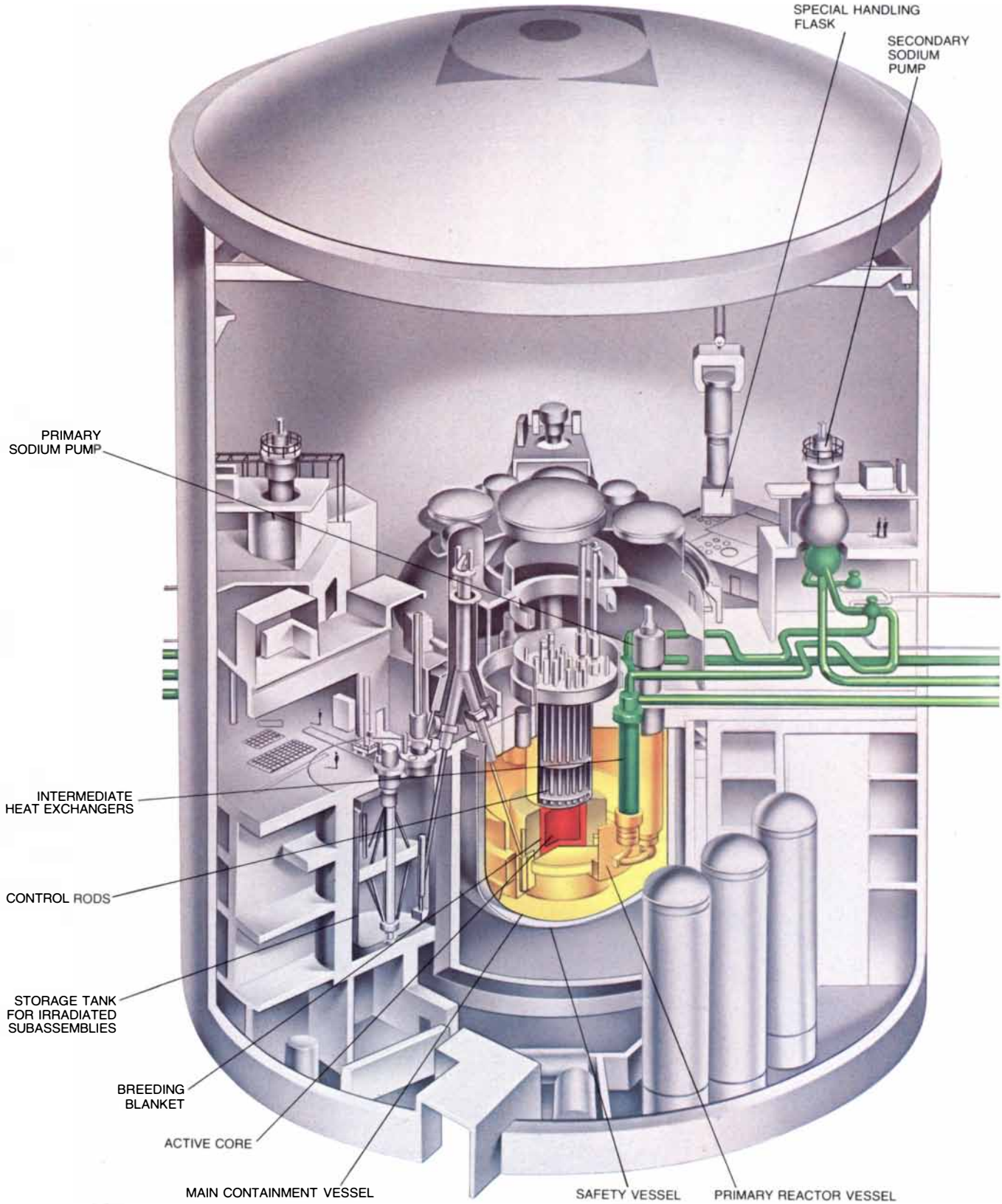
In a fast-neutron reactor the initial fuel load of plutonium is needed to start the fission chain reactions and the pro-

duction of power. During this period plutonium is bred from natural uranium (or from uranium depleted in uranium 235) in the reactor core and in the surrounding "breeding blanket." When the fuel subassemblies that make up the core and the blanket have undergone prolonged neutron irradiation, they must be reprocessed chemically in order to separate and remove the fission products. In each reprocessing operation more plutonium is recovered than existed at the start of the irradiation. The excess plutonium is set aside and is replaced in the reactor by natural or depleted uranium. Everything proceeds as though the reactor were consuming only natural or depleted uranium and simultaneously furnishing new plutonium as a by-product of the plant's operation.

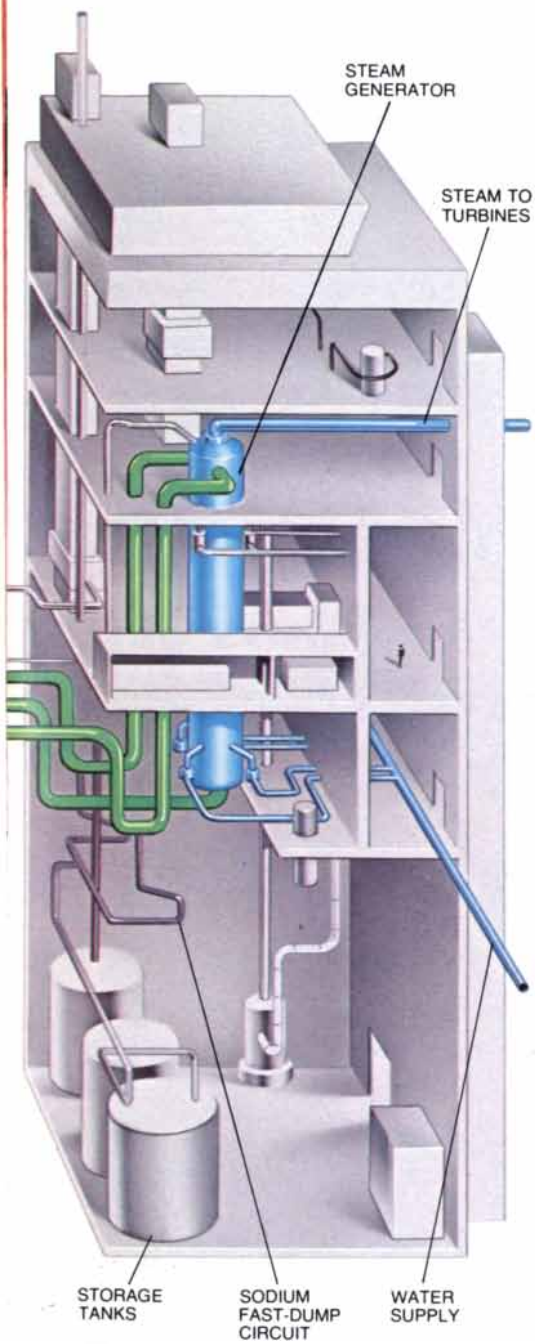
The time required for a breeder reactor to produce enough plutonium to fuel a second identical reactor is called the reactor's doubling time. This time factor is inversely proportional to the reactor's breeding ratio. In the future it is expected that breeding ratios on the order of



HORIZONTAL SECTION of the proposed Superphénix breeder-reactor power station shows the overall layout of the plant, which will consist essentially of a large circular reactor building with four steam-generator buildings laid out radially around it. The central reactor building, which is designed to house all the plant's nuclear components, will be built of reinforced concrete one meter thick; the building will have an inside diameter of 64 meters and a height of about 80 meters. Each steam-generator building will serve one segment of the secondary sodium circuit. (The associated turbogenerator building is not shown in this view.) The site selected for Superphénix is at Creys-Malville in France. Construction of the plant has received the backing of a consortium of European utilities, representing France (51 percent), Italy (33 percent), West Germany (11.04 percent), the Netherlands (2.36 percent), Belgium (2.36 percent) and the United Kingdom (.24 percent). The color coding adopted for this drawing and the ones on the next three pages is given in the key at the bottom.



VERTICAL SECTION of the Superphénix reactor building and one of the four identical steam-generating buildings shows the main operating components of the plant in somewhat greater detail. Superphénix is classified as a pool-type breeder reactor, which means that the active core, the primary sodium pumps and the intermediate heat exchangers are all located within a single large vessel; in this particular design the main steel containment vessel, which is hung from a steel-and-concrete upper slab, is 21 meters across and is filled with 3,300 tons of molten sodium. A cylindrical structure welded to the main vessel supports the control-rod mechanism and the fuel subassemblies, which constitute the active core of the reactor. The four primary pumps convey the sodium upward through the core. The primary reactor vessel separates the "cold" sodium, which enters at the bottom of the subassemblies at a temperature of 395 degrees Celsius, from the "hot" sodium, which leaves at the top at 545 degrees C. The hot sodium then flows downward through the eight intermediate heat



exchangers, which form part of a secondary circuit of nonradioactive sodium, inserted for reasons of safety between the primary sodium circuit and the water-steam circuit. Each of four secondary loops consists of two intermediate heat exchangers, a secondary pump installed inside spherical expansion tank and a steam generator in the adjacent building.

1.4 or so will be achieved, in part by exploiting the concept of the heterogeneous core [*see illustration on next page*]. The corresponding doubling times will then be between 10 and 20 years. Since it is unlikely that the consumption of electricity will double at shorter intervals toward the end of the century, a doubling time in this range will enable fast-neutron reactors to cope with the rising demand for energy unaided, by virtue of their self-fueling feature.

The breeding ratios of the fast-neutron reactors built today are not significant, since for several years the plutonium produced by light-water reactors will constitute the major, if not the exclusive, source of the initial fuel for fast-neutron reactors. Thus a remarkable complementarity exists between these two types of nuclear reactor. Over a fairly long period a two-pronged strategy of nuclear-power generation can be established, with the light-water plants leading the way for the gradual penetration of the market by the fast breeder plants.

Although fast-neutron plants are capable of producing more plutonium than they consume, that potential can be exploited or not. At the discretion of the user plutonium production can be higher or lower than consumption. The amount of plutonium available can be matched exactly to the demand, whether the latter rises, remains stable or even declines; hence a stock of unused plutonium need never be created. In the absence of fast-neutron reactors, on the other hand, it would be impossible to completely burn the plutonium and its transplutonium derivatives produced by the slow-neutron plants. These highly radioactive elements would constitute wastes that would have to be set aside and stored for thousands of years.

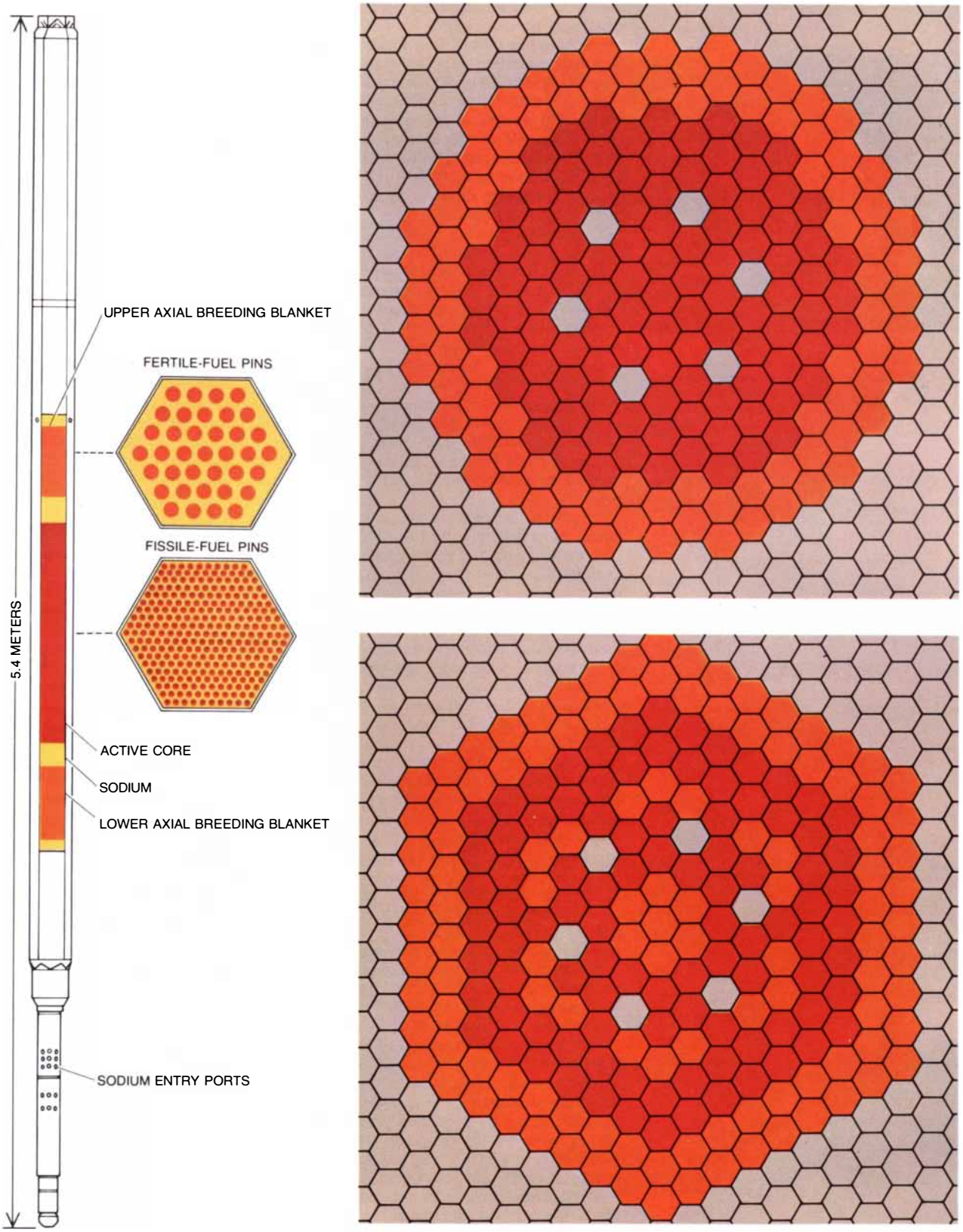
The fact that fission reactions are caused by fast neutrons in a breeder reactor makes the dimensions of the core very compact; the core volume of a 1,000-megawatt fast-neutron plant does not have to exceed 10 cubic meters. Fast-neutron reactors by their very nature generate a great deal of heat per unit of volume. To remove this intense heat output from the fuel subassemblies that make up the reactor core it is necessary to use a coolant endowed with outstanding thermal properties. Water is unsuitable because hydrogen is a powerful neutron moderator, and any material of that kind must be avoided.

Of all conceivable fluids molten sodium is the one that combines the most attractive array of properties. A liquid at 98 degrees Celsius, it boils at 882 degrees C. at atmospheric pressure. Since the maximum sodium temperature in the reactor core never exceeds 550 degrees in normal operation, it is not necessary to pressurize the vessels and the circuits that contain the sodium. Moreover, the excellent thermal properties of sodium mean that the steam produced in the steam generators has characteristics equivalent to those required to drive the turbines of the most modern fossil-fuel power plants. The overall efficiency of a fast-neutron reactor is equal to or greater than 40 percent, whereas the efficiency of a typical light-water power plant does not exceed 33 percent; the comparatively high efficiency of fast-neutron reactors is a positive feature with respect to thermal discharges into the environment.

Every fast-neutron reactor that has been or is being built in the world today calls for molten sodium as the coolant. The fact that all the countries with active breeder programs (including the U.S., the U.S.S.R., France, Britain, West Germany, the Benelux nations, Italy, Japan and India) have made the same basic technological choice is a very favorable factor. It avoids the spreading of effort mounted along divergent lines and enhances overall efficiency. The approach followed has been much the same in all the countries involved. Reactors built and planned during the still prevailing development phase belong to three categories that follow in a logical succession: experimental reactors, demonstration plants and prototype power stations.

In line with this logical sequence the forerunners of Superphénix were Rapsodie and Phénix. The experimental reactor Rapsodie (the name associates the words *rapide* and sodium) was commissioned in 1967. Its power level is low (40 megawatts of thermal output) and it does not produce any electricity. Nevertheless, its main features are representative of the breeder regime from a technical standpoint with respect to temperature and other factors. Rapsodie has operated in a satisfactory manner for almost 10 years, with an average availability of nearly 90 percent during the operating runs. It is in continuous use as a test facility for investigating the effects of prolonged irradiation on various fuel assemblies.

One year after Rapsodie went into operation the decision was made to build the Phénix demonstration plant, named for the mythological bird that was reborn from its own ashes. The achievement of a high breeding ratio was not of particular concern in the design stage. The principal purpose of Phénix was to confirm the validity and reliability of the entire system by demonstrating the possibility of building a fast-neutron power plant within a reasonable period of time and of running it satisfactorily. Phénix was put into regular operation in July, 1974. The record of the first two years is particularly gratifying. These excellent results do not mean that the demonstration is over. The day-to-day operation of the reactor is being closely watched, and unforeseeable incidents

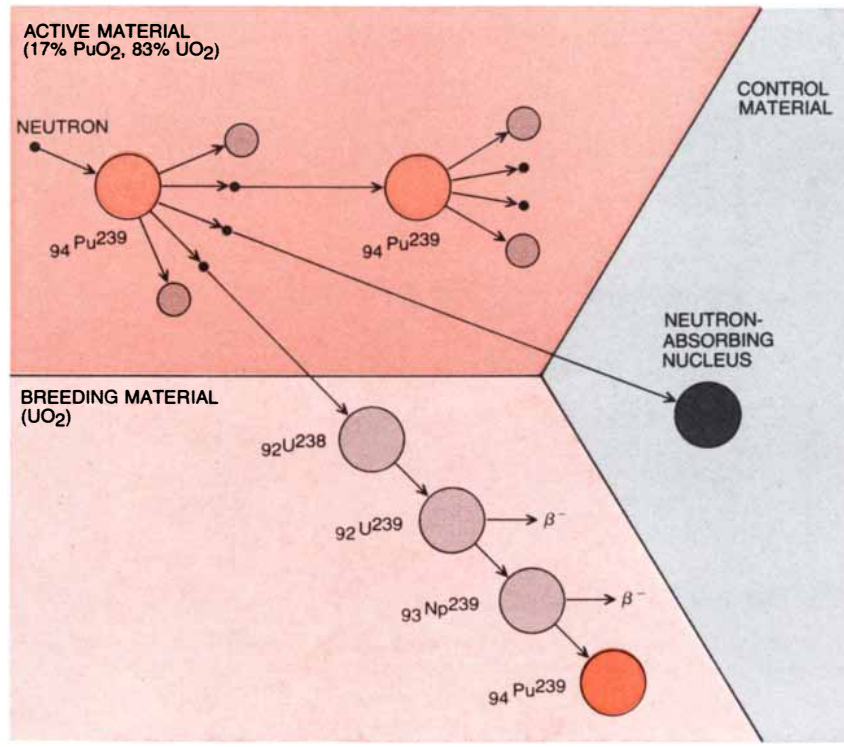


NUCLEAR-FUEL SUBASSEMBLY of the Superphénix reactor is shown in the cutaway vertical diagram at left. In each active-core subassembly the fuel is subdivided into 271 long, thin pins along which the sodium of the primary circuit (*yellow*) flows; the fissile material (*red*) occupies the central portion of the pin, fertile material (*orange*) being placed at both ends. (The fertile subassemblies contain fewer, larger pins.) Two alternative core designs under consideration for the breeder reactors of the future are represented by the schematic horizontal sections at the right. The two designs differ in the arrangement of the stainless-steel subassemblies: in the conventional core design (*top*) the central zone of fissile subassemblies is surrounded by an outer "breeding blanket" of fertile subassemblies; in the new heterogeneous core design (*bottom*) fertile material is inserted into core in the form of clusters of fertile subassemblies. Gray hexagons are control rods. Designs are idealized here; in reality active core and breeding blanket will account for a total of about 600 subassemblies.

could still occur. Small sodium leaks detected during the summer of 1976 in two intermediate heat exchangers have led to the temporary shutdown of the plant for repairs to the observed defects, which are minor and do not call the design into question. The initial results are considered encouraging enough to proceed with confidence.

Superphénix, the next step in the development sequence, will be the prototype for the commercial breeder power plants of the future. In design it is very similar to Phénix. It was thought essential for overall efficiency and success to maintain the continuity of technological choices as far as possible. In spite of this constraint continuous progress in acquired know-how led in some cases to significant changes with respect to Phénix, if only to meet increasingly stringent safety criteria. Creys-Malville, where Superphénix will be built, is in the upper Rhône valley, not far from the electric-power grids of Italy and Germany. The site selected for the plant, on the banks of the Rhône 40 miles east of Lyons, is in a sparsely populated farming region where no other major industrial projects are planned.

From the geological standpoint the Creys-Malville site is in a low-seismicity zone of Degree VI on the international macroseismic scale (which has a range of 12 degrees, with an interval of one degree corresponding to a factor of two in ground acceleration). The Superphénix plant is designed to continue operating after being subjected to a Degree VI earthquake, which corresponds to the maximum intensity already observed in the region. Furthermore, the design guarantees that all essential safety functions of the plant, such as the neutron shutdown of the reactor, the removal of residual power from the core and the integrity of the containment, will be maintained in the event of an earthquake of Degree VII in intensity.



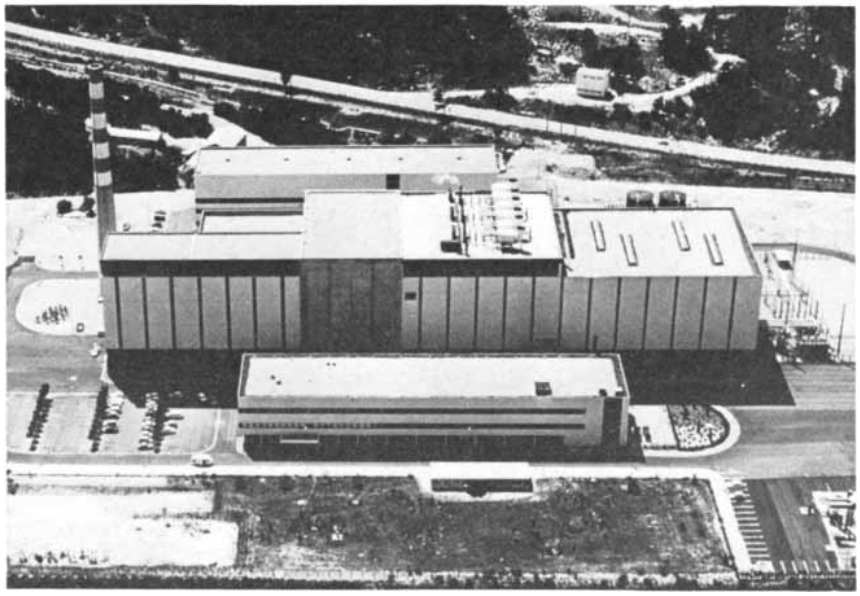
NUCLEAR REACTIONS that take place in the core of a breeder reactor are diagrammed. As in any nuclear reactor, the fissile active-core material (in this case plutonium 239) undergoes a self-perpetuating chain of fission reactions, yielding both the heat energy produced by the reactor and the neutrons that sustain the chain reaction. In the breeder fertile nuclei in the core and in the blanket material (in this case uranium 238) can also be transmuted into fissile nuclei by capturing a stray neutron each, thereby creating new fissile material. The absorption of neutrons by nuclei in the control rods can be adjusted to regulate the rate of the reactions (*right*). In a "fast" breeder the neutrons are not slowed down by a moderating substance such as water.

The Superphénix power station will be designed to adapt its operation to variations in demand on the electric-power grid. It will be operated as a base-load plant. The gross power output of the plant has been set at 1,200 megawatts of electricity, which is similar to the power level of light-water nuclear plants scheduled for construction at the same time. In 1985, 1,200 megawatts will represent between 1.5 and 2 percent of the total installed power of the French grid. The choice of this figure for Superphénix results from a compromise. On the one hand there is a trend toward large nuclear power plants on the grounds of economics; on the other hand extrapolation from Phénix to Superphénix must remain within reasonable limits.

A fast breeder plant does not differ greatly in its general layout and operating scheme from any other nuclear power station. Heat produced in the reactor core is conveyed by a fluid (molten sodium in this case) to water, producing steam, which feeds an electric-generating turbine unit. In order to avoid any accidental reaction between the radioactive sodium and the water an intermediate, or secondary, loop of nonradioactive sodium is inserted between the primary circuit conveying the sodium through the core and the water-steam circuit. Instead of one 1,200-megawatt turbogenerator of advanced design, two 600-megawatt units will be used in parallel, incorporating only conventional equipment that has proved its reliability in many oil-fired plants.

The design of the core and the fuel assemblies is a key factor in the realization of Superphénix. The core, as the seat of heat generation, is the most highly stressed of all the parts of a nuclear plant. This is all the more true in a fast-neutron core, where the heat production per unit of volume is exceptionally high (up to 500 kilowatts per liter) and all the structures are subjected to an intense flux of fast neutrons (6×10^{15} neutrons per square centimeter per second). To ensure that the heat is evacuated without giving rise to excessive temperatures the nuclear fuel is subdivided into long, thin pins (less than a centimeter in diameter) along which the sodium flows at a speed of six meters per second. The fissile material is located in the central portion of the pin, fertile material being placed at both ends. A cluster of 271 pins are fastened together within the hexagonal stainless-steel structure known as a fuel subassembly. All together 364 subassemblies, packed in a regular array, constitute the reactor's active core, which is in turn surrounded by 232 similar subassemblies containing larger pins of fertile material, representing the breeding blanket. The sodium flows upward, entering at the bottom of the subassemblies at 395 degrees C. and leaving at the top at 545 degrees. At the center of the active core 450 watts of heat energy is generated per centimeter of fuel pin.

A fuel mixture with a mean composition of 17 percent plutonium oxide (PuO_2) and 83 percent uranium oxide (UO_2) has been selected as the fissile material; the fertile material consists of uranium oxide alone. Long and satisfactory experience with these materials has been gained in the operation of Rapsodie and Phénix. Of the 25,000 PuO_2-UO_2 fuel pins that have been irradiated so far in Rapsodie, 3,000 survived a burnup of 80,000 megawatt-days per ton and some have reached 150,000 megawatt-days per ton. Less than one pin per 1,000 irradiated failed. So far



PHÉNIX, a 250-megawatt demonstration breeder-reactor power station, is located on the Rhône River near Avignon. The plant began generating electricity at full power in July, 1974.

15,000 fuel pins have been irradiated in the Phénix core. At present subassemblies are taken out of the reactor as soon as they reach a burnup of 50,000 to 65,000 megawatt-days per ton. Not a single pin has failed while in service in Phénix.

Development work is also being devoted to new carbide and nitride fuels, which are likely to exhibit breeding characteristics superior to those of the oxides of plutonium and uranium currently called for. It remains to be seen whether this potential advantage will be offset by increased difficulties in fabrication, irradiation behavior and chemical reprocessing. The use of carbide and nitride fuels in Superphénix is not contemplated at this time.

Another important technical problem concerns the choice of the material for the hexagonal structure of the subassemblies and for the pin tubes, which must meet very stringent requirements. They must maintain good mechanical strength at temperatures approaching 650 degrees C. Furthermore, the internal pressure in the pin tubes may be as high as 30 kilograms per square centimeter, owing to the buildup of gaseous fission products. The pins are also subjected to considerable thermal stresses. Last but not least, they are exposed to a peculiar phenomenon: under prolonged irradiation by fast neutrons, vacancies form in the crystal lattice of the metal and grow into tiny cavities, causing the metal to swell. Some idea of the intensity of neutron bombardment in a high-power fast-neutron reactor can be gained from the fact that every atom of the material "cladding" the fuel pins is struck or at least caused to vibrate once every 100 hours on the average by the passage of a neutron or another atom recoiling from a neutron collision. Another impressive figure is the cumulative number of fast neutrons crossing any given square centimeter of the cladding material after irradiation in the reactor core: this figure approaches one full gram of neutrons! The swelling of metallic alloys under neutron irradiation must be kept low enough to avoid deformation of the subassembly, which is liable to raise problems in the operation of the reactor, particularly in the fuel-handling maneuvers. A great deal of research-and-development work has already been accomplished but more is required in order to find a complete solution to the problem.

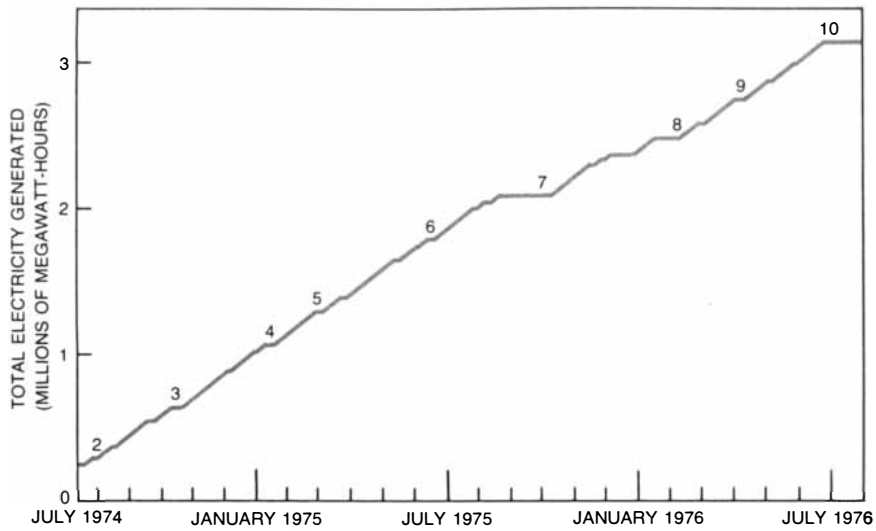
The different types of fast-neutron reactor are distinguished essentially by the organization of the primary sodium circuit. In the pool design the reactor core, the intermediate heat exchangers and the primary sodium pumps are all within a single large vessel. In the loop design only the reactor core is housed within the vessel and the intermediate heat exchangers and pumps are connected to it by loops. It must be stressed that the two systems rely on the same technology, that most development work on components is common to both and that the differences between the two concepts are much less than those between, say, pressurized-water and boiling-water reactors. In most countries loop-type reactors were built first, since the separation of components facilitated construction, operation and maintenance, justifying such a choice at an early stage of development. The first pool-type breeder reactor in the world was built in the U.S. more than 10 years ago. Following the loop-type construction of Rapsodie, the pool concept was adopted for Phénix and, owing to the excellent record of that plant, it was maintained fundamentally unchanged for Superphénix.

It is clear that both the pool system and the loop system can be built and run, and that both have advantages and drawbacks only long operating experience can distinguish. Among the main reasons for the selection of the pool system for Phénix and Superphénix, following a meticulous comparison with the loop system, was a safety consideration. For a large plant, say 1,000 megawatts or more, it was thought the integrity of the primary sodium circuit could be maintained in all reasonably foreseeable circumstances more readily by enclosing it within a single vessel of simple design than by dispersing it in a highly intricate system of pipes and vessels involving many hundreds of meters of piping up to one meter in diameter. Although the main pool-type vessel is larger than the loop-type reactor vessel (roughly 20 meters in diameter as against 10), the pool-type vessel is much more straightforward in design. As a result construction, inspection and maintenance are far easier. The main problem encountered in the pool design concerned the cover of the main vessel. The solution implemented in Phénix could not be extrapolated to the dimensions of Superphénix. It was decided to hang the steel main vessel directly from the steel-and-concrete upper slab, and to put under the slab a layer of metallic thermal insulation that is in contact with the argon atmosphere above the sodium. The tests performed to date indicate that this arrangement is entirely satisfactory.

Experience with nuclear power plants of every type has shown that the steam generator is a crucial component. In fast-neutron reactors particular care must be taken in design and construction to prevent any violent chemical reaction between the sodium and the water, which would result from a leak in the exchanger tubes. The steam-generator model selected for Phénix, the only one with which extensive experience had been gained at the time, was subdivided into 36 low-power modules (17 megawatts each). The subdivision made it possible to subject three complete full-scale modules to thorough tests in simulated operational conditions. Although this approach was justifiable for an initial project, it could not be maintained for a large power plant because of its prohibitive cost. Research for Superphénix was therefore oriented toward units of different design, with a higher power per unit (several hundred megawatts). The problems presented by the

fabrication and the operation of these units did not appear to grow with size, but the large modules do have certain drawbacks, the main ones being the near impossibility of conducting full-scale tests prior to installation in the power plant and the increased electric-power loss in case of the unavailability of a unit.

The tests performed under normal and accidental conditions on two "once through" mock-ups, one with straight ferritic steel tubes and the other with helical Incoloy tubes, provided complete satisfaction and showed good agreement with the design forecasts. The helical-tube model was finally selected for Superphénix, with each secondary loop including a steam generator with a thermal power of 750 megawatts. A steam-reheat stage can be added with either sodium or steam. The sodium system was employed for Phénix, raising the net efficiency of the plant to 42 percent. The steam system was adopted for Superphénix, simplifying the steam generator and the associated circuits, because a cost study showed that the lowered investment cost offset the loss in efficiency.



OPERATING RECORD OF PHÉNIX during its first two years was considered encouraging enough to proceed with Superphénix, the next stage in the French breeder-development program. The level parts of this cumulative electric-generating curve correspond to plant shutdowns; the numbers refer to refueling shutdowns. An extended work period scheduled after one year of operation took place during seventh refueling shutdown. In its first two years of operation Phénix generated electricity at full power for the equivalent of 530 days, a better performance than is typical of present-day light-water reactors in their first two years of operation.

It is obviously important to prevent the development of even the smallest leak in the tubes separating the water from the sodium and to minimize the effects of any contact of the two fluids that may nevertheless occur. Ultrasensitive hydrogen detectors (capable of detecting a leak of as little as two milligrams of hydrogen per second) will be housed in several locations in each steam generator. Automatic systems designed to limit the consequences of an incipient reaction are also available. Two such systems consist of automatic valves that immediately shut off the sodium circuits and discharge systems designed to remove the products of the reaction and to limit the ensuing pressure surge.

The maneuvers required to convey the fuel subassemblies to their core positions and to withdraw them from the reactor after irradiation will be conducted exclusively during plant shutdown. They will be carried out by a series of devices that manipulate the assemblies in sodium at all times, in order to allow the removal of the residual heat released by the fission products. Two eccentric rotating plugs housed in the upper slab of the reactor will make it possible to position the device that grips the subassembly heads above any point of the core and blanket. This system, which allows direct service above each subassembly, also copes with potential deformations of the subassemblies due to swelling under irradiation. One of the main drawbacks of sodium is its opacity, which makes it impossible to follow the progress of the handling operations visually. Ultrasonic transmitter-receiver units, employing the principle of sonar, have been developed to surmount this obstacle. These devices, operating within the sodium itself, provide a guarantee that the subassemblies manipulated will occupy their correct positions at all times, without the risk of colliding with the handling devices.

The principles underlying the control of a fast-neutron reactor are identical with those of any other nuclear reactor. The existence of delayed neutrons gives the mechanisms acting on core reactivity the time to act smoothly, whether to raise or lower the power of the plant or to keep it stable. These operations are performed by means of control rods containing a suitable neutron-absorbing material, which move in channels parallel to the fuel subassemblies. Superphé-

	PHÉNIX	SUPERPHÉNIX
GROSS ELECTRICAL RATING (MEGAWATTS)	264	1,240
THERMAL RATING (MEGAWATTS)	590	3,000
GROSS EFFICIENCY (PERCENT)	44.75	41.5
VOLUME OF CORE (LITERS)	1,227	10,820
LENGTH OF FUEL ASSEMBLIES (METERS)	4.3	5.5
NUMBER OF FUEL PINS PER ASSEMBLY	217	271
OUTSIDE DIAMETER OF FUEL PINS (MILLIMETERS)	6.6	8.65
MAXIMUM LINEAR POWER (WATTS PER CENTIMETER)	430	450
RATE OF FUEL BURNUP (MEGAWATT-DAYS PER TON)	50,000	70,000
MAXIMUM TOTAL NEUTRON FLUX (NEUTRONS PER SQUARE CENTIMETER PER SECOND)	7.2×10^{15}	6.2×10^{15}
BREEDING RATIO	1.12	1.24
NOMINAL CLADDING TEMPERATURE (DEGREES CELSIUS)	650	620
INTERVAL BETWEEN REFUELING OPERATIONS (MONTHS)	2	12

PHÉNIX AND SUPERPHÉNIX are compared in this table. The continuity of technological choices was maintained as far as possible in going to the larger plant, although a number of significant changes were incorporated in Superphénix design, partly to meet new safety criteria.

PRIMARY CIRCUIT	NUMBER OF PUMPS	4
	TEMPERATURE AT CORE INLET	395 DEGREES CELSIUS
	TEMPERATURE AT CORE OUTLET	545 DEGREES CELSIUS
	OVERALL SODIUM FLOW RATE	16.4 METRIC TONS PER SECOND
	WEIGHT OF SODIUM	3,300 METRIC TONS
SECONDARY CIRCUIT	NUMBER OF PUMPS	4
	NUMBER OF INTERMEDIATE HEAT EXCHANGERS	8
	TEMPERATURE AT INLET OF INTERMEDIATE HEAT EXCHANGER	345 DEGREES CELSIUS
	TEMPERATURE AT OUTLET OF INTERMEDIATE HEAT EXCHANGER	525 DEGREES CELSIUS
	OVERALL SODIUM FLOW RATE	13.2 METRIC TONS PER SECOND
	TOTAL WEIGHT OF SODIUM IN THE FOUR LOOPS	1,700 METRIC TONS
TERTIARY CIRCUIT	NUMBER OF STEAM GENERATORS	4
	TEMPERATURE OF SUPERHEATED STEAM	490 DEGREES CELSIUS
	PRESSURE OF SUPERHEATED STEAM	180 BARS
	OVERALL STEAM FLOW RATE	1.36 METRIC TONS PER SECOND

SPECIFICATIONS for the steam-generating system of Superphénix are summarized in this table. The use of sodium in the reactor's primary and secondary circuits is expected to give the new electric station a comparatively high overall thermal efficiency of at least 40 percent.

nix will be provided with a highly redundant system of control rods, divided into three independent groups. One of these is specially designed to penetrate the core even in the extreme and improbable case of its undergoing a large deformation. Uninterrupted monitoring of the Superphénix core is provided by a diversified set of detectors, whose output is processed and correlated by computer. The temperature of the sodium leaving each subassembly is measured by three thermocouples; two of them are of the chromel-alumel type and the third uses a sodium-steel couple and responds almost instantaneously. The boiling-sodium detectors, flowmeters and devices for the detection and localization of any cladding failures are improved versions of those employed in Phénix. The neutron detectors and electronic instruments for measuring variations in reactivity have proved their reliability through long experience with them.

The many precautions implemented in the design of Superphénix were subjected to detailed scrutiny by the licensing authorities before their approval was secured. These safety measures reduce the probability of an accident to an extremely low level. The procedure followed went to the extent of considering the case in which a total shutdown of forced sodium circulation through the core at full power is not accompanied by any action of the many control systems designed to shut down the fission chain reaction and energy production immediately. Even in this case the considerable thermal inertia represented by the large mass of sodium present in the primary vessel (3,300 tons) and the interval of several hundred degrees C. between the sodium temperature in normal operation and its boiling point furnish a substantial time interval for manual emergency action. Nevertheless, it is necessary to ensure that even in the highly improbable case of a serious accident involving a core meltdown the consequences will be contained in such a way that no significant quantity of plutonium or fission products can escape into the environment.

The containment system for Superphénix therefore consists of a series of successive enclosures, which can withstand both internal reactor accidents and external aggression such as an airplane crashing into the power plant. Finally, special arrangements have been made to prevent potential sodium fires and to limit their spread should they occur. Sodium fires would not actually jeopardize the safety of the installation, but it is nonetheless necessary to take full precautions to maximize the reliability and the availability of the power station.

In all areas, not just in the priority area of safety, a considerable research-and-development effort has preceded the design and construction of Superphénix. This program, which calls for full-scale tests in sodium of all components where innovations have been made, will continue to back up construction of the reactor in the coming years.

Phénix was built in slightly more than four years. Preliminary site preparation began in the fall of 1968, and the filling of the primary and secondary circuits with 1,400 tons of sodium was started before the end of 1972. For Superphénix a building schedule spread over 68 months has been adopted. Construction deadlines are comparable to those set for other types of nuclear power plant. The fact that breeder reactors are not pressurized and that their components, even the large ones, are made of comparatively thin stainless-steel sheet and pipe makes it possible to perform most of the final assembly on the site. The Phénix experience clearly showed the advantages of this approach and the flexibility that it engendered in adherence to the construction schedule.

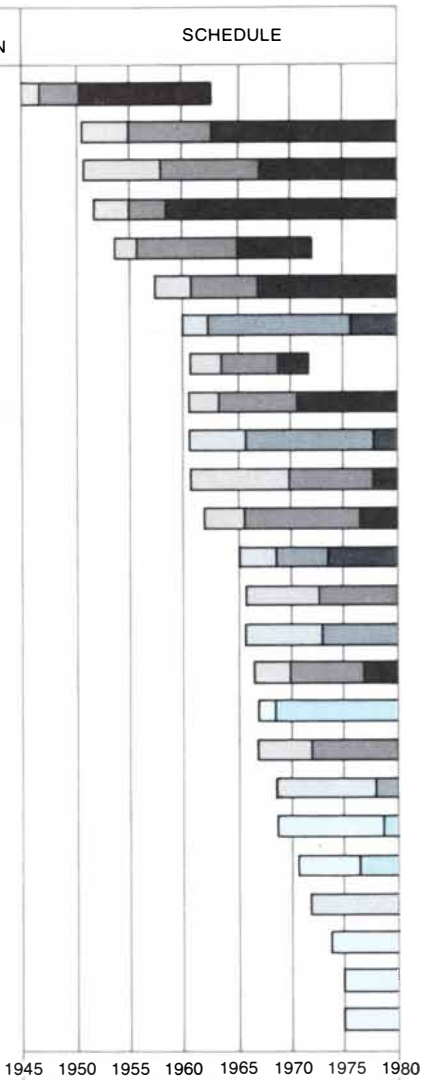
The investment cost of the Superphénix power plant will significantly exceed that of a light-water plant of comparable output. This cost differential is unavoidable since Superphénix is the first plant of its kind, whereas light-water plants (of which more than 100 have been built to date in the world) have amply profited from the fruits of technical progress and above all of series production. In spite of the fact that Superphénix is a prototype, it should be emphasized that the cost of a kilowatt-hour of electricity produced by Superphénix will be in the same range as that produced by an oil-fired power station. It is probable that the investment cost of fast breeder plants, which will progressively decline as larger numbers are constructed, will remain for some time higher than that of light-water plants, if only because of the larger amounts of stainless steel employed and the presence of an intermediate sodium circuit added for safety reasons. Economic competitiveness with light-water plants will derive from a cheaper fuel cycle, made possible by fuel breeding, and this factor will become increasingly important with the foreseeable increase in the price of natural uranium.

The expansion program of the French national utility company Électricité de France (EDF) already calls for a series of breeder plants, employing plutonium provided by a large number of pressurized-water reactors built simultaneously. It is reasonable to expect that two pairs of fast-neutron plants will be initiated in France between 1980 and 1985, representing, together with Superphénix, about 8,000 megawatts of electric-generating capacity in service in the early 1990's. Commitments may grow to 2,000 megawatts per year after 1985, so that by the year 2000 fast-neutron plants may account for about a fourth of the installed capacity and a third of the total

energy output of all the nuclear plants in France. Simultaneous with the successive launching of these plants will be the construction of plants for the fabrication and reprocessing of fast-breeder fuels, thus closing the fuel cycle. The latter will be high-capacity plants (with an output of about 200 tons of oxides per year) aimed at achieving a low overall fuel-cycle cost.

The importance of Superphénix must be gauged in relation to the coming generation of power plants derived from it. It is in a way the culmination of the technological development phase and the final stage before the commercial series, the technical definition of which will rely directly on the Superphénix experience. If everything proceeds as planned, by the mid-1980's, thanks to Superphénix, one may expect to have at least a preliminary operating record with a large fast-neutron power plant. This experience, which will be shared by several large electric utilities, symbolizes the joining of efforts by the European countries involved in aiming at the earliest possible commercial launching of a type of reactor that is indispensable to their economies.

REACTOR	LOCATION	POWER (MEGAWATTS)	FORM OF OUTPUT	BASIC DESIGN
EXPERIMENTAL BREEDER REACTOR 1	U.S.	.2	ELECTRICAL	LOOP
DOUNREAY FAST REACTOR	U.K.	15	ELECTRICAL	LOOP
EXPERIMENTAL BREEDER REACTOR 2	U.S.	20	ELECTRICAL	POOL
B.R. 5	U.S.S.R.	10	THERMAL	LOOP
ENRICO FERMI FAST BREEDER REACTOR	U.S.	66	ELECTRICAL	LOOP
RAPSODIE	FRANCE	40	THERMAL	LOOP
B.N. 350	U.S.S.R.	150	ELECTRICAL	LOOP
SOUTHWEST EXPERIMENTAL FAST OXIDE REACTOR	U.S.	20	ELECTRICAL	LOOP
B.O.R. 60	U.S.S.R.	12	THERMAL	LOOP
PROTOTYPE FAST REACTOR	U.K.	250	ELECTRICAL	POOL
FAST-FLUX TEST FACILITY	U.S.	400	THERMAL	LOOP
K.N.K. 2	GERMANY	20	ELECTRICAL	LOOP
PHÉNIX	FRANCE	250	ELECTRICAL	POOL
P.E.C.	ITALY	116	THERMAL	LOOP
S.N.R. 300	GERMANY	300	ELECTRICAL	LOOP
JOYO	JAPAN	100	THERMAL	LOOP
B.N. 600	U.S.S.R.	600	ELECTRICAL	POOL
FAST BREEDER TEST REACTOR	INDIA	15	ELECTRICAL	LOOP
CLINCH RIVER BREEDER REACTOR	U.S.	380	ELECTRICAL	LOOP
COMMERCIAL FAST REACTOR	U.K.	1,300	ELECTRICAL	POOL
SUPERPHÉNIX	FRANCE	1,200	ELECTRICAL	POOL
MONJU	JAPAN	300	ELECTRICAL	LOOP
S.N.R. 2	GERMANY	1,300	ELECTRICAL	
B.N. 1500	U.S.S.R.	1,500	ELECTRICAL	
PROTOTYPE LARGE BREEDER REACTOR	U.S.	1,200	ELECTRICAL	



TYPE OF FACILITY	DESIGN PHASE	CONSTRUCTION PHASE	OPERATION PHASE
EXPERIMENTAL REACTOR			
DEMONSTRATION PLANT			
COMMERCIAL PROTOTYPE			

WORLD SURVEY of progress in the development of liquid-metal-cooled fast breeder reactors includes all facilities with a thermal-power output greater than one megawatt. The plants are listed in chronological order according to the beginning of their design stage. Different colors are used to distinguish the three main categories of reactors built or planned so far: experimental reactors, demonstration plants and prototype commercial power stations. Different intensities of color denote design, construction and operation phases. Projections are made only to 1980; bars that stop short of the present represent decommissioned facilities. S.N.R. 300, Superphénix and S.N.R. 2 are multinational European projects. The German K.N.K. reactor has been in operation since 1968 with a slow-neutron core; beginning in 1977 it will be run with a fast-neutron core as K.N.K. 2. The British Prototype Fast Reactor at Dounreay in Scotland and the Russian B.N. 350 at Shevchenko, two demonstration plants comparable in size to the French Phénix, are both completed but have not yet been run at full power, owing to difficulties with their steam-generating equipment. Preliminary site work is about to begin on the closest comparable U.S. plant, the 380-megawatt Clinch River project near Oak Ridge, Tenn., which is expected to be completed in the early 1980's.

Waves in the Solar Wind

The solar wind is the sun's outer atmosphere expanding into space. Variations in its speed are interpreted as velocity waves that evolve with distance, much as ocean waves steepen as they approach a beach

by J. T. Gosling and A. J. Hundhausen

The corona, the tenuous outer atmosphere of the sun, has been familiar to observers for centuries as the faint silvery glow surrounding the black disk of the moon during a total eclipse of the sun. Although the corona was first analyzed spectroscopically in the 19th century, its most remarkable characteristic remained unrecognized until the 1940's. At that time previously unidentified emission lines in the coronal spectrum were found to be radiation emitted by common elements, such as oxygen and iron, that were stripped of most of their outer electrons. The presence in the corona of such highly ionized atoms can be explained only if the temperature of the corona is higher than a million degrees Kelvin, more than 100 times hotter than the underlying visible surface of the sun. The thermal energy of the ionized coronal gas is in fact so great that even the enormous gravitational field of the sun cannot retain the corona in a static bound atmosphere. Instead the coronal gas expands outward from the sun at tremendous speed, filling interplanetary space with a solar wind.

Satellites and spacecraft probing interplanetary space over the past 15 years have directly sampled the solar wind and have monitored its flow in the vicinity of the orbit of the earth. The observations reveal that the solar wind flows almost radially outward from the sun at an average speed of 450 kilometers per second, faster than a million miles per hour. The flow of the solar wind, however, is far from uniform; indeed, it is extremely gusty. Its speed commonly varies by several hundred kilometers per second over a period of about a week. The largest of the variations can be thought of as velocity waves that evolve in form as they propagate outward from the sun to the outer reaches of the solar system. Other types of waves are also present in the solar wind, but they will not concern us here.

Like the corona the solar wind is a highly rarefied plasma, or ionized gas, consisting primarily of free protons and electrons. Since it expands almost radially away from the sun at nearly constant velocity, its density decreases at a rate that is approximately proportional to the inverse square of its distance from the sun. At the orbit of the earth the density is some 10 particles per cubic centimeter; at the orbit of Jupiter it should be about .4 particle per cubic centimeter. As the solar wind expands it cools; near the orbit of the earth its temperature has fallen from a million degrees to about 100,000 degrees. At that temperature the speed of sound through the gas is about 50 kilometers per second. Since at the orbit of the earth the solar wind is traveling some 450 kilometers per second, the average speed of the gas is nine times faster than the speed of sound, that is, the flow is highly supersonic. Spacecraft observations further reveal that embedded in the gas is a weak remnant of the sun's magnetic field; near the orbit of the earth the average value of the field is .00005 gauss. (The strength of the earth's magnetic field near the poles is .6 gauss.)

In order to understand how waves in the solar wind form and evolve it is useful to think of the solar wind as being an ordinary compressible fluid. Such an approach does, however, ignore the details of the complex internal kinetic processes making such a tenuous gas behave as a fluid. Near the orbit of the earth the density of the gas is so low that any proton in the solar wind may collide with one of its neighbors on the average of only once in two or three days. Therefore processes other than direct collisions between particles must be responsible for causing the dilute gas to act as a fluid. Those other processes are not yet completely understood, although it appears likely that inherent instabilities of a plasma in the presence of the weak magnetic field in interplanetary space are ultimately responsible.

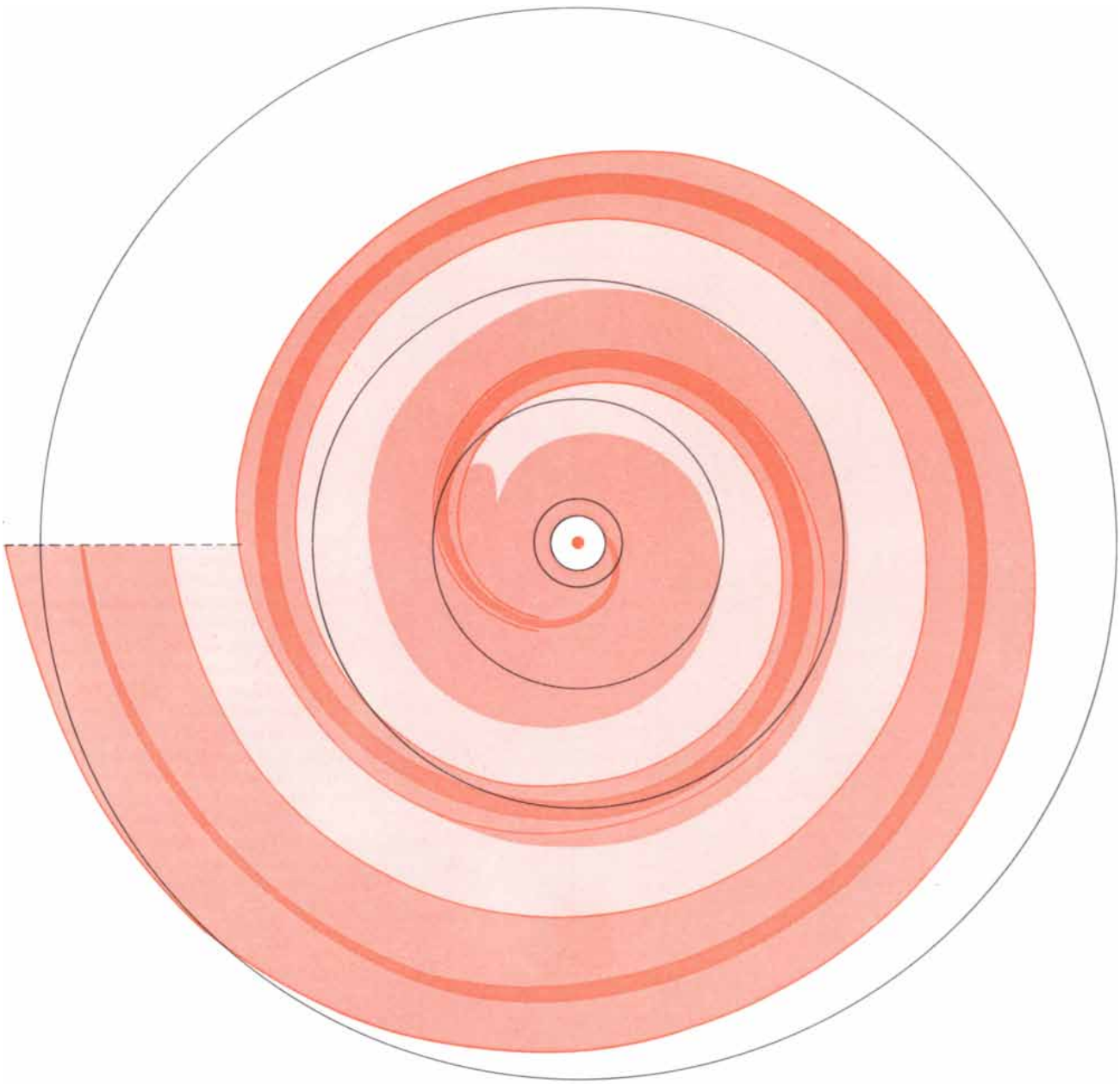
Observations from spacecraft near the orbit of the earth have revealed that the main variations of the solar wind's speed, density and pressure conform to a characteristic pattern. That pattern has come to be known as the solar-wind stream structure. At the leading edge of a solar-wind stream the speed of the wind's flow rises rapidly; at the trailing edge of the stream the speed declines much more slowly back to low values. At the leading edge where the speed rises the density of the particles of the gas also rises, quickly attaining a maximum that is several times greater than the average density. At the trailing edge where the speed falls the particle density also falls, attaining abnormally low values. The pressure across the stream, which is proportional to the product of the density and the temperature, also reaches a maximum as the speed rises, but it quickly falls again and remains nearly constant throughout much of the trailing portion of the stream. That characteristic pattern of variation near the orbit of the earth is readily explained if one thinks of the stream as a velocity wave that gets steeper as it travels away from the sun.

How do velocity waves first develop in the solar wind? What makes them get steeper? The lack of homogeneity in the speed of the solar wind stems ultimately from the nonuniformity of the temperature and density of the corona. This coronal nonuniformity arises primarily from the fact that the sun's magnetic field, which permeates the corona, is itself nonuniform from one place to another on the sun. Thus different portions of the corona escaping into space to become the solar wind expand outward from the sun at different rates.

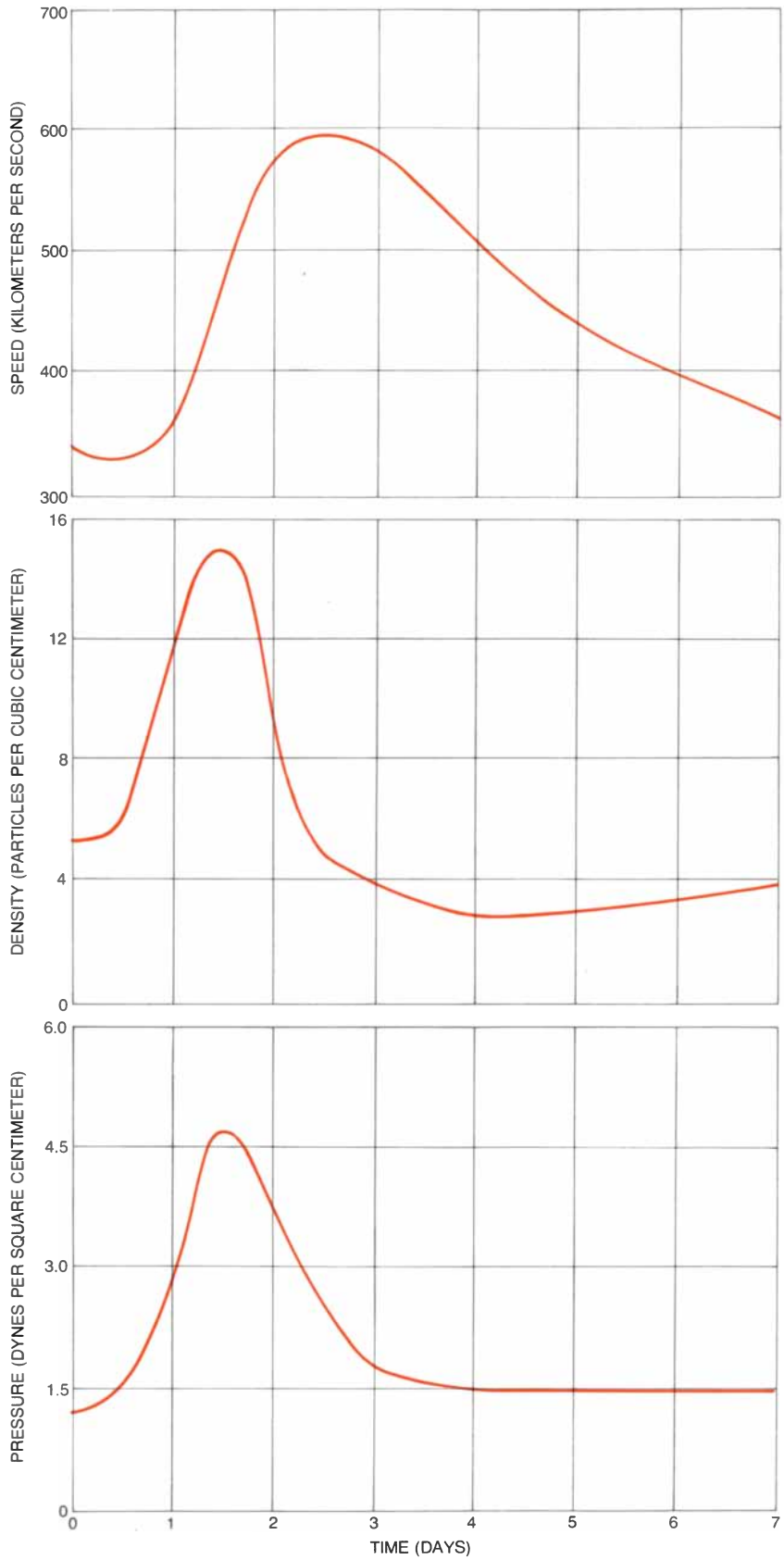
Consider now what happens at the earth as the sun rotates once every 27 days. Alternately slow, then fast, then slow gas is directed toward the earth. The variation in speed from low values to high ones and back to low ones again constitutes one complete stream or wave in the solar wind. The faster-moving gas in the wave will overtake the slower-moving gas ahead of it and will collide with it while simultaneously running away from the slower-moving gas behind it. Hence the wave will evolve toward a sawtooth form at greater distances from the sun. Since the gas is

compressible, the material within the wave is rearranged as the wave steepens. Individual parcels of gas near the leading edge of the wave are compressed and heated, and parcels on the trailing edge are rarefied and cooled.

Some of the slower-moving gas ahead is swept up by the wave and accelerated. The resulting transfer of momentum and energy from the faster gas to the slower will slightly decrease the wave's amplitude: the maximum difference in the velocity of the gas from one point to another within the wave. An observer near the orbit of the earth would see the evolved velocity wave and its accompanying regions of compression and rarefaction sweep past his observation point. The temporal variations in the speed, density and pressure associated with streams in the solar wind can be



SOLAR-WIND STREAM (*broad band in color*) expands outward from the sun (*colored dot*), pulled by the sun's rotation into a pattern resembling a spiral. Such a stream can be regarded as a wave in the solar wind created when the expanding gas of one region of the sun's corona, or outer atmosphere, travels faster than that of other regions. At any one fixed longitude in space the faster gas of the wave will gain on the slower gas in front of it while simultaneously running away from the slower gas behind it. Thus the gas at the leading edge of the wave will be compressed (*narrow band in darkest color*) and the gas at the trailing edge will be rarefied (*lightest color*). At a distance of about three astronomical units from the sun (one astronomical unit is the mean distance from the sun to the earth) this particular solar-wind wave develops two shocks (*lines in solid color*) that diverge from the region of greatest compression. The wave itself widens as it travels farther from the sun, and its amplitude, the difference between the velocity of the fastest gas and that of the slowest gas, decreases. Here the development of the wave is traced out to a radius of 20 astronomical units; the thin concentric black circles mark the orbits of the earth (one astronomical unit from the sun), Mars (1.5 astronomical units), Jupiter (5.2 astronomical units), Saturn (9.5 astronomical units) and Uranus (19.2 astronomical units). On the scale of this illustration it is not possible to depict the evolution of the wave inside the orbit of the earth. If the solar-wind wave were continued out beyond 20 astronomical units for one more revolution, the shock at the trailing edge of the next revolution (not shown) would intersect the shock at the leading edge of the wave's outermost revolution shown here.



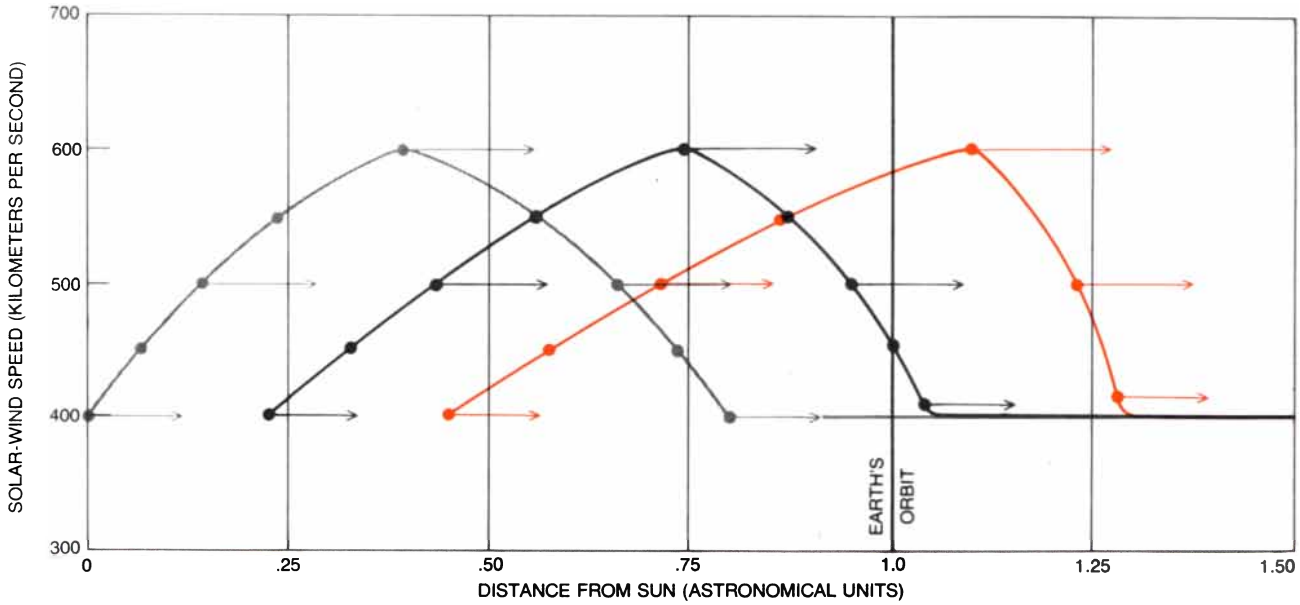
IDEALIZED PROFILE OF A SOLAR-WIND STREAM has been determined from many observations of solar-wind waves sweeping by a satellite in orbit around the earth. The leading edge of the stream's profile is at the left side of each panel. Within any such wave the speed of the solar wind rises rapidly to a maximum and then declines more slowly back to a low value (*top*). As the faster gas encounters the slower gas in front of it the density of the gas at the leading edge sharply increases (*middle*), leaving behind an area of rarefaction. As density increases, the pressure in gas also builds up, decreasing once again in trailing portion of wave (*bottom*).

readily identified with such evolved velocity waves.

Solar-wind velocity waves steepen with increasing distance from the sun in a manner analogous to the way ocean waves steepen as they move toward a beach. There are, of course, differences between waves in water and waves in the solar wind. First, water is incompressible, whereas the solar wind is easily compressed. Second, water waves convey energy and momentum but not mass, whereas the solar wind transfers energy, momentum and mass from the sun to the outer reaches of the solar system. In spite of these differences the analogy is close.

In the case of ocean waves the speed of the wave in shallow water is proportional to the square root of the water's depth. Hence the crests of ocean waves approaching a beach travel faster than the troughs, which are at a lesser depth, and the waves get steeper. Eventually large ocean waves running toward a beach steepen so much that the crests literally outrun the troughs. At that point the waves topple over and break. What happens in the solar wind when at large distances from the sun the crest of a velocity wave begins to outrun the trough? In other words, what is the interplanetary analogy of a breaker?

Let us consider in more detail the situation of a single parcel of gas halfway up the leading edge of a solar-wind velocity wave that has an amplitude of v_0 [*see illustration on page 40*]. Gas in the trough ahead of the wave is moving slower than the parcel, and thus the parcel is gaining on it. Gas on the crest of the wave behind the parcel is moving quicker than the parcel, and thus is overtaking it. An observer moving with the parcel would thus see gas streaming toward the parcel from both the direction of the sun (the faster-moving gas in the crest) and the direction opposite the sun (the slower-moving gas in the trough) at speeds ranging up to a maximum of $\frac{1}{2}v_0$. He would not see the gas streaming into the parcel, however, because the weak magnetic field embedded in the solar wind keeps the parcels of gas from interpenetrating. Hence pressure builds up in the gas around the parcel and resists the inflow from both sides. To the observer moving with the parcel the gas coming from both sides is decelerated, compressed and heated as it encounters the region of high pressure. To a stationary observer watching the solar-wind wave from a fixed point in space the deceleration of the gas coming from both directions will appear only as a deceleration of the high-speed gas near the crest and an acceleration of the low-speed gas near the trough. As a result momentum and energy are transferred from the fast-moving gas at the crest to the slow-moving gas at the trough.



SOLAR-WIND VELOCITY WAVES STEEPEN as they travel away from the sun much as ocean waves steepen as they run toward a beach: the crest of the wave moves faster than the trough and so gradually catches up with it. In this superposition of three successive "snapshots" of an idealized solar-wind velocity wave made one day apart, it can be better seen how individual parcels, or small volumes, of gas (*dots*) on the leading edge of wave are compressed into region of high density as those on trailing edge are spread out into a region of low density.

It is at this point that the supersonic nature of the solar wind becomes important. The speed of sound, s, is the characteristic speed with which all pressure signals of small amplitude propagate in a gas. Therefore as the solar-wind wave gets steeper the region of high pressure expands into the gas ahead of it and behind it at the speed of sound. As the gas in the trough ahead senses the pressure signal it will be accelerated at the expense of the speed of the gas at the crest. In fact, as long as the peak speed of the wave is such that the speed of the converging gas ($\frac{1}{2}v_0$) is less than the speed of sound (s) the pressure signal will always be sensed by parcels of gas both in the trough of the wave and at the crest before the crest overruns the trough. That is, as long as the value of $s - \frac{1}{2}v_0$ is greater than zero, the solar-wind wave will gradually be damped out.

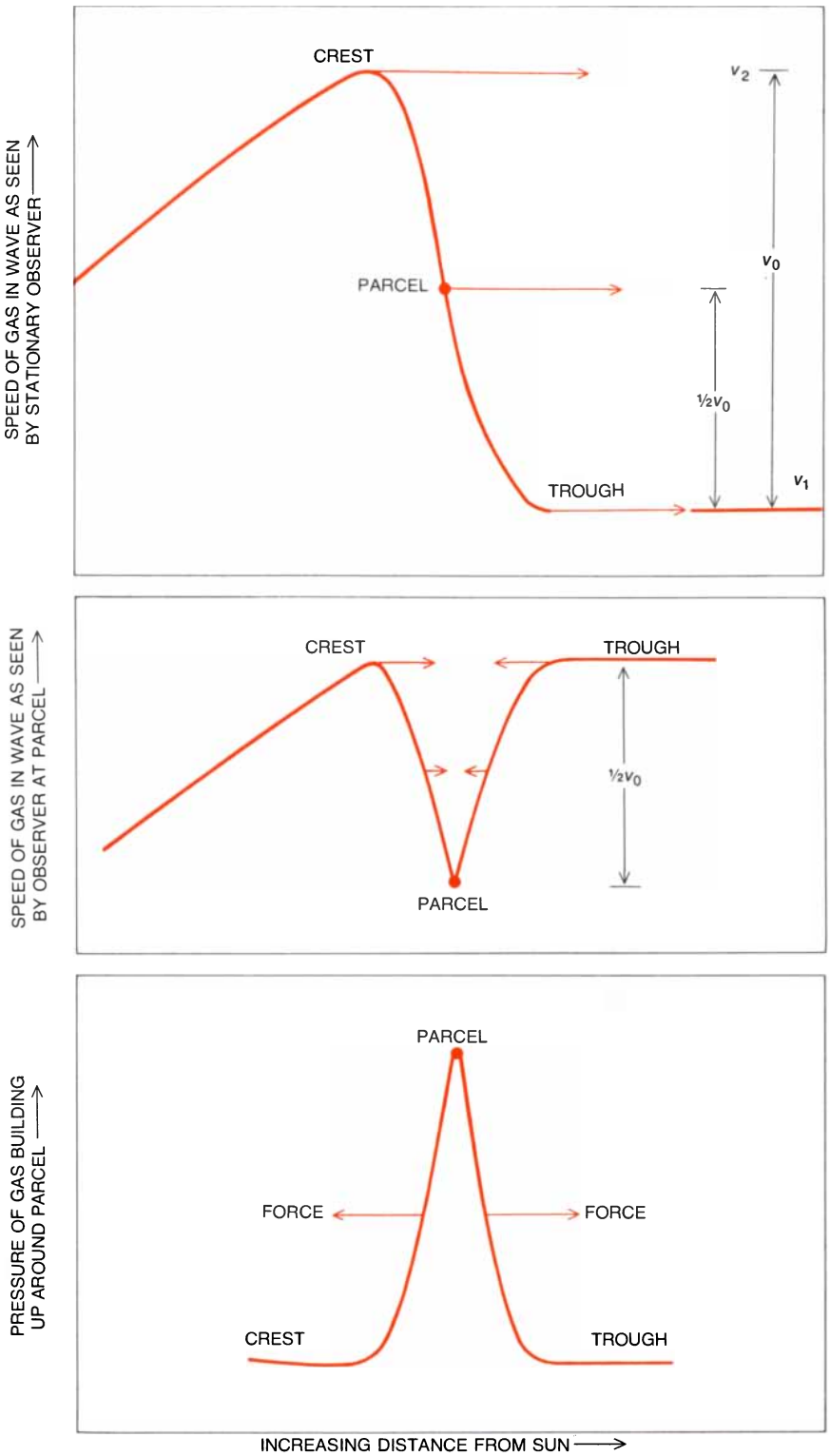
What happens when a wave in the solar wind has such a large amplitude that the speed of the gas converging on the parcel is greater than the speed of sound? In that case the high-pressure region, which attempts to expand outward in both directions at the speed of sound, cannot do so rapidly enough to warn the gas in the trough ahead (and at the crest behind) of the impending arrival of the wave. What happens instead is that the pressure within the wave at the site of the parcel grows very rapidly as the wave continues to steepen, so rapidly, in fact, that a pair of shock waves form, one on each side of the high-pressure region. One of the shocks propagates backward from the central high-pressure region toward the crest of the wave; the other propagates forward toward the trough. Shocks are pressure waves of large amplitude in which pressure, density and flow speed change rapidly. They travel through a gas faster than the speed of sound. In fact, the speed at which both shocks travel through the solar wind is precisely determined by the magnitude of the rise in pressure within the solar-wind wave, and that speed will be just what is necessary to keep the wave from toppling over and breaking.

Although one of the two shocks propagates backward from the region of maximum pressure, a stationary observer would actually see both shocks being transported outward from the sun by the very high bulk flow speed of the solar wind, a flow that is still highly supersonic and is still greater than the speed at which a shock propagates backward through the gas. The speed of sound in the solar wind decreases with the wind's increasing distance from the sun and the accompanying drop in temperature, so that virtually all waves in the solar wind should eventually get steep enough to form the pair of shocks.

Several quantitative models of how waves in the solar wind get steeper have been developed. One of us (Hundhausen) has devised a one-dimensional quantitative model that ignores some physical effects that are probably present in the solar wind yet describes the evolution of the waves fairly well. Such models have proved to be valuable in interpreting observations of solar-wind streams near the earth; they have also provided appealing explanations of why variations in the speed, density and pressure of the gas within such streams are coupled to one another.

Solar-wind streams have been observed almost entirely at one point in space, namely near the earth. Accordingly it has not been possible until recently to directly detect the steepening of the waves in the solar wind because that would call for measurements at two widely separated distances from the sun. With measurements obtained only near the earth, the steepening of velocity waves could be inferred only from the coupled variations in the solar wind. Moreover, almost all the solar-wind streams observed near the earth do not have a double-shock structure because the waves have not yet got steep enough for the shocks to form. The models do predict, however, that virtually all solar-wind streams should have shocks in them at distances beyond two or three astronomical units from the sun. (One astronomical unit is the mean distance between the sun and the earth.) Furthermore, the models predict that the amplitude of all the waves should decrease very slowly as they proceed toward the outer reaches of the solar system.

With the launching of the spacecraft *Pioneer 10* in 1972 and *Pioneer 11* in 1973, both of which drifted past Jupiter some five astronomical units from the sun, it became possible for the first time to directly test the concept (and specific models) of the evolution of velocity waves in the solar wind. The first data available from the Ames Research Center's plasma probe on *Pioneer 10* as it traversed interplanetary space were the hourly values of the speed of the wind. It was apparent from the data that beyond the earth each significant stream in the solar wind exhibited a rapid rise in speed followed by a much slower decline. In addition, superposed on the rising por-



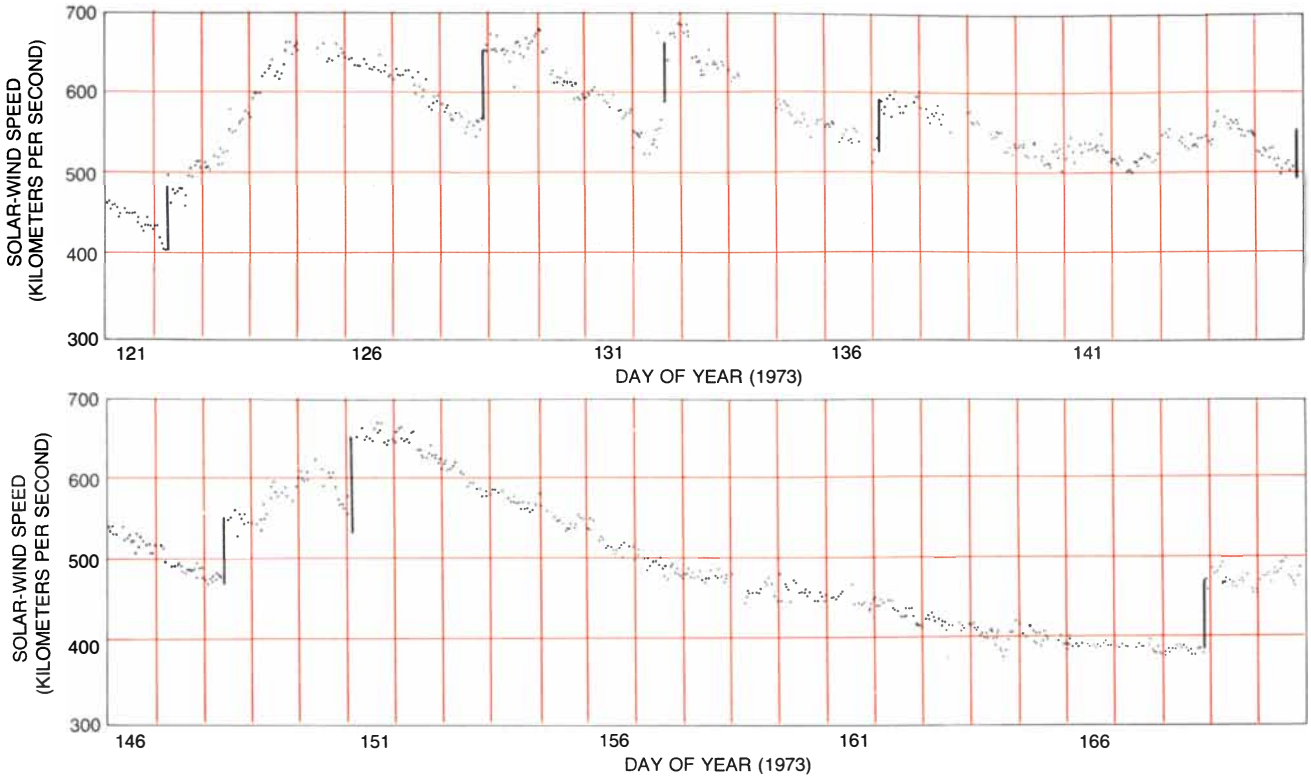
STEEPENING OF SOLAR-WIND VELOCITY WAVES IS LIMITED by pressure building up on the leading edge of the wave. The panel at the top shows a steepened wave of amplitude v_0, where v_0 is equal to the difference between the velocity of the crest of the wave (v_2) and the velocity of the trough (v_1). A parcel of gas halfway up the wave moves at $\frac{1}{2}v_0$ with respect to the trough. An observer moving with the parcel would see gas streaming toward the parcel from both the faster-moving gas in the crest and the slower-moving gas in the trough at speeds ranging up to a maximum of $\frac{1}{2}v_0$ (*middle*). The converging gas causes the pressure to build up around the parcel; thus the inflow from both sides is resisted (*bottom*). A stationary observer watching the wave from a fixed point in space would see the slow gas in the trough being accelerated and the fast gas in the crest being decelerated. When speed of gas moving toward the central parcel ($\frac{1}{2}v_0$) is greater than speed of sound, two shocks form around region of high pressure and diverge from it, one traveling back toward the crest and the other traveling ahead toward trough. Shocks prevent wave from toppling over and "breaking" like an ocean wave.

tion of the wave were abrupt rises in speed on the time scale of an hour, such as would be expected if shocks were present. The characteristic pattern of a fast rise and a slow decline in the speed of the solar wind can be found over a large range of amplitudes and periods, and it is precisely the sawtooth signature expected for evolving velocity waves. In the absence of measurements of the solar wind's density, pressure and magnetic field it is not possible to conclusively demonstrate that the abrupt jumps in speed on the rising portion of the streams in the *Pioneer 10* data were shocks. The solar-wind-speed data are consistent with such an interpretation, however, and reports of *Pioneer 10*'s measurements of later variations in the interplanetary magnetic field provide supporting evidence.

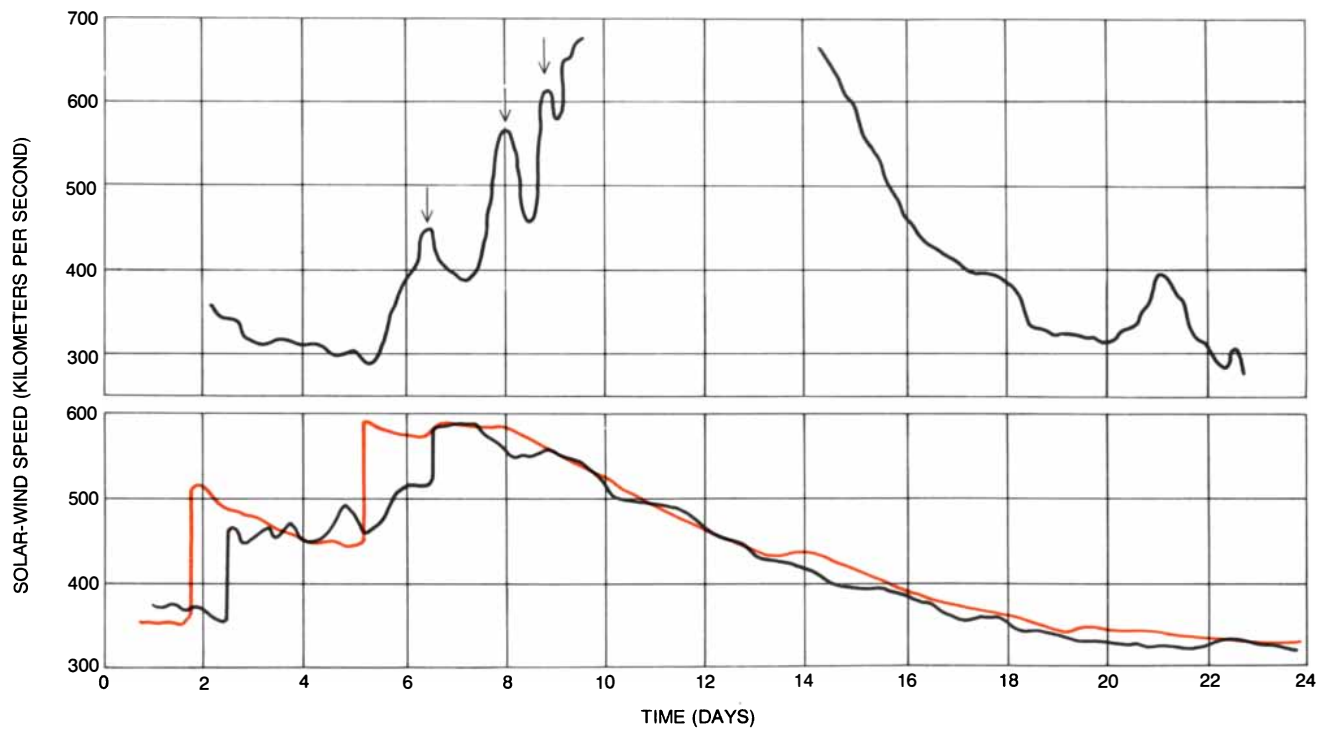
The trajectory of *Pioneer 10* between the earth and Jupiter was such that in August and September of 1973 the spacecraft was nearly aligned with the sun and the earth. That two-month period presented an excellent opportunity to directly observe the steepening of velocity waves between two points widely spaced along a nearly common line extending radially from the sun. The bottom illustration on the opposite page shows a sequence of measurements of the speed of the solar wind obtained by the Los Alamos Scientific Laboratory's plasma probe on the *IMP 7* satellite (the seventh Interplanetary Monitoring Platform) in orbit around the earth and another sequence obtained by *Pioneer 10* 17.5 days later at a distance of 4.7 astronomical units from the sun. The two sets of data have been shifted with respect to each other in such a way that the leading edge of the wave measured by each spacecraft would coincide if the gas on the wave's leading edge had propagated from *IMP 7* out to *Pioneer 10* at a constant speed.

It can be seen that in fact the leading edges of the waves in the two samples do not coincide: *Pioneer 10* detected the onset of the wave approximately three days early. The wave arrived early because the high-speed gas at the crest of the wave swept up the slower-moving gas in front of it, accelerating that gas as expected. Moreover, at 4.7 astronomical units from the sun the wave was clearly more sawtoothed than it was near the earth, and near the leading edge there were abrupt increases in speed that were not evident when the wave passed the earth. The measurements of the wave by *Pioneer 10* show that the wave indeed had the profile that would be expected if it had steepened and had developed a pair of shocks bounding a region of compression at its leading edge.

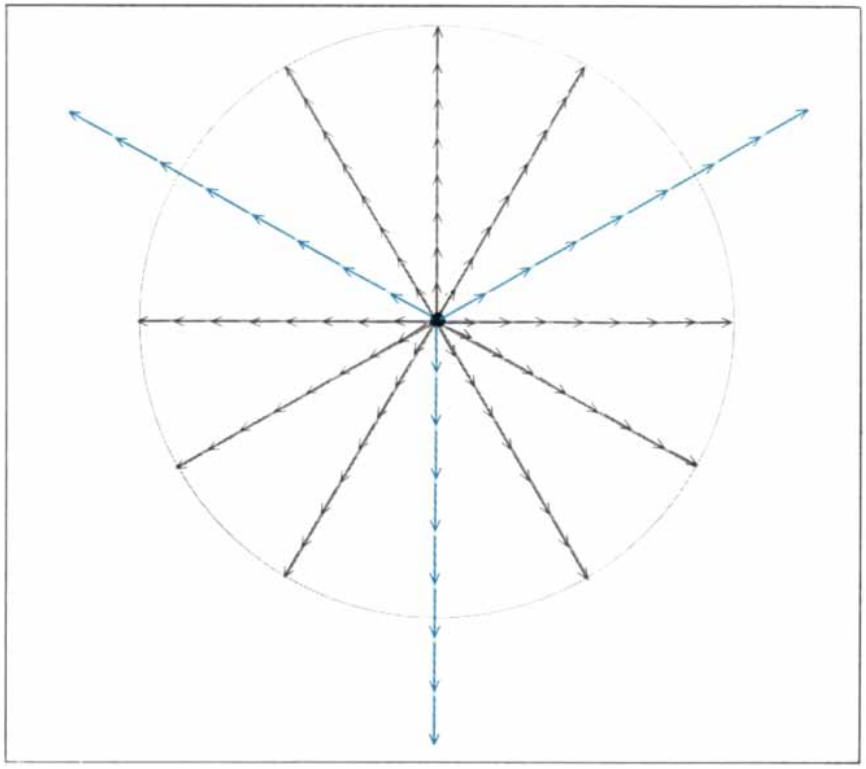
The favorable comparison between the model and the observations is even more striking when one independently utilizes measurements from *IMP 7* of



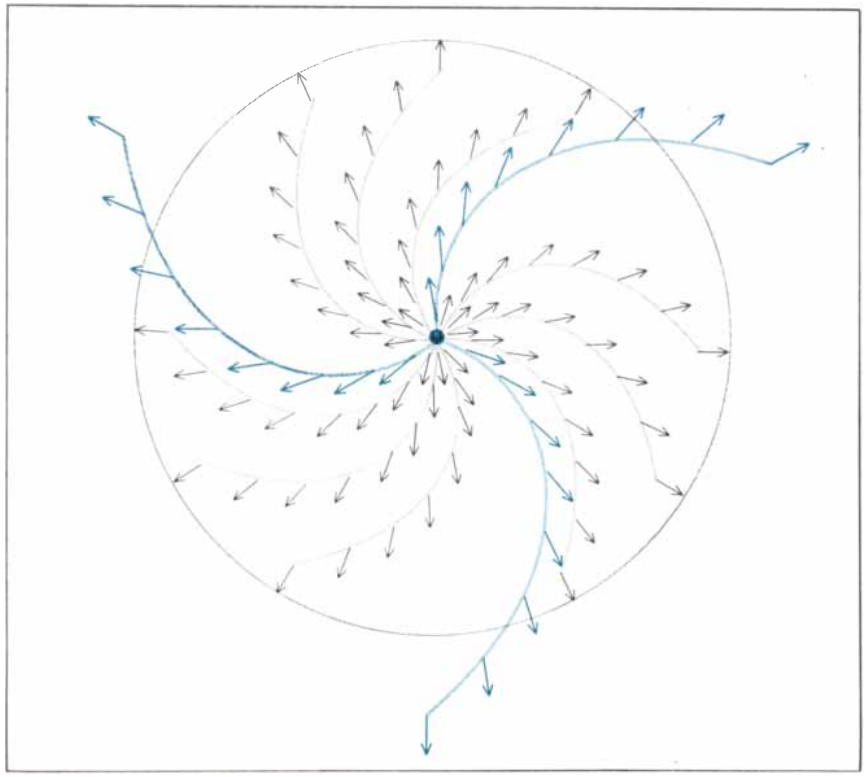
SPEED OF THE SOLAR WIND NEAR JUPITER was monitored hourly (*dots*) by the *Pioneer 10* spacecraft during a 50-day period in May and June, 1973. During that time the spacecraft moved from 4.03 astronomical units to 4.23 astronomical units. At that distance from the sun virtually all large-amplitude solar-wind waves have steepened into a sawtooth form. The abrupt jumps in speed on the rising portion of each velocity wave encountered are shocks; their location has been emphasized in the illustration by the vertical lines.



EVOLUTION OF A SOLAR-WIND WAVE at a distance of 4.7 astronomical units radially from the sun in the direction of the earth was predicted from the shape of the wave observed at the earth in August and September, 1973. The wave was observed near the earth (*top*) by instruments aboard *IMP 7* (the seventh Interplanetary Monitoring Platform) in orbit around the earth. The gap in the data was caused by the passage of the satellite into the magnetosphere of the earth, where the solar wind cannot be directly sampled. From the observations the form of the wave at 4.7 astronomical units was predicted by means of a theoretical model (*bottom panel, colored line*). Measurements made of the wave there by *Pioneer 10* show that the wave had steepened and had developed two shocks on its leading edge (*bottom panel, black line*); those features were predicted reasonably well. The overall amplitude of the wave at 4.7 astronomical units is considerably less than it was at one astronomical unit. Moreover, the large-amplitude short-period fluctuations (*arrows*) in the speed of the solar wind on the rising portion of the wave near the earth have been damped out, a fact that the model also predicted reasonably well.



STREAMLINES IN SOLAR WIND trace out the locus of all parcels of gas (*arrows*) having a common origin in one region of the sun. Black arrows indicate gas traveling at 350 kilometers per second; colored arrows indicate gas traveling at 500 kilometers per second. Gray circle indicates position of the earth's orbit. The solar wind travels nearly radially outward from the sun. If, as in the example shown here, the sun did not rotate, the streamlines would be straight.



SOLAR-WIND STREAMLINES ARE CURVED in the solar system because the sun rotates, bending the streamlines into spirals. The amount of curvature of each streamline depends on the speed of the gas traveling away from the sun. Where the streamlines converge the gas is compressed and forms a solar-wind stream; where the streamlines diverge the gas is rarefied.

the solar wind's speed, density and pressure near the earth to predict by means of the one-dimensional model the structure of the wave at *Pioneer 10*. The agreement between the predicted waveform and the observed waveform is quite good, particularly for the overall shape of the wave, the amplitude of the variation in the speed of the gas across the wave and the time when the wave reached *Pioneer 10*. The model did less well in predicting the amplitude of the abrupt increases in speed at the shock fronts. Without going into the details of the model, it can be said that its weaknesses apparently lie in the simplified assumptions that must be made to render it amenable to numerical computations on a computer. In particular the model neglects the effects of magnetic forces, the conduction of heat and the fact that the high-pressure region in the solar-wind wave eventually induces the flow of the gas to become nonradial.

So far we have emphasized the evolution of waves in the solar wind at only one fixed heliocentric longitude: the line from the sun to the earth. The evolution of a wave along any one longitude, however, is merely a small part of a much larger wave pattern that is evolving throughout the entire solar system. The pattern can be best indicated by streamlines: imaginary lines that trace out the locus of all parcels of gas originating at a common position on the sun but at different times. If the sun did not rotate, the streamlines would be straight lines extending radially outward from it. The sun's rotation causes the streamlines in its equatorial plane to be bent into spirals. The slower the parcels of gas expand, the more tightly the spiral streamlines are wound. Narrow bands of compression (high density) form in areas where the streamlines converge and broad regions of rarefaction (low density) form where the streamlines diverge. When the emission of gas from the sun is constant with respect to time, the entire pattern of curved streamlines rotates with the sun. It is worth emphasizing that only the pattern rotates; each parcel of gas in the solar wind moves outward almost radially.

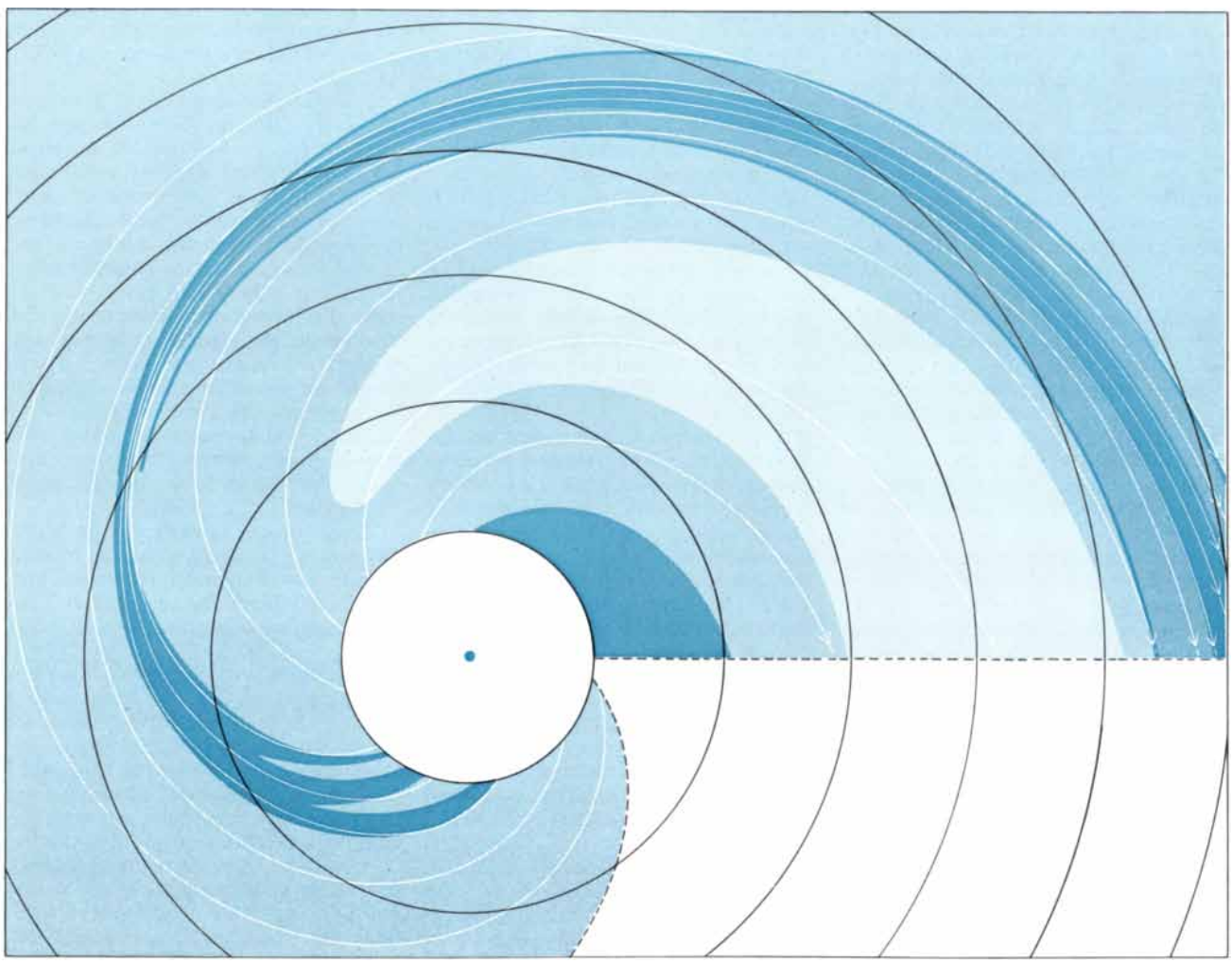
From the data of the solar-wind stream observed in 1973 by *IMP 7* and *Pioneer 10* and the aid of a theoretical model one can deduce the overall pattern of the solar wind in the plane of the earth's orbit as it might have appeared to a stationary observer situated far above the north pole of the sun. The stream was most compressed at a distance of about three astronomical units from the sun. At that point the compressed region contained all the streamlines that were originally at the leading edge of the wave. Beyond three astronomical units the region of compression slowly expanded. The expansion was hastened as the pair of forward-reverse

shocks formed by the steepened wave propagated away from the center of compression. Near the orbit of Jupiter streamlines that were originally separated by more than 180 degrees in solar longitude were contained within the region of compression bounded by the shocks. A broad region of rarefaction constituted the remainder of the pattern.

The concept of steepening velocity waves thus seems to explain the observed changes in solar-wind structure between one astronomical unit and about five astronomical units. The structure can be theoretically extended into the still more distant reaches of the solar system as well [*see illustration on page 37*]. The velocity wave is expected to continue to steepen in much the same way out to beyond 20 astronomical units, where the spiral pattern will begin to close back in on itself; at that point the shock at the trailing edge of the stream's next revolution will collide with the forward shock at the leading edge of the revolution just inside it. Out to this distance and for some unknown distance beyond it, the evolution of solar-wind velocity waves will continue to drive the solar wind toward an increasingly inhomogeneous state; for example, at 20 astronomical units the relative difference between the density of the gas in the compressed region and that in the rarefied region will be greater than it is at the orbit of the earth.

It has recently been shown that the shocks formed by steepened waves in the solar wind beyond the orbit of Mars give rise to a large number of cosmic-ray protons of low energy (between one million and 50 million electron volts). In fact, it is possible that almost the entire flux of low-energy cosmic-ray protons detected at the earth is produced by such shocks and is modulated by the magnetic fields associated with streams in the solar wind at large distances from the sun. Moreover, recent observations indicate that Jupiter acts as an electron gun, accelerating electrons to the energy range of one million to 10 million electron volts. The electrons fill interplanetary space; their propagation through the solar wind, however, is strongly affected by the pairs of forward-reverse shocks, which form a barrier the electrons cannot penetrate.

The study of the production and propagation of energetic particles in interplanetary space is only one field of investigation for which knowledge of waves in the solar wind is important. Indeed, the highly evolved nature of solar-wind streams at large distances from the sun promises to have an important bearing on the study of a host of diverse problems, ranging from the dynamics of the magnetosphere of the outer planets to the eventual interaction of the solar wind with the interstellar gas.



SOLAR-WIND STREAM OBSERVED in the plane of the earth's orbit in August and September, 1973, by *IMP 7* and *Pioneer 10* can be reconstructed from theory and from the spacecraft observations of the solar wind. Black circles indicate distances from the sun ranging from one astronomical unit to six astronomical units. Different shades of color represent contours of density of the solar wind, each one differing from the next by a factor of four. (Contours are adjusted to take into account the fact that the actual density of the gas decreases with increasing distance from the sun.) Thin curved arrows in white represent streamlines. Thick curved lines in solid color represent shocks. Although pattern of the gas inside the orbit of the earth is not known, three compressions associated with the large-amplitude, short-period fluctuations in speed observed by *IMP 7* on the leading edge of the main wave can be seen passing orbit of the earth.

Opiate Receptors and Internal Opiates

Morphine and its derivatives seem to exert their effects by binding to specific receptor sites on cells in the brain and the spinal cord. Morphinelike substances within the body may also act at those sites

by Solomon H. Snyder

In 1680 the English physician Thomas Sydenham wrote: "Among the remedies which it has pleased Almighty God to give man to relieve his sufferings, none is so universal and so efficacious as opium." Over the intervening centuries the admiration of the medical community for the pain-relieving effects of opium and its derivatives has been tempered by an awareness of their toxicity and addictiveness. This awareness, together with the lack of any other known class of drugs that exert the powerful analgesic action of opiates, stimulated an intensive search for synthetic opiates with the good properties of morphine and without the bad ones. More recently natural morphinelike substances have been found in the brain. These substances promise to open new avenues to an understanding of precisely where in the body opiates act, how they do so and why they are addictive.

Opium has been used as a drug at least since classical Greek times, not only, of course, because it deadens pain but also because it gives rise to euphoria. The word comes from the Greek *opion,* poppy juice; the drug is present in the milky exudate obtained by incising the unripe seedpod of the poppy *Papaver somniferum*. In 1803 a young German pharmacist, Friedrich Sertürner, isolated an opium alkaloid that he named morphine, after Morpheus, the Greek god of dreams, and by the middle of the 19th century the use of pure morphine rather than crude opium preparations had spread widely. The toxicity and addictiveness of morphine were recognized only after the drug had become an established feature of clinical medicine. The administration of opiates to wounded Civil War soldiers made opiate addiction a significant social problem in the U.S. and prompted the search for nonaddictive synthetic opiates.

The goal of a nonaddictive opiate has remained frustratingly elusive. Time and time again the medical community has enthusiastically greeted the development of a supposedly nonaddictive opiate, only to be disappointed when, after enough patients had taken the drug for a long enough time, it turned out to be addictive, in some cases even more so than morphine. In the 1890's, for example, the Bayer company introduced heroin, a morphine derivative with two added acetyl groups ($COCH_3$), as a nonaddictive analgesic. Similarly, in the 1940's meperidine (Demerol) became the most popular opiate analgesic in American medical practice because it was thought to be nonaddictive. The growing number of Demerol addicts soon convinced the Bureau of Narcotics that it was otherwise. Indeed, until quite recently it seemed that the analgesic effects of opiates and their addictiveness are inextricably linked.

Drugs, hormones and neurotransmitters (chemicals released by nerve endings that modulate the firing of other nerve cells, or neurons) all produce their highly selective effects at very low concentration. It is usually assumed that they act at specific receptor sites consisting of large molecules and located on the external surface of cells in the target organs. For the opiates a variety of evidence supports the receptor concept.

First, all opiate agonists, or analgesically active substances, show basic similarities in their molecular architecture. Morphine and most other opiates have a rigid *T*-shaped structure with two broad water-repelling surfaces at right angles to each other, a hydroxyl group (OH) capable of hydrogen bonding and a positively charged nitrogen atom that can form an ionic bond, all suggesting possible noncovalent interactions with a geometrically and chemically complementary receptor site. Moreover, synthetic opiate agonists have been devised that, although they are similar in their basic structure, are considerably more potent than natural agonists. Etorphine, for example, is 5,000 to 10,000 times more potent than morphine. It gives rise to euphoria and relieves pain in a dose as small as .0001 gram, making it even more powerful than LSD, which is often cited as the most potent mind-altering substance. Surely for a drug to act in such small doses it must seek out highly selective receptor sites.

Second, most opiates exist in at least two optical isomers: mirror-image molecules that are identical in chemical composition but, like left and right hands, cannot be superposed in space. Optical isomers can be distinguished by the direction in which a solution of the isomer rotates the plane of polarization of light in a polarized beam. Usually only the levorotatory isomer, the one that rotates the plane of polarization to the left, can relieve pain, elicit euphoria or any of the other actions associated with opiates. This stereospecificity of opiate action supports the model of a highly specific receptor that can distinguish the "handedness" of the opiate molecule.

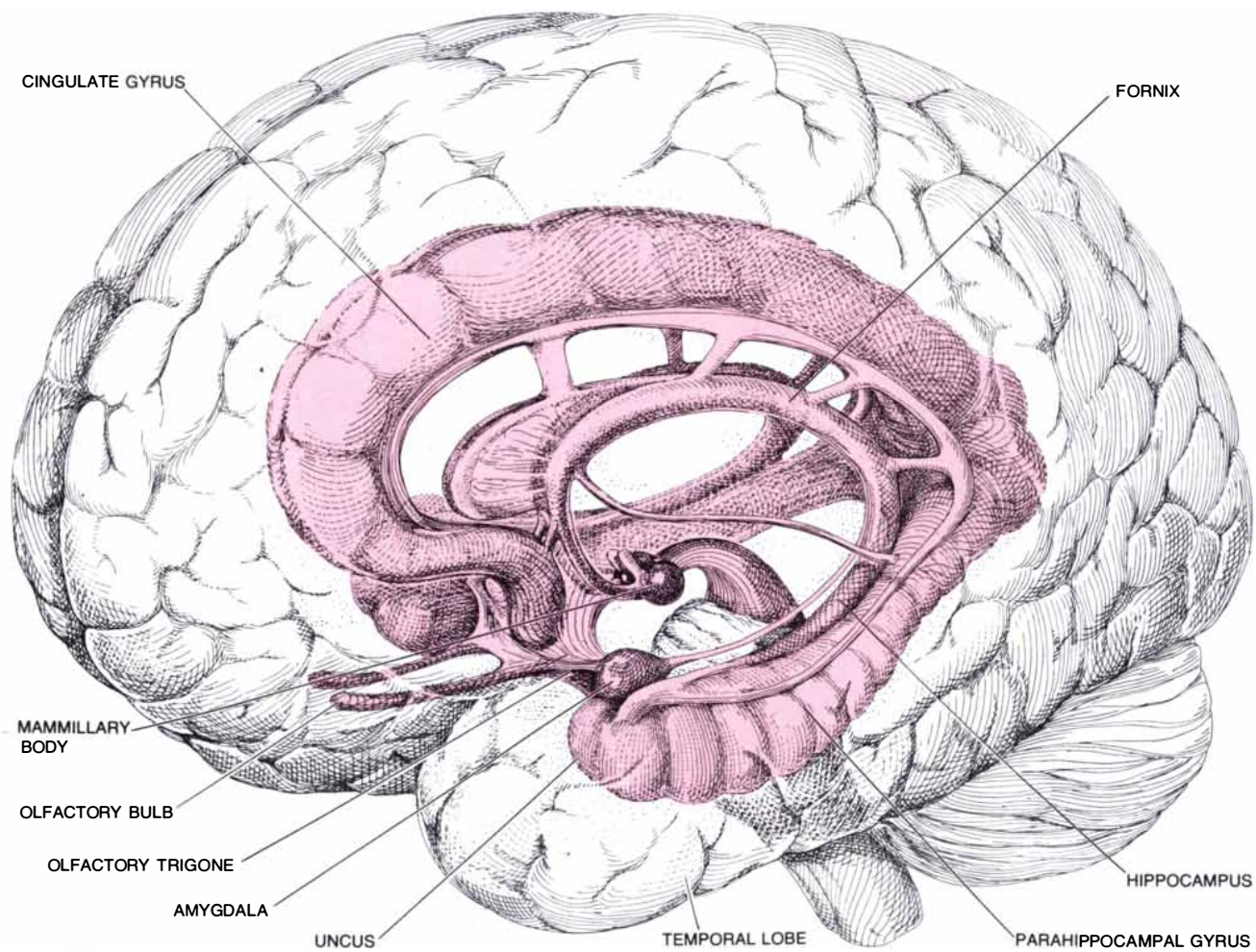
Third, opiate agonists can be transformed by very slight molecular modifications into antagonists, substances that specifically block the analgesic and euphoric actions of agonists without eliciting any such effects themselves. For example, the substitution of an allyl group ($CH_2-CH=CH_2$) for the methyl group (CH_3) on the nitrogen of morphine converts it into nalorphine, a potent antagonist that blocks all the pharmacological effects of morphine. An experimental animal or a person at the point of death from morphine poisoning can be revived almost instantaneously by much smaller amounts of nalorphine. An effect so rapid implies a common site of action. It seems as if the antagonists occupy opiate receptor sites, doing nothing themselves but blocking access to opiate agonists.

For these reasons pharmacologists long assumed that specific opiate receptors existed in the brain and possibly in other tissues. By synthesizing an entire series of morphine derivatives and test-

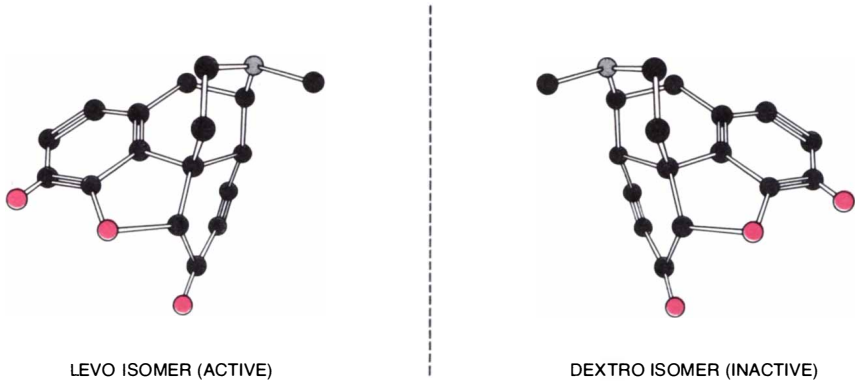
ing their pharmacological effects in animals investigators began to build up a picture of what parts of the opiate molecule are responsible for what physiological effects. In the process they visualized what the physical form of the postulated opiate receptor might be like, but unequivocal identification of the receptor itself remained elusive.

The great stumbling block in characterizing the receptor was the fact that, like most other compounds, opiates will bind to almost any biological or nonbiological membrane. Thus nonspecific binding, that is, binding not associated with the receptor, greatly exceeds receptor binding, which therefore could not even be detected. Avram Goldstein and his colleagues at the Stanford University School of Medicine attempted to differentiate between specific and nonspecific binding of opiates to fragments of membrane from homogenized brain cells by applying the criterion of stereospecificity: they compared the ability of a pharmacologically active optical isomer to inhibit the binding of a radioactively labeled opiate with the ability of the inactive optical isomer to do so. Although the technique was an important advance, Goldstein and his colleagues found that only 2 percent of the total binding of opiates to brain-cell membrane in the test tube was stereospecific. This result was discouraging, since it implied that isolating the small amount of specific binding to the postulated receptor from the very large amount of nonspecific binding would be an almost impossible task.

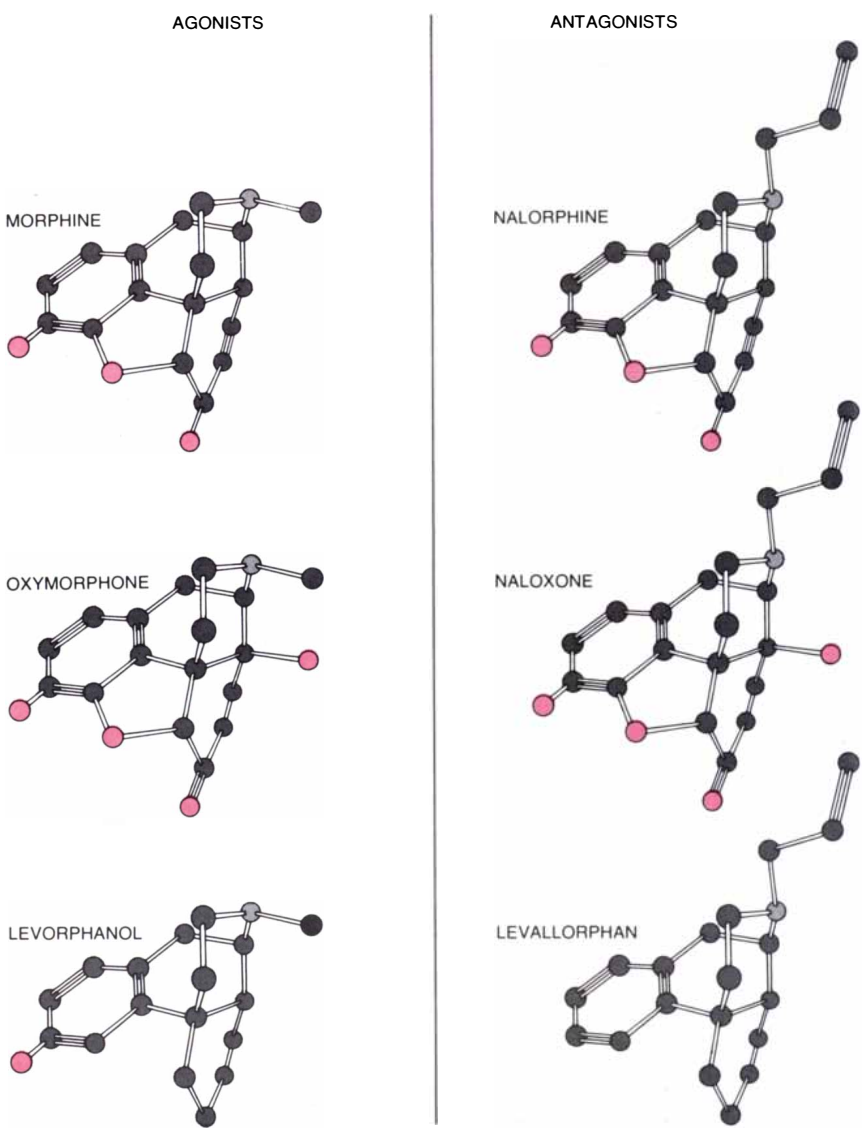
In my laboratory at the Johns Hopkins University School of Medicine, however, Candace B. Pert (then a graduate student) and I reasoned that simple technical maneuvers might amplify specific receptor binding enough for it to be measurable. To this end we made use of the methodology pioneered by Pedro Cuatrecasas of Johns Hopkins in characterizing the receptor for the hormone insulin. Cuatrecasas had managed to detect specific binding by making the labeled hormone highly radioactive, so that he could apply very low concentrations to fragments of target-cell membrane supported on filters. The hormone had a strong affinity for the receptor site, which meant that applying it in low concentrations would favor specific binding over nonspecific binding. Cuatrecasas also washed the membrane fragments thoroughly but very rapidly after hormone binding to selectively remove molecules that were bound nonspecifically. With these techniques we were able to identify high-affinity binding sites for opiates in fragments of cell membrane from rat brain and guinea-pig intestine with the aid of radioactively labeled naloxone, a potent opiate antagonist. To investigate the specificity of the binding we compared the ability of the active and inactive optical isomers of the opiates to compete with radioactively labeled molecules of naloxone for binding to the receptor. We found that the active isomers of both agonists and



OPIATE RECEPTORS IN THE BRAIN were identified by measuring the specific binding of radioactively labeled opiate drugs to cell fragments from different brain areas. The largest amount of binding was found in cells from the limbic system (*color*), a series of evolutionarily primitive regions at the core of the brain that are primarily involved with smelling in lower vertebrates and with the arousal of emotions in man. The high concentration of receptors in the limbic system suggests it is here opiates exert their euphoria-producing actions, and also that one internal opiatelike substance or more may play some normal role in modifying emotional component of pain.



OPTICAL ISOMERS, or mirror-image molecular forms, of the opiate analgesic (pain-killing) drug morphine have very different pharmacological activities. Only the levorotatory (*levo*) isomer, which in solution rotates plane-polarized light to the left, produces the characteristic analgesic effects of the drug; the dextrorotatory isomer is totally inactive. This stereospecificity of opiate action supports concept of a receptor that can distinguish "handedness" of a drug molecule. In illustration carbon is black, nitrogen is gray, oxygen is color. Hydrogen is not shown.



SLIGHT CHEMICAL MODIFICATIONS in the molecular structure of opiate analgesics (also known as agonists), such as the replacement of a methyl group (CH_3) by an allyl group ($CH_2 - CH = CH_2$), can convert them into antagonists that block the analgesic effects of agonists in very small doses. Antagonists appear to work by binding to an inactive form of the receptor.

antagonists could displace naloxone that was already bound to membrane, and that the pharmacologically inactive isomers had practically no effect on such binding.

Stereospecific binding alone, however, is not sufficient to establish an association between an opiate and a receptor. Symmetrically arranged constituents of brain tissue other than the receptor might be expected to discriminate between the isomers of opiates. Horace H. Loh of the University of California School of Medicine in San Francisco has found that a major class of brain lipids (fatty substances) can discriminate between the optical isomers of opiates. In my own laboratory Gavril W. Pasternak has found that opiates are bound stereospecifically even by some glass filters.

To ensure that the binding we had observed in fragments of homogenized brain represented opiate-receptor interactions that operate in real life we tested many drugs to determine whether their affinity for the receptor closely paralleled their pharmacological activity. This turned out to be the case: potent opiates such as levorphanol and morphine have a much greater affinity for the receptor than weak ones such as meperidine and propoxyphene (Darvon). Even apparent discrepancies serve as exceptions that prove the rule. For example, etorphine, an analgesic that is some 6,000 times more potent than morphine, has only about 20 times morphine's affinity for receptor. Etorphine, however, is 300 times more soluble in lipid and is correspondingly more efficient in penetrating the brain from the blood. The combination of a 20-fold greater affinity and a 300-fold greater penetration can account for etorphine's 6,000-fold greater potency.

Similarly, codeine, an opiate widely used for the relief of pain, cough and diarrhea, is about 20 percent as potent as morphine in man but is less than .1 percent as potent as morphine in binding to the opiate receptor. The discrepancy is readily explained by the fact that codeine is metabolized by enzymes in the liver to morphine, which then enters the brain and accounts for the pharmacological actions of codeine. E. Leong Way of the University of California School of Medicine in San Francisco has shown that although codeine can effectively relieve pain in rats when it is injected intravenously, it has no analgesic activity when it is injected directly into the brain. These are conditions under which morphine is quite effective.

Ideally one should compare potency and receptor binding in the same system. This is possible with opiates because their ability to inhibit electrically induced contractions of guinea-pig intestine is known to closely parallel their analgesic activity. Moreover, the affini-

ty of both agonists and antagonists for receptor binding sites is closely similar in guinea-pig intestine and in the brain. Working with a simple assay system in my laboratory, Ian Creese measured the ability of different opiates to inhibit contraction of guinea-pig intestine and compared it with their ability to inhibit naloxone binding to the same tissue. He found that receptor affinity and pharmacological potency correlated remarkably well for both agonists and antagonists in guinea-pig intestine, suggesting that the stereospecific binding site was indeed an authentic receptor.

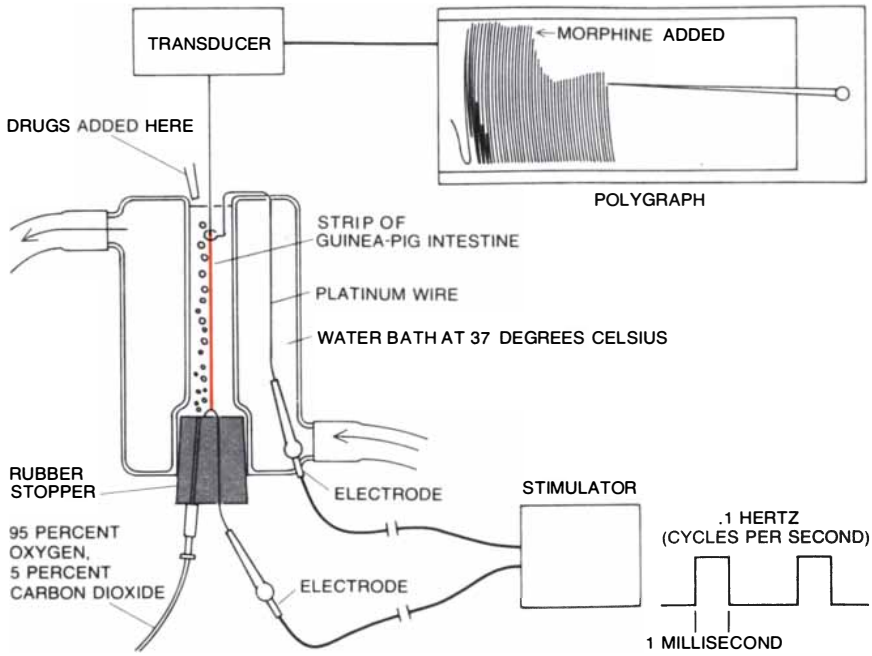
There was still one major source of mystification: in the binding experiments antagonists and their corresponding agonists appeared to have the same affinity for the receptor, but it was well known that in living organisms antagonists are much more potent than opiate agonists. This paradox was resolved when we slightly modified our technique for measuring the receptor. The conditions under which we had conducted our first experiments were unnatural in that they lacked any of the ions that are found abundantly in the body, particularly sodium. As soon as we incorporated sodium into our procedure dramatic differences in the behavior of agonists and antagonists appeared. Concentrations of sodium that are found normally in the body enhanced the binding of opiate antagonists and greatly diminished the binding of agonists. These effects were selective for sodium. They were mimicked somewhat by lithium, which has a similar atomic radius, but not by other positively charged ions such as potassium, rubidium and cesium.

In order to evaluate a wide range of opiates we measured the extent to which sodium would alter the ability of a given drug to inhibit the binding of naloxone to the opiate receptor. Our "sodium index" represents the ratio of the concentration of the drug required to inhibit naloxone binding by 50 percent in the presence of sodium to the concentration of the drug required in the absence of sodium. The larger the number is, the more sodium lowers the ability of the drug to inhibit naloxone binding. In other words, a sodium index of 2 means that a drug's ability to bind to the opiate receptor (measured by its inhibition of naloxone binding) is reduced by half in the presence of sodium.

The sodium index quite accurately predicts the extent to which a drug is an opiate agonist or antagonist. Pure opiate antagonists, such as naloxone and naltrexone, have a sodium index of 1 or less, meaning that sodium does not decrease their ability to bind to the opiate receptor and in fact sometimes increases it. Other opiate antagonists are slightly "contaminated" by agonist activity. These drugs can relieve pain but, presumably because of their antagonist properties, they do not seem to cause addiction as readily as "pure" opiate agonists. Early drugs with these mixed agonist-antagonist properties were not clinically useful, however, because they had a number of adverse side effects: they gave rise to anxiety, agitation and sometimes hallucinations. The best example of such a drug is nalorphine, which has been used clinically for many years to treat an overdose of morphine or heroin. It has a sodium index of about 2.5, meaning that its binding to the receptor becomes two to three times weaker in the presence of sodium. At the other end of the spectrum pure opiate agonists, which have no antagonist properties, become 12 to 60 times weaker in the presence of sodium. It is unclear just what accounts for the variation among the sodium indexes of agonists. None of them have obvious antagonist effects, but perhaps these drugs vary in the purity of their agonist properties. A purer agonist might be a more effective pain reliever, but it might also be more addictive.

The most interesting drugs are those with sodium indexes between 3 and 7. These are the combined agonist-antagonist drugs that offer the greatest promise as nonaddictive pain relievers. Most of these relatively nonaddictive analgesics belong to the novel class of compounds known as the benzomorphans, which were developed by Everette L. May of the National Institute of Arthritis, Metabolism and Digestive Diseases and other workers. The prototype is pentazocine (Talwin), which is widely used in the U.S. and is the only powerful opiate analgesic that is not subject to stringent "dangerous drug" regulations. Because its antagonist properties presumably balance its agonist ones, pentazocine is not detected in the hot-plate test, a traditional pharmacological screening procedure in which drug-treated mice are placed on an electric hot plate and the analgesic effect of the drug is measured by how soon they jump off. In spite of this, pentazocine is an effective analgesic in man and was ultimately shown to relieve pain in other animal tests. In many animals it is difficult to detect the development of tolerance to pentazocine and physical dependence on it. In man the drug is much less likely to be addictive than other opiates are, although there are reported cases of addiction.

Because of the peculiar pharmacological properties of agonist-antagonist analgesic drugs it has been difficult for drug companies to make them. There is no obvious chemical design that will ensure their desirable properties. One cannot even be confident that a given drug will possess the desired "negligible" addictiveness until it has been evaluated in man at vast expense. Measuring the binding of a drug to the opiate receptor in brain-cell fragments in the presence and absence of sodium greatly simplifies the entire procedure. Instead of synthesizing the many grams of a compound needed for animal tests one can employ only milligram quantities. So far the



BIOASSAY APPARATUS for determining the pharmacological potency of opiate drugs was set up in the laboratory of José M. Musacchio at the New York University Medical Center. A strip of guinea-pig ileum (part of the intestine), which is known to possess opiate receptors resembling those in the brain, is connected by means of a transducer to a polygraph that records its contractions. The potency of opiate agonists is measured by their effectiveness in inhibiting rhythmic contractions of the intestinal smooth muscle induced by an electric stimulator.

simple opiate-receptor assay has been able to predict the clinical potential of agonist-antagonist analgesics as effectively as much more expensive screening procedures in monkeys and men.

The selective effect of sodium on opiate-receptor binding appears to be an integral feature of receptor function. The presence of the ion seems to increase the number of antagonist receptors and decrease the number of agonist receptors. Since opiate agonists and antagonists are so similar chemically, it seems more reasonable to postulate that both drugs bind to the same receptor and that the opiate binding site can vary as the receptor shifts back and forth between two different molecular conformations. Sodium, which is responsible for the transition between the two states, differs in structure from the opiates and so presumably acts at a different site on the receptor molecule. In our model the binding of sodium would fix the receptor in the antagonist conformation, for which antagonists have a high affinity and agonists have a low one. By the same token agonists would have a substantial affinity for the agonist, or "no sodium," form of the receptor, whereas antagonists would not. Agonist-antagonists would have intermediate affinities for the two forms.

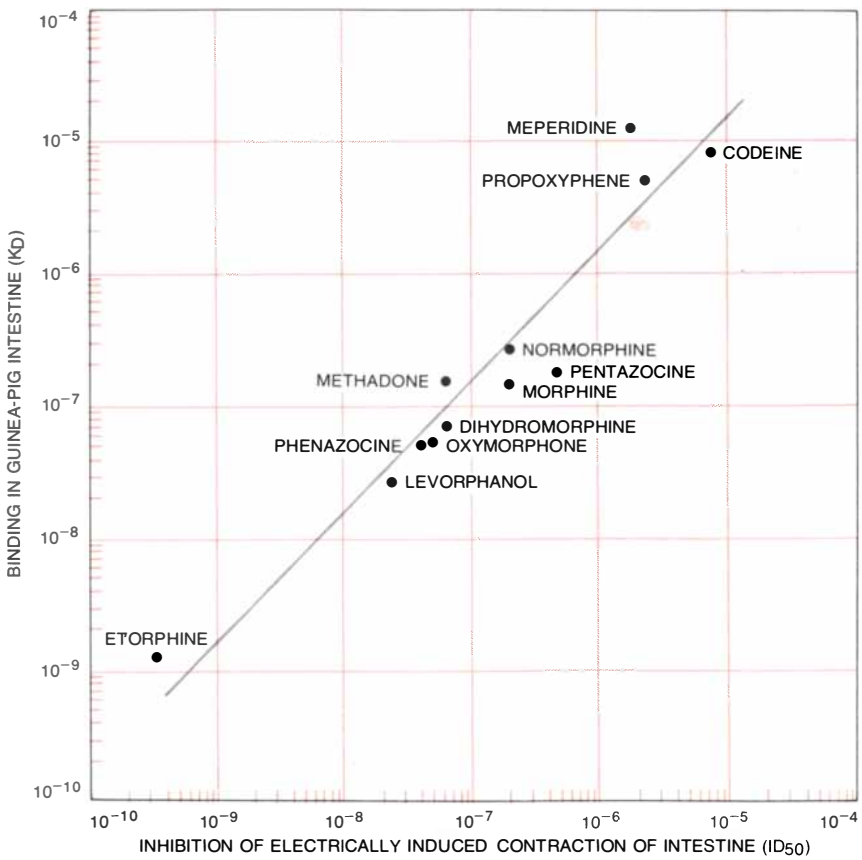
Typical morphine effects, such as analgesia and euphoria, would be observed only if drugs bind to the agonist state of the receptor, a condition that provides a simple molecular mechanism to explain opiate antagonist activity. By binding to the receptor in the sodium-binding state opiate antagonists reduce the number of receptors that are capable of mediating morphine effects. Because the fluid bathing the cellular membranes of the brain is rich in sodium one may expect that the opiate receptor normally exists in the sodium-binding antagonist state. This is consistent with the fact that in living organisms antagonists have a much greater potency than agonists. In an environment devoid of sodium one would predict that opiate agonists and antagonists would be equally potent.

With a simple, sensitive and specific opiate-receptor assay we were able to ask numerous questions. Many brain functions and neurotransmitter systems are distributed throughout the brain. If the distribution of the opiate receptor mirrored the distribution of some specific brain property, such a property might be implicated in opiate actions. Since opiates elicit analgesia, the brain structures involved in the perception of pain were natural suspects.

Two major brain pathways have been implicated in the perception of pain. Sharp, localized pain is poorly relieved by opiates and appears to be conveyed by a pathway that evolved late, consisting of a series of clusters of cells in each side of the thalamus, which is an important center for integrating sensory information. Duller, more chronic and less localized pain is quite effectively relieved by opiates and appears to be conveyed by a pathway that evolved earlier, consisting of many interconnected nerve cells, most of which lack a fatty myelin insulation and therefore conduct impulses rather slowly. Called the paleospinothalamic system, this pathway ascends along the midline of the brain; its way stations include the central gray matter of the brain stem and the central part of the thalamus.

Together with Michael J. Kuhar, Pert and I measured the distribution of the opiate receptor in many regions of the brain of monkeys and man, using both direct receptor-binding techniques and autoradiography of brain sections to which radioactively labeled agonist had been bound. The map of the distribution of the receptor in the brain strikingly parallels the paleospinothalamic pain pathway. There is also a high density of opiate-receptor binding in the amygdala, the corpus striatum and the hypothalamus, all of which are parts of the limbic system, a group of brain regions that largely mediate emotional behavior. Although these regions are not traditionally associated with the perception of pain, when parts of the limbic system in animals are stimulated electrically, the animal appears frightened, as if it were suffering pain. These brain regions therefore seem to be concerned more with the emotional component of pain and hence with the euphoric effects of opiates than with the analgesic effects.

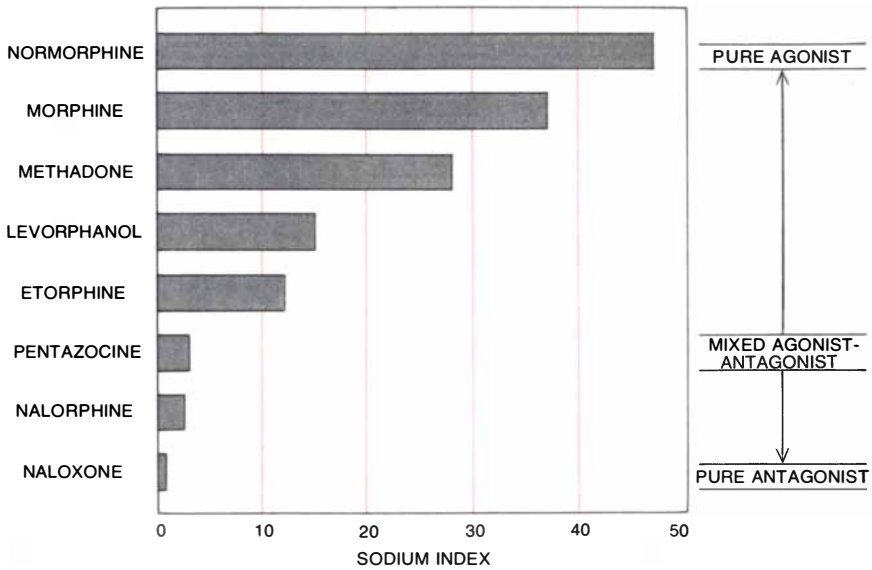
Within the spinal cord opiate receptors are localized in a dense band corresponding to the substantia gelatinosa, an important way station for the upward conduction of sensory information relating to pain. This observation bears on the long-standing controversy over whether opiate analgesia is mediated only in the brain or also in the spinal cord. It now seems likely that it is mediated in both. Receptors are also localized in the substantia gelatinosa of the caudal trigeminal nucleus of the spinal cord. This region receives unmyelinated fibers that carry painful stimuli from the face and hands, providing a pathway for the regulation of pain sensations from those parts of the body. Within the brain stem opiate receptors are highly localized in what are called the solitary nuclei, which may account for how opiates depress the cough reflex and reduce gas-



CLOSE CORRELATION between the pharmacological potency of opiate agonists in the guinea-pig intestine and their affinity for the opiate receptor in the same tissue shows that the receptor is genuine. Here the ID_{50} of several opiates (the concentration required to inhibit by 50 percent the electrically induced contraction of the intestine) is plotted against the K_D of these drugs (the concentration required to inhibit by 50 percent the stereospecific binding of radioactive naloxone). The values for the drugs tested cluster around a theoretical correlation line.

tric secretion, and in the area postrema, which contains the site where opiates apparently induce nausea and vomiting.

To further explore the general biological significance of the opiate receptor we measured receptor binding in a wide range of animal species. Binding was detected in the brain of all the vertebrate animals we examined but was strikingly absent in invertebrates. Surprisingly, there was no evolutionary trend. In the most primitive vertebrates we evaluated, such as the hagfish and the dogfish shark, there was as much opiate-receptor binding as there was in monkeys and man. Moreover, the opiate receptor of these primitive fishes displayed virtually the same drug specificity as the opiate receptor of mammals, indicating that few if any changes in the chemical structure of the receptor had occurred in the course of vertebrate evolution. This suggested that the opiate receptor is normally concerned with receiving some molecule that has remained the same throughout evolution.



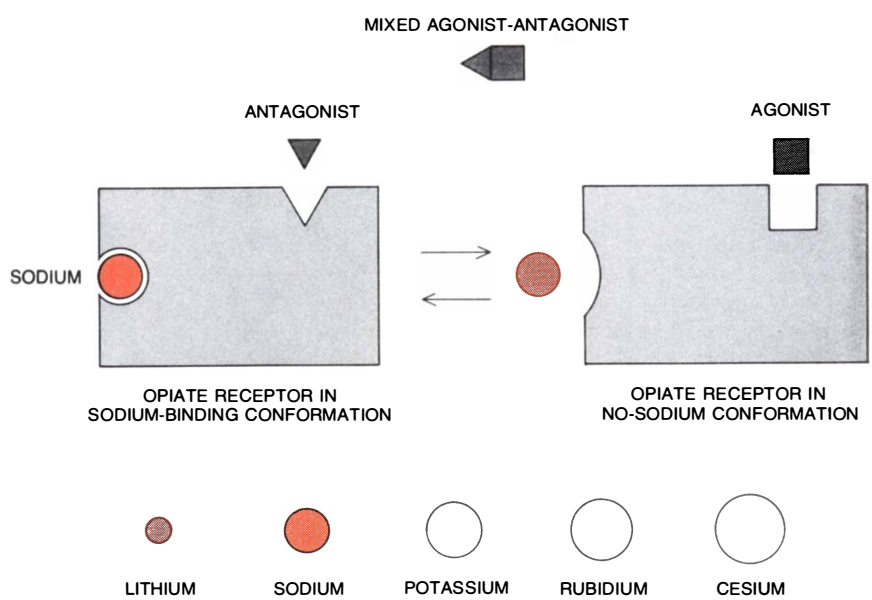
PRESENCE OF SODIUM IONS dramatically decreases the affinity of the opiate receptor for agonists without changing its affinity for antagonists. Bars show the ratio of the concentration of a drug required to inhibit by 50 percent the binding of radioactively labeled naloxone to the receptor in the presence of sodium to the concentration of the drug required in the absence of sodium. Lower values indicate a lesser reduction of potency in the presence of sodium: "pure" antagonists have a sodium index of 1 or less, whereas "pure" agonists have high values. The sodium index of a new opiate drug is thus a good indicator of its agonist or antagonist properties.

Clearly man was not made with morphine inside him. The existence in all vertebrates of specific opiate receptor sites strongly indicated the presence of a natural morphinelike substance in the brain, possibly a neurotransmitter, that acts at these sites. This speculation was supported by experiments that examined the localization of opiate receptors within nervous tissue.

Opiates, like most other drugs that affect the mind, are thought to act primarily at synapses in the brain, the specialized regions where the terminal of a nerve fiber makes a junction with the outer membrane of another nerve cell and chemically modulates its activity. It is possible to fractionate brain tissue into different subcellular components so as to localize the receptors at the submicroscopic level. When brain is homogenized in a solution with certain concentrations of the sugar sucrose, the nerve endings pinch off and seal up to form intact sacs called synaptosomes. Within the synaptosomes are the synaptic vesicles, which store the nerve ending's neurotransmitter content, together with several mitochondria, which provide energy for its activities. When synaptosomes are prepared by conventional procedures, the membrane of the other cell participating in the synapse adheres to many of them. Thus one synaptosome may constitute an entire synapse.

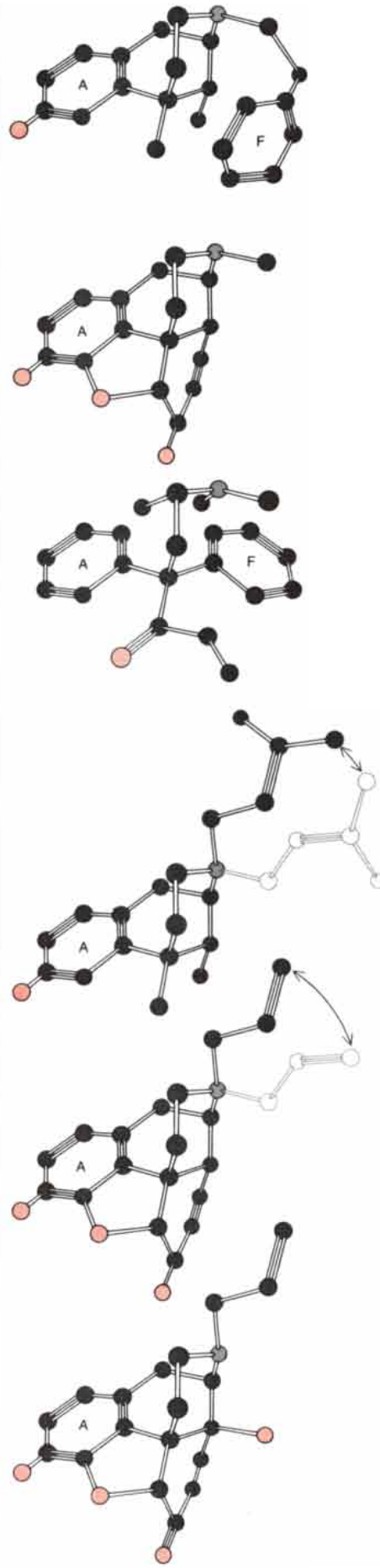
Synaptosomes can be ruptured by the osmotic shock of putting them in water. The synaptic vesicles and mitochondria within the nerve ending are released, and the membranes across the synapse break off together. By spinning these fragments in a centrifuge one can separate the synaptic vesicles, the mitochondria and the synaptic membranes. When we applied all these subcellular fractionation procedures to brain tissue and measured the opiate-receptor binding in the various fractions, we found that the receptor binding was largely confined to the synaptosome fractions. When the synaptosomes were ruptured, the receptor was recovered primarily from the synaptic membranes; little if any of it was associated with the synaptic vesicles. Hence it appears that the opiate receptor is associated with synaptic regions of the brain.

Since neurotransmitters act at synapses, the opiate receptor appeared to function very much like a receptor site for a natural neurotransmitter substance in the brain. Indeed, experiments conducted by John Hughes and Hans W. Kosterlitz of the University of Aberdeen provided direct evidence for the existence of such a morphinelike transmitter. For years Hughes and Kosterlitz had been studying the effects of morphine on involuntary muscle such as that of the guinea-pig intestine or the mouse vas



MODEL OF OPIATE RECEPTOR proposed by the author explains the sodium effect by postulating that the receptor can exist in two different conformations: a sodium-binding form with a high affinity for antagonists and a no-sodium form with a high affinity for agonists. Sodium effect is selective; of the other positively charged ions, only lithium has a similar effect.

STRONG AGONIST	PHENAZOCINE
AGONIST	MORPHINE
WEAK AGONIST	METHADONE
MIXED AGONIST-ANTAGONIST	PENTAZOCINE
ANTAGONIST CONTAMINATED WITH SOME AGONIST ACTIVITY	NALORPHINE
PURE ANTAGONIST	NALOXONE



deferens, the contracting tube that carries sperm from the testes to the urethra. As in the guinea-pig intestine, electrically induced contractions of the mouse vas deferens are suppressed by opiates. Hughes and Kosterlitz discovered that extracts of brain tissue could mimic the effect. The effect was both stereospecific and inhibited by low concentrations of the antagonist naloxone, so that the morphinelike brain factor was clearly binding to opiate receptors. Lars Terenius of the University of Uppsala and Pasternak and I independently identified the same substance in brain extracts by monitoring its ability to compete with radioactive opiates for binding to the opiate receptor.

Hughes and Kosterlitz subsequently isolated the morphinelike factor from the brain of pigs. They found that it consisted of two closely related short peptides, both made up of five amino acid units. They coined the name enkephalin for the two peptides, from the Greek for "in the head." One of the peptides, methionine-enkephalin, has the sequence NH_2-tyrosine-glycine-glycine-phenylalanine-methionine-COOH. The other, leucine-enkephalin, has the same first four amino acids and leucine in place of methionine. Hughes and Kosterlitz found that pig brain contains about four times more methionine-enkephalin than leucine-enkephalin. In my laboratory, however, Rabi Simantov and I independently purified the two enkephalins from the brain of cattle and found that the ratio of the two peptides is reversed. Since the genetic code for leucine and methionine differs by only one nucleotide out of three, the discrepancy between species may be simply a result of random genetic drift. On the other hand, the ratio between the two enkephalins has been shown to vary

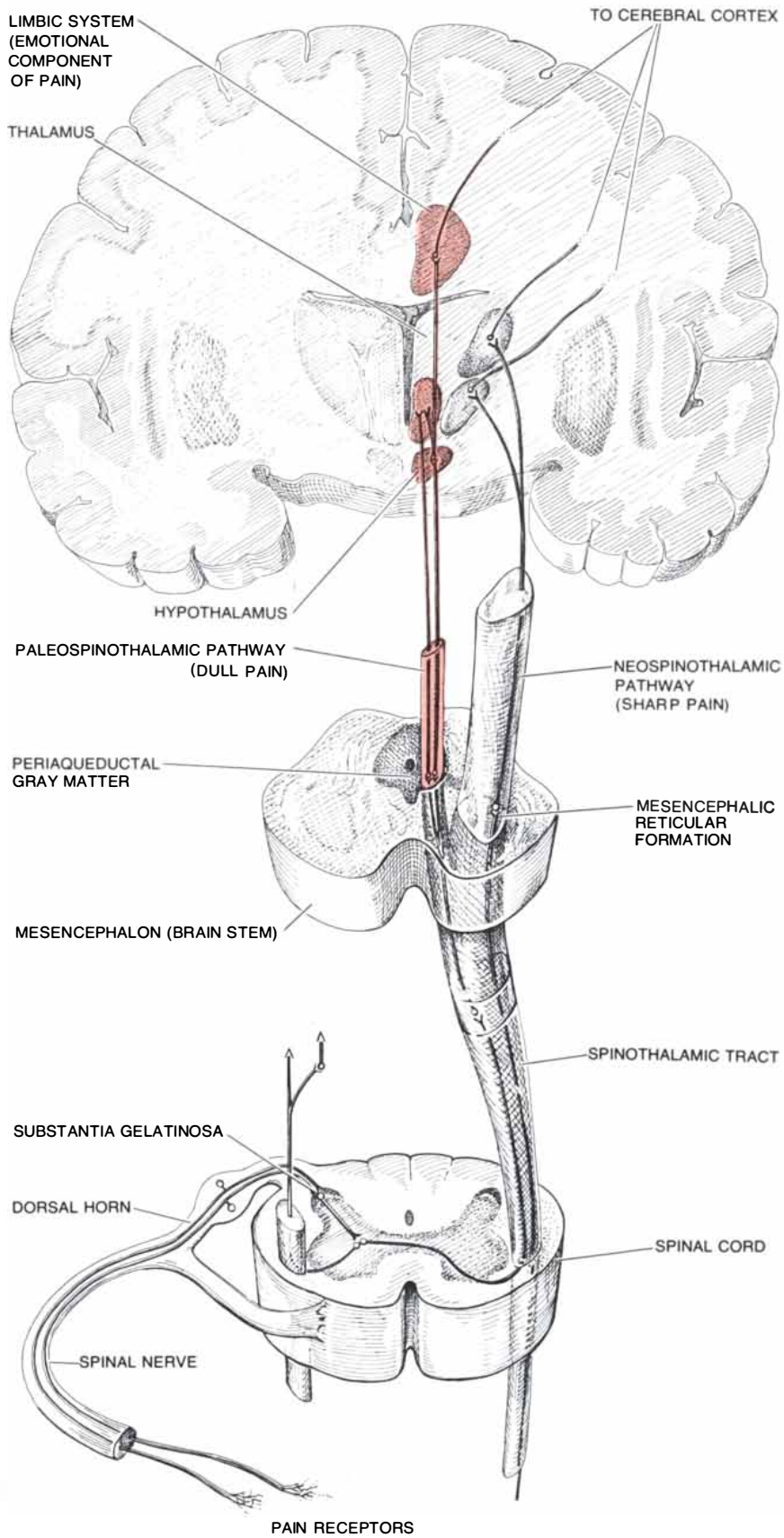
PROTOTYPICAL OPIATE DRUGS shown here suggest possible relations between molecular structure and pharmacological activity. Benzene ring *A* and the nitrogen (*gray*) are universal and seemingly crucial for all opiate actions, whether the receptor is in the agonist conformation or the antagonist one. In addition the agonist form appears to have a specific binding site for the ring *F* of potent agonists such as phenazocine, which stabilizes the receptor in the agonist conformation. Morphine is only moderately potent because it lacks ring *F*. Methadone has rings *A* and *F* but is a weak agonist because it cannot assume the critical orientation of the two rings required for strong binding to the agonist form of the receptor. The side chain of antagonists appears to bind to a special site on the antagonist form of the receptor. In mixed agonist-antagonists such as nalorphine or pentazocine this chain is flexible and rotates freely, so that at any given time some molecules are in the antagonist form and some are not. In the "pure" antagonists such as naloxone the presence of an adjacent hydroxyl group (OH) reduces the free rotation of the side chain, thereby fixing it spatially in the postulated antagonist position.

from one brain region to the next, suggesting that the slight structural difference between these peptides may serve some regulatory function.

A variety of evidence now indicates that the enkephalins are neurotransmitters of specific neuronal systems in the brain that mediate the integration of sensory information having to do with pain and emotional behavior, and that subserve unidentified functions as well. Regional variations in enkephalin levels tend to parallel the distribution of opiate receptors. Moreover, enkephalins appear to be localized in nerve endings. In my laboratory Simantov purified the two enkephalins from cattle brain and made antibodies to both of them by injecting them into guinea pigs. He and Kuhar, and also Tomas Hökfelt of the Royal Caroline Institute in Stockholm, then located enkephalins in nervous tissue by affixing a fluorescent dye to the antibody molecules so that the antibody-enkephalin complexes would glow when they were irradiated with ultraviolet. In such experiments Simantov and Kuhar observed that fluorescence was limited to the nerve endings of brain cells, which was consistent with the hypothesis that enkephalin is a neurotransmitter. The distribution of enkephalin nerve terminals as determined by the fluorescent-antibody technique is closely similar to that of opiate receptors as revealed by autoradiographic mapping: the terminals are concentrated in the substantia gelatinosa of the spinal cord, the amygdala of the limbic system and the central part of the thalamus.

Outside the brain enkephalin has been detected throughout the gastrointestinal tract in many species, but it has not been demonstrated in substantial amounts in any other tissues. Its striking localization in the intestines and the brain resembles the disposition of other peptides that appear to serve hormonelike roles in the intestines and to act as neurotransmitters in the brain, such as somatostatin, gastrin, vasoactive intestinal peptide and substance *P*. The limited distribution of these peptides may well be related to the fact that the gastrointestinal tract and the nervous system arise from the same layer in the developing embryo.

How might enkephalin act on brain cells? When opiates are applied to neurons that possess specific opiate receptors, the rate at which the cells fire is generally inhibited. Such inhibition may be due to any one of a number of mechanisms. Neurotransmitters are conventionally thought to bind to receptors on the membrane of the receiving neuron and then to trigger some alteration of the membrane's properties, such as a change in its permeability to ions. In its resting state the membrane is electrically polarized: its outside is positive with respect to its inside. Most well-known

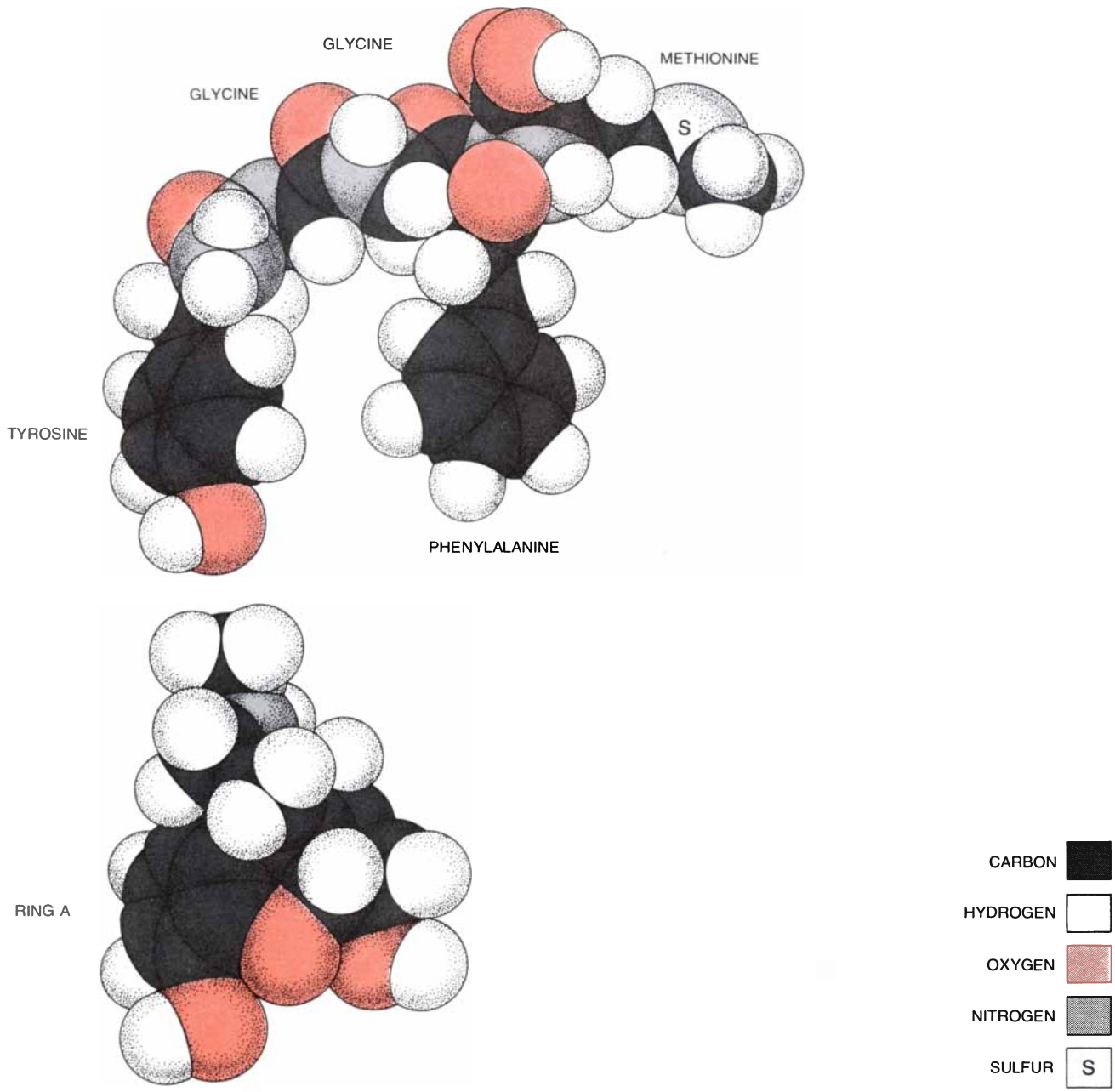


PAIN PATHWAYS carrying information from the periphery of the nervous system to the brain are separated into two types: the laterally located neospinothalamic pathway that transmits sharp, localized pain and the medially located paleospinothalamic pathway (*color*) that transmits less localized burning pain. Burning pain is best relieved by opiates, and opiate receptors have been found to be concentrated in the substantia gelatinosa and in the central thalamus.

excitatory transmitters, such as acetylcholine and glutamate, facilitate the firing of the receiving neuron by producing a depolarizing flow of positively charged sodium ions across the cell membrane. Clearly identified inhibitory transmitters, such as glycine and GABA, make the membrane more resistant to depolarization by excitatory transmitters. They do so by increasing the permeability of the membrane to negatively charged chloride ions, thereby hyperpolarizing it.

Surprisingly, the inhibitory action of enkephalin seems to be quite different from that of either glycine or GABA. Walter Zieglgansberger and Albert Herz of the Max Planck Institute for Psychiatry in Munich have shown by inserting electrodes into the receiving cell that it is not hyperpolarized by enkephalin or opiates; instead inhibition is caused by a reduction in the depolarizing flow of sodium across the membrane. Zieglgansberger and Herz concluded that opiates and enkephalin inhibit neuronal activity in some systems by blocking the sodium influx elicited by excitatory neurotransmitters, apparently by acting directly at the channels in the membrane of the receiving cell through which sodium passes.

New evidence suggests that enkephalin may have an additional mechanism of inhibition. A recent experiment conducted by Carol C. LaMotte, Pert and me showed that opiate receptors are not limited to receiving cells. We severed the dorsal root of the spinal cord of the monkey so that the incoming sensory nerves degenerated. Their target cells in the substantia gelatinosa of the spinal cord, however, remained unchanged. There was then a significant reduction in the number of measurable opiate receptors, indicating that receptors must be localized on the nerve terminals as well as on the receiving cells. The discovery



SPACE-FILLING MOLECULAR MODELS of morphine (*bottom*) and methionine-enkephalin (*top*), a morphinelike substance in the brain, have some structural features in common. The benzene ring *A* of morphine that bears a hydroxyl group is in precisely the same orientation as the benzene ring of the amino acid tyrosine at one end of the enkephalin peptide chain, suggesting that this group binds to the opiate receptor in both cases. The second ring on the enkephalin molecule is that of the amino acid phenylalanine, and it appears to interact with the agonist conformation of the opiate receptor in much the same way that the benzene ring *F* of potent opiate agonists does.

of such receptors suggested another model for enkephalin inhibition.

It is well known that the release of neurotransmitter from a nerve terminal is triggered by the depolarization of the membrane of the terminal, which occurs when the nerve impulse propagates to the end of the nerve fiber. The amount of transmitter released appears to be proportional to the degree of depolarization; the more the membrane is depolarized, the more transmitter is secreted. It is possible that the neurons that release enkephalin form synapses on the terminals of excitatory neurons. Enkephalin released at such synapses would bind to the opiate receptors on the excitatory nerve terminal, increasing the conductance of sodium across the membrane of the terminal and partially depolarizing it. Then when a nerve impulse reached the terminal, the net depolarization generated by it would be reduced, and there would be a corresponding decrease in the amount of excitatory transmitter released. Although here the action of enkephalin would itself be excitatory, increasing the flow of sodium across the nerve-terminal membrane, the ultimate effect on the cell receiving the excitatory nerve terminals would be inhibitory, since the amount of excitatory transmitter influencing its activity would be reduced.

This latter model fits reasonably well with what has been learned about the function of the opiate receptor from test-tube experiments. As I have mentioned, sodium dramatically alters the binding of opiates, increasing the affinity of the opiate receptor for antagonists and decreasing its affinity for agonists. In living animals the system presumably works in reverse, with the binding of opiate or enkephalin to the receptor changing its affinity for sodium and hence altering the permeability of the membrane to the ion. In the nerve-terminal-inhibition model increased sodium conductance would depolarize the terminal of an excitatory neuron, reducing the secretion of excitatory neurotransmitter. In the systems studied by Zieglgansberger and Herz sodium conductance would be decreased. Both effects on sodium conductance fit the biochemical evidence relating sodium and the opiate receptor.

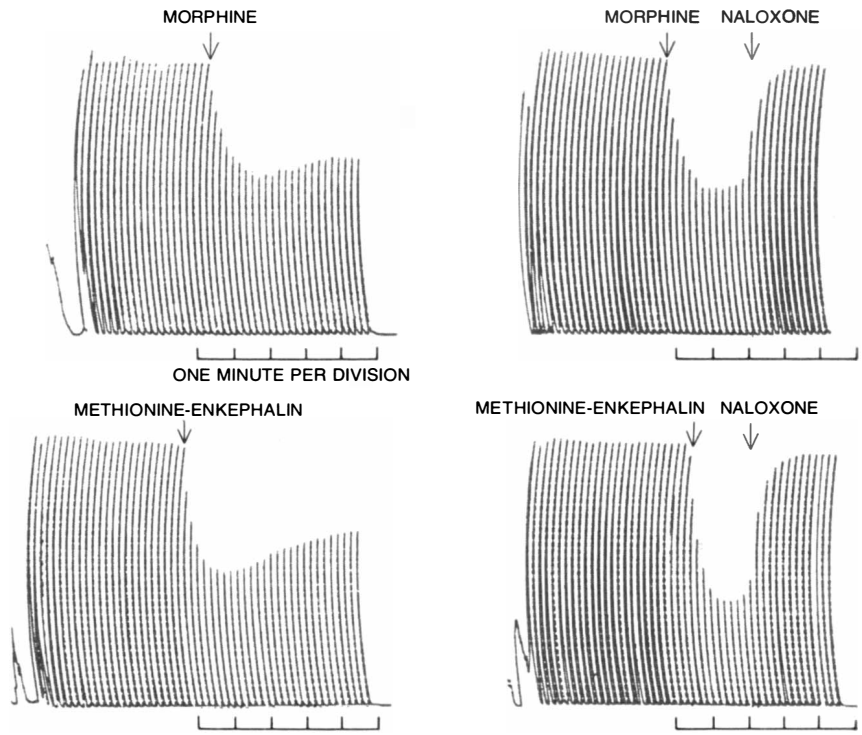
It is a significant fact that when rats addicted to morphine are given morphine, the activity of their brain cells is not inhibited. Brain-cell activity may even be excited. Since all these phenomena occur at the level of the cell membrane possessing the opiate receptor, it would seem that the mechanism of opiate addiction is to be found at that level.

The term addiction is not easily defined, but in all cases it manifests itself in the phenomena of tolerance and physical dependence. Tolerance refers to the situation wherein after the repeated administration of a drug higher doses are required to elicit effects that were previously elicited by much smaller doses. It can arise with many drugs, nonaddicting as well as addicting. One mechanism of tolerance is metabolic: the drug stimulates the synthesis of enzymes in the liver that destroy it. Then larger doses are required to achieve the same blood and tissue levels that were once achieved with smaller doses. A more important form of tolerance, particularly for addictive drugs, is termed cellular tolerance. After continued exposure to drugs such as alcohol, barbiturates and opiates, even high brain levels of the drug are no longer effective, so that progressively larger doses are required. Although cellular tolerance does not necessarily imply addiction, all known addictive drugs do elicit such tolerance.

Physical dependence is even more intimately associated with addiction. It is said to exist if when the administration of a drug is terminated, the drug user develops severe withdrawal symptoms. For the extreme alcoholic the symptoms include the tremulousness, convulsions and visual hallucinations of "delirium tremens"; for the opiate addict they include stomach cramps, diarrhea, sleeplessness and nervous excitation with widely dilated pupils and gooseflesh (hence the term "cold turkey").

There is also the phenomenon of cross-tolerance between drugs of a given class. Addicts tolerant to one opiate will be tolerant to another, so that physicians can treat withdrawal from heroin with other opiates, such as methadone. Alcohol, barbiturates and antianxiety drugs, such as chlordiazepoxide (Librium) or diazepam (Valium) seem to share brain mechanisms for tolerance and dependence. Thus the delirium tremens of alcoholic withdrawal can be relieved not only by alcohol but also by barbiturates and antianxiety drugs. There is no cross-tolerance, however, between the class of opiates and the class of alcohol, barbiturates and antianxiety drugs. Nevertheless, in a formal sense the processes of tolerance and physical dependence are the same for most classes of drugs. Hence if one could understand the biochemical mechanisms involved for one class, such as the opiates, one would know something about what was happening with drugs in other classes.

The discovery of enkephalin has suggested how enkephalin-containing neurons might play a role in opiate addiction. A simple model can explain how changes in the firing rate of enkephalin neurons could account at least in part for the behavioral manifestations of tolerance and physical dependence. Under resting conditions opiate receptors are exposed to a certain basal level of enkephalin. Administered morphine binds to usually unoccupied receptors, thereby potentiating the analgesic effects of the enkephalin system. On sustained treatment with morphine cells that have opi-



ENKEPHALIN MIMICS MORPHINE, inhibiting the electrically induced contractions of the guinea-pig intestine, in these polygraph readings obtained with the apparatus shown on page 47. The inhibition of enkephalin, like that of morphine, is blocked by the antagonist naloxone. Similarity in the effects of the two compounds suggests that they act at same receptor.

ate receptors find themselves overloaded with opiatelike material and convey, by some hypothetical neuronal feedback loop, a message to the enkephalin neurons to stop firing and releasing enkephalin. When this happens, the receiving cells are exposed only to morphine, so that they can tolerate more of it to make up for the enkephalin they are no longer receiving. When the administration of morphine is stopped, the opiate receptors find themselves with neither morphine nor enkephalin, and this lack initiates a sequence of events that result in withdrawal symptoms.

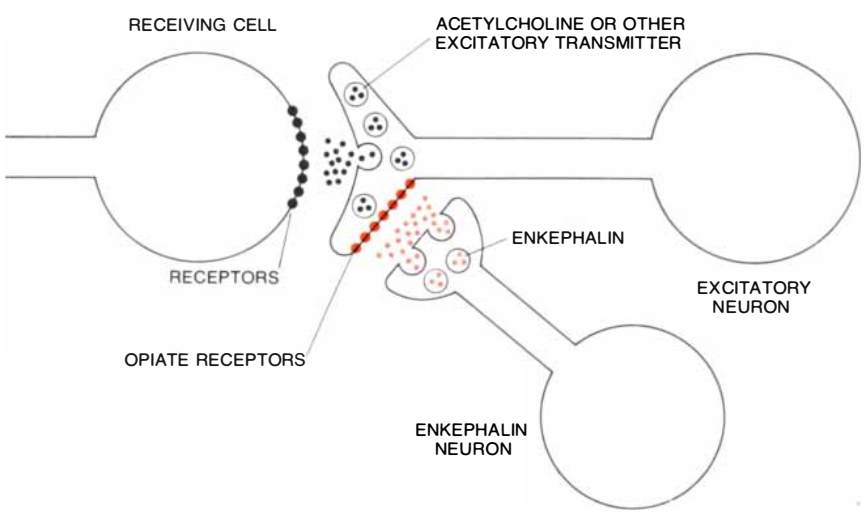
Recent experiments have supported this model. In other neurotransmitter systems of the brain the cessation of neuronal firing is associated with a buildup in the neurotransmitter content of the brain simply because the transmitter continues to be synthesized although it is no longer released. If the same thing happens with enkephalin systems, one would expect enkephalin levels to rise during the gradual development of tolerance and physical dependence. After animals have been withdrawn from morphine and the behavioral manifestations of abstinence have subsided, enkephalin levels should return to normal. These expectations were confirmed in experiments by Simantov and me. In rats given morphine regularly in the course of becoming addicted the enkephalin levels double. When abstinence is precipitated by treatment with naloxone, the animals exhibit a variety of withdrawal symptoms that subside in about an hour, at which time enkephalin levels return to normal.

How might information about the presence or absence of opiates or enkephalin at opiate receptors be conveyed inside a neuron? A change in sodium conductance may be part of the mechanism. The substances known as cyclic nucleotides may also play a role, as they do for certain other neurotransmitters and hormones, mediating the intracellular effects of these compounds after they have bound to receptors on the outside of the cell. The cyclic nucleotides cyclic adenosine monophosphate (cyclic AMP) and cyclic guanosine monophosphate (cyclic GMP) are respectively synthesized by the enzymes adenylate cyclase and guanylate cyclase. These two enzymes are thought to float freely in the double layer of lipid molecules that make up the cell membrane. When they encounter a receptor protein to which a molecule of the appropriate neurotransmitter is bound, they are activated and start to synthesize cyclic AMP or cyclic GMP. The two cyclic nucleotides often appear to antagonize each other in mediating the intracellular effects of hormones that have opposing effects on the same system. For example, acetylcholine, which causes involuntary muscle to contract, produces in its target cells an increase in the level of cyclic GMP and a decrease in the level of cyclic AMP. Epinephrine, on the other hand, which causes involuntary muscle to relax, produces in the same cells an increase in cyclic AMP and a decrease in cyclic GMP. Both transmitters presumably work by modifying the relative activities of adenylate cyclase and guanylate cyclase.

Several drugs that alter the intracellular levels of cyclic nucleotides affect opiate actions and can even mimic opiate withdrawal. Moreover, in the brain opiates themselves depress the level of adenylate cyclase and also inhibit the stimulation of the synthesis of the enzyme by certain hormones. Significantly, the relative potencies of opiates in inhibiting the synthesis of adenylate cyclase parallel their effect on opiate-receptor binding in the same cells. These drugs also simultaneously elevate the intracellular level of cyclic GMP, presumably by selectively activating guanylate cyclase. Thus, as happens in other biological systems, opiates appear to affect the two principal cyclic nucleotides in opposite directions.

The experiments of Marshall W. Nirenberg and Werner A. Klee at the National Institutes of Health with cultures of cancerous nerve cells have suggested that the effects of opiates on cyclic-nucleotide levels may account for some aspects of addiction. Brain cells do not divide and multiply like most other kinds of cells, so that they cannot be grown in culture in their normal state. Cancerous nerve cells, however, obtained from the nervous system tumor known as a neuroblastoma will grow in culture. For their experiments Nirenberg and Klee used hybrid cells created by fusing neuroblastoma cells with glial cells obtained from another type of brain tumor. (Glial cells, which make up 90 percent of the cells in the brain, closely surround the nerve cells and provide them with essential nutrients.) Working with these glia-neuroblastoma hybrids, Nirenberg and Klee observed the effects on intracellular cyclic-nucleotide levels of chronic exposure to morphine or enkephalin. They found that adenylate cyclase is inhibited by opiates but that over long periods of exposure the cells compensate by synthesizing more molecules of the enzyme. As a result higher concentrations of opiates are required to produce decreases in cyclic AMP because the additional molecules of adenylate cyclase make up for those inhibited by opiates.

The system thus appears to be "tolerant": morphine concentrations that lower cyclic-AMP levels in cells do not change those levels after sustained treatment with morphine or enkephalin. On the withdrawal of morphine cyclic-AMP levels rise markedly because the newly synthesized molecules of adenylate cyclase are not being inhibited. This excessive production of cyclic AMP could conceivably be a biochemical correlate of withdrawal symptoms. In the same cultured cells, however, sustained treatment with opiates does not affect the number of receptors or the binding of opiates to the receptor, indicating that the receptor itself is probably not



MECHANISM OF ENKEPHALIN INHIBITION may be indirect. Instead of acting directly on the receiving nerve cell the substance may block the release of excitatory neurotransmitters such as acetylcholine and glutamate, thereby reducing the receiving cell's excitatory input. According to the author's model, enkephalin released from a neuron binds to opiate receptors on the terminal of an excitatory neuron, partially depolarizing the terminal membrane and reducing the net depolarization produced by the arrival of a nerve impulse. The amount of neurotransmitter released from the terminal is proportional to the net depolarization, so that less excitatory transmitter is released. The receiving cell is then exposed to less excitatory stimulation and reduces its firing rate. Such an enkephalin inhibitory system may modulate the activity of the ascending pain pathways in the spinal cord and the brain. Opiate drugs would act by binding to unoccupied enkephalin receptors, thereby potentiating the effects of the system.

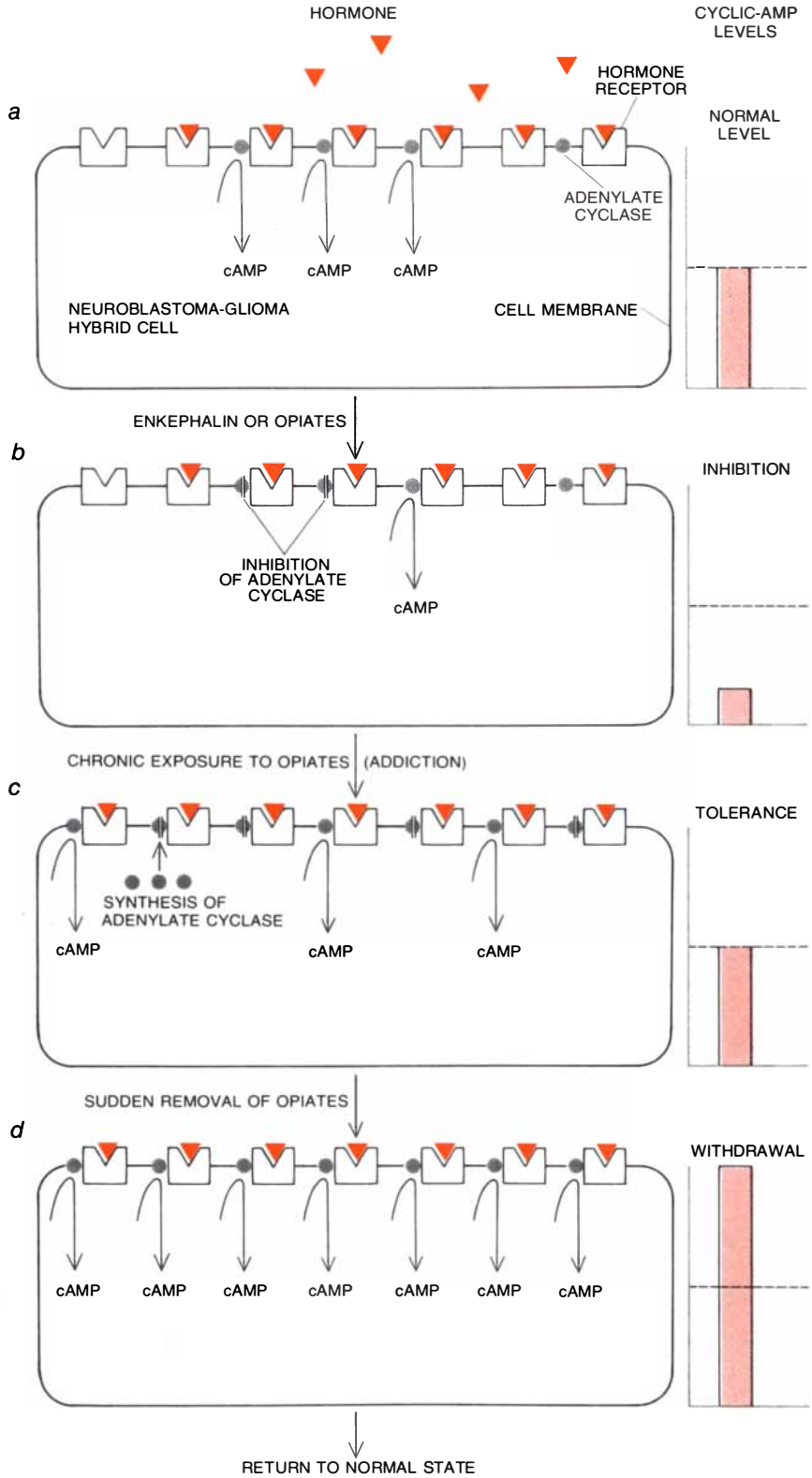
the site of the fundamental biochemical modification in addiction.

Cyclic-nucleotide changes in addiction have not yet been demonstrated in brain tissue, but opiates and enkephalin do elevate cyclic GMP and lower cyclic AMP in slices of brain tissue in the test tube. It is therefore reasonable to suppose chronic effects in addiction might be similar to those in Nirenberg and Klee's nerve-cell cultures. Changes in the disposition of cyclic nucleotides in neurons that possess opiate receptors could conceivably serve as the postulated feedback signal that might alter the firing rate of enkephalin neurons, initiating the development of tolerance and physical dependence.

Such models of addiction have been complicated somewhat in recent years by the discovery that the enkephalins are not the only natural opiatelike peptides. Even before the discovery of enkephalin Goldstein reported that the pituitary gland at the base of the brain, which is embryologically and structurally distinct from the nervous system, contains a factor with opiatelike effects. This finding became particularly intriguing when Hughes and Kosterlitz noticed a curious connection: the amino acid sequence of methionine-enkephalin matched a segment of the pituitary peptide hormone beta-lipotropin, which is normally associated with stimulating the breakdown of fat. Choh Hao Li of the University of California School of Medicine in San Francisco had previously isolated from the pituitary gland of the camel a fragment of beta-lipotropin 31 amino acids long whose function was somewhat of a mystery: it had very little fat-metabolizing activity but it did incorporate the sequence of methionine-enkephalin. Noting this, Goldstein obtained samples of Li's peptide and found that it had opiatelike effects.

Li named his peptide beta-endorphin, for endogenous morphine. It had a high degree of analgesic activity: it was more than 48 times more potent than morphine when it was injected directly into the brain of experimental animals and three times more potent when it was injected intravenously. Moreover, its activity was blocked by naloxone, a specific opiate antagonist. Roger Guillemin of the Salk Institute went on to isolate two other peptides, as well as beta-endorphin itself, from a mixture of hypothalamus and pituitary tissue from pigs. One of them, named alpha-endorphin, has a sequence corresponding to amino acids 61 through 76 of beta-lipotropin, and it has analgesic and tranquilizing effects in animals. The other, designated gamma-endorphin, has a sequence corresponding to amino acids 61 through 77 of beta-lipotropin. Injected into rats, it induces violent behavior.

The biological function of the pituitary endorphins is something of a mys-



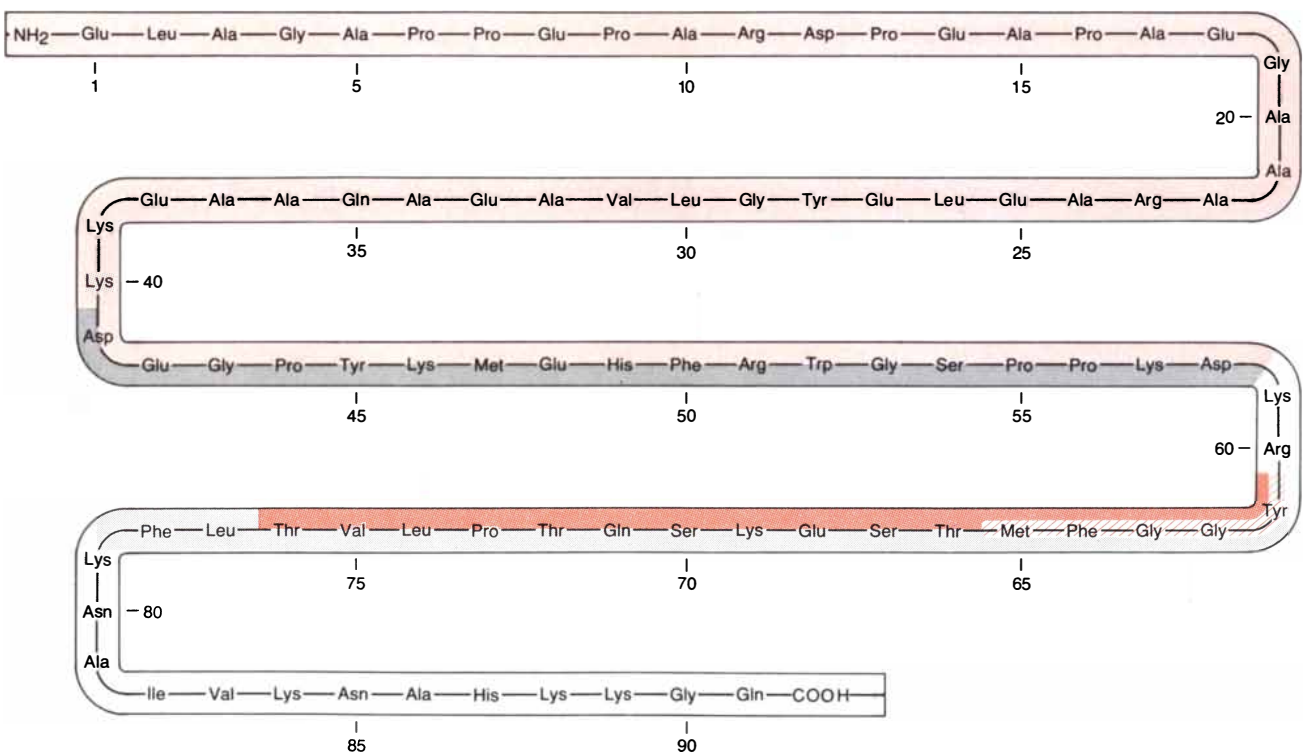
MODEL OF ADDICTION is provided by the biochemical changes that accompany the administration of morphine to cancerous nerve cells grown in cell culture. On stimulation from certain hormones the membrane-embedded enzyme adenylate cyclase synthesizes the intracellular messenger molecule cyclic adenosine monophosphate (cyclic AMP, or cAMP), which then mediates the physiological effects of the hormone (*a*). The administration of opiates inhibits adenylate cyclase, reducing cyclic-AMP levels (*b*). On sustained exposure to opiates the cell adapts by synthesizing more molecules of the enzyme, so that a normal amount of cyclic AMP is produced (*c*). The cell is now "tolerant" to the original dose of opiate. When drug administration is stopped, all the enzyme molecules become active and synthesize an excess of cyclic AMP (*d*). This excess may trigger a sequence of events leading to withdrawal symptoms.

tery. It is unlikely that they pass from the pituitary into the brain, since there is no direct communication between the two organs. Conceivably they regulate pituitary functions that are known to be altered by opiates. For example, morphine releases antidiuretic hormone (ADH) from the posterior pituitary. The fact that the pituitary gland possesses opiate receptors similar in their properties to the brain's opiate receptors suggests that the influence of opiates on the function of the posterior pituitary, such as the release of ADH, may involve the pituitary's own opiate receptors.

What are the relations among beta-lipotropin, the endorphins and enkephalin? Biologically active peptides are usually synthesized by being cleaved from larger precursor peptides. One might speculate that beta-lipotropin or some partial sequence of it serves in the brain as the precursor of enkephalin. The fact is, however, that beta-endorphin can be detected in brain tissue only in low amounts, and beta-lipotropin has never been reported in the brain. It is therefore likely that the larger peptides such as beta-endorphin play their major role in the pituitary and that enkephalin is confined to the brain and the spinal cord. Since the analgesic and euphoric effects of opiates are mediated only in the brain and the spinal cord it seems likely that the analgesic actions of endorphins are fortuitous and unrelated to their normal function.

Enkephalin is degraded very rapidly by protein-cleaving enzymes, so that it is difficult to evaluate its analgesic potential when it is injected into an animal intravenously or even directly into the cerebral ventricles. With compounds analogous to enkephalins that resist metabolic breakdown, however, it has been possible to show that enkephalins have an analgesic potency similar to that of morphine.

The fact that enkephalins can produce analgesia suggests that appropriate analogues might be the long-sought nonaddictive analgesics. Surely man is not addicted to himself. Or is he? Surprisingly, when the cerebral ventricles of rats are repeatedly injected with enkephalin or beta-endorphin, as they were in experiments by Eddie T.-F. Wei and Loh at the University of California School of Medicine in San Francisco, they develop symptoms of tolerance and physical dependence. Whether or not the rats are addicted in the same way that human heroin addicts are is difficult to evaluate. Since subjective responses cannot be elicited from the rats, one can only measure withdrawal symptoms such as shaking and diarrhea. Conceivably the nature of the apparent dependence elicited by sustained exposure to the natural opiatelike peptides is qualitatively different from the nature of that elicited by opiates. If that is the case, mixed agonist-antagonists developed from the enkephalins might offer promise as nonaddictive analgesics. In any event continued efforts to characterize the opiate receptor and its associated peptide neurotransmitters should lead to both therapeutic benefits and a deeper understanding of how the brain works.



GAMMA-LIPOTROPIN

ALPHA-ENDORPHIN

BETA-MELANOTROPIN

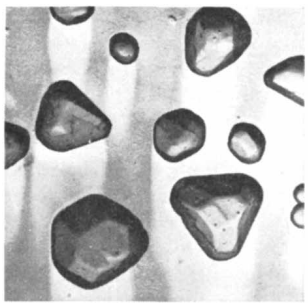
METHIONINE-ENKEPHALIN

BETA-ENDORPHIN

Ala	ALANINE	Gly	GLYCINE	Phe	PHENYLALANINE
Arg	ARGININE	Ile	ISOLEUCINE	Pro	PROLINE
Asn	ASPARAGINE	His	HISTIDINE	Ser	SERINE
Asp	ASPARTIC ACID	Leu	LEUCINE	Thr	THREONINE
Gln	GLUTAMINE	Lys	LYSINE	Trp	TRYPTOPHAN
Glu	GLUTAMIC ACID	Met	METHIONINE	Tyr	TYROSINE
				Val	VALINE

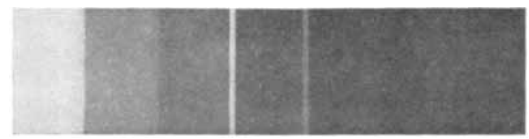
BETA-LIPOTROPIN, a pituitary peptide hormone 91 amino acids long, has amino acid sequences with several distinct physiological functions. The peptide chain as a whole induces the metabolism of fat, as does the segment termed gamma-lipotropin (amino acid units 1 through 58). The sequence 41 through 58 is that of the hormone beta-melanotropin, which plays a role in skin pigmentation. The sequence 61 through 91 is that of beta-endorphin, a pituitary peptide that has analgesic effects when it is injected intravenously or is injected directly into the brain. A second pituitary peptide, designated alpha-endorphin (61 through 76), has similar but less potent effects. The beta-lipotropin sequence 61 through 65 is identical with that of methionine-enkephalin, a morphinelike peptide found in the brain, the spinal cord and the intestines. The relation between the opiatelike peptides and beta-lipotropin is not known.

Kodak instant print film: it's quite different in the way it captures color

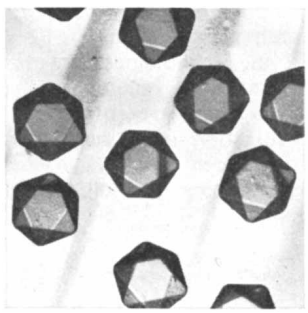


1 μm

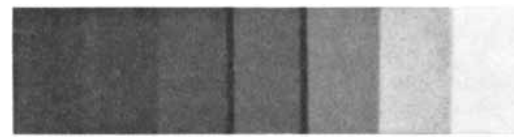
Silver halide crystals behave so:



INCREASING EXPOSURE →



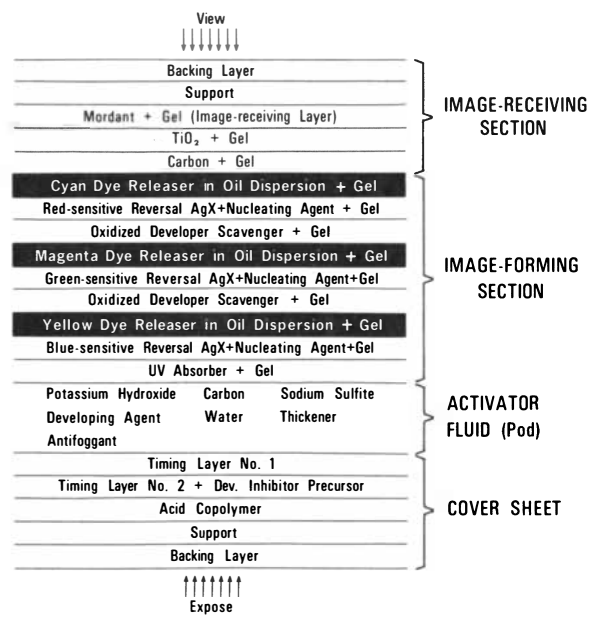
except the kind in Kodak instant print film.



In the first kind, several photons from the scene hitting within a short time mark a crystal for development. The second kind are constructed in a way that makes them sensitive internally instead of externally. Photoelectrons get trapped inside. During processing, electrons supplied chemically go where the photoelectrons went instead of triggering development on the surface, as happens to unstruck crystals with the kind of developer employed. Light thus serves to mark a crystal for nonconversion to silver, and the unstruck crystals undergo development—the reverse of what happens in conventional silver halide photography.

In color photography, though, it's not the silver image that's wanted, but a color picture. In Kodak instant print film, reacted developer—now thought of as an electron transfer agent—is formed in proportion to the silver. This now electron-avid stuff leaves the silver behind and, by regenerating itself with electrons from a compound in the adjacent layer, changes that compound so that it will hydrolyze to release a dye built into its molecular structure. The relatively unreactive dye migrates to the receiving layer, where it binds itself.

All this happens in triplicate, for each third of the visible spectrum. For the red and green, sensitizing dyes are required, and each of the three light-sensitive layers requires its own separate image-dye-releaser layer. Then there are the layers where oxidized developer is scavenged before it can cause release of the wrong dye, the solution-permeable opaque layers that keep light out of the built-in darkroom in the film, the titanium dioxide that reflects light for the finished picture, and the layers that time the ballet that starts when the pod contents are released. Then the action stops. The time scale has slowed down from the microseconds of electron flow in the silver halide crystal, through the few minutes for the dyes to migrate where they belong, to the distant days when we look back to a moment in our lives.



The theorists construct their models of nature, the scientists select the principles to use, and the engineers lay down the layers of many components.

Kodak

SCIENCE AND THE CITIZEN

Breeding in the U.S.

Since 1948 the U.S. has invested a total of about $2.8 billion in a research program aimed at developing a virtually inexhaustible alternative energy source for the 21st century: the liquid-metal-cooled fast breeder reactor (LMFBR), a type of nuclear power plant capable of producing more fissionable fuel than it consumes. In the fiscal year 1977 that cumulative expenditure will be increased by an additional $686 million, an outlay that exceeds the preceding fiscal year's funding level by about a third. The breeder program remains the largest single item in the U.S. budget for energy-related research and development, a priority ranking reflecting the widely held belief that the fuel economy of the breeder still represents this country's best prospect for satisfying its long-term energy needs.

Yet in spite of the early U.S. lead in the development of experimental breeder reactors and the continuing substantial appropriation of public funds for further research and development, it is now generally acknowledged that the U.S. breeder program lags behind parallel programs in several other advanced industrial countries. In France, for example, where the Phénix breeder reactor, a 250-megawatt prototype installation, has been in operation since 1974, the decision has been made to proceed with the construction of Superphénix, a 1,200-megawatt power station (see "Superphénix: A Full-Scale Breeder Reactor," by Georges A. Vendryes, page 26). Between now and 1985 French nuclear authorities expect to receive orders for four more domestic breeder-reactor power plants on the scale of Superphénix or larger, with the result that by the early 1990's French breeder reactors could account for a total of some 8,000 megawatts, a substantial fraction of that country's total installed capacity for generating electricity by means of nuclear power.

Meanwhile in the U.S. preliminary site preparation is just beginning on the first American prototype breeder, the 380-megawatt Clinch River demonstration plant near Oak Ridge, Tenn. According to the latest timetable announced by the Energy Research and Development Administration (ERDA), a final decision on whether or not to go ahead with the widespread deployment of full-scale commercial breeders in the U.S. will not be made until 1986, when the Clinch River plant will have been completed and fully tested. On that schedule the first commercial U.S. breeder, comparable in scale to Superphénix, cannot be expected to begin operation until 1993, roughly 10 years after Superphénix is scheduled to go critical. Nevertheless, ERDA forecasts that by the year 2000 the U.S. could be obtaining as much as 60,000 megawatts of electricity from breeder-reactor power stations.

There is evidence that even this much-revised, drawn-out timetable may be hard to meet. According to a recent report prepared by the General Accounting Office (GAO), the investigative arm of Congress, a review of what will be needed for the commercial development of the LMFBR, including the reactor's necessary support facilities, suggests that "following a decision in the mid- to late 1980's to commercialize the LMFBR ... the scenario most likely to be met ... would result in four to six commercial-size LMFBR's in operation by the year 2000." The comparatively small scale of this deployment contrasts sharply with a forecast made by ERDA in March, 1975, which envisioned 128 breeders in operation by 1998. "This scenario," the GAO report says, "is clearly not likely to be met."

A somewhat less ambitious schedule that "closely approximates current ERDA projections" is also considered in the GAO report, which concludes that even this "optimistic scenario ... could only be met if the research, development and demonstration for fuel-fabrication and plutonium-reprocessing technologies were accelerated and Government participation extended into the early stages of commercial operation."

The GAO's own "conservative scenario" is based on the general assumption that successful commercialization of the LMFBR will require the development not only of the reactor technology but also of three supporting technologies that constitute the LMFBR fuel cycle: fuel fabrication, fuel reprocessing and radioactive-waste disposal. "Major private investment commitments," the GAO analysis finds, "are not likely to be forthcoming until LMFBR's and the three supporting technologies can be demonstrated to be licensable and operable routinely" on a commercial scale. Until recently, the report continues, "the LMFBR program has placed greatest emphasis on reactor development," whereas in the future "more attention must also be given to the relationship between reactor development and the timing and rate of introduction of the three supporting fuel-cycle technologies." In this connection, the GAO analysts point out, "public acceptance and institutional adaptation may be more difficult than technical considerations. Any degree of LMFBR commercialization before the end of this century will require the active support of diverse interest groups if the long lead-time decisions and commitments are to be made in time for scale-ups of the required technologies."

Assuming that "basic uncertainties of safety, safeguards and environmental effects are resolved early and forthrightly," the GAO report concludes, "the start to LMFBR commercialization can be made by the mid-1990's." Even the level of commercialization of LMFBR's envisioned in the GAO's conservative projection, however, would require that research programs for the supporting fuel-cycle technologies "run on a schedule parallel to reactor development. Thus, if any degree of LMFBR commercialization is to occur before the end of this century, RD&D for the fuel-cycle technologies ... will have to be accelerated."

Shell Game

A basic principle of current American military policy is that the U.S. requires a triad of strategic-weapons systems to deter any potential aggressor: long-range bombers, intercontinental ballistic missiles (ICBM's) and submarine-launched ballistic missiles (SLBM's). The sea-based force, consisting of 41 Polaris submarines, each carrying 16 SLBM's, has been steadily upgraded with missiles of increasing range and accuracy, most of which (496 out of a total of 656) have now been fitted out with 10 to 14 multiple independently targetable reentry vehicles (MIRV's). The Navy's goal is to replace up to 18 boats of the present fleet with a new class of Trident submarines, each carrying 24 SLBM's with a range of 4,600 miles and with multiple warheads that would be not only independently targetable but also maneuverable during reentry.

The Air Force plan to build 244 B-1 bombers, at a total cost of about $24 billion, to replace its present fleet of B-52's has been well publicized. Less well known is the Air Force plan to replace the present force of 1,000 Minuteman ICBM's with a more powerful and more accurate missile, "Missile X," which would not be confined to an identifiable launching silo but could be moved around in a random pattern, creating what has been called "the ultimate shell game." The estimated cost of developing and installing Missile X is upward of $30 billion.

The rationale for Missile X is the assumption that a Russian force now ap-

The New 810

Datsun introduces the family cars with a 240-Z engine.

Fuel injected response.
Finally, an economical family car that doesn't skimp on performance. Because the six-cylinder, overhead cam engine gives you all the get-up-and-go you've been missing. Plus the power you need to go up hills and get onto highways.

Suddenly, there's a family car that performs and handles like a 240-Z.

Sedan or station wagon.
Both give you the economy and durability you expect from a Datsun. The famous 240-Z engine has been proven time after time against the toughest competition. The all-steel unibody is welded for quiet strength. And the fully independent suspension in the sedan assures you extra smoothness and stability.

Suddenly, a new car with a built-in reputation for durability and long life.

Comfort of a mid-size car.
Comfort you can feel: In the lush interior, tilt steering wheel, six-way adjustable driver's seat, plush cut-pile carpeting that goes all the way through the trunk. And the soothing sound of the AM/FM stereo radio.

Comfort for your peace of mind: Thanks to steel belted radials, fully appointed dash. And automatic maintenance warning system.

And comforts you can opt for: Like automatic transmission, power steering and air conditioning.

All at a comfortable price. Compared to the six-cylinder Volvo or BMW, for example, you could save at least $2,000.

The new Datsun 810. Suddenly comfort and performance come together in an economical family car.

Suddenly it's going to dawn on you.

DATSUN SAVES

proaching 1,600 ICBM's will soon be able to carry enough accurate warheads to fatally cripple two-thirds of the U.S. strategic triad, the ICBM's and the long-range bombers, in a surprise attack. According to the "worst case" scenario, the U.S.S.R. might achieve this result and still keep a few hundred ICBM's in reserve, which would not be threatened by our surviving SLBM's because the accuracy of the SLBM's is not sufficiently great. The counterforce doctrine states that it takes an ICBM to kill an ICBM, hence the need for the mobile Missile X.

According to an article in *Air Force Magazine* by Lieutenant General Alton D. Slay, the Strategic Air Command began defining the specifications of Missile X late in 1973 and started building prototype hardware last October. "The essence of deterrence," writes General Slay, "rests in large measure with the ability to absorb a first strike and then retaliate with the appropriate level of force.... [Although] U.S. [ICBM] silos are survivable today and will remain so for some time ..., the U.S. cannot afford to discount the eventual emergence of a Soviet capability that could reduce silo survivability to an unacceptably small number. Therefore an alternative to silo basing must be developed."

The Air Force has been studying three alternatives: air launching from large cargo planes, multiple aboveground shelters for transportable missiles, which would be moved around at random, and continuous hardened tunnels five to 20 miles long in which a "missile transporter/launcher moves at random intervals ..., providing total uncertainty as to missile location." The Air Force demonstrated more than two years ago that a Minuteman could be dropped by parachute from a C-5 cargo plane and subsequently launched. Since air launching has fairly obvious drawbacks (such as lower missile accuracy and the need for adequate warning to get the launching fleet airborne), it seems likely that one of the two ground-based systems will be selected.

Missile X itself will weigh twice as much as a Minuteman (68,000 kilograms compared with 34,000) and will have "several times" the payload of Minuteman III (which is reported to carry three independently targetable warheads, each with a yield of about 175 kilotons). The accuracy of Missile X, according to General Slay, "will be a quantum advance over today's Minuteman III." The accuracy of the present Minuteman III, that is, the CEP, or "circular error probable" (the radius of a circle within which the warhead has a 50 percent chance of landing), is said to be about 400 meters. If "quantum" is used in the usual sense of a tenfold change, the CEP of Missile X could be as small as 40 meters.

Curiously, little has been said about

the potential impact of Missile X on the new round of strategic-arms-limitation talks (SALT). Although the SALT I agreement due to expire in October does not explicitly forbid the deployment of mobile ICBM's, the U.S., in an appended unilateral statement, said it would look with disfavor on such missiles because their number could not be verified by "national technical means," that is, satellite reconnaissance. Most discussions of the obstacles to a new agreement have centered on what restrictions, if any, to place on a new intermediate-range Soviet bomber ("Backfire") and on sea-launched cruise missiles being developed by the U.S. with a range of more than 600 kilometers (see "Cruise Missiles," by Kosta Tsipis; SCIENTIFIC AMERICAN, February). Guidelines established at Vladivostok in November, 1974, would, if converted into a formal SALT II treaty, set a limit of 2,400 on the total number of strategic delivery vehicles—ICBM's, SLBM's and heavy bombers—that each side could deploy through 1985, and also set a ceiling of 1,320 on the total number of ICBM's and SLBM's that could be fitted out with MIRV's. The Carter Administration has already announced that it will push for a lower limit on the total number of delivery vehicles. It has also indicated that it would slow down the development of the sea-launched cruise missile. In his first news conference President Carter suggested that the U.S. and the U.S.S.R. might reach a "mutual agreement" not to deploy mobile ICBM's, but he did not say the U.S. would take the initiative by suspending development of Missile X.

Persistent Pox

Smallpox persists, although barely. The disease, which seemed last summer to be making its last stand in the southern desert of Ethiopia, was apparently carried by an infected Somali nomad into neighboring Somalia, where it spread slowly and was first confirmed at the end of September. As of the first week in February 39 cases had been discovered in Somalia, almost all of them in a limited area of Mogadisho, the capital and major port; what appeared to be a limited outbreak (four cases) in a nearby area of Kenya had been traced to Mogadisho.

The outbreak in Somalia has been a frustrating turn of events in the final phase of a 10-year campaign conducted by the World Health Organization to eradicate the most devastating pestilence in human history (see "The Eradication of Smallpox," by Donald A. Henderson; SCIENTIFIC AMERICAN, October, 1976). Ethiopia was the last country in which the disease was endemic, and there the campaign appears to have been successful: in spite of active surveillance no new case has been discovered in all Ethiopia since last August 9. Since the general level of vaccination is high in Somalia, there seems to be little chance of a widespread outbreak there. With active containment measures now being taken—isolation of patients and vaccination of contacts—the thread of transmission should be broken soon. For two years after the last case is recorded, surveillance will continue in the region. Then, after certification by a WHO commission, smallpox will be considered to have been eradicated from the earth.

There are several reasons why eradication—the first such achievement in medical history—is feasible in the case of smallpox. The disease is serious and was much feared, so that there was motivation for a worldwide effort against it. There is a highly effective vaccine against the smallpox virus. The disease is made visible by its rash and pustules, so that undetected, asymptomatic cases are rare; the few cases that are so mild as to cause no rash are apparently too mild for transmission. Most important, the virus has no animal host (as the plague bacillus has in the rat, for example); the disease is transmitted only from person to person. (There is a similar disease, monkeypox, that has sometimes been transmitted to people, but it seems not to spread from person to person.) When no human being harbors the virus, there should remain only one reservoir: the stocks in research and diagnostic laboratories. The WHO is now surveying all the world's laboratories to locate such stocks and to urge that they be destroyed; the hope is that only a few laboratories will need to retain the virus for research, and that they will enforce rigid safety precautions.

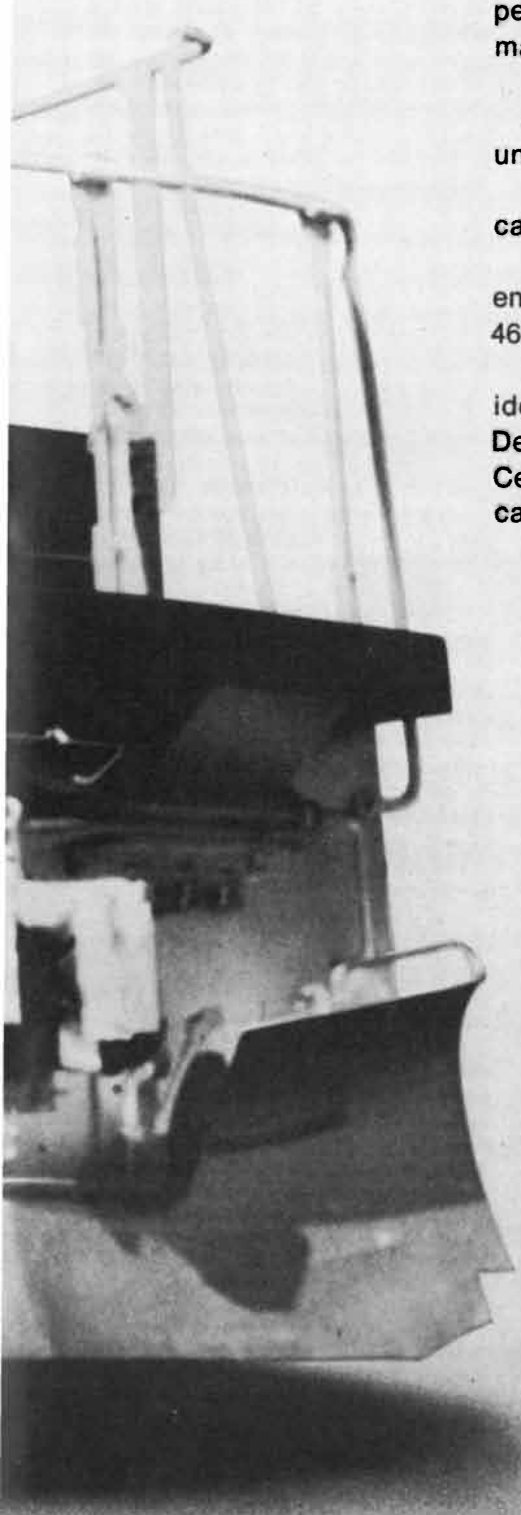
Making Fermions out of Bosons

Can two wholes add up to a half? In the arithmetic of subatomic physics it appears that perhaps they can.

All particles in nature can be sorted into two classes according to their "statistics" (which describe their behavior in groups) and their spin angular momentum (which is closely related to statistics). One class of particles, the fermions, obey a rule forbidding more than one particle from occupying a given state; the spin angular momentum of the fermions, when measured in fundamental units, is invariably a half integer (such as 1/2, 3/2, 5/2). The other class of particles, the bosons, can be brought together in unlimited numbers; their spin is always an integer (such as 0, 1, 2). The proton, the neutron and the electron are all fermions; the bosons include the photon, or quantum of electromagnetic radiation, and the pions, kaons and other mesons.

Composite structures (such as atoms)

The best ideas in diesel-starting batteries come from ESB.



The idea: develop a diesel-starting battery that is superior to its competitors. In capacity and storage life. And in ease of handling and maintenance. Then make it durable and reliable.

The result: our Exide UD 466 Ironclad® battery.

It's 11% more powerful than its predecessor. Yet because it's one unit (not the standard four), it takes just a fraction of the time to install.

Our amazing new battery is easier to handle, too. Simpler to take care of. Safer to work with. And a lot more difficult to damage.

ESB technology and research have made us one of the largest energy storage companies in the world. And they've made the Exide UD 466 the best diesel-starting battery there is.

If you'd like to learn more about our ideas in energy, write: Communications Department, ESB Incorporated, 5 Penn Center Plaza, Philadelphia, PA 19103. Or call: (215) 972-8000.

Ideas in energy.

ESB

made up of these particles are either fermions or bosons according to whether they include an odd or an even number of fermions. It should be apparent that a composite made up only of bosons should always be a boson. A remarkable theory contradicting this principle has recently come to light. It seems that two integer-spin bosons can combine to make a particle with half-integer spin and the statistics of a fermion.

The bosons required for this operation are rather special. One must bear an electric charge; the other must be a magnetic monopole, a particle with a magnetic charge (either "north" or "south"). Actually it has been known for many years that the interaction of a monopole and a point electric charge would generate "extra" angular momentum, just as the interaction of electric and magnetic fields sets an electric motor spinning. What has now been established is that such a spontaneously spinning object could arise as a solution to a quantum field theory and that the object would be a fermion.

The magnetic monopole was introduced into physics to fill a vacant niche in James Clerk Maxwell's theory of electromagnetism. It was supposed to be an elementary particle, an object without substructure or antecedent; such elementary monopoles have not been observed. Three years ago it was shown that in the context of certain gauge theories, which are generalizations of Maxwell's theory, monopoles could appear that were not elementary but were composed of other particles bearing only electric charges. This discovery was made independently by Gerhard 't Hooft of the University of Utrecht and by Alexander M. Polyakov of the Landau Institute for Theoretical Physics near Moscow.

More recently it has been observed that the monopoles created in gauge theories might form bound states with a point electric charge. The possibility of such composites was suggested independently and from different premises by 't Hooft and Peter Hasenfratz of Utrecht and by Roman W. Jackiw and Claudio Rebbi of the Massachusetts Institute of Technology. Both groups report their findings in *Physical Review Letters.*

As in the case of point monopoles, a spinless composite monopole bound to a spinless electric charge could have a combined spin angular momentum of 1/2. The statistics problem for such an object was analyzed by Alfred S. Goldhaber of the State University of New York at Stony Brook. Also writing in *Physical Review Letters,* he showed that the new particle would indeed be a fermion—a fermion made up entirely of bosons.

These theoretical developments show how a fermion could be made out of suitable bosons, but they provide no evidence that any of the known fermions are constructed in that way. Indeed, for familiar fermions such as the electron or the proton such a structure is impossible, since the strong magnetic field of the monopole is not observed near those particles. A more plausible set of candidates might be the quarks, the presumed constituents of protons and neutrons and related particles. Quarks themselves have not been detected in isolation and their magnetic properties are not known.

The distinction between fermions and bosons is profoundly important because the two kinds of particles seem to have different roles in nature. In a sense the world is made of fermions, which interact with one another by exchanging bosons. It now seems possible, although it is very far from certain, that this model of the universe could be simplified by eliminating one of the fundamental categories and describing fermions as composites made of bosons.

Reverse Uncertainty

The uncertainty principle, introduced by Werner Heisenberg, states that the product of the uncertainties in the values of certain related variables, such as the position and momentum of a subatomic particle, is greater than or equal to Planck's constant. Over the years the principle has become a metaphor for the fact that the observer inevitably disturbs what is being observed. Recently David V. Forrest, James H. Ryan and Phillip Zeidenberg of the New York State Psychiatric Institute encountered a curious instance of the reverse effect: what was being observed disturbed the observer. Writing in *The American Journal of Psychiatry,* they describe a clinical study of marihuana intoxication conducted at the institute's Drug Abuse Research Unit in the course of which the investigators became increasingly "distracted, irritable and nauseated."

The cause of their symptoms became evident when an analysis of urine from a control subject given a placebo revealed the presence of metabolic products of Δ -9-tetrahydrocannabinol, the active ingredient of marihuana. Both the control subjects and the investigators had succumbed to a passive drug intoxication (a "contact high") simply by breathing the air in the room where the subjects were smoking. Once identified, the interference caused by the "reverse Heisenberg effect" was eliminated by a change in the experimental design: the investigators began observing the marihuana-smoking subjects with closed-circuit television cameras connected to monitors in a remote station.

CLOCKWISE FROM ABOVE:

Owl at 20 feet
Celestron 750mm, f/6

Rose at 15 feet
Celestron 750mm, f/6

Peacock at 50 feet
Celestron 750mm, f/6

Orion Nebula
Celestron 8

Rings of Saturn
Celestron 14

Squirrel at 75 feet
Celestron 1250mm, f/10

We Give You New Worlds

The basic Celestron 8 ($790), shown here on equatorial wedge and tripod.

Beyond the front end of a Celestron Telescope lie entirely new worlds — bigger, brighter, sharper, more detailed worlds.

Our sunlit Earth — marvelous in its intricacy, overwhelming in its color, profound in its immediacy. And worlds beyond — in a universe of majestic proportions and breathtaking beauty. A universe where planets roam the outer reaches of our solar system and their moons shuttle to and fro. Where great spheres of stars — light-centuries in diameter — ride loftily above the disk of our galaxy. Where remote clouds of gases cast their eerie filamentary glow onto the waters of infinity.

These are the worlds we offer.

And next week at this time, you could have in your hands the means to explore them.

For a closer look at the worlds we offer, send for our free color catalog. Write to Celestron International, 2835 Columbia St., Box 3578-BC, Torrance, Calif. 90503, USA. Or telephone (213) 328-9560.

Celestron

Bringing Man and Nature Closer Together

WHAT OUR TWO UNCOMMON

Porsche Turbo Carrera: A $28,000* rear-engine, rear-wheel drive ultimate sports car.

Audi Fox: A $5,445* front-engine, front-wheel drive sedan.

What could they possibly have in common?

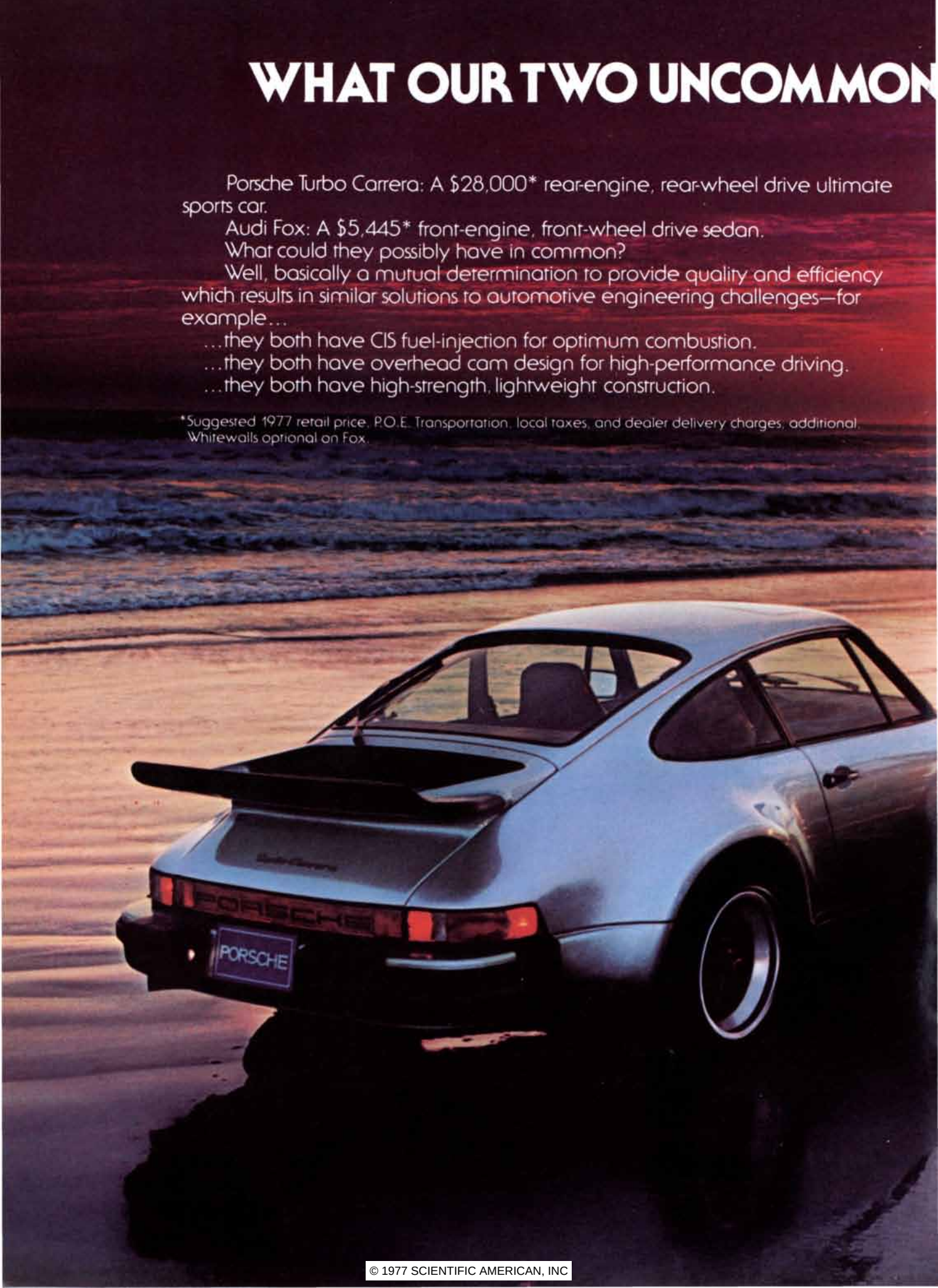
Well, basically a mutual determination to provide quality and efficiency which results in similar solutions to automotive engineering challenges—for example...

...they both have CIS fuel-injection for optimum combustion.

...they both have overhead cam design for high-performance driving.

...they both have high-strength, lightweight construction.

*Suggested 1977 retail price. P.O.E. Transportation, local taxes, and dealer delivery charges, additional. Whitewalls optional on Fox.



Biological Nitrogen Fixation

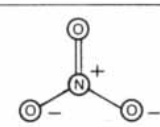
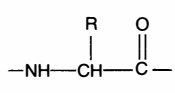
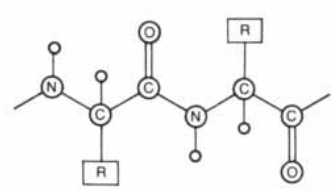
Only a few bacteria and simple algae have the cellular equipment needed to "fix" the nitrogen of the atmosphere into ammonia. They are the major suppliers of this limited agricultural resource

by Winston J. Brill

Antoine Laurent Lavoisier gave nitrogen the name "azote," meaning "without life," because it differed from the other main component of air (oxygen) in being unable to sustain the metabolism of living organisms. The name has turned out to be an ironic one. Nitrogen is an essential constituent of proteins, and we now know that large amounts of it are required by all forms of life. Indeed, for both plants and animals nitrogen is probably the commonest limitation to growth, and an inadequate supply of nitrogen for agriculture is an important factor contributing to human hunger.

Nitrogen is at once an abundant element, making up almost 80 percent of the earth's atmosphere, and a scarce nutritional resource. The paradox is easily explained: the form of nitrogen in air is so inert that it is useless to the vast majority of organisms. Nitrogen can enter biological systems only when it has been "fixed," or combined with certain other elements, such as hydrogen or oxygen. Today the fixation can be accomplished industrially, through the manufacture of ammonia from hydrogen and atmospheric nitrogen. The making of ammonia and of other chemical fertilizers derived from it is now a major industry, but the bulk of all fixed nitrogen is of biological origin.

In nature, nitrogen fixation is a faculty reserved to a few genera of bacteria. (Included among these bacteria are several blue-green algae, a group of organisms that today are generally classified with the bacteria under the name cyanobacteria.) No higher organisms have developed the capability, although several participate indirectly by forming symbiotic associations with nitrogen-fixing bacteria. The best-known of these relations is the one between the plants called legumes and various bacteria of the genus *Rhizobium*. Other nitrogen-fixing bacteria associate with other host plants, and many live freely in the soil or in water. A few are photosynthetic; some require oxygen; others can grow only when oxygen is excluded from their environment. All these organisms apparently share a common mechanism for nitrogen fixation; as in the industrial process, the initial product is ammonia. They also share a unique enzyme: nitrogenase. We are just beginning to unravel the structure of nitrogenase, to learn how it functions and how it is regulated and to understand what characteristics distinguish the organisms that possess it. The potential benefits of this knowledge can be reckoned in a higher worldwide standard of living. The cost of fertilizer has increased dramatically in the past few years, affecting the cost of food in the more affluent countries and restricting the supply of food in the less affluent ones. If the nitrogen-fixing activities of bacteria can be understood, they might also be improved, and ultimately they might be conferred on other organisms, perhaps including cereal crops. The result would be reduced dependence on nitrogenous fertilizer.

SUBSTANCE	FORMULA	STRUCTURE
MOLECULAR NITROGEN	N_2	
NITRATE ION	NO_3	
AMMONIA	NH_3	
PEPTIDE (SUBSTRUCTURE OF PROTEINS)		

FIXATION OF NITROGEN is the conversion of the abundant but nearly inert molecular gas N_2 into compounds useful to living organisms. In biological fixation and in the industrial Haber process the immediate product is ammonia. Nitrates are another source of fixed nitrogen common in soil and in chemical fertilizers. Most of the nitrogen that enters biological systems is made into proteins, molecules made up of amino acids linked by a carbon-nitrogen bond.

Nitrogen Chemistry

The nitrogen in the atmosphere is a diatomic gas, that is, it consists of molecules made up of two atoms each, denoted N_2. Molecular nitrogen is nearly inert because the chemical bond joining the atoms is exceptionally strong and stable; it is a triple bond and a large quantity of energy must be supplied in order to break it.

In industrial fixation the required energy is provided by fossil fuels. In a process developed in the early years of the 20th century by Fritz Haber and Karl Bosch atmospheric nitrogen is

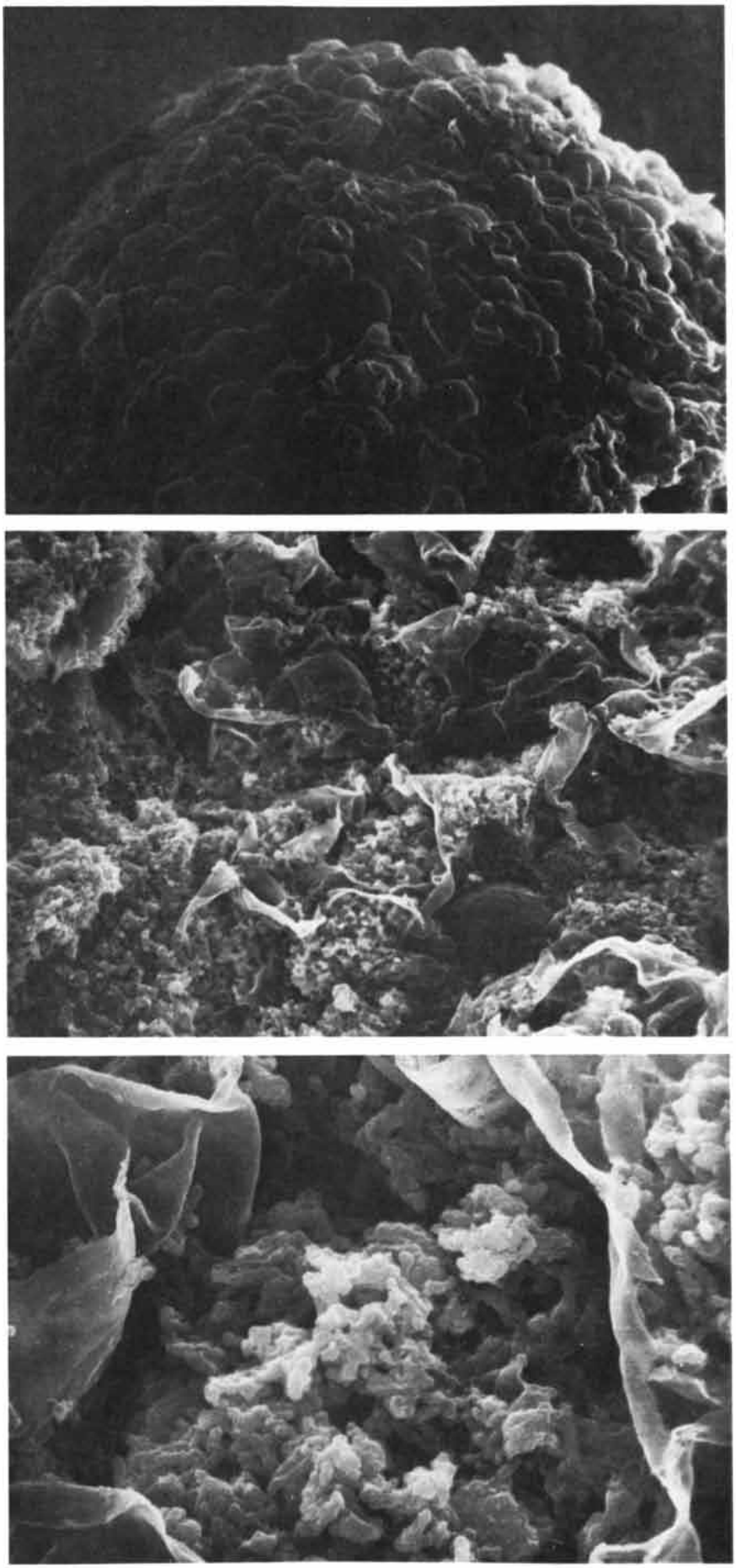
combined with hydrogen at high temperature and pressure in the presence of a catalyst containing iron. The product is ammonia (NH_3), which is itself an effective fertilizer and which can be converted into other useful nitrogen compounds, such as urea and nitrates.

The energy cost of the Haber process becomes apparent when the source of the required hydrogen is considered: it is extracted from natural gas or petroleum. For this reason the cost of nitrogenous fertilizers is closely correlated with the cost of fossil fuels.

In the soil fixed nitrogen from industrial fertilizer or from natural sources is taken up by plant roots and is ultimately employed in the synthesis of biological molecules. By far the largest share is incorporated in the structure of proteins, the versatile molecules that are responsible for metabolism in living cells. Proteins are made up of amino acids, all of which contain at least one atom of nitrogen. A typical protein might be composed of a few hundred amino acid units. A crucial structural element is the peptide bond, which links one amino acid to the next; the bond connects a nitrogen atom in one amino acid to a carbon atom in another.

Through plant and animal wastes and dead tissues fixed nitrogen is returned to the soil, where much of it can be recycled. Proteins are dismantled to yield amino acids, and these are often broken down further to ammonia or nitrate; these substances can then be absorbed anew by living roots. Competing with plant roots, however, are a class of decay bacteria, the denitrifiers, that break down fixed nitrogen and return it to the form of a diatomic gas. The denitrifying bacteria thus complete the nitrogen cycle by returning the element to the atmosphere. As a result of their activities the reservoir of fixed nitrogen in the soil must be continually replenished. In agriculture there is an even greater drain on the supply of fixed nitrogen: each time a crop is harvested for market, the nitrogen it contains is lost to the soil.

Small quantities of fixed nitrogen are



SOYBEAN ROOT NODULE consists of plant cells crammed full of nitrogen-fixing bacteria. A single nodule is seen at progressively greater magnification in this sequence of scanning electron micrographs made in the author's laboratory at the University of Wisconsin. At the top is the pebbly surface of a small root nodule. In the middle photograph the nodule has been sliced open to expose the interior. At the bottom one region of the interior surface is enlarged, showing a dense mass of bacteria spilling out of a ruptured plant cell. The bacteria are of the species *Rhizobium japonicum*, which associates exclusively with soybeans; other plants of the family Leguminosae harbor other *Rhizobium* species. Associations of legumes and *Rhizobium* account for 40 percent of all nitrogen fixed by biological means, and for virtually all that fixed by crop plants.

added to the biosphere each year by inorganic processes, such as the formation of nitrates in automobile engines, in lightning discharges and in volcanic emissions. A somewhat larger amount—about a fourth of the total world production of fixed nitrogen—is ammonia made by the Haber process. The remainder is the bacterial product. It is estimated to amount to 150 million metric tons per year.

Legumes

The activities of nitrogen-fixing bacteria were harnessed for human benefit centuries before either the bacteria or the nitrogen-fixing process was discovered. Farmers have long experience in the growing of legumes, such as soybeans, peanuts, alfalfa, beans, peas and clover. These crops can revitalize the soil, an effect that is now understood to result from nitrogen fixation by species of *Rhizobium* that form nodules in the roots of the legumes. Commercial *Rhizobium* inoculants are employed in planting legumes to ensure that the soil contains the appropriate bacteria. Legume crops are often grown in rotation with nonlegumes, such as corn. In this way nitrogenous substances from one season's legume crop help to fertilize the next season's grain crop. For a maximum yield of grain industrial fertilizer must still be applied, but the amount needed is reduced.

The *Rhizobium* enters the legume root through a root hair, a cell on the surface of the root that is specialized for absorption. The cell wall of the root hair invaginates to form an infection thread, which contains many proliferating *Rhizobium* cells. Most of these infections are abortive, but a few grow back to the base of the root hair and, by repeating the invagination process, enter the cortical cells of the root. Eventually the tip of the infection thread ruptures, releasing the bacteria into the cortical cells, which then develop into a bulbous enlargement: a root nodule. The nodule consists of enlarged plant cells, most of which are packed with bacteria. Ammonia produced by the bacteria is combined with carbon compounds derived from plant photosynthesis to yield amino acids, which are then incorporated into plant proteins.

As the term "infection thread" implies, the introduction of *Rhizobium* into legumes resembles a disease process, but it is one in which the plant cooperates. The welfare of the plant is served best if it encourages infection by *Rhizobium* but excludes all other bacteria, some of which might be pathogens or parasites. This discrimination is accomplished through a system of chemical markers by which the plant and the bacterium recognize each other.

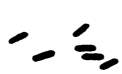
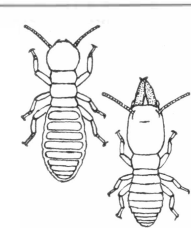
Each legume is associated with a distinct species of *Rhizobium*. The bacteria that form nodules in soybeans, for example, will not infect alfalfa. In 1974 the first element of the recognition mechanism responsible for this specificity was discovered by Benjamin B. Bohlool and Edwin L. Schmidt of the University of Minnesota. They identified a protein from soybean that binds to cells of *Rhizobium japonicum,* the bacterial species that infects soybeans, but not to any other *Rhizobium* species. Frank B. Dazzo and David H. Hubbell of the University of Florida subsequently found another protein that seems to bear the same relation to clover and to *Rhizobium trifolii,* the bacterium that infects clover roots. They named the protein trifoliin.

Dazzo has continued his work at the Center for Studies of Nitrogen Fixation at the University of Wisconsin. He has recently shown that trifoliin is found on the surface of clover root hairs, the initial site of infection. Furthermore, he has shown that trifoliin binds to a polysaccharide on the surface of the infecting *Rhizobium trifolii* but not to surface polysaccharides from other *Rhizobium* species.

A plausible hypothesis derived from these experiments is that trifoliin acts as a link between bacterium and plant. Further studies, employing labeled antibody molecules, have provided preliminary information on the sites where trifoliin binds to the plant root and to the bacterial surface. Remarkably, the two binding sites are antigenically similar, that is, they have an affinity for the same antibody molecules. The significance of this similarity is not yet understood, but analogies are known. For example, the surfaces of some pathogenic bacteria are structurally similar to the surfaces of mammalian cells. As a result of this mimicry host responses that would normally eliminate an invading microorganism can sometimes be deceived.

Nonrhizobial Fixation

The symbiotic association between legumes and *Rhizobium* is the most highly developed and most sophisticated system for biological nitrogen fixation, but it is not the only one. For exam-

	FREE-LIVING BACTERIA				
NITROGEN-FIXING ORGANISM	 *AZOTOBACTER VINELANDII*	 *CLOSTRIDIUM PASTEURIANUM*	 *KLEBSIELLA PNEUMONIAE*	 *RHODOSPIRILLUM RUBRUM*	 *CITROBACTER FREUNDII*
ASSOCIATED ORGANISM	NONE	NONE	VARIOUS	NONE	 TERMITE
NATURAL HABITAT	AEROBIC SOILS	ANAEROBIC SOILS	AEROBIC AND ANAEROBIC SOILS; WATER; ALSO IN ASSOCIATION WITH PLANTS, MAN	SURFACE OF POLLUTED PONDS (A PHOTOSYNTHETIC BACTERIUM)	TERMITE GUT

NITROGEN-FIXING ORGANISMS include several genera of bacteria and cyanobacteria, or blue-green algae. These are among the simplest and presumably the most primitive organisms, being distinguished by the absence of cell nuclei. Among the nitrogen-fixing or-

ple, the alder tree, a hardwood species common in the northwestern U.S., encapsulates nitrogen-fixing bacteria in root nodules not unlike those of the legumes. Another symbiosis involves a small aquatic fern, *Azolla,* and a cyanobacterium that is capable of both photosynthesis and nitrogen fixation. Occupying cavities in the fern leaves, the cyanobacterium supplies nutrients that enable the fern to propagate in waters deficient in fixed nitrogen. Farmers in Vietnam have made use of the alga's capabilities by allowing *Azolla* to grow in flooded rice paddies; the fern might also be grown in ponds and harvested as a nitrogen-rich mulch.

Another apparent symbiosis, although it is probably a rather loose one, was discovered by Johanna Dobereiner of the Agricultural Research Institute in Brazil. She found nitrogen-fixing bacteria growing in association with the roots of certain tropical grasses. For example, the grass *Digitaria* was found to support populations of the bacterium *Spirillum lipoferum,* which is known to fix nitrogen. The bacteria do not form specialized structures such as nodules but simply grow on the surface of the roots. It is notable that most important grain crops, including wheat and corn, are genetically derived from tropical grasses.

A further intriguing discovery was made when a colleague of Dobereiner's noticed that among corn plants growing in nitrogen-deficient fields a few were taller than the rest. When the exceptional plants were dug up, Dobereiner found *Spirillum lipoferum* associated with the roots. It was a finding of great potential importance, since it implied that corn might be grown without fertilizer if the bacterium-root association could be reliably established. In subsequent experiments in several other laboratories, however, attempts to increase the yield of corn by inoculation with *Spirillum lipoferum* have had variable results. The nature of the association is still under investigation.

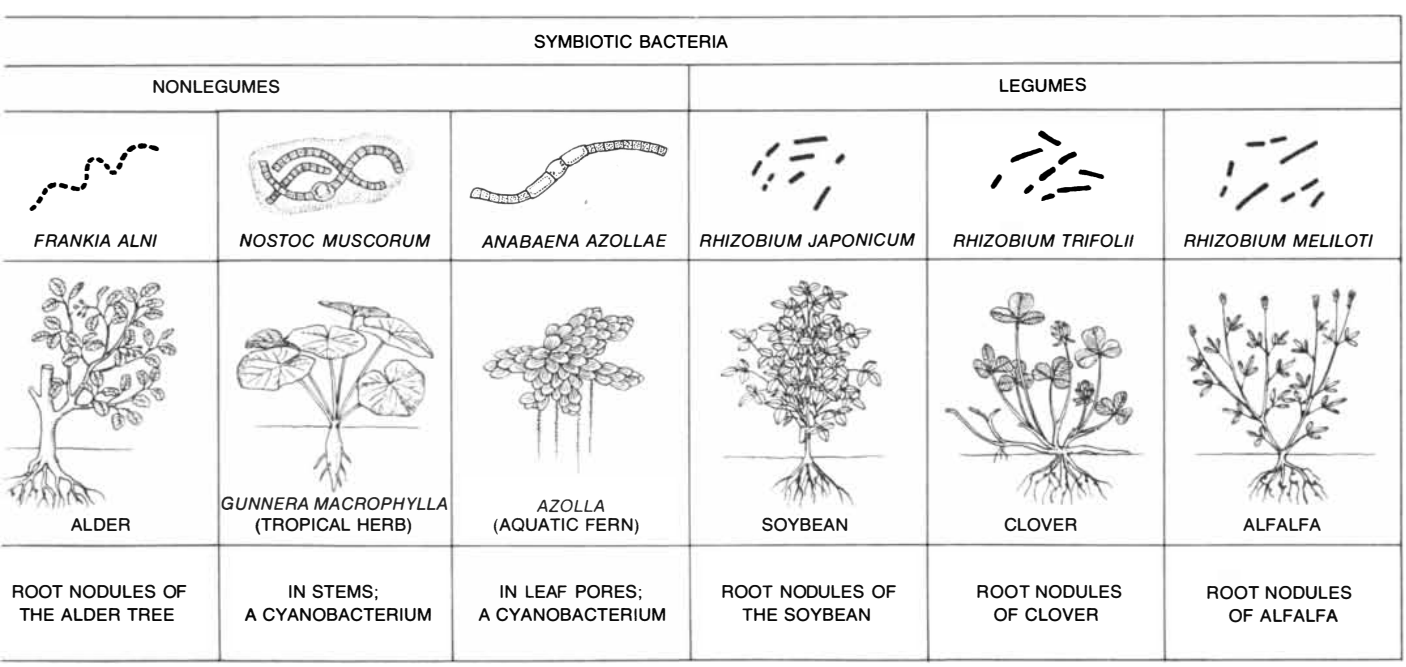
Among the more unusual symbioses are those involving nitrogen-fixing bacteria that colonize termites and shipworms. These pest species, which survive on a diet of wood, have long been known to harbor microorganisms that secrete enzymes for the digestion of cellulose. It has now been established that they also receive dietary assistance from another population of microorganisms. Wood is a poor source of biological nitrogen; bacteria living in the intestines of the termites and shipworms provide a supplement derived from atmospheric nitrogen.

In addition to obligatory symbionts, there are a number of organisms that fix nitrogen while living independently. Many cyanobacteria, for example, fix atmospheric nitrogen while growing freely at the surface of ponds. In many cases these blue-green algae constitute a nuisance: the organic nitrogen, released from the dead algal cells, promotes the growth of aquatic weeds and contributes to the process called eutrophication. On the other hand, the same algae might be cultivated and harvested as fertilizer or animal feed.

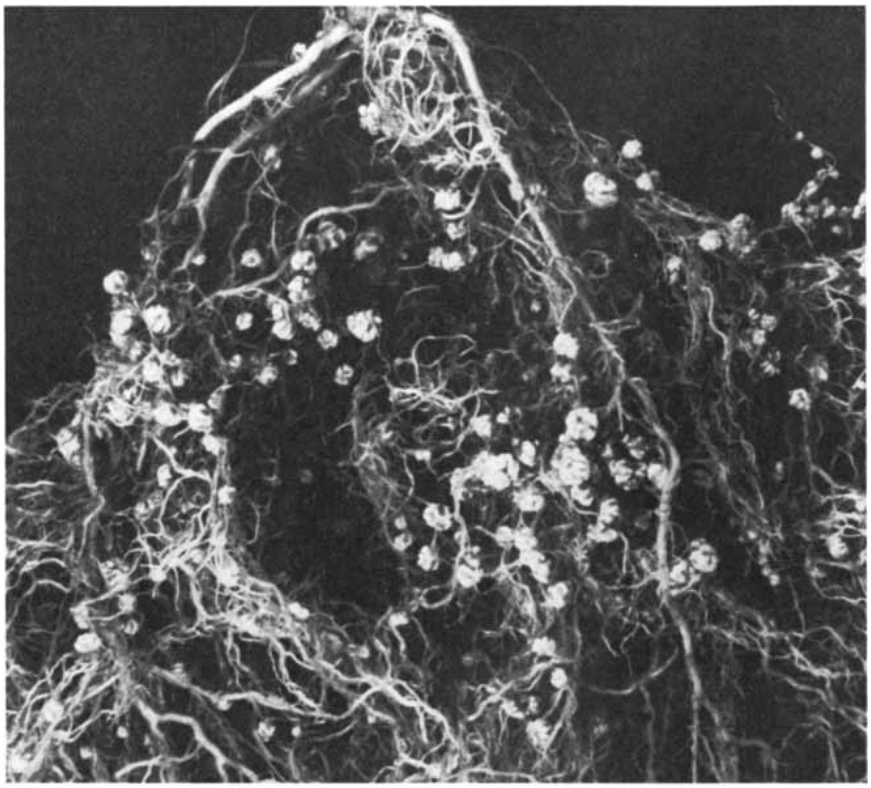
There are also free-living bacteria with the capacity to fix nitrogen. They include members of the genus *Clostridium,* which are anaerobic bacteria; they cannot grow in the presence of oxygen. Others, members of the genus *Klebsiella,* can grow either with or without oxygen and are found both free-living and in association with plants and animals. Finally, there is the genus *Azotobacter,* a group of aerobic bacteria whose name derives from Lavoisier's term for nitrogen. The contribution of these free-living bacteria to global supplies of fixed nitrogen is probably modest, but their contribution to the study of nitrogen fixation has been quite large. Because the complications of symbiosis are avoided, free-living bacteria are favored organisms for investigations of the biochemistry of fixation.

Nitrogenase

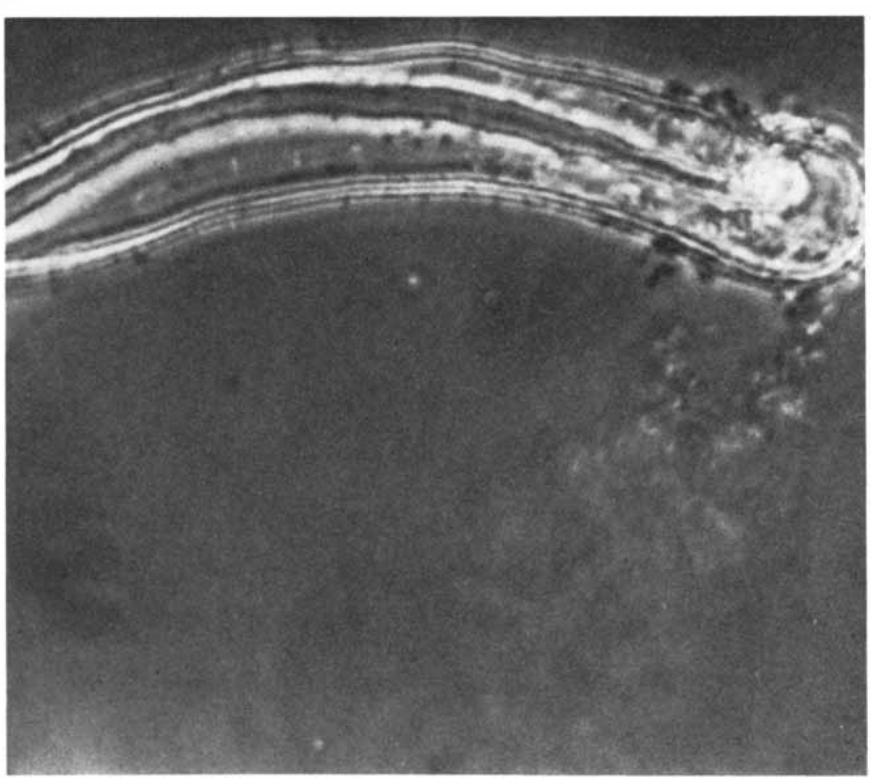
The overall chemical reaction of nitrogen fixation is the same whether it is achieved by the Haber process or takes place in the living cell. First the triple bond of the N_2 molecule must be broken; then three hydrogen atoms must be bound to each nitrogen atom. In the Haber process the hydrogen is supplied as a molecular gas; in most nitrogen-fixing bacteria it is extracted from organic molecules such as glucose, the principal carbohydrate product of photosynthesis. Hydrogen atoms are transferred from glucose to nitrogen through a network of intermediate molecules. Actually it is only the electrons that are actively transported; the aqueous medium of the cell is a sea of protons, or hydrogen nuclei, and these are readily supplied to complement free electrons. A transfer of electrons between two substances is called an oxidation-reduction reaction; the donor of electrons is said to be oxidized by the reaction and the acceptor to be reduced. Thus in nitrogen



ganisms are free-living forms and those that thrive only in symbiotic association with higher plants and animals; a few can adopt either mode of life. The bacteria that live in an anaerobic habitat cannot survive exposure to oxygen; aerobic bacteria need oxygen for growth.



SOYBEAN ROOTS have many bulbous nodules that result from infection with *Rhizobium*. In this symbiotic association the plant supplies carbohydrate to the bacterium, which in turn supplies plant with fixed nitrogen. In legumes the plant also protects the bacteria from exposure to oxygen, which permanently denatures nitrogenase, the crucial enzyme in nitrogen fixation.



INFECTION WITH RHIZOBIUM takes place through the root hairs, cells at the periphery of roots. In this photomicrograph, made by Frank B. Dazzo of the University of Wisconsin, many bacteria (*fuzzy black objects*) adhere to a root hair. In addition the bacteria have already entered the cell through an infection thread, the long tube extending through the root hair. The infection thread is formed by the invagination of the cell wall; the eventual release of the bacteria from the thread into the cortical cells of the root leads to the development of a nodule.

fixation glucose is oxidized and nitrogen is reduced. Of course the two molecules do not interact directly; the pathway that connects them is a complex one, and some segments of it have not yet been thoroughly explored.

One of the mystifying features of the nitrogen-fixing reaction is the amount of energy consumed. The energy, which is derived from the breakdown of glucose or other carbohydrates, is supplied in the form of adenosine triphosphate (ATP), the universal energy currency of the cell. The conversion of one molecule of N_2 into two molecules of ammonia requires from 12 to 24 molecules of ATP. Part of this energy must be applied to breaking the stable bond between the nitrogen atoms, but far more is required than a naïve analysis would suggest. The likely explanation for the inefficiency is that not all the ATP goes toward reducing nitrogen; some of it may be diverted to competing reactions. In any case it appears that nitrogen fixation has a high cost in energy for the cell, just as it does for the industrial chemist.

The key molecule in the nitrogen-fixation pathway is the enzyme nitrogenase. All nitrogen-fixing organisms contain nitrogenase, which does not seem to vary significantly in structure from one species to another. No organisms that lack nitrogenase are able to fix nitrogen.

The enzyme consists of two proteins, labeled Component I and Component II. Component I has a molecular weight of 220,000 and is made up of four subunits, each of which is a single strand of amino acids; in addition Component I contains 24 iron atoms and two molybdenum atoms. Component II has a molecular weight of 55,000, consists of two protein subunits and includes four iron atoms.

Rather little is known of how this array of proteins and metal atoms is assembled to make an enzyme. The structure of the entire molecule may eventually be mapped through X-ray-diffraction analysis. In the meantime less direct methods have provided some information on the chemical environments of the various metal atoms. It is probably not a coincidence that the most effective catalysts in the Haber process are transition metals such as iron and molybdenum. In nitrogenase the role of the two molybdenum atoms is particularly fascinating because they seem to be part of the active site of the enzyme. The molybdenum is bound not to the large proteins of Component I but to a small cofactor, which Vinod K. Shah, working in my laboratory at the University of Wisconsin, has purified. A surprising recent finding is that the cofactor also contains some of the iron atoms associated with Component I. The molecule Shah isolated is capable of reactivating nitrogen fixation in a mutant strain of bacteria that lacks the cofactor. Edward I. Stiefel of the Charles F. Kettering Research

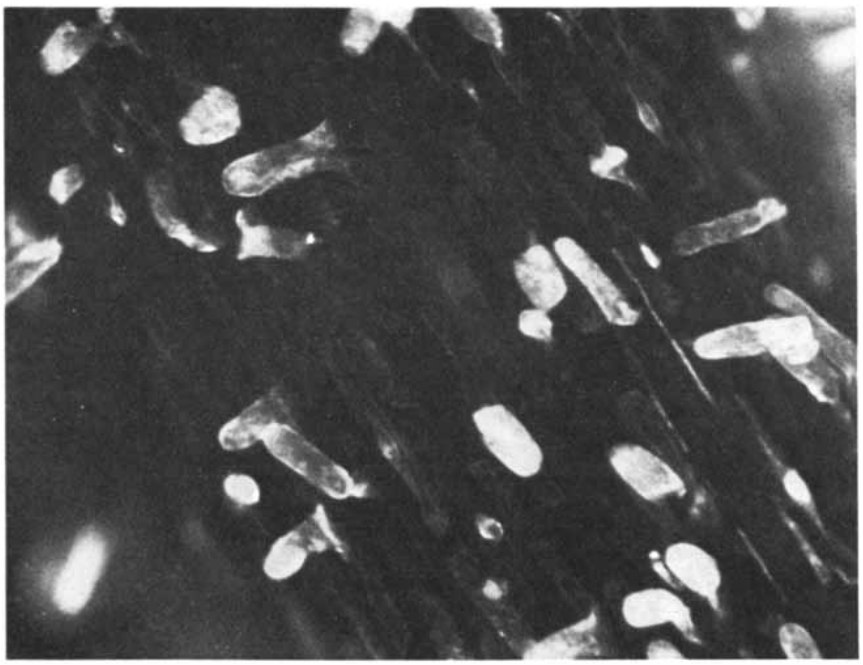
Laboratory has pointed out that among all the metals found in enzymes molybdenum is uniquely capable at both higher and lower oxidation states of transferring two electrons and two protons, and that may be its role in nitrogenase. Only a few other enzymes contain molybdenum; one of them, interestingly, is nitrate reductase, an enzyme required by plants for the conversion of nitrates into amino acids.

Much of the progress made in the past 15 years toward understanding the biochemistry of nitrogen fixation can be ascribed to two experimental techniques of uncommon importance. One of these techniques is the acetylene-reduction assay. In 1965 Robert Schöllhorn and Robert H. Burris of the University of Wisconsin and Michael J. Dilworth of Murdoch University in Australia discovered that the gas acetylene inhibits nitrogenase activity. Subsequent experiments showed that the enzyme reduces acetylene (C_2H_2) to another gas, ethylene (C_2H_4). Nitrogenase activity can therefore be evaluated simply by incubating an organism with acetylene and then measuring the production of ethylene by gas chromatography. Earlier methods for estimating nitrogen fixation involved tracing heavy isotopes of nitrogen or measuring an increase in the nitrogen content of an organism, a time-consuming procedure. The acetylene-reduction assay is both quick and comparatively accurate.

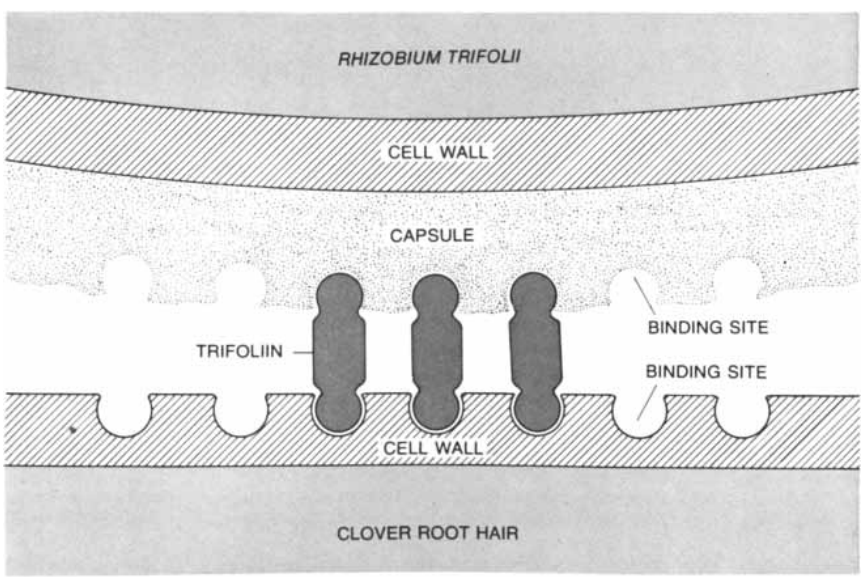
The other improvement in experimental methods was the development of a system for observing nitrogen fixation in vitro, that is, in the absence of living cells. Extracts of bacterial cells can be prepared in which the nitrogenase is not destroyed, but ordinarily the enzyme cannot function without its associated cellular machinery. In the early 1960's it was discovered that cell-free extracts could be made to fix nitrogen if small amounts of ATP were added along with a strong reducing agent, such as sodium dithionite. With such an in vitro system nitrogen fixation becomes a laboratory process that can be manipulated and measured with comparative ease.

What is the progress that has been made in understanding nitrogenase biochemistry? The present state of knowledge is summarized in recent findings of William H. Orme-Johnson of the University of Wisconsin, Leonard E. Mortenson of Purdue University and Barry E. Smith and his colleagues at the University of Sussex. They have shown that the first event in the sequence that leads to fixation is the reduction of enzyme Component II by an electron-transport protein external to the nitrogenase. The reduced Component II reacts with ATP and then reduces Component I. Finally Component I reduces molecular nitrogen, eventually forming ammonia. The same sequence of events can be described in another way. Component II first accepts an electron from a transport protein; the electron is then transferred to Component I and finally to nitrogen. No substances intermediate between nitrogen and ammonia have been discovered, so that all the intermediate states must remain bound to the nitrogenase. There is evidence corroborating the intuitive supposition that the electrons are transferred by iron and molybdenum atoms, but the actual mechanism of transfer remains a mystery.

It is the final event in this sequence—the actual reduction of nitrogen—that is at once the most interesting and the most baffling. The process probably will



EXCLUSIVE RELATION between a legume and a nitrogen-fixing bacterium is demonstrated by the binding of *Rhizobium trifolii* to the root hairs of clover. A fluorescent dye was first linked to a polysaccharide in the bacterial capsule, an amorphous sheath that surrounds the cell wall. The labeled polysaccharide was then incubated with clover roots. Fluorescence of the clover root hairs indicates that capsule polysaccharide has preferentially bound to them. Similar experiments show that polysaccharide molecules from other *Rhizobium* species do not bind to the clover roots and that *Rhizobium trifolii* does not bind to the roots of other legumes.



RECOGNITION OF RHIZOBIUM by a legume seems to be mediated by a protein that links the bacterium to the root hair. In the case of clover the protein has been given the name trifoliin. The binding sites for trifoliin in the cell wall of the plant and in the bacterial capsule are antigenically related; presumably they are also similar in structure. One interpretation of this surprising discovery is that bacterial binding site developed through imitation of plant; such mimicry may have helped to defeat the plant's defenses against invasion by foreign organisms.

not be understood in detail until the structure of the active site of nitrogenase has been determined. One hypothesis is that one of the bound intermediate states is a diimide, a molecule with the structure HN = NH. It is supposed, in other words, that at least one hydrogen atom is bound to each nitrogen before the bond between the nitrogen atoms is completely dissolved. There is an appealing symmetry between this process and the presumed mechanism by which nitrogenase reduces acetylene to ethylene, but there is little evidence either proving it or disproving it.

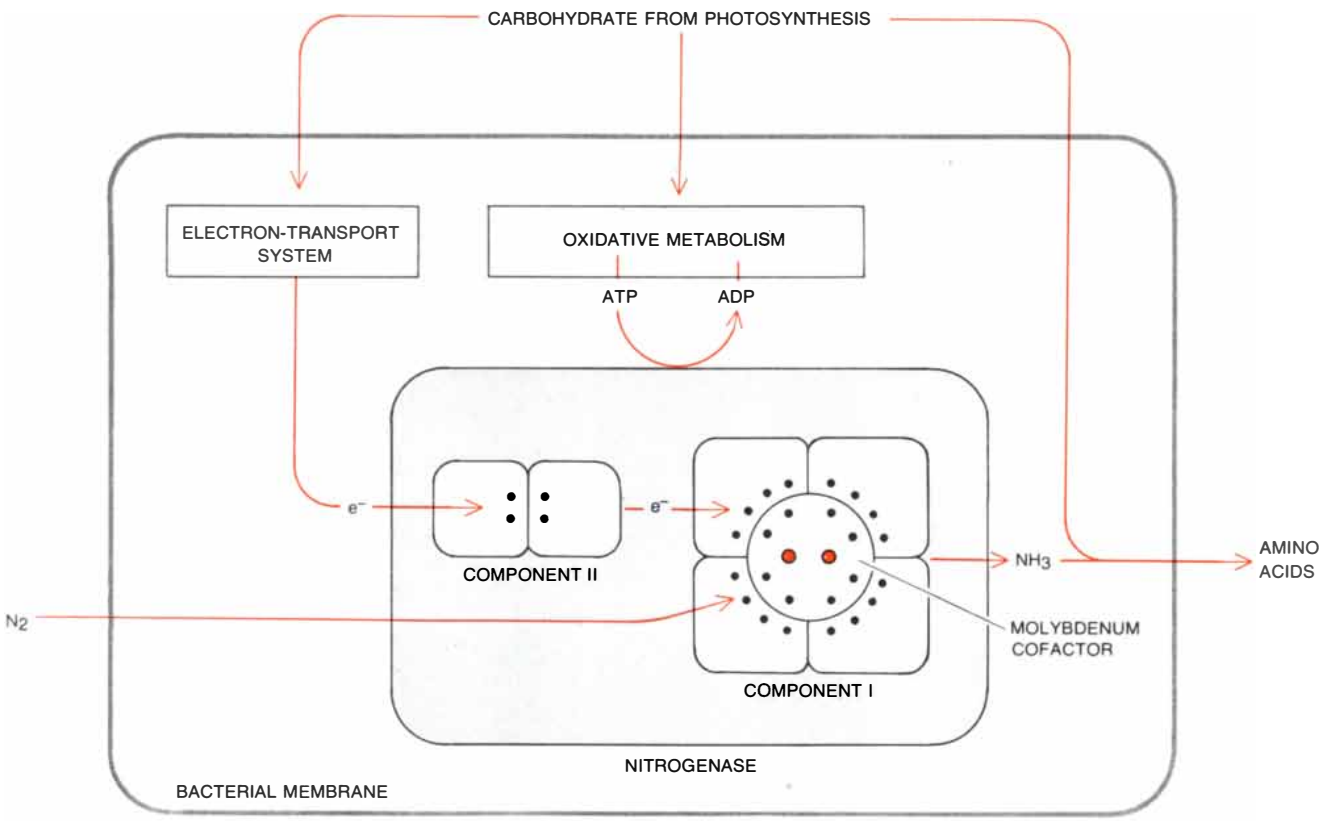
Competing Reactions

One discovery made possible by the development of in vitro nitrogenase systems was the peculiar behavior of the enzyme in the absence of its substrate. When nitrogenase is supplied with ATP but is isolated from nitrogen, hydrogen gas is evolved. Apparently the transport of electrons to the enzyme proceeds normally in the absence of nitrogen; when the electrons reach the active site of the enzyme, they merely recombine with protons. In fact, even in a normal atmosphere some electrons and ATP seem to be wasted in forming hydrogen. Karel Shubert and Harold J. Evans of Oregon State University have determined that most *Rhizobium*-legume associations waste about half of the electrons reaching the nitrogenase. They estimate that the annual U.S. soybean crop produces a volume of hydrogen gas with an energy equivalent to that of 300 billion cubic feet of natural gas. Plainly the efficiency of fixation could be improved if the parasitic production of hydrogen could be eliminated. An encouraging finding is that at least two symbioses—those involving the cowpea and the alder tree—do not lose electrons by forming hydrogen gas. In all likelihood the hydrogen is created at the nitrogenase active site as it is in other organisms but is recycled before it can escape the cell.

An alternative to suppressing hydrogen production is to exploit it. For example, ponds of cultivated blue-green algae might convert solar energy into both fixed nitrogen and free hydrogen. The principal difficulty would be in collecting the hydrogen.

A peculiarity of all nitrogenase systems is that both protein components of the enzyme are denatured by contact with oxygen. The oxygen poisoning is irreversible; the activity of the enzyme cannot be restored, even by removing the oxygen or by adding strong reducing agents. This sensitivity to oxygen is vexing to the biologist studying nitrogenase, since all his experimental apparatus must be designed to exclude oxygen; it would seem to present as great a challenge to the nitrogen-fixing organism. Indeed, those organisms have adopted a variety of strategies for protecting their enzymes.

The problem of oxygen-labile enzymes is a trivial one for the anaerobic nitrogen-fixing bacteria. For these organisms oxygen is a toxic gas in any case, and they live only in those environments, such as the deeper strata of soil, from which oxygen is naturally excluded. The bacterium *Klebsiella pneumoniae* is able to grow with or without oxygen as long as ammonia or nitrate is present. It is able to grow on atmospheric nitrogen, however, only in the absence of oxygen; apparently the bacteri-



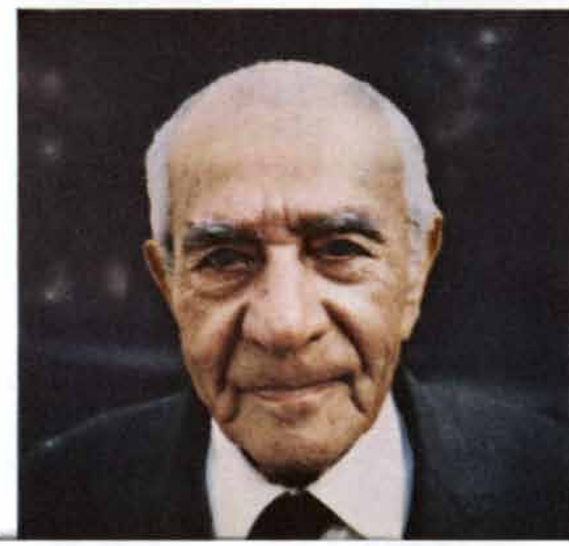
BIOCHEMISTRY OF FIXATION involves the transfer of hydrogen atoms from carbohydrates, such as glucose, to nitrogen. The site of transfer is the enzyme nitrogenase, a complex protein with two main components. The smaller component, Component II, has two subunits and contains a total of four iron atoms. The larger component, Component I, consists of four protein molecules together with 24 iron atoms; it also has a small cofactor containing two molybdenum atoms. Vinod K. Shah in the author's laboratory has recently shown that the cofactor also contains some of the Component I iron atoms. In order to transfer hydrogen atoms active transport is required only for the electrons; protons, or hydrogen nuclei, can be freely deposited in and withdrawn from the aqueous medium of the cell. Electrons derived from carbohydrate are donated first to Component II, then to Component I, where the actual reduction of nitrogen to ammonia takes place. The functioning of both components requires energy, which is supplied in the form of adenosine triphosphate (ATP). Mechanism of ammonia synthesis is apparently the same in all nitrogen-fixing species, although in photosynthetic bacteria the electrons and the ATP are transferred directly from the photosynthetic apparatus to the enzyme nitrogenase without the intermediate stage of carbohydrates.

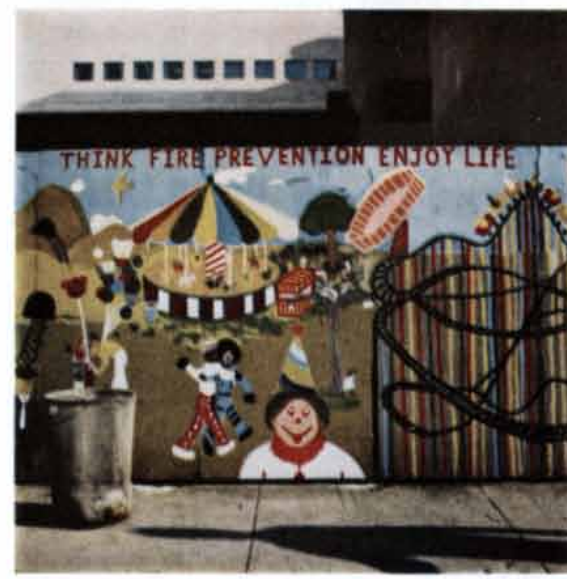
You can tell a lot about an individual by what he pours into his glass.

Bushmills.
The world's oldest whiskey.
What individuals have poured into their glass since 1608.

The "Yachtsman" glass created for the Bushmills Collection by Henry Halem

A blend of 100% Irish Whiskies. 86 Proof. Bottled in Ireland. The Jos. Garneau Co., New York, N.Y. © 1974





The lens of the SX-70 Land camera is a window on unnoticed worlds.

Through its frame you can discover the imprint of Nature on a single leaf. Or the inexplicable beauty of a great machine. Or nobility in a care-worn face.

The SX-70 is automated to an exquisite degree so it never diverts your concentration from your subject. And its instant picture tells you, on the spot, whether you have captured the elusive quality that illuminates the commonplace.

Polaroid Corporation

Everyone should know some body language.

"5 Year/60,000-Mile Dura-Built Engine Guarantee"

Every 1977 Vega comes with the strongest engine guarantee ever offered by Chevrolet.

Vega's Dura-Built guarantee covers every repair needed due to defects in material or workmanship in the cylinder block, cylinder head, intake and exhaust manifolds, water pump, and *all* internal engine parts for five years or up to 60,000 miles.

Should something go wrong with your engine during that time, your Chevy dealer will fix it. For free.

Naturally enough, our Dura-Built guarantee doesn't cover repairs needed because of accident, misuse, or lack of proper maintenance. But if it's our fault, it gets fixed. For five years. Or up to 60,000 miles.

Your Chevy dealer has the complete Dura-Built guarantee statement. Ask him for it. Because it's one of today's Vega's strongest features.

GM
MARK OF EXCELLENCE

"33/24 mpg (EPA)"

With the Dura-Built 140-cubic-inch, four-cylinder engine and standard 4-speed transmission, the 1977 Vega carries EPA estimates of 33 mpg on the highway and 24 mpg in the city.

Obviously though, the EPA figures are only estimates. Your actual mileage may vary depending on the kind of driving you do, your driving habits, the condition of your car and its available equipment. Also, California figures are lower. See your Chevy dealer.

"Zincrometal Protection"

To help fight corrosion, every 1977 Vega is protected with a patented coating system called Zincrometal.*

Unlike galvanizing, the Zincrometal system applies a zinc coating to the inner body surfaces of the car only.

Its advantages are 1) it helps fight corrosion, and 2) it allows a smoother application of our high-quality primers and paints.

All in all, it's yet another of the Vega's strong features for 1977.

(*Zincrometal is a registered trademark of the Diamond Shamrock Corporation.)

"Vegabond"

That's what a lot of Vega owners have become. A Vegabond is a person who enjoys getting away from the humdrum with a certain amount of style. A certain amount of style is standard equipment on your Vega. The getaway part is your responsibility. Take it away, Vegabond.

Chevrolet

Today's Vega

5 year/60,000-mile engine guarantee.

um has no means of protecting its nitrogenase from deactivation.

Many of the blue-green algae that fix nitrogen have specialized, thick-walled cells, called heterocysts, that contain the nitrogenase. Presumably the heterocysts isolate the enzyme from atmospheric oxygen. Bacteria that fix nitrogen in an aerobic habitat have developed still another means of defense. They possess enzymes that reduce oxygen to water as soon as it enters the cell.

The most sophisticated oxygen barrier is the one found in *Rhizobium*-legume symbioses. Oxygen is trapped before it can reach the bacteria by an oxygen-binding protein, leghemoglobin, that is synthesized by plant tissue in the root nodules. It is the only form of hemoglobin found in the plant kingdom, and like the more familiar animal hemoglobins it has the ability to bind oxygen tightly and to give it up on demand. As a result the *Rhizobium* can adopt an efficient, aerobic metabolism while still protecting the nitrogenase from oxygen. This arrangement may be one of the principal benefits to the bacterium of the symbiotic way of life. In a larger context the oxygen lability of nitrogenase, along with the large amount of energy required for its activity, may be responsible for confining nitrogen fixation to a relatively few species.

The large energy requirement of the biological fixation process argues that a parsimonious organism should not fix nitrogen unless it is necessary for growth. That is indeed the behavior observed: if fixed nitrogen is present in the bacterial environment, ammonia production is suppressed. Fertilizer applied to a legume crop reduces the number of root nodules and hence the amount of nitrogen fixed by *Rhizobium*.

Regulation of Nitrogenase

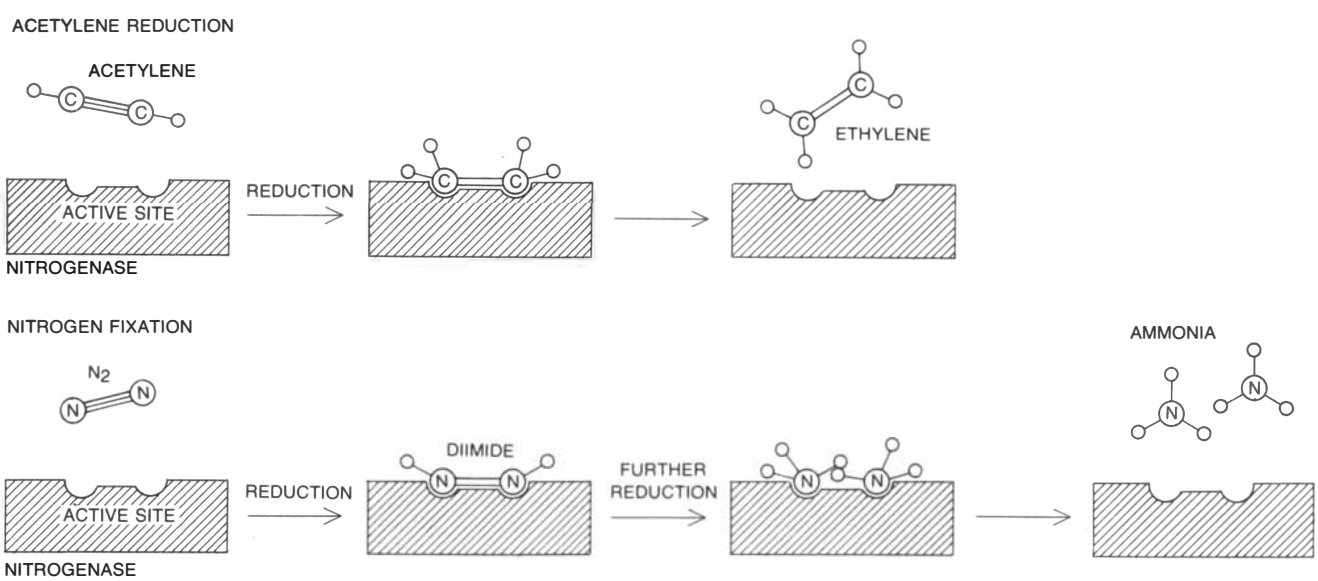
The responsiveness of bacteria to fixed nitrogen in the environment implies that the fixation system is under metabolic control. The basic mechanism of control has been determined: fixed nitrogen suppresses further fixation by halting the synthesis of nitrogenase. As in many other biological systems, control is exercised by the repression of gene expression. The regulation of nitrogenase has been studied in greatest detail in *Klebsiella pneumoniae;* in that bacterium the crucial molecule in the regulatory pathway is an enzyme, glutamine synthetase.

In several species of bacteria glutamine synthetase participates in important aspects of nitrogen metabolism. Its primary role is to catalyze the first step in the synthesis of amino acids. Ammonia, whether it is derived from nitrogen fixation or from some other source, first enters the biochemical pathway by reacting with glutamate, an amino acid, to form another amino acid, glutamine; it is this reaction that is mediated by glutamine synthetase. Most of the other amino acids are then made by transferring the nitrogen from glutamine to other compounds. Boris Magasanik and his colleagues at the Massachusetts Institute of Technology have shown, in bacteria that do not fix nitrogen, that glutamine synthetase also regulates the synthesis of enzymes that degrade certain nitrogenous substrates.

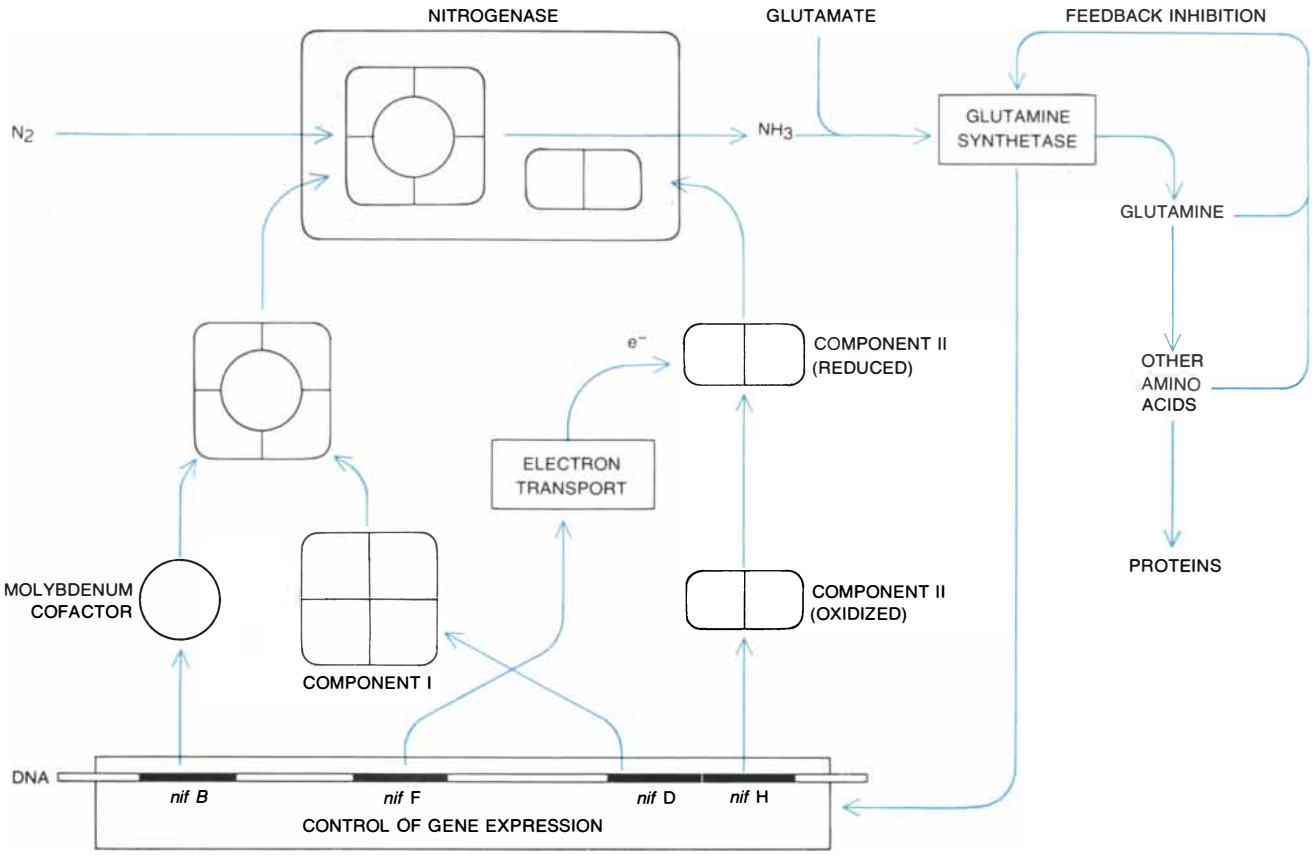
Glutamine synthetase is itself regulated by feedback inhibition from several of the ultimate products of amino acid synthesis. High concentrations of glutamine or of some other amino acids diminish the activity of the enzyme and hence suppress the production of additional amino acids. A plausible interlocking mechanism for the regulation of nitrogen fixation is easily imagined: fixation might be encouraged by the presence of active glutamine synthetase, since that would imply a relative deficiency in the ultimate nitrogen-containing products, the amino acids. The inactivation of glutamine synthetase, on the other hand, would suppress fixation, since the enzyme would be inactivated only when amino acids were abundant.

This hypothesis seemed to be confirmed by studies of mutant strains of *Klebsiella pneumoniae* with defective genes coding for glutamine synthetase. These strains do not synthesize nitrogenase, indicating that glutamine synthetase must play a key role in the regulation of nitrogenase synthesis in this organism. It is not yet understood, however, exactly how the one enzyme controls the synthesis of the other. Moreover, the regulatory mechanism in *Klebsiella pneumoniae* is now known to be even more complex, in that both molybdenum and oxygen have an influence on the synthesis of nitrogenase.

The regulation of nitrogenase synthe-



HYPOTHETICAL SEQUENCE of events in nitrogen fixation assumes that all intermediate states remain bound to nitrogenase. The sequence is based in part on the presumed mechanism of the acetylene-reduction assay, a test of nitrogenase activity. In the absence of nitrogen, nitrogenase converts acetylene to ethylene; a triple bond is reduced to a double bond and two atoms of hydrogen are added. In shape and in the presence of a triple bond molecular nitrogen resembles acetylene, and it is plausible to suppose nitrogenase acts on the two molecules approximately the same way. If that is the case, an intermediate compound in nitrogen fixation would be a molecule called a diimide. When acetylene is the substrate of nitrogenase, the reduction halts after a single step and the product is released as ethylene. The action of the enzyme on nitrogen is obviously different: each nitrogen atom is reduced three times, acquiring three electrons and becoming a molecule of ammonia. Just as this proposed mechanism is hypothetical, so the nitrogenase active site that is shown here is an arbitrary one; little is known about the actual structure of the active site. Neither the diimides nor other intermediates have been detected.



CONTROL OF NITROGEN FIXATION is apparently achieved by regulating the genes, labeled *nif*, that code for nitrogenase. The regulatory molecule is an enzyme, glutamine synthetase. Ammonia made by nitrogen fixation is combined with glutamate, in a reaction catalyzed by glutamine synthetase, to form the amino acid glutamine. Other amino acids are made by modifying glutamine, and high concentrations of several of these have been shown to inhibit glutamine synthetase. It follows that the activity of that enzyme is a measure of the cell's need for fixed nitrogen. In logical confirmation of this finding the enzyme glutamine synthetase seems to regulate the rate of fixation, if only indirectly, by turning the *nif* genes on and off. The mechanism of fixation control, however, is still not fully understood.

sis in other free-living bacteria has been studied through experiments with a substance that is structurally similar to glutamate. This glutamate analogue, methionine sulfoximine, is a powerful inhibitor of enzymes involved in the metabolism of ammonia. A bacterial growth medium that contains ammonia will ordinarily inhibit the synthesis of nitrogenase. Joyce K. Gordon in my laboratory has shown that ammonia does not prevent nitrogen fixation in the presence of methionine sulfoximine. When bacteria are grown with this analogue of glutamate, they accumulate high concentrations of nitrogenase and even excrete ammonia. A similar response has been observed in many kinds of bacteria, including cyanobacteria and other photosynthetic bacteria. This consistent result is evidence that all the cells regulate nitrogenase synthesis through some enzyme involved in the assimilation of ammonia or through the amino acid products of such an enzyme.

Improved Efficiency

Although our understanding of biological nitrogen fixation is still far from perfect, it is not presumptuous to consider improving on nature. Several methods of enhancing fixation might be practical now, and for some of the more radical approaches it is at least apparent what must be learned and achieved.

One of the most obvious methods of increasing the production of biological nitrogen is by improving the efficiency of the *Rhizobium*-legume association. Perhaps the simplest way to achieve that is by genetic screening of both plants and bacteria to determine the combinations best suited for a particular environment. The development of the acetylene-reduction assay makes possible rapid screening of plants for nitrogen-fixing capacity.

Through such a technique Robert Maier, a graduate student in my laboratory, isolated a mutant strain of *Rhizobium japonicum,* the soybean-nodulating bacterium, that fixed more nitrogen in a laboratory assay than did its parent strain. We wondered if the mutant bacteria would also give superior performance in the field. It is often difficult to introduce a new *Rhizobium* strain to a cultivated field because of competition from bacteria already present, and so the bacteria were tested on a plot of land at the University of Hawaii that was free of soybean-nodulating *Rhizobium* species. The soybeans were inoculated with the mutant and with the parent strain; preliminary experiments have achieved greater yields with the mutant. The problem of competition from indigenous bacteria remains, however, if such "superstrains" are to be adopted on a large scale. One possible solution would be to introduce the genes responsible for superior fixation into the strains that are already successful in the field.

The vigor and efficiency of the host plant in legumes also has a strong influence on the amount of nitrogen fixed. Ralph W. F. Hardy and his colleagues at E. I. du Pont de Nemours and Company have found that adding carbon dioxide gas to small plots of soybeans greatly improves the yield of beans, presumably because of greater production of carbohydrate by photosynthesis. Moreover, nitrogen fixation was found to proceed at a higher rate and to continue longer into the life cycle of the plant. It is not practical to flood farmers' fields with carbon dioxide, but the same effect might be achieved by breeding plants with a higher photosynthetic efficiency.

Another improvement in efficiency would result from the selection of *Rhi-*

zobium-legume associations that do not dissipate photosynthetic energy by evolving hydrogen gas. As I have mentioned, two such plants are known: the cowpea and the alder. Perhaps it would be possible to transfer whatever mechanism suppresses hydrogen production in these species to some of the more desirable crop plants.

Rhizobium-legume associations are responsible for about 40 percent of the nitrogen fixed by biological means, and for virtually all the nitrogen fixed by cultivated plants. The best prospects for an immediate improvement in the supply of biological nitrogen are almost certainly to be found among these organisms. There are more than 10,000 species of Leguminosae, of which only about 10 percent have even been examined for nodulation. Fewer than 50 species are cultivated. There may well be other plants in the family that could be exploited for agriculture.

Another approach to increasing the nitrogen supply is through manipulating the biochemical mechanisms that regulate fixation in bacteria. For example, Gordon has isolated mutant strains of *Azotobacter* that continue to fix nitrogen and even excrete ammonia in the presence of nitrogenous fertilizer. Such bacteria could be cultured in ponds on a substrate of cheap organic carbon, such as paper-mill wastes. If ammonia-excreting strains of blue-green algae were available, even the organic carbon could be eliminated, since photosynthesis would supply carbohydrate. The contents of the pond would be harvested as fertilizer.

Ammonia-excreting bacteria might also be adapted to living in the soil near the roots of plants such as wheat or corn. Perhaps an artificial symbiosis could be contrived by selecting plant varieties whose roots exude carbon-rich substances that would nourish the bacteria. The bacteria in turn would enrich the soil in fixed nitrogen.

Altering the bacterial regulatory mechanism could even benefit legumes. If *Rhizobium* species could be made insensitive to the concentration of ammonia, legume crops might supply more fixed nitrogen to the soil, rather than depleting what is already present.

Genetic Manipulation

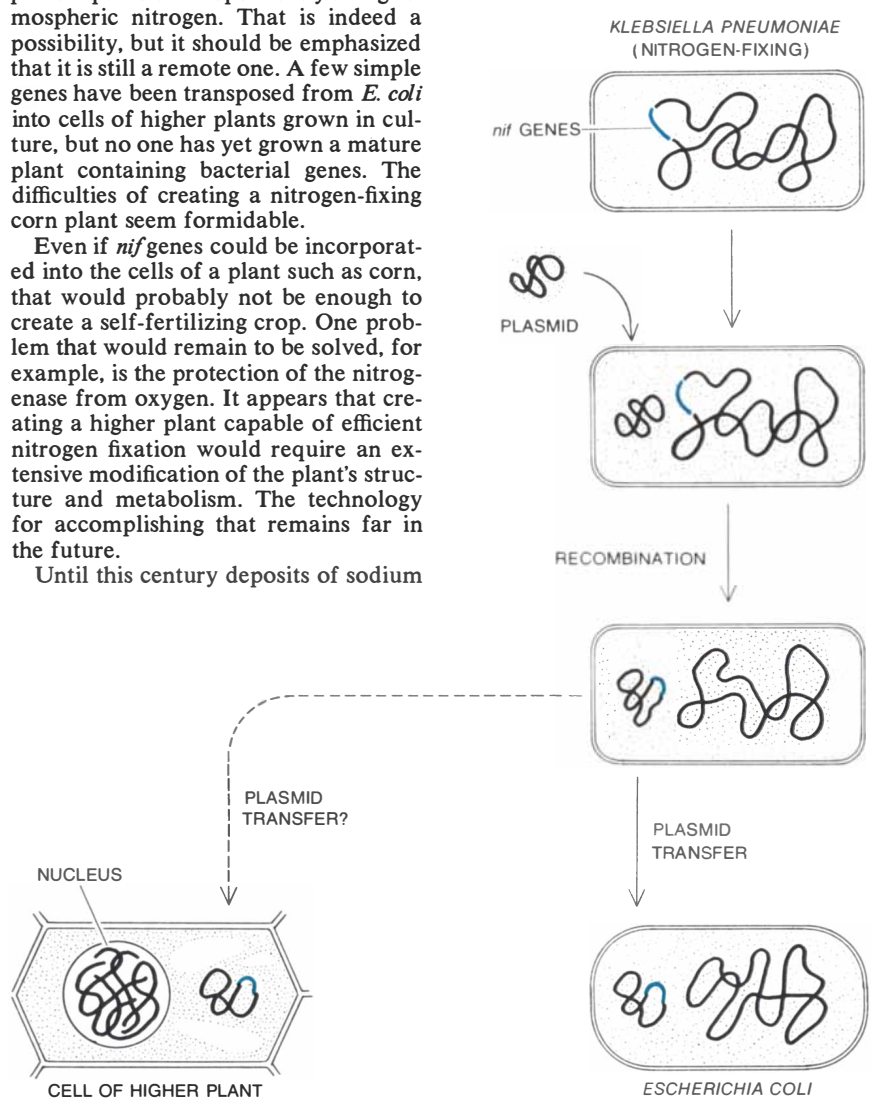
Certainly the most ambitious program for raising nitrogen output is one based on genetic modification, and in particular on the transfer of genes from one organism to another. In at least one bacterium, *Klebsiella pneumoniae,* a substantial number of the genes involved in fixation (the *nif* genes) have already been mapped in the laboratory of Raymond C. Valentine of the University of California at Davis and in my laboratory. They are clustered in a small region of the bacterial chromosome.

Ray A. Dixon and John R. Postgate of the University of Sussex have transferred the cluster of *nif* genes in *Klebsiella pneumoniae* to another bacterium, *Escherichia coli,* a favored pet of biologists and a common inhabitant of the human gut. The transfer was accomplished by first incorporating the *nif* genes in a plasmid, a bit of extrachromosomal DNA, and then introducing the plasmid into *E. coli* cells. *E. coli* is not a nitrogen-fixing bacterium. The new strain created by addition of the plasmid was found to synthesize nitrogenase, and was able to fix nitrogen provided that it was protected from oxygen.

The success of this experiment has justifiably aroused excitement. It suggests the possibility, by a further genetic transfer, of creating a corn or wheat plant capable of independently fixing atmospheric nitrogen. That is indeed a possibility, but it should be emphasized that it is still a remote one. A few simple genes have been transposed from *E. coli* into cells of higher plants grown in culture, but no one has yet grown a mature plant containing bacterial genes. The difficulties of creating a nitrogen-fixing corn plant seem formidable.

Even if *nif* genes could be incorporated into the cells of a plant such as corn, that would probably not be enough to create a self-fertilizing crop. One problem that would remain to be solved, for example, is the protection of the nitrogenase from oxygen. It appears that creating a higher plant capable of efficient nitrogen fixation would require an extensive modification of the plant's structure and metabolism. The technology for accomplishing that remains far in the future.

Until this century deposits of sodium nitrate, principally the nitrate deposits of Chile, were the major source of fixed nitrogen for agriculture and incidentally for munitions. In 1893 Sir William Crookes warned the British Association for the Advancement of Science that the Chilean deposits were approaching depletion. It was the knowledge of impending scarcity that provided the incentive for the development of industrial ammonia manufacture. It now appears that the oil and natural gas required for fertilizer production are also being depleted, and yet another source for agricultural nitrogen must be found. Biological processes, which are already the major contributors to the world nitrogen cycle, are obvious candidates.



TRANSFER OF GENES from a nitrogen-fixing bacterium to some other organism, such as a crop plant, offers the most spectacular means of increasing the world supply of fixed nitrogen. A first stage in such a transfer has already been accomplished: *nif* genes from *Klebsiella pneumoniae* were incorporated into a plasmid, a bit of extrachromosomal DNA, and implanted in *Escherichia coli,* a bacterium that has no nitrogenase. A similar transfer into the cells of higher plants, however, would be far more difficult. Moreover, the possession of *nif* genes would not by itself ensure that a plant would fix nitrogen. The *nif*-containing *E. coli,* for example, are able to manufacture the enzyme nitrogenase but do not fix nitrogen because the enzyme cannot be protected from oxygen. A nitrogen-fixing cereal plant thus seems a remote possibility.

The Acoustics of the Singing Voice

The voice organ is an instrument consisting of a power supply (the lungs), an oscillator (the vocal folds) and a resonator (the larynx, pharynx and mouth). Singers adjust the resonator in special ways

by Johan Sundberg

Clearly there is something quite unusual about the voice of a first-class opera singer. Quite apart from the music, the intrinsic quality of such a voice can have a forceful impact on the listener. Moreover, a well-trained singer produces sounds that can be heard distinctly in a large opera house even over a high level of sound from the orchestra, and can do so week after week, year after year. If a second-rate singer or a completely untrained one tried to be heard over an orchestra, the result would be a scream and the singer's voice would soon fail. Is it only training that makes the difference? Or is the instrument that produces an excellent singer's voice itself different from other people's?

Let us begin with a description of that instrument. The voice organ includes the lungs, the larynx, the pharynx, the nose and the mouth. The main voice function of the lungs is to produce an excess of air pressure, thereby generating an airstream. The air passes through the glottis, a space at the base of the larynx between the two vocal folds (which are often called the vocal cords but are actually elastic infoldings of the mucous membrane lining the larynx). The front end of each vocal fold is attached to the thyroid cartilage, or Adam's apple. The back end of each is attached to one of the two small arytenoid cartilages, which are mobile, moving to separate the folds (for breathing), to bring them together and to stretch them. The vocal folds have a function apart from that of producing sound: they protect the lungs from any small objects entrained in the inspired airstream. Just above the vocal folds are the two "false" vocal folds, which are engaged when someone holds his breath with an overpressure of air in the lungs. The vocal folds are at the bottom of the tube-shaped larynx, which fits into the pharynx, the wider cavity that leads from the mouth to the esophagus. The roof of the pharynx is the velum, or soft palate, which in turn is the door to the nasal cavity. When the velum is in its raised position (which is to say during the sounding of all vowels except the nasalized ones), the passage to the nose is closed and air moves out through the mouth.

The larynx, the pharynx and the mouth together constitute the vocal tract, a resonant chamber something like the tube of a horn or the body of a violin. The shape of the tract is determined by the positions of the articulators: the lips, the jaw, the tongue and the larynx. Movements of the lips, jaw and tongue constrict or dilate the vocal tract at certain sites; protruding the lips or lowering the larynx increases the length of the tract.

Now consider the voice organ as a generator of voiced sounds. Functionally the organ has three major units: a power supply (the lungs), an oscillator (the vocal folds) and a resonator (the vocal tract). With the glottis closed and an airstream issuing from the lungs, the excess pressure below the glottis forces the vocal folds apart; the air passing between the folds generates a Bernoulli force that, along with the mechanical properties of the folds, almost immediately closes the glottis. The pressure differential builds up again, forcing the vocal folds apart again. The cycle of opening and closing, in which the vocal folds act somewhat like the vibrating lips of a brass-instrument player, feeds a train of air pulses into the vocal tract. The frequency of the vibration is determined by the air pressure in the lungs and by the vocal folds' mechanical properties, which are regulated by a large number of laryngeal muscles. In general the higher the lung pressure is and the thinner and more stretched the vocal folds are, the higher is the frequency at which the folds vibrate and emit air pulses. The train of pulses produces a rapidly oscillating air pressure in the vocal tract; in other words, a sound. Its pitch is a manifestation of the vibratory frequency. Most singers need to develop full control over a pitch range of two octaves or more, whereas for ordinary speech less than one octave suffices.

The sound generated by the airstream chopped by the vibrating vocal folds is called the voice source. It is in effect the raw material for speech or song. It is a complex tone composed of a fundamental frequency (determined by the vibratory frequency of the vocal folds) and a large number of higher harmonic partials, or overtones. The amplitude of the partials decreases uniformly with frequency at the rate of about 12 decibels per octave. The "source spectrum," or plot of amplitude against frequency, for a singer is not very different from that for a nonsinger, although the spectrum does tend to slope more steeply in soft speech than it does in soft singing.

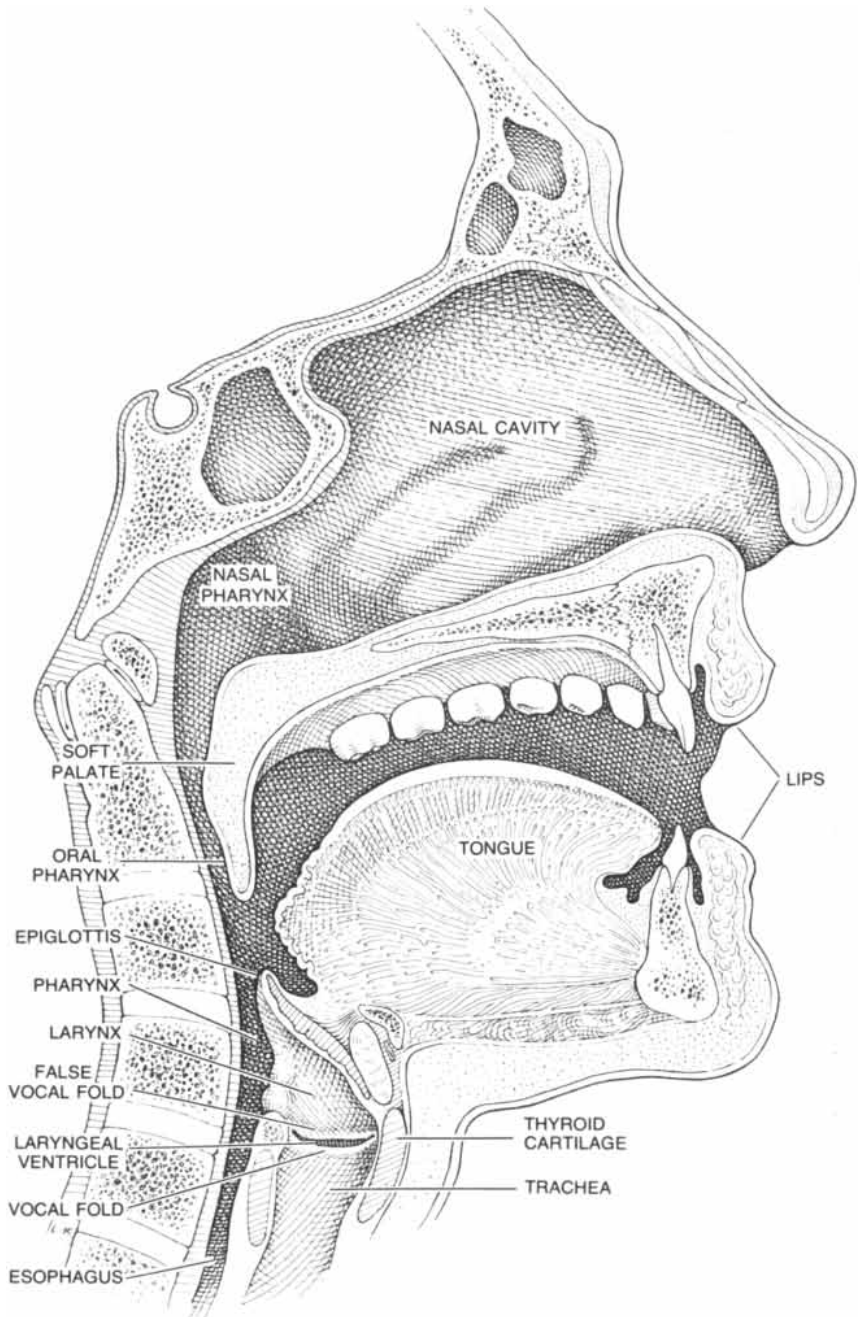
The vocal tract is a resonator, and the transmission of sound through an acoustic resonator is highly dependent on frequency. Sounds of the resonance frequencies peculiar to each resonator are less attenuated than other sounds and are therefore radiated with a higher relative amplitude, or with a greater relative loudness, than other sounds; the larger the frequency distance between a sound and a resonance is, the more weakly the sound is radiated. The vocal tract has four or five important resonances called formants. The many voice-source partials fed into the vocal tract traverse it with varying success depending on their frequency; the closer a partial is to a formant frequency, the more its amplitude at the lip opening is increased. The presence of the formants disrupts the uniformly sloping envelope of the voice-source spectrum, imposing peaks at the formant frequencies. It is this perturbation of the voice-source envelope that produces distinguishable speech sounds: particular formant fre-

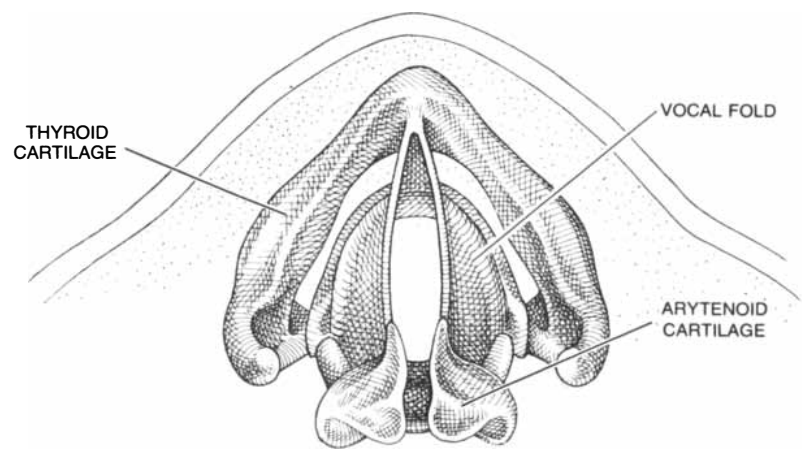
quencies manifest themselves in the radiated spectrum as peaks in the envelope, and those peaks are characteristic of particular sounds.

The formant frequencies are determined by the shape of the vocal tract. If the vocal tract were a perfect cylinder closed at the glottis and open at the lips and 17.5 centimeters (about seven inches) long, which is about right for the average adult male, then the first four formants would be close to 500, 1,500, 2,500 and 3,500 hertz (cycles per second). Given a longer or shorter vocal tract, these basic frequencies are somewhat lower or higher. Each formant is associated with a standing wave, that is, with a static pattern of pressure oscillations whose amplitude is at a maximum at the glottal end and near a minimum at the lip opening [*see illustration on page 86*]. The lowest formant corresponds to a quarter of a wavelength, which is to say that a quarter of its wavelength fits within the vocal tract. Similarly, the second, third and fourth formants correspond respectively to three-quarters of a wavelength, one and a quarter wavelengths and one and three-quarters wavelengths.

Any change in the cross section of the vocal tract shifts the individual formant frequencies, the direction of the shift depending on just where the change in area falls along the standing wave. For example, constriction of the vocal tract at a place where the standing wave of a formant exhibits minimum-amplitude pressure oscillations generally causes the formant to drop in frequency; expansion of the tract at those same places raises the frequency.

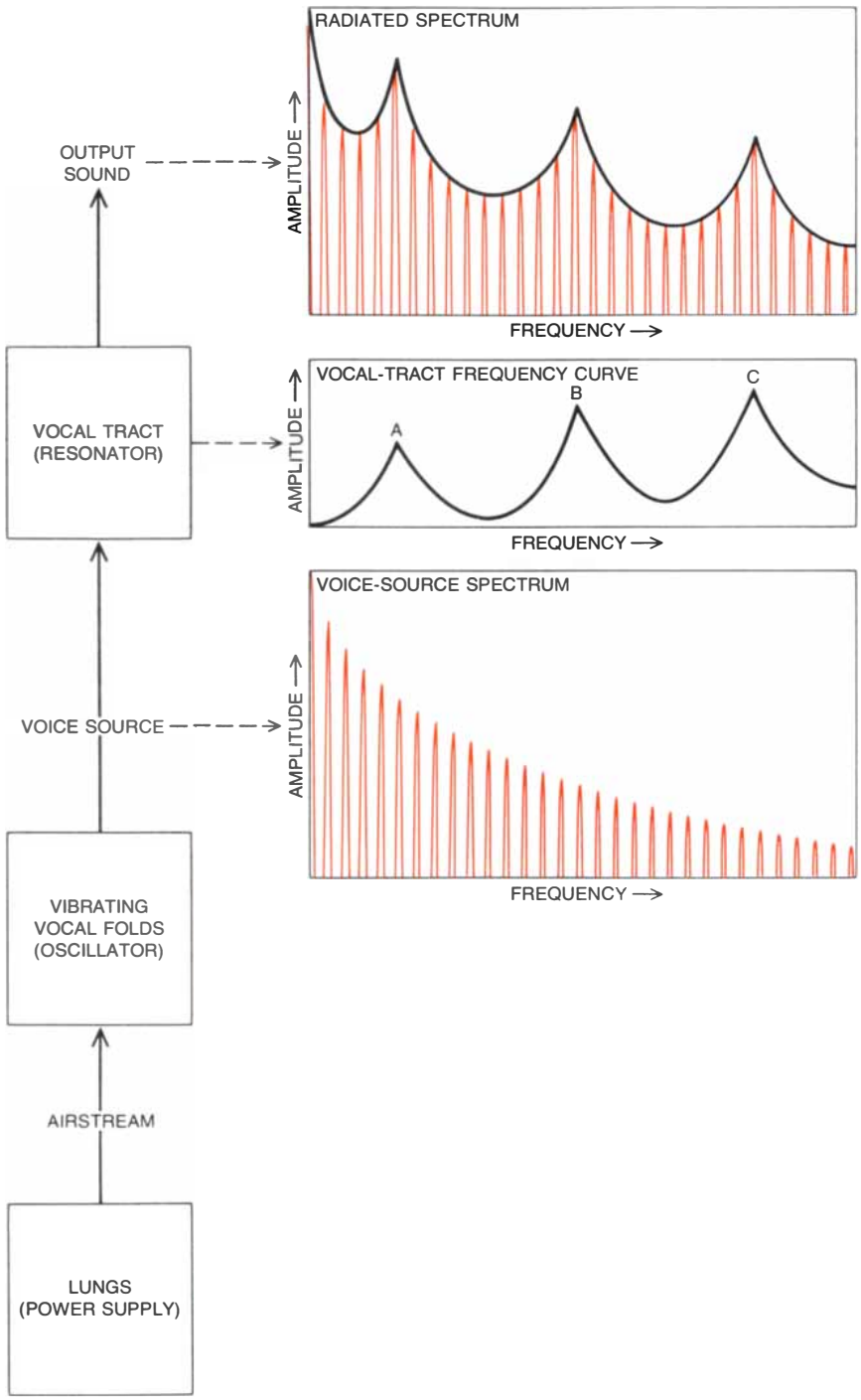
The vocal tract is constricted and expanded in many rather complicated ways, and constricting it in one place affects the frequency of all formants in different ways. There are, however, three major tools for changing the shape of the tract in such a way that the frequency of a particular formant is shifted in a particular direction. These tools are the jaw, the body of the tongue and





VOICE ORGAN is composed of the lungs and the larynx, pharynx, mouth and nose, shown in longitudinal section (*top*). The larynx is a short tube at the base of which are twin infoldings of mucous membrane, the vocal folds. The larynx opens into the pharynx; the opening is protected during swallowing by the epiglottis. The larynx, pharynx and mouth (and in nasal sounds also the nose) constitute the vocal tract. It is a resonator whose shape, which determines vowel sounds, is modified by changes in the position of the articulators: the lips, the jaw, the tip and body of the tongue and the larynx. The vocal folds, seen from above in a transverse section (*bottom*), are opened for breathing and are closed for phonation by the pivoting arytenoid cartilages.

the tip of the tongue. The jaw opening, which can constrict the tract toward the glottal end and expand it toward the lip end, is decisive in particular for the frequency of the first formant, which rises as the jaw is opened wider. The second-formant frequency is particularly sensitive to the shape of the body of the tongue, the third-formant frequency to the position of the tip of the tongue. Moving the various articulatory organs in different ways changes the frequencies of the two lowest formants over a considerable range, which in adult males averages approximately from 250 to 700 hertz for the first formant and from 700 to 2,500 hertz for the second. Moving the articulatory organs is what we do when we speak and sing; in effect we chew the standing waves of our formants to change their frequencies. Each articulatory configuration corresponds to a set of formant frequencies, which in turn is associated with a particular vowel sound. More specifically, the formant frequencies enhance voice-source partials of certain frequencies and thus manifest themselves as the peaks characterizing the spectrum envelope of each vowel sound.



VOICE ORGAN is composed functionally of a power supply, an oscillator and a resonator. The airstream from the lungs is periodically interrupted by the vibrating vocal folds. The resulting sound, the voice source, has a spectrum (*right*) containing a large number of harmonic partials, the amplitude of which decreases uniformly with frequency. The air column within the vocal tract has characteristic modes of vibration, or resonances, called formants (*A, B, C*). As the voice source moves through the vocal tract each partial is attenuated in proportion to its distance from formant nearest it in frequency. The formant frequencies thus appear as peaks in the spectrum of the sound radiated from the lips; the peaks establish particular vowel sounds.

All the elements and functions of the voice organ that I have been describing are common to singers and nonsingers alike. Do singers bring still other faculties into play or manipulate the voice instrument in different ways? Let us begin by comparing normal male speech and operatic singing. Careful attention to a singer's voice reveals a number of modest but very characteristic deviations in vowel quality from those of ordinary speech. For example, the *ee* sound of a word such as "beat" is shifted toward the umlauted *ü* of the German "*für*"; the short *e* of "head" moves toward the vowel sound of "heard." The general impression is that the quality of the voice is "darker" in singing, somewhat as it is when a person yawns and speaks at the same time; voice teachers sometimes describe the effect as "covering."

These shifts in vowel quality have been found to be associated with peculiarities of articulation. In "covered" singing the larynx is lowered, and X-ray pictures reveal that the change in the position of the larynx is accompanied by an expansion of the lowest part of the pharynx and of the laryngeal ventricle, the space between the true vocal folds and the false ones. It is interesting to note that voice teachers tend to agree that the pharynx should be widened in singing, and some of them mention the sensation of yawning. In other words, a low larynx position and an expanded pharynx are considered desirable in singing.

What we recognize as a darkened voice quality in singing is reflected very clearly in the spectrum of a sung vowel sound. A comparison of the spectra of the vowel in "who'd" as it is spoken and sung shows that the two lowest formant frequencies are somewhat lower in the sung version and that the spectral energy, or amplitude, is considerably higher between 2,500 and 3,000 hertz [*see top illustration on page 89*]. This spectral-envelope peak is typical of all voiced sounds sung by professional male singers. Indeed, its presence, regardless of the pitch, the particular vowel and the dynamic level, has come to be consid-

New
Vivitar Series 1 90mm f2.5 macro lens

The Vivitar Series 1 90mm f2.5 macro lens may well be the sharpest lens in 35mm photography. The lens utilizes the floating group concept to maintain optimum performance throughout its focusing range. Tests for resolution and contrast from infinity to life-size (1:1 reproduction) give the Series 1 90mm f2.5 lens some of the highest overall axial to corner ratings obtained for macro lenses.

A 90mm macro lens yielding good performance could have been produced using a normal double Gauss design. To achieve and maintain very high levels of performance from infinity to life-size, however, Vivitar Series 1 designers used a unique 8 element/7 group configuration to bring aberrations to an absolute minimum and to stabilize them throughout focusing distances from a reproduction ratio of 1:2 to infinity. The extremely stringent performance demands to eliminate aberrations in the lens also required the use of optical glass of a very high index of refraction and some uncommonly thick elements.

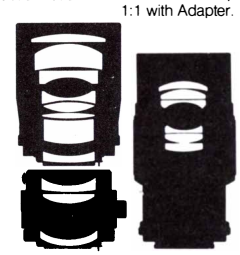
Using the concept of a null lens, borrowed from astronomical optics, the designers created a 3 element macro corrector-lens adapter that achieves a true flat-field image, high resolution and excellent contrast in the 1:2 to 1:1 reproduction range. The macro corrector-lens adapter is not a magnifying lens. Its sole function is to compensate aberrations produced when the lens is moved away from the film plane for life-size photography.

The selection of 90mm as the focal length of this lens provides two distinct benefits to photographers. It is an ideal focal length for portraiture and general purpose photography. When used with its macro adapter the lens allows life-size photography at a greater working distance from the subject than shorter focal length macro lenses. This greater working distance increases the photographer's options in illuminating macro subjects and lessens chances of disturbing live subjects.

As with all Vivitar Series 1 lenses, the mechanical configuration has been as carefully engineered and manufactured as the optics. The lens engravings give the photographer maximum information and legibility. The lens barrel styling is entirely functional, all controls being placed in the most appropriate positions for precise, comfortable operation.

Optical Specifications

Elements/Groups:	
Main lens:	8 elements, 7 groups.
Macro Adapter:	3 elements, 3 groups.
Lens coating:	VMC Vivitar multicoating.
Angle of acceptance:	27°
Aperture range:	f2.5 to f22
Minimum focus distance from film plane:	39.3cm (15.5 in.) without Adapter, 35.5cm (14 in.) with Adapter.
Maximum reproduction ratio:	1:2 without Adapter, 1:1 with Adapter.



Vivitar Series 1 design *conventional design*

Mechanical Specifications

Length at infinity:	90mm (3.5 in.) without Adapter, 138mm (5.4 in.) with Adapter.
Maximum barrel diameter:	70mm (2.8 in.)
Weight	
Main lens:	644 gms. (23 oz.)
with Adapter:	936 gms. (33 oz.)
Filter size:	58mm
Lens case:	Semi-hard, 2 compartment case.

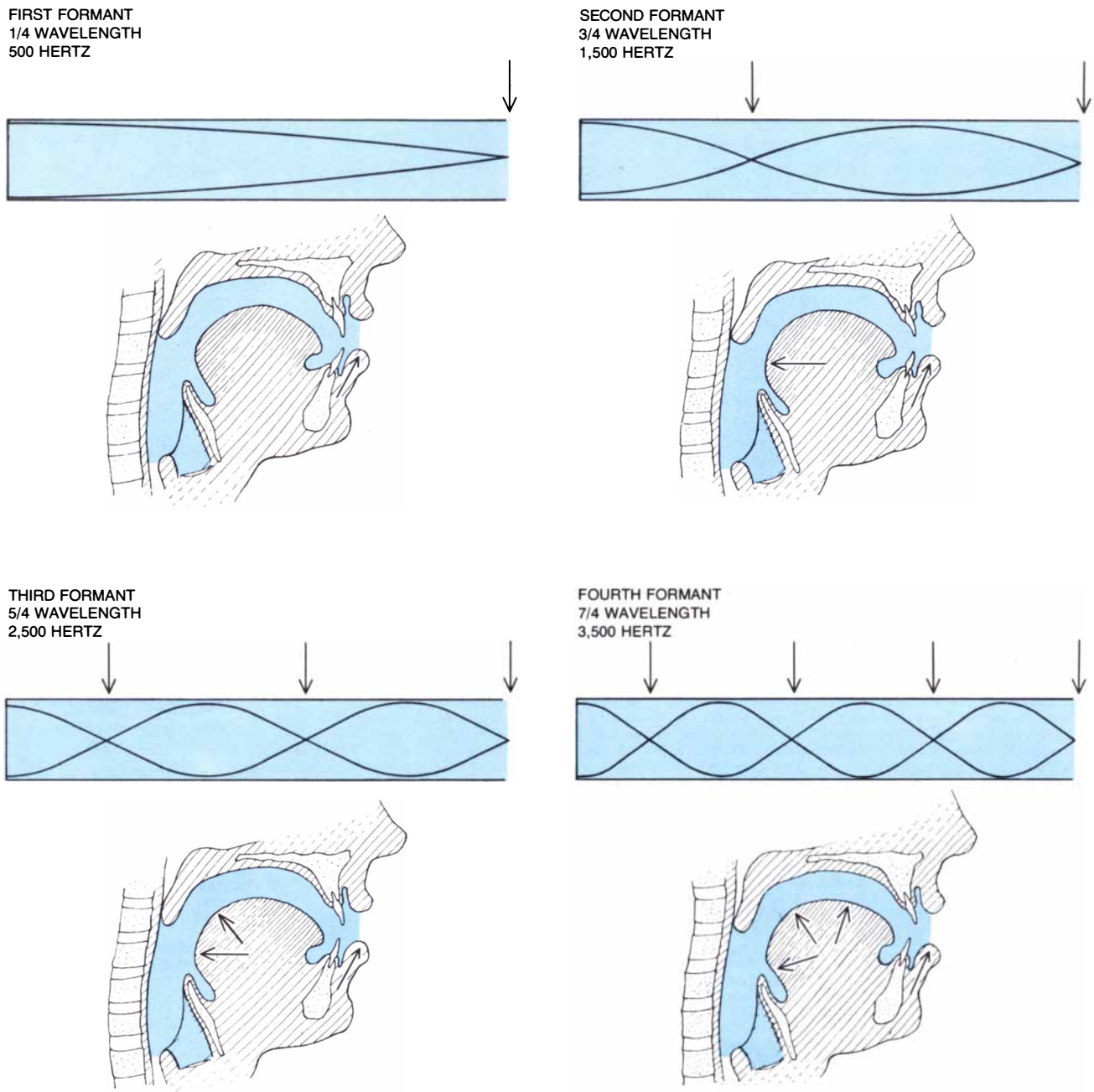
Available in mounts to fit Nikon, Canon, Minolta, Olympus OM and Universal Thread Mount Cameras.

Vivitar®

Marketed in the U.S.A. by
Ponder & Best, Inc.
Corporate Offices: 1630 Stewart St.,
Santa Monica, CA 90406.
In Canada:
Vivitar Canada, Ltd./Ltée., Quebec.

ered a criterion of quality; the extra peak has been designated the "singing formant."

What is the origin of the singing-formant peak? The peaks in the spectrum envelope of a vowel normally stem, as I have explained, from the presence of specific formants. The insertion of an extra formant between the normal third and fourth formants would produce the kind of peak that is seen in the spectrum of a sung vowel [*see bottom illustration on page 89*]. Moreover, the acoustics of the vocal tract when the larynx is lowered are compatible with the generation of just such an extra formant. It can be calculated that if the area of the outlet of the larynx into the pharynx is less than a sixth of the area of the cross section of the pharynx, then the larynx is acoustically mismatched with the rest of the vocal tract; it has a resonance frequency of its own, largely independent of the remainder of the tract. The one-sixth condition is likely to be met when the larynx is lowered, because the lowering tends to expand the bottom part of the pharynx. I have estimated on the basis of X-ray pictures of a lowered larynx that this lowered-larynx resonance frequency should be between 2,500 and 3,000 hertz, that is, between the frequencies of the normal third and fourth



FORMANTS correspond to standing waves, or static patterns of air-pressure oscillations, in the vocal tract. Here the first four formants are shown as standing waves in cylindrical tubes, the schematic equivalent of the vocal tract (*colored areas in drawings*). The sine waves represent the amplitude of the pressure differential, which is always maximal at the glottal end and minimal at the lips. For the lowest formant a quarter of a wavelength is within the vocal tract and, if the tract is 17.5 centimeters long, the formant's frequency is about 500 hertz (cycles per second). The second, third and fourth formants are 3/4, 5/4 and 7/4 of a wavelength, and their frequencies vary accordingly. If the area of the vocal tract is decreased or increased at a place where the formant's pressure amplitude is at a minimum (*arrows*), that formant's frequency is respectively lowered or raised; the same change in area has the opposite effect if it is at a pressure maximum.

Noise—clamorous companion of man's progress—is becoming a significant environmental problem.

At the General Motors Research Laboratories we are seeking to reduce noise at its source through increased understanding of the mechanisms of noise generation. Simultaneously, we are studying human responses to traffic sounds.

Considerable effort has been focused on tires, a major source of noise in both cars and trucks. Interestingly, air flow around the tire is not a significant noise source. But aerodynamic pumping between treads can be, depending upon tread pattern. Also important are tread vibrations in the vicinity of the contact patch.

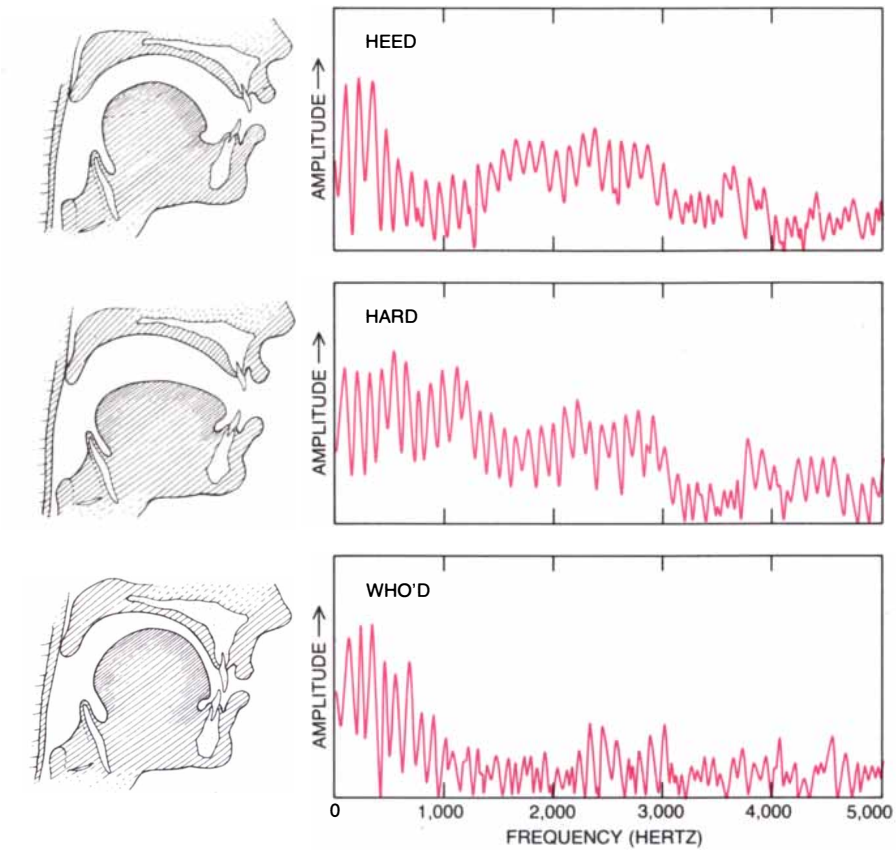
In another study, we are using signal coherence analysis to relate cylinder combustion pressure to noise radiation. This is part of an overall effort to learn in detail how engine structures transmit combustion-related noise.

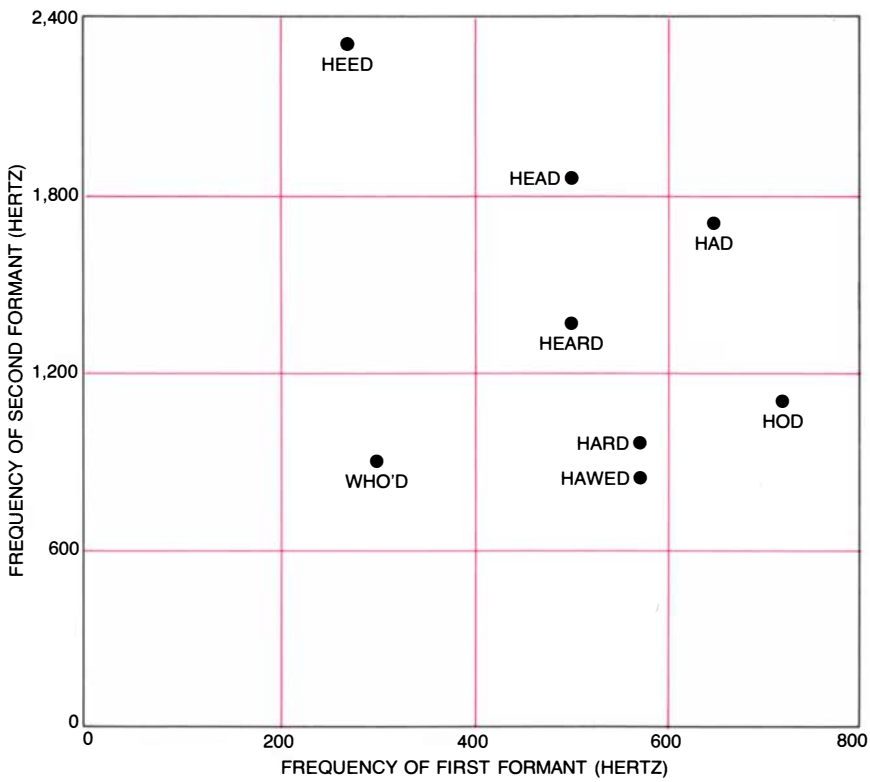
Is it possible to quantify the annoyance associated with traffic noise? Psychological studies in one suburban area established that an L_{eq} measurement (average amount of sound energy reaching the ear per unit time) of 60 dB was the approximate threshold above which people were willing to pay for decreased annoyance.

These research programs—and others being conducted at the Laboratories—are aimed at restoring one of life's more precious qualities . . . a quiet environment.

We're working on several hush-hush projects.

General Motors Research Laboratories
Warren, Michigan 48090





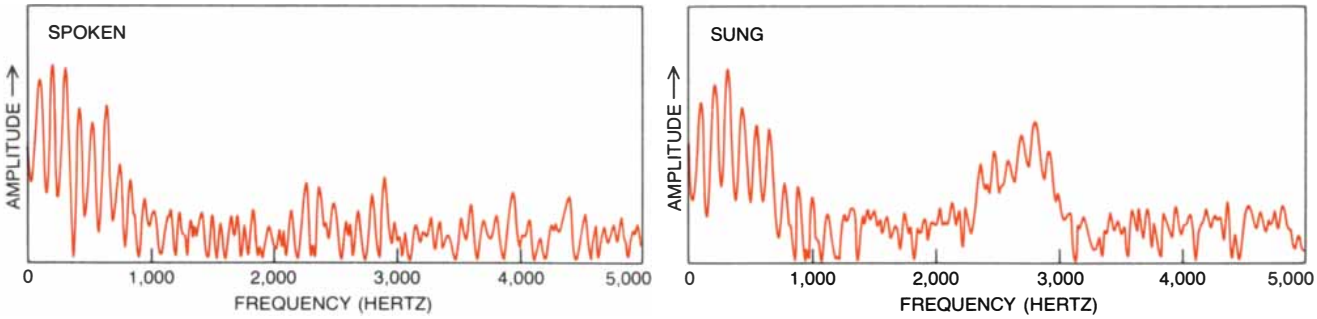
MOVEMENT OF ARTICULATORS changes the cross section of the vocal tract, shifting formant frequencies. Three articulatory configurations are shown (*top*) together with the spectrum of the vowel sound produced by each; the peaks in the spectrum envelope reflect the formant frequencies. The chart (*bottom*) gives the frequencies of the first and second formants in some English vowel sounds as spoken by an average male. For a female or a child the envelope pattern would be about the same but the peaks would be shifted somewhat higher in frequency.

formants and just where the singing-formant peak appears. The lowering of the larynx, in other words, seems to explain the singing-formant peak.

It also accounts for something else. Acoustically the expansion of the lowermost part of the pharynx is equivalent to an increase in the length of the vocal tract, and the lowering of the larynx adds still more to the length. The result is to shift downward all formant frequencies other than the larynx-dependent extra formant. This lowering of frequency is particularly notable in formants that depend primarily on the length of the pharynx. Two examples of such formants are the second formant of the vowels in "beat" and "head," and a drop in the frequency of those formants moves their vowels respectively toward those of *"fur"* and "heard." The lowering of the larynx, then, explains not only the singing-formant peak but also major differences in the quality of vowels in speech and in singing.

To explain the singing formant's articulatory and acoustic origin is not enough, however. Why, one wonders, is it desirable for singers to lower the larynx, producing the singing formant and darkening the quality of their vowels? A plausible answer to the question has been found. It is related to the acoustic environment in which opera and concert singers have to work: in competition with an orchestra. Analysis of the average distribution of energy in the sounds of an opera or symphony orchestra shows that the highest level of sound is in the vicinity of 450 hertz; above that the amplitude decreases sharply with frequency. Now, normal speech develops maximum average energy at about the same frequency and weakens at higher frequencies. A singer who produced sounds with the energy distribution of ordinary speech would therefore be in trouble: the orchestra's much stronger sounds would drown out the singer's. The average sound distribution of a trained singer, on the other hand, differs from that of normal speech—and of an orchestra—mainly because of the singing-formant effect. We have shown that a singer's voice is heard much more easily against recorded noise that has the same average energy distribution as an orchestra's sound if the voice has a singing formant. Not only is the formant almost invariably audible, because its frequency is in a region where the orchestra's sound is rather weak, but also it may help the listener to "imagine" he hears other parts of the singer's spectrum that are in fact drowned out by the orchestra.

The singing formant is at an optimal frequency, high enough to be in the region of declining orchestral-sound energy but not so high as to be beyond the range in which the singer can exercise

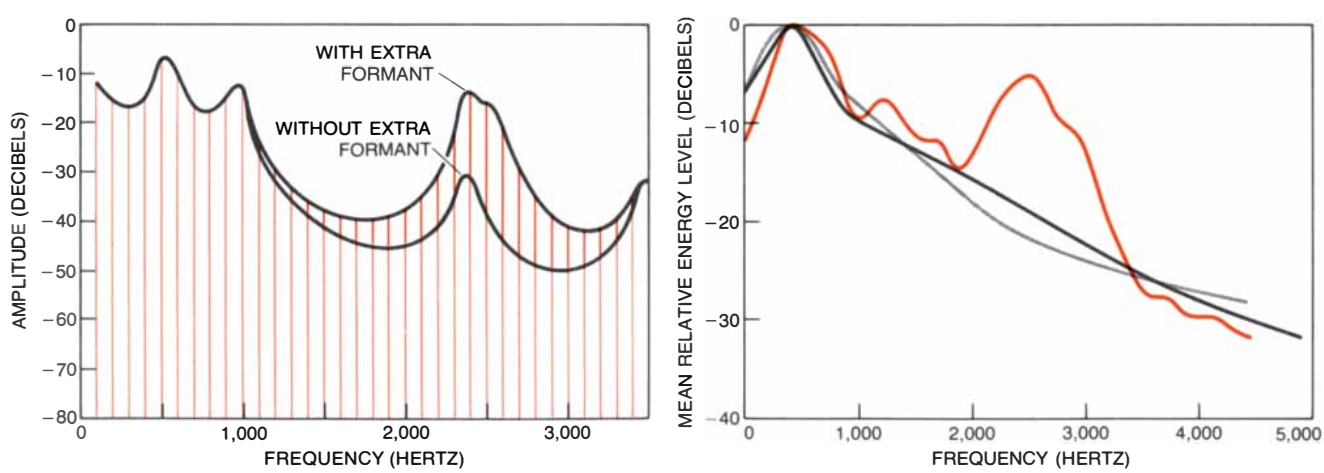


VOWELS SOUND DIFFERENT in speech and in singing and the difference is visible in their recorded spectra. Here the spectra of the vowel in "who'd" as spoken (*left*) and as sung (*right*) by a male opera singer are compared. What is significantly different about the sung spectrum is the spectral-energy peak that appears in it between about 2,500 and 3,000 hertz. The new peak is called the singing formant.

good control. Because it is generated by resonance effects alone, it calls for no extra vocal effort; the singer achieves audibility without having to generate extra air pressure. The singer does pay a price, however, since the darkened vowel sounds deviate considerably from what one hears in ordinary speech. In some kinds of singing that price is too high: the ideas and moods expressed in a "pop" singer's repertoire, for example, would probably not survive the deviations from naturalness that are required to generate the singing formant. And pop singers do not in fact darken their vowels; they depend on electronic amplification to be heard.

In cartoons a female opera singer is almost invariably depicted as a fat woman with her mouth opened very wide. In a study of female singers I have found that the way in which the jaw is manipulated is in fact quite different in ordinary speech and in singing. In speech the size of the jaw opening varies with the particular vowel, but in female singing it tends to depend also on the pitch of the tone that is being sung: the higher a soprano sings, the wider her jaw is opened. This suggested to me that a soprano must vary the frequency of her first formant according to the pitch at which she is singing. Analysis of formant frequencies confirmed that the articulation was being varied in such a way as to raise the first-formant frequency close to the frequency of the fundamental of the tone being sung. I noted such a frequency match whenever the frequency of the fundamental was higher than the frequency of a vowel's first formant in ordinary speech.

The reason becomes clear when one considers that the pitch frequency of a soprano's tones is often much higher than the normal frequency of the first formant in most vowels. If a soprano sang the vowel *ee* at the pitch of her middle C and with the articulation of ordinary speech, her first formant would be in the neighborhood of 270 hertz and the pitch frequency (the frequency of her lowest spectrum partial) would be almost an octave higher, at 523 hertz. Since a sound is attenuated in proportion to the distance of its frequency from a formant frequency, the fundamental would suffer a serious loss of amplitude. The fundamental is the strongest partial in the voice-source spectrum, and the higher its pitch is, the more important the fundamental is for the loudness of the tone, and so the singer's *ee* would be rather faint. Assume that her next sound was the *ah* sound of "father," to be sung at the pitch of high *F*. The fundamental, at 698 hertz, would be very close to the frequency of the first formant, about 700 hertz, and so the tone would be loud. The loudness of the singer's tones would vary, in other words, according to a rather unmusical determinant: the frequency distance between first formant and fundamental. In order to modulate the loudness accord-



SINGING FORMANT'S ORIGIN (*left*) and its utility in singing (*right*) are demonstrated. An extra formant was inserted between the usual third and fourth formants in an experiment with an electronic resonator that behaves like the vocal tract (*left*). The new formant increased the amplitude of the partials near it by more than 20 decibels; similarly, an extra formant (achieved by lowering the larynx) supplies the high-frequency peak in the spectrum of a sung vowel. The three curves (*right*) show the averaged distribution of energy in the sound of orchestral music (*black*), of ordinary speech (*gray*) and of the late tenor Jussi Björling singing with an orchestra (*colored*). The distribution is very similar for speech and the orchestra at all frequencies; it is the singer's voice that produces the peak in the colored curve between 2,000 and 3,000 hertz. In that frequency region a singer's voice is loud enough, compared with an orchestra's sound, to be discerned.

ing to the musical context, the singer would need to continually vary her vocal effort. That would strain her vocal folds. (Experiments with synthesized vowel sounds suggest that it would also produce tones more characteristic of a mouse under severe stress than of an opera singer!)

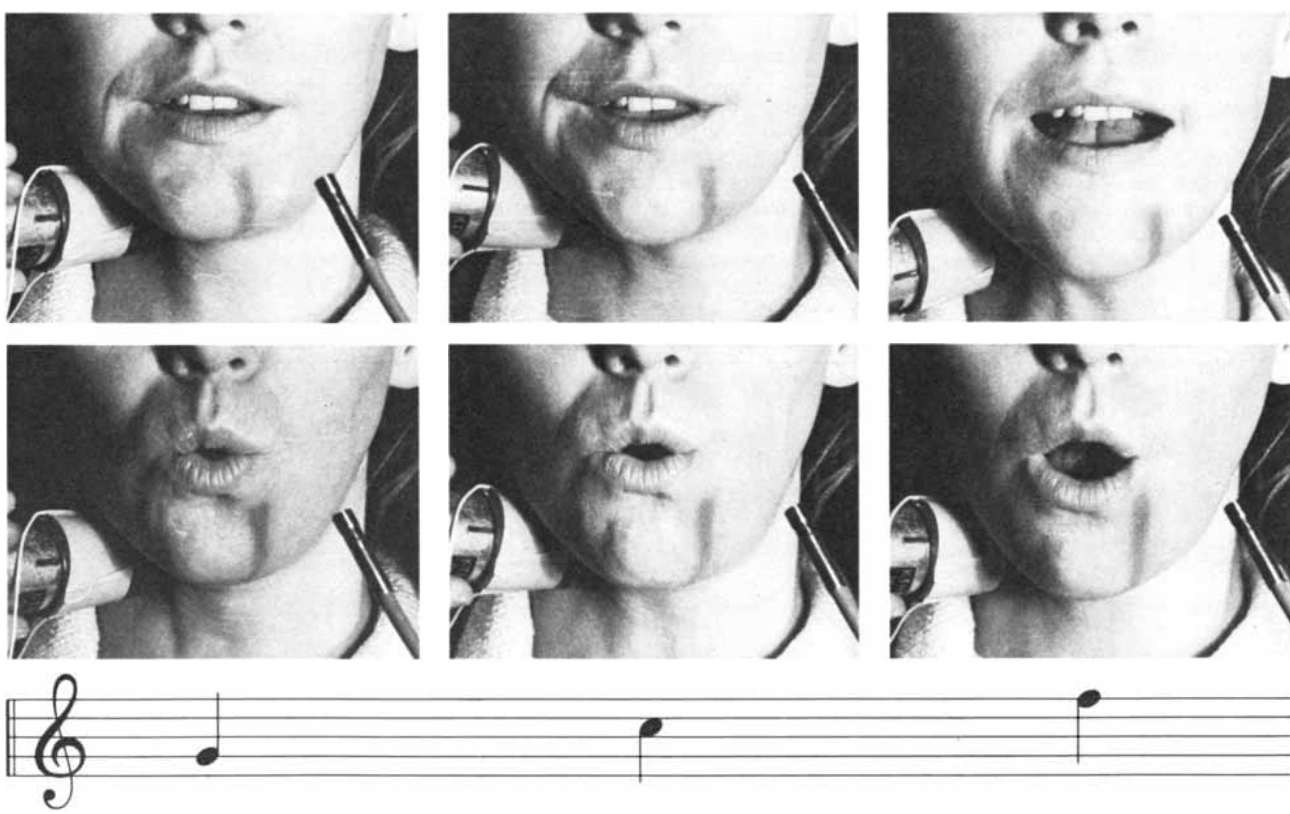
The soprano's solution is to move the first formant up in frequency to match the frequency of the fundamental, thus allowing the formant always to enhance the amplitude of the fundamental. The result is that there is minimal variation in loudness from pitch to pitch and from vowel to vowel. Moreover, changing the size of the jaw opening in this way provides maximum loudness at the lowest possible cost in vocal effort. The strategy is probably resorted to not only by sopranos but also by other singers whose pitch range includes frequencies higher than those of the first formants of ordinary speech: contraltos, tenors and occasionally even baritones.

It can be hard for a student of singing to learn this special way of regulating the jaw opening, and particularly hard if the jaw muscles are under constant tension. That may explain why many singing teachers try to get their students to relax the jaw. Another frequent admonition is: "Hear the next tone within yourself before you start to sing it." That could be necessary because proper manipulation of the jaw opening requires some preplanning of articulation for particular vowels and for the pitch at which they are to be sung. Opening the jaw, however, is not the only way to raise the first-formant frequency. Shortening the vocal tract by drawing back the corners of the mouth serves the same purpose, and that may be why some teachers tell their students to smile when they sing high tones.

Since formant frequencies determine vowel quality, shifting the first-formant frequency arbitrarily according to pitch might be expected to produce a distorted vowel sound, even an unintelligible one. It does not have this effect, largely because we are accustomed to hearing vowels produced at various pitches in the ordinary speech of men, women and children with vocal tracts of very different lengths; if a vowel is high-pitched, we associate it with relatively high formant frequencies. The correlation is so well established in our perceptual system that we may perceive a change of vowel when we hear two sounds with identical formant frequencies but different pitches; if a singer raises her first-formant frequency with the pitch, some of that rise is actually required just to maintain the identity of the vowel. It is true that when the pitch is very high, our ability to identify vowels deteriorates, but that seems to be the case no matter what the formant frequencies are. The soprano, in other words, does not sacrifice much vowel intelligibility specifically as a result of her pitch-dependent choice of first-formant frequency. (Incidentally, composers of vocal music are conscious of the problem of vowel identification at high pitches and generally avoid presenting important bits of text only at the top of a soprano's range; often the text is repeated so that the words can be well understood at a lower pitch.)

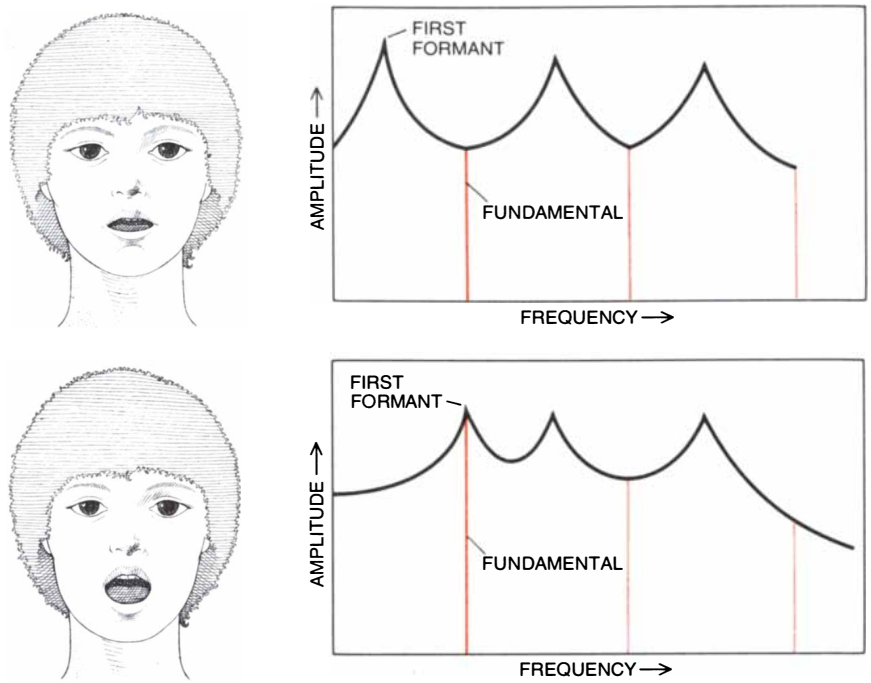
It is clear that a good deal of the difference between spoken and sung vowels can be explained by the singer's need for economy of vocal effort. The general idea is the same, whether in being heard over the orchestra or in maintaining loudness at high pitch: to take advantage of vocal-tract resonance characteristics so as to amplify sounds. The importance of these resonances, the formants, is paramount.

Confirmation of the importance of



SOPRANOS and other singers of high tones tend to open their mouth wider with rising pitch. The tendency is demonstrated in these photographs of a soprano singing the vowel sounds of "heed" (*top*) and of "who'd" (*middle*) at successively higher pitches, shown in musical notation (*bottom*). When these photographs were made, the singer held a vibrator against her neck and a small microphone was placed near her lips. She began to sing each vowel at a specified pitch and then, with the vibrator turned on, she stopped singing but maintained the positions of the articulatory organs. The vibrator now supplied a steady, low-pitched sound that was influenced by the singer's vocal tract just as her own voice source would have been but that was more suitable for analysis than a high voice tone, which has few partials.

the formants was provided by a recent study of how male voices are classified as bass, baritone or tenor. Obviously the singer's frequency range is ultimately the determinant, but even when the true range (which is established primarily by the shape, size and musculature of the vocal folds) has not yet been developed, a good voice teacher can often predict the classification after listening to a student's voice. How is that possible? Thomas F. Cleveland, who was visiting our laboratory at the Royal Institute of Technology in Stockholm and is now at the University of Southern California, analyzed vowels sung by basses, baritones and tenors with respect to formant frequencies and the spectrum of the voice source. Then he had a jury of voice teachers listen to the vowel samples and classify the voices. The teachers tended to classify vowels in which the formant frequencies were comparatively low as having been sung by bass voices and vowels whose formant frequencies were high as having been sung by tenors. Variations in the voice-source spectrum (which varied slightly with the pitch at which a vowel was being sung), on the other hand, did not provide a basis for consistent classification. In a second test the same jury judged a series of synthesized (and therefore clearly defined) sounds and confirmed Cleveland's original impression: the lower the formant frequencies of a given vowel were, the lower the singer's voice range was assumed to be.



NEED FOR WIDE JAW OPENING arises from the fact that a soprano must often sing tones whose fundamental (lowest partial) is far higher in frequency than the normal first formant of the vowel being sung. When that is the case (*top*), the amplitude of the fundamental is not enhanced by the first formant and the sound is weak. Opening the jaw wider raises the pitch of the first formant. When the first-formant frequency is raised to match that of the fundamental (*bottom*), the formant enhances the amplitude of the fundamental and the sound is louder.

Cleveland found that typical bass and tenor voices differ in formant frequencies very much as male and female voices do. The formant-frequency differences between males and females are due mainly to vocal-tract length, and so the bass-tenor differences are probably also largely explained by the same physical fact. Formant frequencies are determined, however, not only by the individual's vocal-tract morphology but also by habits of articulation, which are highly variable. Be that as it may, vocal-tract morphology must set limits to the range of formant frequencies that are available to a singer.

At this point the reader who knows and cares about music may be rather disappointed. I have failed to mention a number of factors that are often cited as determinants of excellence in singing: the nasal cavity, head and chest resonances, breathing and so on. These factors have not been mentioned simply because they seem to be not relevant to the major acoustic properties of the vowel sounds produced in professional operatic singing. Our research suggests that professional quality can be achieved by means of a rather normal voice source and the resonances of the vocal tract.

Our implied model may not be perfect, to be sure. It is just possible, for example, that the nasal cavity has a role in the singing of vowels that are normally not nasalized. If that is so, we have attributed its effect to the voice source, thus compensating for one error by making another. Moreover, we have dealt only with sustained vowel sounds, whose production is important but is certainly not the only acoustic event in singing.

Resonances outside the vocal tract, such as in the head or the chest, cannot contribute appreciably to the singer's acoustic output in view of the great extent to which sound is attenuated as it passes through tissues. This is not to say that such resonances may not be important to the singer, who may receive cues to his own performance not only from what he hears but also from felt vibrations. As for breathing, it is clear that the vocal folds would vibrate no matter by what technique an excess of air pressure is built up below the glottis. Breathing and laryngeal manipulation are likely to be physiologically interdependent, however, since the larynx is the gatekeeper of the lungs. Probably different ways of breathing are associated with different adjustments of the larynx, and probably some ways are effective for singing and others are inadequate or impractical.

Finally we return to the original question: What is so special about a singer's voice? The voice organ obeys the same acoustic laws in singing that it does in ordinary speech. The radiated sound can be explained by the properties of the voice-source spectrum and the formants in singing as in speech. From an acoustical point of view singers appear to be ordinary people. It is true that there is a major difference between the way formant frequencies are chosen in speech and the way they are chosen in singing, and hence between the way vowels are pronounced in singing and the way they are pronounced in speech. A man with a wide pharynx and with a larynx that will resonate at a frequency of between 2,500 and 3,000 hertz is likely to be able to develop a good singing voice more readily than a person who lacks those chacteristics. And his progress may be facilitated if his vocal folds give him a range that agrees with his formant frequencies. As for a female singer, she should be able to shift the first formant to join the pitch frequency in the upper part of her range; that requirement may bar some women with a long vocal tract from having a successful career as a coloratura soprano. There are, in other words, a few morphological specifications that probably have some effect on the ease with which someone can learn to sing well. There are other conditions that may be more important, however. It is in the complex of knowledge, talent and musical instinct that is summed up as "musicality," rather than in the anatomy of the lungs and the vocal tract, that an excellent singer's excellence lies.

The Oldest Rocks and the Growth of Continents

Did most of the continental crust emerge early in the earth's history, or was much of it added later by tectonic processes? Recent advances in the dating of ancient rocks lend strong support to the latter view

by Stephen Moorbath

The story of the earth began 4.6 billion years ago, when a great cloud of gas and dust contracted and condensed to form the entire solar system, including the earth, in something like its present shape and form. It is believed that the earth condensed rapidly—possibly within only a few million years—and that the process of gravitational accretion, combined with the release of energy by the decay of short-lived radioactive isotopes, supplied enough heat for the earth to melt and undergo rapid chemical differentiation into a largely fluid core composed of iron and nickel and a partly solid overlying mantle mainly made up of silicates and oxides of all the chemical elements.

No terrestrial rocks closely approaching an age of 4.6 billion years have yet been discovered. The evidence for the age of the earth is circumstantial, being based on the following indirect reasoning. First, radioactive-isotope dating methods, such as those based on the decay of uranium to lead and rubidium to strontium, show that meteorites solidified about 4.6 billion years ago. Meteorites are currently believed to represent debris left over after the accretion of the solar system. Second, similar measurements on the oldest rocks and soils from the moon also yield an age close to 4.6 billion years, which is interpreted as the time when the moon began to differentiate into its own core, mantle and crust.

Even though no 4.6-billion-year-old rocks have been found on the earth, the measured growth through geologic time in the amount of the radiogenic isotopes lead 206 and lead 207, formed respectively by the radioactive decay of uranium 238 and uranium 235, has a simple pattern indicating that 4.6 billion years ago the isotopic abundance ratios of lead were identical in the parent material of the meteorites and of the earth. When the growth curve of the two lead isotopes is extrapolated backward in time from the oldest-known deposits of lead ore, it passes through the lead-isotope composition found in samples of the uranium-free 4.6-billion-year-old iron meteorite that made Meteor Crater in Arizona. Since lead 204 is a nonradiogenic isotope, it has always been constant in amount and therefore provides a convenient reference with which the other lead isotopes can be compared. The ratio of lead 207 to lead 206 changes with time because uranium 235 decays much more rapidly than uranium 238. Their half-lives are respectively .71 and 4.51 billion years; thus nearly all the uranium 235 present when the earth was formed has already decayed to lead 207, so that the ratio of lead 207 to lead 204 no longer increases very much. On the other hand, the ratio of lead 206 to lead 204 is still increasing rapidly in the earth because only about half of the uranium 238 originally present has so far decayed.

It should be said that although it is highly probable that the earth was formed 4.6 billion years ago, the phrase "age of the earth" strictly speaking refers to that point in time 4.6 billion years ago when the isotopic composition of lead in the earth was the same as that in the parent body of the meteorites. The earth, unlike the moon, has continued as an internally active body because it is large enough to retain the heat generated by the decay of the long-lived radioactive isotopes of such elements as uranium, thorium, potassium and rubidium present in minute concentrations in its mantle.

During the 1960's the earth sciences underwent a major revolution with the general acceptance of global tectonics, a unified hypothesis of earth history that accounts satisfactorily for most of the major surface features of the globe, such as continents, ocean basins, mountain ranges, the distribution of volcanoes and earthquakes, and the distribution in time and space of the main types of rocks found at or near the earth's surface. The hypothesis involves such major processes as the creation of new oceanic crust at mid-ocean ridges through the upwelling of materials derived from the mantle; ocean-floor spreading; continental drift; the subduction (pushing under) of oceanic crust under lighter continental crust; the creation of new igneous rocks by the partial melting of subducted oceanic crust and the consequent addition of new granitic material at the margins of continents, and the collision of continents following the complete subduction of oceanic crust and the closing of oceans, leading to the thickening of continental crust and to the creation of mountain ranges such as the Alps and the Himalayas.

The present configuration of the continents is, we now understand, a consequence of the latest episode of continental drift, which began only about 200 million years ago. About 600 million or perhaps one billion years ago the ancestors of all the present continents were evidently combined into one immense supercontinent, named Pangaea, which may have come into existence as much as 2.7 billion years ago. The grandeur of the concept is quite overwhelming. Only a few pioneers in the preceding generation of earth scientists saw what was coming. My own hero among them is the British geologist Arthur Holmes, who had an abiding faith in the underlying coherence and interdependence of all geological phenomena and was also a pioneer in the development and application of the isotopic dating techniques that are central to this discussion.

My own research interests have centered on the origin of the continental crust and its development and growth through geologic time. This has involved a search for the oldest rocks on the earth and the detailed comparison of ancient rocks with much younger ones whose origin and significance are well understood. At the end of the 18th century James Hutton introduced into geology the doctrine of uniformitarianism, which states that the present is the key to

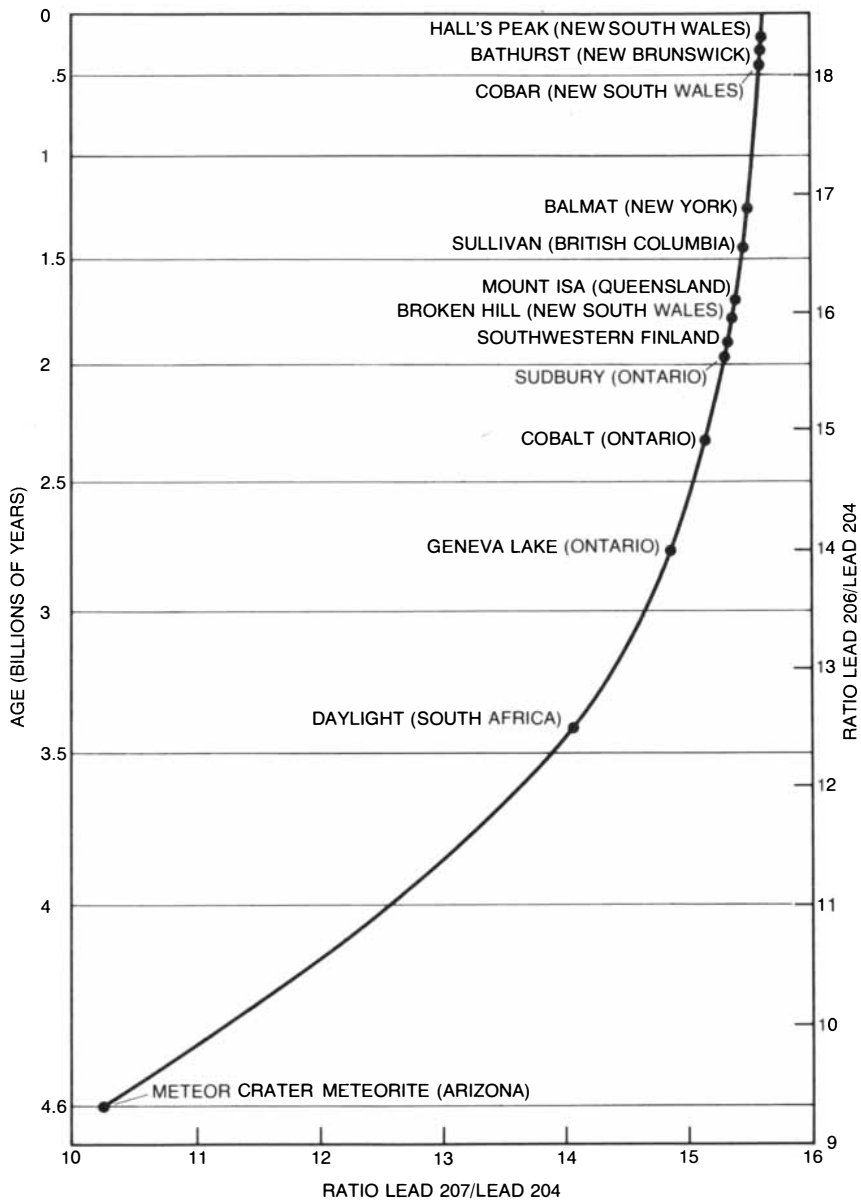
the past. Since we lack a Wellsian time machine to carry us into the past, the best that geologists can do is to test Hutton's doctrine against the rocks. Can the concepts and hypotheses of modern global tectonics, so well established for the past few hundred million years of earth history, be extended to the remote geological past, perhaps to the beginning of earth history itself? I believe a modified form of uniformitarianism can indeed be extended further back in time than has hitherto been thought possible.

About three-tenths of the earth's solid surface is occupied by continental crust, the stuff of the continents and their surrounding continental shelves. The remainder of the solid surface is oceanic crust, which is quite different in character. Continental crust has an average density of 2.7 grams per cubic centimeter; for oceanic crust the comparable figure is three grams, and for the upper mantle it is 3.4 grams. Continental crust is much thicker than oceanic crust down to the junction with the mantle, averaging 35 to 40 kilometers and reaching 60 to 70 kilometers in mountain ranges; oceanic crust averages only six kilometers. The age of continental crust is much greater, ranging in places up to more than 2.7 billion years; the age of oceanic crust is nowhere greater than 200 million years. Continental crust also has a complicated geological structure and a variable chemical composition; oceanic crust has a comparatively simple layered structure and uniform composition.

The upper 10 to 20 kilometers of continental crust consists chiefly of igneous and metamorphic rocks, which are overlain in many places by sedimentary rocks. The upper assemblage of continental rocks has the average chemical composition of the common igneous rock granodiorite, which consists of chemically combined oxides in the following approximate proportions: silicon oxide (SiO_2), 66 percent; potassium oxide (K_2O), 3 percent; sodium oxide (Na_2O), 4 percent; calcium oxide (CaO), 5 percent; magnesium oxide (MgO), 2 percent, and iron oxides (FeO and Fe_2O_3), 4 percent. Other oxides are present in smaller amounts.

AMONG THE OLDEST TERRESTRIAL ROCKS are a sequence of metamorphosed and highly deformed igneous rocks, which have been named the Amîtsoq (Greenlandic for barren) gneisses. A section of the gneiss appears as an outcrop in the right foreground in this photograph made on the west coast of Greenland near Godthaab, the capital. Radioactive-dating measurements made in the author's laboratory at the University of Oxford show that the rock crystallized 3.75 billion years ago. The large dark mass at the left, which cuts through the gneiss, is part of a dike of congealed lava. Both gneiss and dike are cut by a narrow light-colored dike about 2.6 billion years old.

The upper continental crust grades downward into the increasingly metamorphosed rocks known as gneisses. Broadly speaking, gneisses are the higher-temperature and higher-pressure equivalents of the rocks above them, but their mineralogy is simpler and their chemical composition is more basic, that is, it is poorer in silicon oxide. The average chemical composition resembles that of the common igneous rocks diorite and tonalite, which compared with the upper continental crust contain less silicon, potassium and sodium and more calcium, magnesium and iron. Hence gneisses are intermediate in chemical composition between average upper continental crust and average oceanic crust, which is relatively poor in silicon (49 percent SiO_2), potassium (1 percent K_2O) and sodium (3 percent Na_2O) and relatively rich in calcium (11 percent CaO), magnesium (8 percent MgO) and iron (9 percent FeO and Fe_2O_3). The gneisses of the lower continental crust are also much depleted in chemically combined water and in certain "incompatible" trace elements, including among many others the radioactive elements uranium, thorium and rubidium. This incompatibility is due to the fact that the sizes and chemical affinities of the atoms of such elements prevent them from being easily accommodated in the dense, tight crystal structures that are stable at the higher pressures and temperatures of the lower continental crust and the underlying mantle.

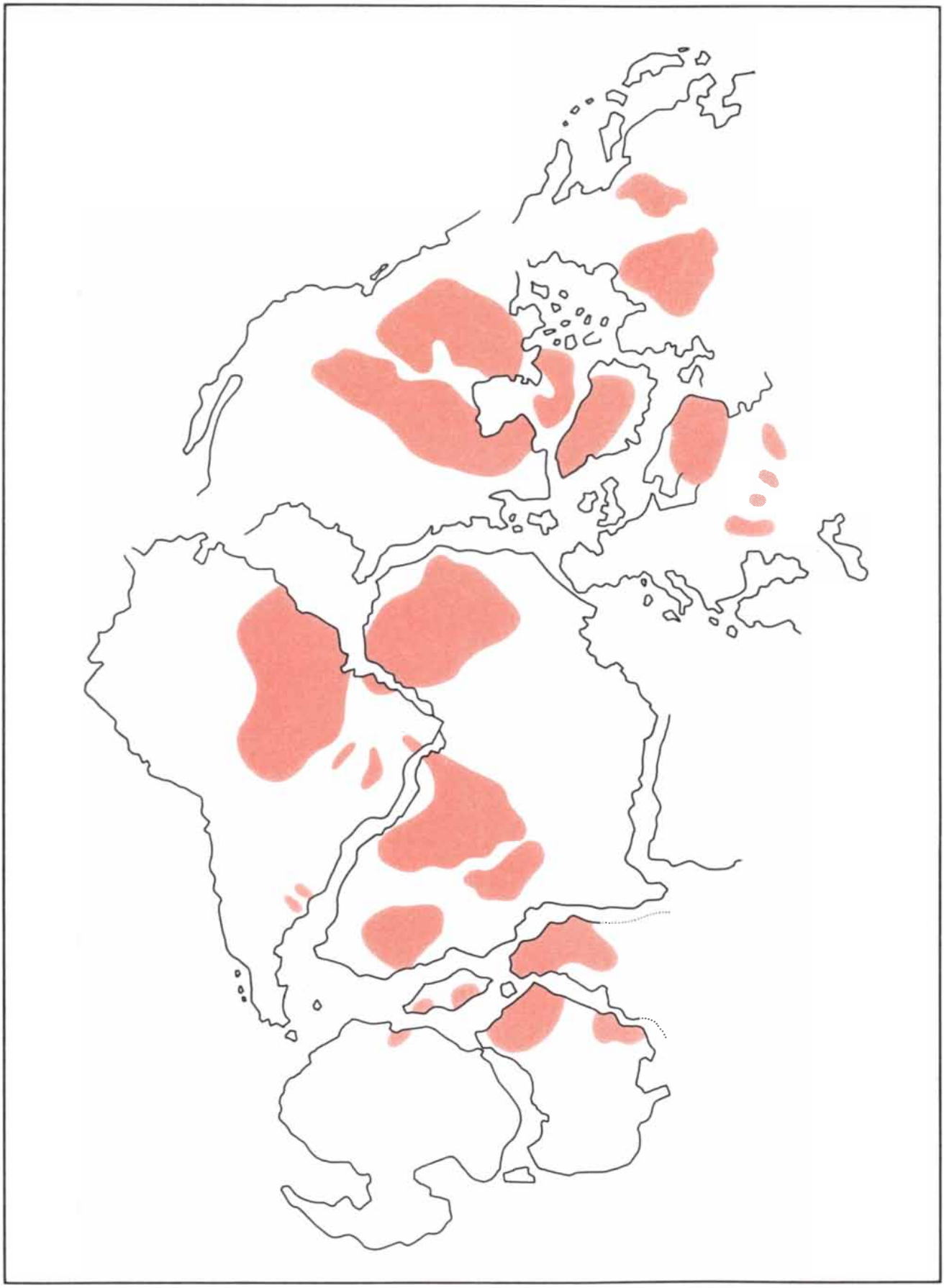


LEAD-LEAD METHOD for determining the age of the earth is based on changing ratios of lead isotopes in the course of geologic time. The reference isotope, lead 204, is not produced through radioactive decay, and so it has always been constant in amount. Lead 206 and lead 207 are formed respectively by the radioactive decay of uranium 238 and uranium 235. Uranium 235 has a much shorter half-life (.71 billion years) than uranium 238 (4.51 billion years). Accordingly most of the uranium 235 that was originally present has decayed, so that the ratio of lead 207 to lead 204 has changed relatively little over the past two billion years. The ratio of lead 206 to lead 204, however, is still increasing significantly. The growth curve for the two ratios can be accurately computed. Plotted on the curve here are the ages of some large deposits of lead ore. When growth curve for ratios is extrapolated backward in time, it yields an age of 4.6 billion years for fragments of the iron meteorite that excavated Meteor Crater in Arizona.

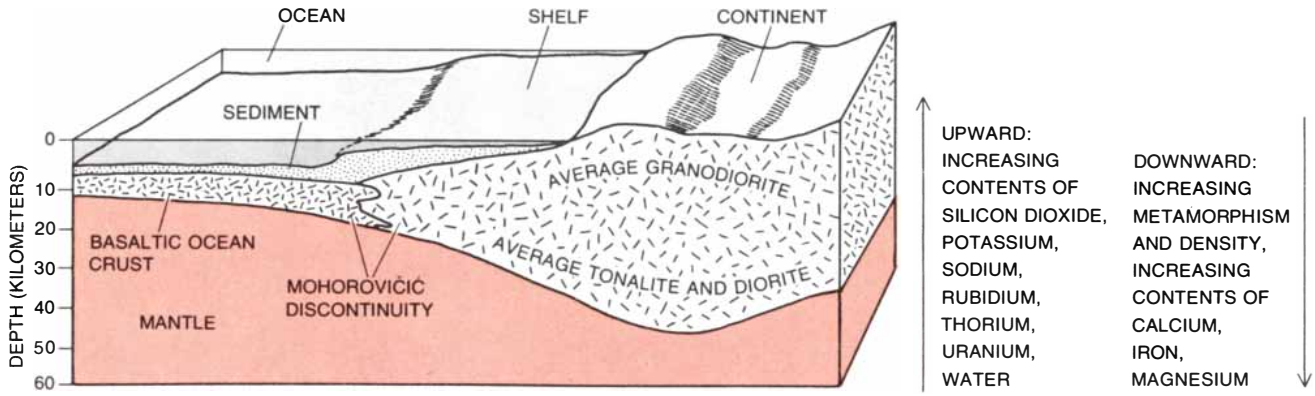
Although the lower continental crust and the mantle contain only minute traces of the incompatible elements, these traces supply the mantle with enough radioactivity to fuel the earth's heat engine, which is what powers global tectonics. There is also a strong tendency for the incompatible elements, including by an extraordinary accident of nature all the radioactive ones, to migrate toward the top of the continental crust, where they fit more easily into the relatively open crystalline structures of the silicates and oxides found there. The vertical gradation in the composition of the continental crust is a phenomenon whose significance has become fully appreciated only in recent years.

There are several conflicting concepts of the origin and evolution of continental crust. One extreme hypothesis is that most if not all of it was made early in the course of the earth's chemical differentiation and that ever since then it has been reworked, that is, heated, melted, recrystallized and deformed. It is true that the effects of such processes can be seen in many parts of the continental crust, particularly in mountain belts and in the exposed roots of ancient mountain belts, but the essence of this view is that the volume of the continental crust has been almost constant since earliest times.

An alternative view, to which I and many other workers adhere, is that the volume and extent of the earliest continental crust were relatively small and have grown throughout geologic time by the irreversible chemical differentiation of the upper mantle, followed by the accretion of the newly differentiated material to the preexisting continents, particularly at or near their margins. The processes by which this happens might be analogous to the one occurring at the present-day western margins of North America and South America, which are moving westward over the denser oceanic crust of the Pacific. As the oceanic crust is subducted to deeper levels it gets progressively hotter until it partially melts and gives rise to lighter, more siliceous rocks of the "calc-alkaline" diorite-tonalite-granodiorite family, which feed the hungry volcanoes all down the west coast of the Americas. The wedge of upper mantle overlying the deeply subducted oceanic crust but



ANCESTORS OF PRESENT CONTINENTS were once part of a supercontinent named Pangaea. Beginning about 200 million years ago the fragments representing the present continents began to drift apart. The regions in color are terrains older than 1.6 billion years.



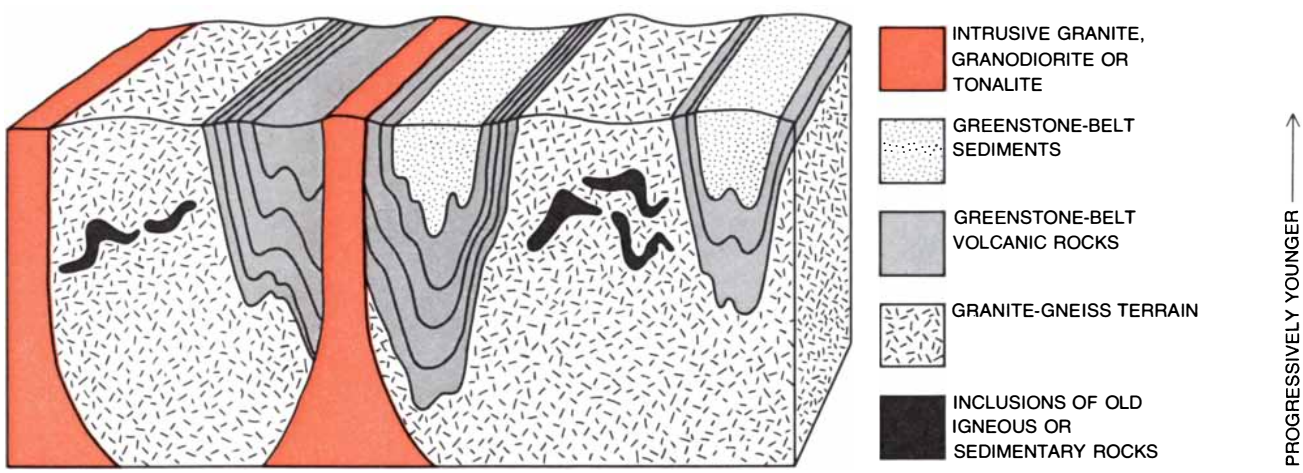
CONTINENTAL CRUST is markedly thicker and less dense than oceanic crust. Oceanic crust is fairly homogeneous in composition; continental crust has a complicated structure and a variable composition. Changes in composition with depth are summarized at the right.

underlying the older continental crust also partially melts to yield new basaltic and calc-alkaline rocks. The remaining dense, unmelted "ultrabasic" material gradually sinks deeper into the mantle and never again takes part in the manufacture of lighter continental crust.

The newly formed calc-alkaline magma crystallizes slowly within the crust to form batholiths: immense bodies of coarse-grained igneous rock. Only a small fraction of the total magma production emerges as lava at the surface. The uplift of parts of the continental crust over the past few tens of millions of years has exposed many batholiths through erosion. The best-known include the Sierra Nevada batholith, the California batholith and the Andean batholith. Many geologists believe such bodies are genuine new additions to the continental crust. This does not exclude the possibility that the new calc-alkaline material has incorporated a small amount of remelted older continental material from the edge of the preexisting continent. Some workers, however, maintain that these great granodiorite and tonalite batholiths are predominantly reworked, partially melted older continental crust. The relevance of this question to the entire debate of "continental reworking" as opposed to "continental growth" will be readily apparent. Before describing a method that can be used to attack this fundamental question let me briefly summarize what rocks can tell us about the earliest discernible history of the earth.

On every continent there are large Precambrian "shield" areas composed of a large variety of igneous and sedimentary rocks, whose ages, on the basis of radioactive-isotope dating, fall between 2.5 and 2.8 billion years. Precambrian rocks older than 2.5 billion years are commonly termed Archean. Perhaps as much as 50 percent of the area of the North American continent was in existence by about 2.5 billion years ago. The situation is probably not very different for other continents. Furthermore, the observed thickness of Archean rock sequences, together with the results of studies of high-pressure, high-temperature mineral assemblages in Archean rocks exposed at the earth's surface by uplift and erosion, leave no doubt that during late Archean times the continental crust had an average thickness of somewhere between 25 and 40 kilometers, which is similar to what is found today. It is quite clear that the most diverse igneous, sedimentary and metamorphic rocks were already being manufactured by a range of geological processes hardly less varied than much more recent ones, except that living organisms played only a minor part in the rock-forming processes.

In the big shield areas of the world there are two principal groups of late Archean rocks: greenstone belts and granite-gneiss terrains. The greenstone belts, so named because of the greenish tinge of many metamorphosed volcanic rocks, consist mainly of volcanic and sedimentary rocks laid down at the surface, either on land or under water, and therefore known collectively as supracrustal rocks. In a typical greenstone belt early volcanic activity has produced ultrabasic and basic lavas, believed to result from a high degree of partial melting in the underlying upper mantle during a period when much heat was being released. In some greenstone belts



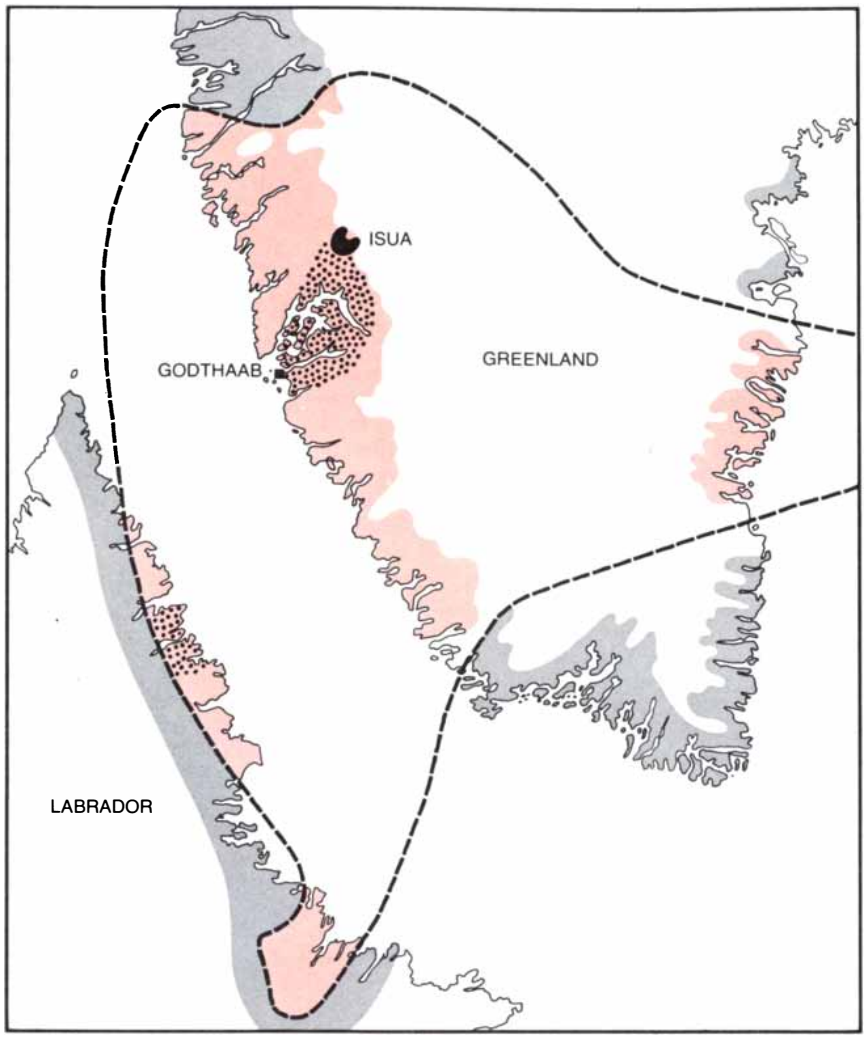
TYPICAL ARCHEAN SHIELD, where the earth's oldest rocks are found, consists principally of greenstone belts embedded in larger areas of granite-gneiss terrains. The key indicates relative ages of five major rock types shown in the cross section. Vertical scale is exaggerated. Greenstone belts consist of volcanic and sedimentary rocks. Granite-gneiss terrains are mostly deformed and banded gneisses of granodiorite and tonalite that were derived from igneous rocks. The inclusions are the broken-up remnants of still older greenstone belts.

successive volcanic rocks in a sequence become progressively more enriched in silicon, potassium, sodium and aluminum, and poorer in magnesium, iron and calcium, giving rise to basalts, andesites and dacites. The overall trend is from "basic" through "intermediate" to "acid," as defined by silicon oxide content. Basalt usually predominates. Farther up in the sequence sedimentary rocks become important, presumably because land was emerging from the sea. The best-studied of the greenstone belts, the Swaziland sequence in the Barberton Mountain Land of southeastern Africa, reaches a thickness of nearly 20 kilometers.

Archean greenstone belts exhibit variable degrees of metamorphic alteration and recrystallization. Such effects can be so weak that the ancient lavas and sediments look almost as if they had just been deposited. Such is the case in parts of the Canadian and southern African shields. There the continent has undergone little vertical movement in the past 2.8 billion years. The rocks were never buried more than 10 kilometers below the surface. In other places the rocks were buried to depths of 10 to 30 kilometers, as is shown by their intense deformation and recrystallization. For example, basalts are converted to their metamorphic equivalent amphibolites, shales to schists and limestones to marbles. Nevertheless, the essential characteristics of the entire supracrustal assemblage can never be destroyed and always remain clearly recognizable.

Many greenstone belts are discontinuous ribbons, eroded remnants of what may once have been a more continuous basin of deposition. The remnants typically occupy an area of several hundred square kilometers. They are generally surrounded by immense "seas" of gneisses, the metamorphic equivalents of igneous rocks such as granodiorites and tonalites. Such gneisses are actually the predominant Archean rock types. Some individual granite-gneiss terrains occupy thousands of square kilometers. In spite of the close juxtaposition of greenstone belts and granite-gneiss terrains it is not always possible to be quite sure which of the terrains is the younger and which the older because there may have been too much deformation for the original contact relations to have been preserved.

There is heated controversy among geologists over the relative ages of greenstone belts and granite-gneiss terrains. Many believe that greenstone-belt rocks were deposited in a primitive oceanic environment and in effect represent ancient oceans, and that the granite-gneiss terrains may be the remnants of ancient continents. It comes down to a chicken-and-the-egg argument: Which came first, the continents or the oceans? How were they related in Archean times? Modern global-tectonic processes can make continents out of oceanic crust; did that happen during the Archean? Furthermore, is there any real evidence that late Archean greenstone belts and granite-gneiss terrains are primordial oceans and continents, as some have suggested? Much research is focused on these problems.



ARCHEAN-SHIELD AREA OF GREENLAND embraces the coastline of Labrador, which was joined to Greenland before the most recent episode of continental drift. The two landmasses are now about three times farther apart than they are seen here. The broken line encloses the early Precambrian (Archean) shield. Area in color consists mostly of rocks between 2.6 and 2.9 billion years old. Black dots in the colored region represent rocks between 3.6 and 3.8 billion years old. The black crescent at Isua, the site of a huge deposit of iron ore, marks the location of an extensive formation of supracrustal rocks (rocks deposited on preexisting crust) that turned out to be 3.8 billion years old. Areas in gray are mostly younger Precambrian rocks.

Let us now consider the ages of the older granite-gneiss terrains that underlie some greenstone belts. Are there any rocks on the earth's surface that are older than 2.8 billion years, and if so, what are they like? How close to the terrestrial limit of 4.6 billion years can we get?

In 1966 V. R. McGregor, a young New Zealand geologist working with the Geological Survey of Greenland, had begun detailed geological mapping in the mountainous area around Godthaab, the capital of Greenland. The great variety and complex structure of the rocks made it no easy task, but after several years of work McGregor was able to recognize a distinctive sequence of geological events. The oldest rocks recognized were the Amîtsoq gneisses, a sequence of metamorphosed and variably deformed igneous rocks of typical continental character, including granodiorites, tonalites and diorites. McGregor's interpretation of the sequence was not at first generally accepted.

A clue to the absolute ages of these rocks came from a potassium-argon measurement carried out some years earlier by Ole Larsen of the University of Copenhagen. Made on a mineral sample separated from the late Qôrqut granite, the measurement gave an age of 2.6 billion years. McGregor had attributed this granite to a late geological

event, No. 10 in his sequence. He reasoned that if the rocks associated with event No. 10 were as old as 2.6 billion years, those associated with event No. 3 might be the oldest yet found on the earth. (At that time the rocks associated with events No. 1 and No. 2 had not been recognized.)

When I first heard of McGregor's findings in 1970, and of the doubts that had been expressed in some quarters, I thought the problem could be solved by carrying out rubidium-strontium and uranium-lead age determinations on samples of Amîtsoq gneiss. McGregor sent a few samples to the University of Oxford, where work by L. P. Black, N. H. Gale, R. J. Pankhurst and me quickly showed that the rocks were much older than any that had been previously reported, probably between 3.6 and four billion years.

We lacked adequate samples to establish a more precise age, and so McGregor and I, with the support of the Geological Survey of Greenland, decided to make a large collection of rocks in the Godthaab area in the summer of 1971, particularly in those areas where McGregor had recognized the Amîtsoq gneisses. Subsequent age measurements at Oxford on all the samples confirmed an age of about 3.75 billion years, which we now interpret as being close to the age when the igneous precursors of the gneisses separated from upper-mantle source regions to form this uniquely ancient but typical continental crust. Useful confirming evidence for this ancient age came in 1973 from Halfdan Baadsgaard of the University of Alberta, who reported a uranium-lead age of 3.7 billion years on the rare mineral zircon (zirconium silicate) extracted from several samples of Amîtsoq gneiss.

Since 1971 my colleagues and I have further established that event No. 8 in the McGregor sequence occurred about 2.85 billion years ago, and we have confirmed that the late Qôrqut granite was emplaced in the crust 2.6 billion years ago. These youngest rocks in the Godthaab area are almost as old as the oldest rocks previously reported from other areas. In fact, it was at about the same time that lunar scientists were beginning to report ages similar to those for the Amîtsoq gneisses for the totally unrelated basalts found on the maria, or "seas," of the moon.

The Amîtsoq gneisses are known to have been reheated and strongly deformed by geological events that took place between 2.6 and 2.9 billion years ago. Many people ask how one can obtain a reliable age of 3.75 billion years for rocks that were so strongly altered almost a billion years later. It is possible because geochronologists nowadays use the "whole rock" method of analysis for the rubidium-strontium and uranium-lead dating schemes. In this method samples of rock weighing at least five kilograms are completely ground to a fine powder, from which small representative samples are taken for laboratory analysis. Such samples give consistent ages of 3.75 billion years for the Amîtsoq gneiss. Minerals separated from the Amîtsoq gneiss, however, such as biotite (a hydrated silicate of magnesium, iron, aluminum and potassium) and hornblende (a hydrated silicate of aluminum, calcium, magnesium and iron, with a small amount of potassium), yield ages of about 2.6 billion years, and some yield ages as low as 1.6 billion years; these ages are interpreted as times when the Amîtsoq gneisses were strongly heated.

EVENT NUMBER	TIME (BILLIONS OF YEARS AGO)	DESCRIPTION
1	?	Formation of terrestrial crust of unknown type; no longer discernible.
2	~3.8	Eruption of volcanic lavas and deposition of sedimentary rocks, some under water, that created the Isua supracrustal belt and the inclusions in the Amîtsoq gneisses of the Godthaab area.
3	~3.75	Intrusion of the parent rocks of the Amîtsoq gneisses: granodiorites, tonalites and diorites.
4	~3.7	Deformation and metamorphism (recrystallization) of the earlier supracrustal rocks and of the igneous rocks. Formation of the Amîtsoq gneisses.
5	?	Intrusion of many basaltic dikes (Ameralik dikes) into all earlier rocks.
6	~2.9+	Eruption of volcanic lavas and deposition of sedimentary rocks onto earlier crust.
7	~2.9	Intrusion of large igneous bodies of anorthosite (calcium aluminum silicate rocks).
8	~2.8	Intrusion of vast amounts of granodiorites and tonalites, the parent rocks of the Nûk gneisses.
9	~2.7	Further deformation and metamorphism of all the above rocks. Formation of the Nûk gneisses.
10	~2.6	Intrusion of the Qôrqut granite.
11	~2.2+	Intrusion of many basaltic dikes.
12	~1.6	Heating of the entire area.

SEQUENCE OF GEOLOGICAL EVENTS in the Archean-shield area along the west coast of Greenland around Godthaab and Isua (*see map on preceding page*) was worked out in the late 1960's by V. R. McGregor. Most of the age measurements that are given here were made in the author's laboratory at Oxford. Events 6 through 10 are not recognizable at Isua.

But why should whole-rock samples give ages that differ from those of mineral samples from the same rock? This seeming paradox has made it possible to sort out the complex sequence of events within a given metamorphosed terrain. Consider the rubidium-strontium whole-rock method, which utilizes the decay of rubidium 87 to strontium 87. The rare element rubidium easily substitutes for its chemically close relative potassium, so that all potassium-bearing minerals contain a trace of rubidium, often not more than a few parts per million. In fact, all the rubidium on the earth is to be found in potassium minerals, because there is not enough to form special rubidium minerals.

In the course of time radiogenic strontium 87 builds up by the decay of rubidium 87 at the appropriate sites in the crystal lattice of the mineral. Rubidium and strontium are chemically quite different because their ions differ in radius and charge, so that radiogenic strontium 87 will diffuse out of a mineral such as biotite at temperatures as low as 200 or 300 degrees Celsius. It does not usually move far, however, being taken up by surrounding grains of plagioclase feldspar (calcium sodium aluminum silicate), in which strontium can easily substitute in crystal-lattice sites for the closely similar element calcium. Plagioclase feldspar is a common constituent of most rocks. As a result rock samples weighing several kilograms and containing many thousands of individual grains of biotite and plagioclase, together with other strontium donors and acceptors, can remain closed systems with regard to strontium 87 even through a strong heating episode. Of course, the rock must also remain a closed system to rubidium. The same basic principles apply to the age methods that are based on the decay of uranium to lead.

There is more to the Greenland story. McGregor soon found evidence that the Amîtsoq gneisses were not the oldest terrestrial rocks. In common with many other granite-gneiss terrains they contain fragmented chunks of older rocks that, although strongly recrystallized, are clearly recognizable as having originally been volcanic and sedimentary rocks of various kinds that were laid

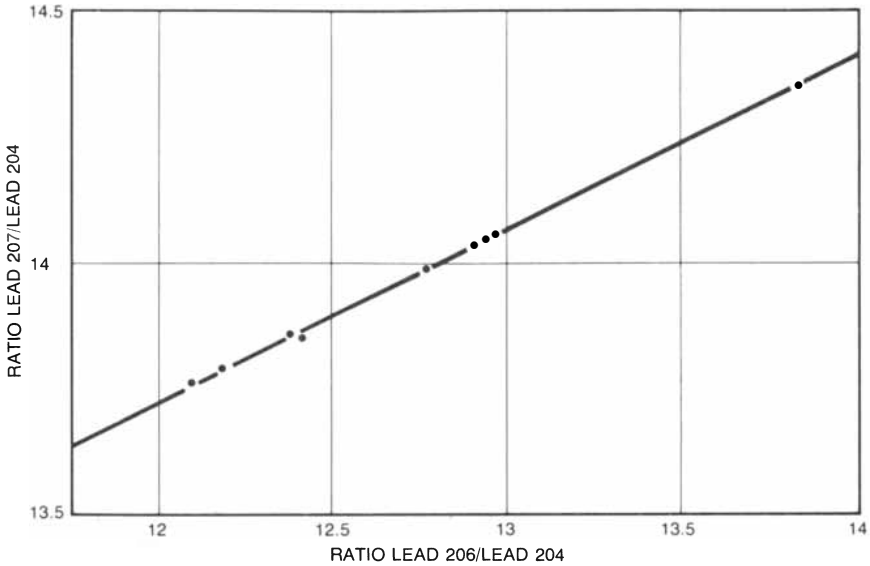
down at the surface and then buried to a depth of several kilometers before the igneous precursors of the Amîtsoq gneisses intruded into them and broke them up. It is not easy to date the pre-Amîtsoq inclusions, and so we do not yet know just how much older they are than the Amîtsoq gneisses.

Now, however, we come to one of the most interesting parts of the search for the oldest rocks in western Greenland. In 1971, while McGregor and I were collecting rocks in the Godthaab area, an unexpected opportunity arose to visit a remote mountainous area named Isua, about 100 kilometers northeast of Godthaab at the very edge of the great inland ice sheet. At that time the Danish mining company Kryolitselskabet Øresund A/S was beginning to explore a huge iron-ore deposit at Isua, which outcrops at a height of some 1,400 meters and is partly covered by the inland ice.

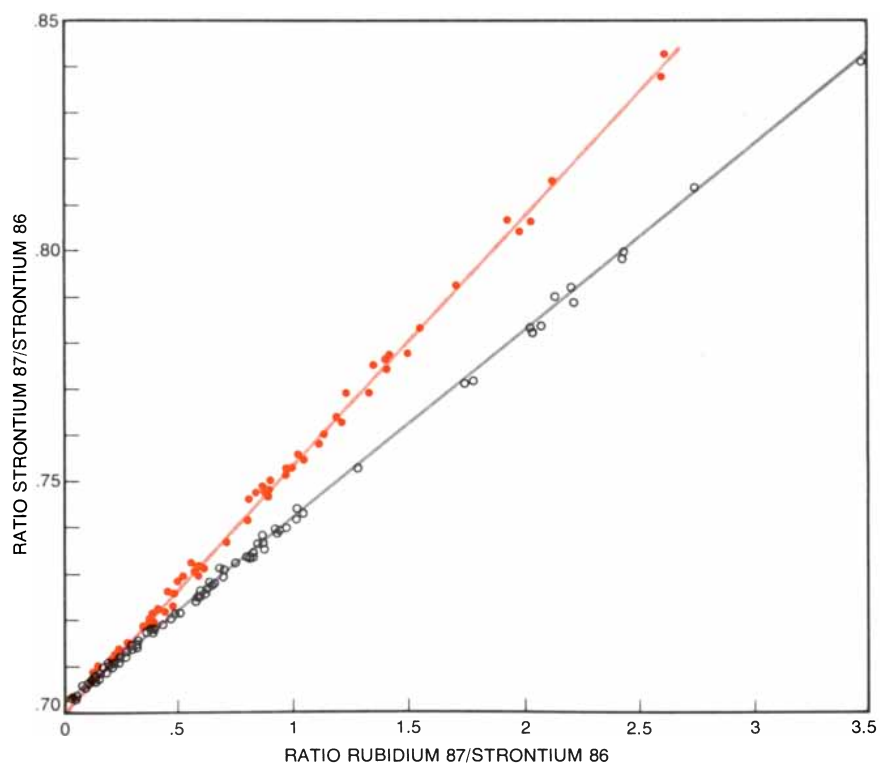
The ore deposit had been discovered a short while before in the course of an airborne-magnetometer survey, and the company geologists had already made a preliminary geological map of the area. The ore body formed a small part of an incomplete oval arc between 12 and 25 kilometers in diameter consisting of a 3,000-meter-thick sequence of variably deformed and strongly metamorphosed volcanic and sedimentary rocks. It was perfectly clear that these rocks were supracrustal, that is, laid down at the surface. The rocks on all sides of the supracrustal belt consisted of typical granite-gneisses cut vertically by numerous dikes of congealed lava, which also cut the supracrustal rocks.

McGregor and I at once suspected that the gneisses and dikes were the less deformed and less metamorphosed equivalents of the Amîtsoq gneisses and Ameralik dikes of the Godthaab area. The rocks in this bleak Arctic upland were almost totally exposed, but the points of contact between the gneisses and the supracrustal rocks were sheared and deformed. There were nonetheless fairly clear signs that the supracrustal rocks were older than the gneisses. In any case our immediate impression was that the Isua rocks were the equivalents of the Godthaab rocks in a much more pristine condition. And it eventually turned out to be so.

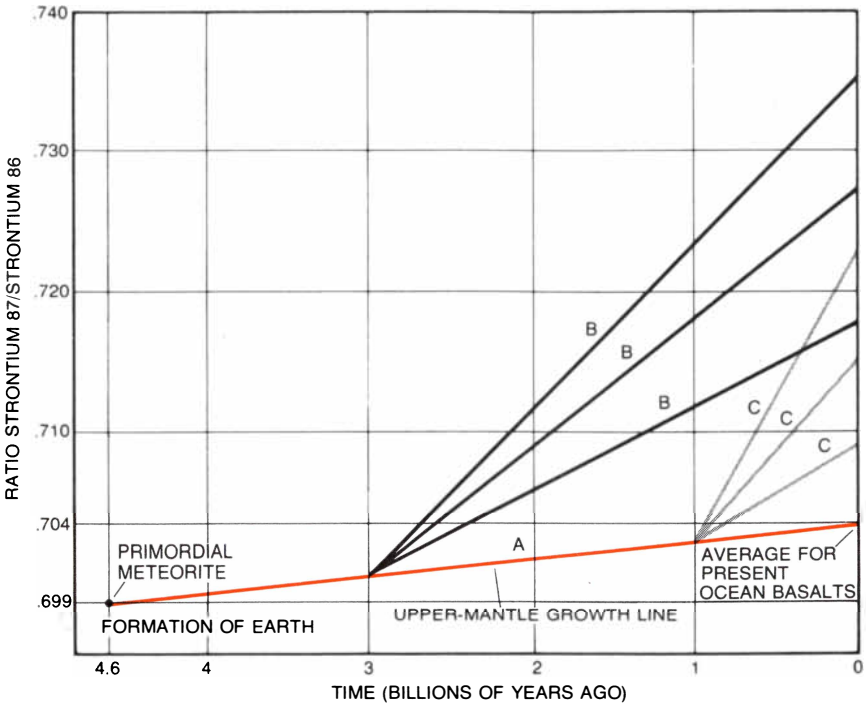
When R. K. O'Nions, Pankhurst and I subjected the gneisses and supracrustal rocks to rubidium-strontium and uranium-lead measurements in Oxford, they all gave ages close to 3.8 billion years. McGregor and I had been walking on permanent continental crust, in all its igneous, sedimentary and metamorphic manifestations, formed some 3.8 billion years ago and since that time virtually undisturbed by any major geological upheaval. As a result of this experience I have come to regard the essential permanence and indestructibility of the



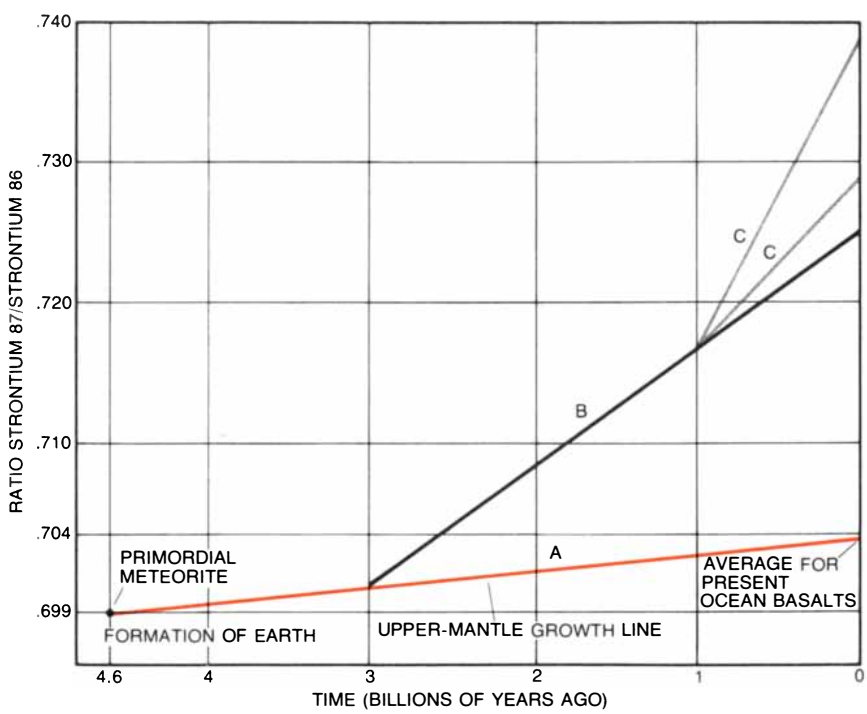
LEAD-LEAD AGE DETERMINATION for samples of a sedimentary banded-iron formation at Isua in western Greenland indicates that the sediments were deposited 3.8 billion years ago. As is indicated in the chart on page 94, lead 206 is produced by decay of the still fairly abundant uranium 238; lead 207 is produced by the decay of the much rarer uranium 235. Since the amount of lead 204 is constant over time, the ratio of lead 206 to lead 204 increases more rapidly than the ratio of lead 207 to lead 204. For Isua samples slope of line produced by two ratios corresponds to 3.8 billion years, the greatest age yet determined for a terrestrial rock.



RUBIDIUM-STRONTIUM ISOCHRON DIAGRAM for two groups of Archean gneisses from Greenland indicates that one group is 900 million years older than the other. As in the lead-lead method, two changing isotope ratios are compared. Here the isotope that remains constant over time is strontium 86. The amount of rubidium 87, with a half-life of 50 billion years, slowly decreases by decaying into strontium 87. Hence the more rubidium 87 there is initially in a sample compared with strontium 86, the faster the ratio of strontium 87 to strontium 86 will increase with time. The two isochron lines plotted here are for whole-rock samples of Amîtsoq gneiss (*color*) and Nûk gneiss (*black*). The steeper the slope, the older the rock. Thus Amîtsoq gneiss is 3.75 billion years old and Nûk gneiss is 2.85 billion. The intercepts on the vertical axis represent initial values of the ratio of strontium 87 to strontium 86 at those dates.



RATIO OF STRONTIUM 87 TO STRONTIUM 86 in the earth's upper mantle has increased from .699 to .704 (*line A*) in 4.6 billion years as rubidium 87 has slowly decayed into strontium 87. As explained in the preceding illustration, rocks with higher ratios of rubidium 87 to strontium 86 exhibit steeper increases in the ratio of strontium 87 to strontium 86. By plotting such ratios and extending them backward in time one can deduce when materials with different ratios had a common origin as well as the initial ratio of strontium 87 to strontium 86. If the latter falls on line *A*, it means that the material was derived from the upper mantle or from mantle-derived oceanic crust and not from remelted continental crust. Thus lines *B* and *C* represent material derived from mantle respectively three billion years ago and one billion years ago.



REWORKING OF CONTINENTAL CRUST is indicated if the slopes of strontium 87/strontium 86 growth lines converge on values significantly greater than the values expected for the upper mantle. In the example shown here mantle material that had been converted into continental crust three billion years ago (*line B*) was remelted and reworked one billion years ago, producing two growth lines labeled *C* beginning with a strontium 87/strontium 86 ratio of .717.

continents as one of the most profound phenomena of geology.

The geology of the Isua area has been closely studied by several other workers who consider the supracrustal rocks to be older than the igneous precursors of the gneisses. The original crust on which the supracrustals were deposited has not, however, been found, and it may no longer be discernible. It is probable that the Isua supracrustal rocks are the geological equivalents of the broken-up pre-Amîtsoq chunks found within the Amîtsoq gneisses of the Godthaab area.

The entire rock assemblage at Isua is quite characteristic of the association between greenstone belts and granite-gneiss terrains. Our age measurements for the supracrustal rocks and the gneisses agree within the analytical error, which is between 50 and 100 million years in this range of ages. This suggests to us that the deposition of the supracrustal rocks and the subsequent emplacement of the igneous precursors of the gneisses occurred within the space of a few tens of millions of years some 3.8 billion years ago.

The character of the sediments at Isua proves that water existed on the surface of the earth 3.8 billion years ago. Is there any evidence that life existed as well? A search for biogenic compounds in the Isua rocks by Bartholomew S. Nagy and Lois A. Nagy of the University of Arizona has so far had negative results. The earliest evidence for biological activity, although it is still controversial, comes from certain greenstone-belt sediments in the Barberton Mountain Land of southeastern Africa, which may have been deposited more than three billion years ago.

Other very old rocks have been discovered in the past two years. The existence of rocks 3.6 billion years old around Saglek Bay on the Labrador coast of Canada has been reported by R. W. Hurst of the University of California at Santa Barbara and G. W. Wetherill of the University of California at Los Angeles, working in collaboration with Kenneth Collerson of the Memorial University of Newfoundland and David Bridgwater of the Geological Survey of Greenland. This area was joined to the Godthaab area of Greenland before Canada and Greenland were separated by the opening of the Davis Strait about 60 million years ago. P. N. Taylor of Oxford has reported 3.5-billion-year-old gneisses from northern Norway. Martha Hickman of the University of Leeds, and also C. J. Hawkesworth, O'Nions, J. F. Wilson and I, have shown that a part of the ancient shield of Rhodesia is about 3.6 billion years old, although most of it is probably only about 2.7 or 2.8 billion years old. Samuel S. Goldich of Northern Illinois University and Carl E. Hedge of the U.S. Geological Survey have reported a preliminary, and still somewhat debated, age of

about 3.6 billion years for the gneisses of the Minnesota River valley in southern Minnesota. It is rumored that other areas with ages of between 3.5 and 3.8 billion years will soon be reported. It all goes to show that typical continental crust already existed only some 800 million years after the formation of the earth, and that the earth's core, mantle and crust had already separated from one another. There is no positive evidence, however, that any of these very ancient rocks represent the very first crust of the earth to be formed.

I am pessimistic about the likelihood of discovering even older rocks on the earth and thus bridging the gap between 3.8 and 4.6 billion years. If a substantial fraction of the heat generated by the earth's accretion had been quickly and efficiently dissipated into space, it would in principle have been possible for the core, the mantle and some form of crust to have segregated 100 or 200 million years after accretion. Otherwise the accreted mass of the earth might have remained too turbulent for effective chemical segregation until 3.8 or 3.9 billion years ago.

There is the further possibility that the earth, like the moon, was heavily bombarded by asteroid-sized bodies, representing the debris left over from the formation of the solar system about four billion years ago. This bombardment could have caused enough mechanical and thermal disturbance to prevent the separation of continental crust from the mantle up to that time. When the bombardment had ended, the processes of chemical segregation could have become established over a period of 100 or 200 million years.

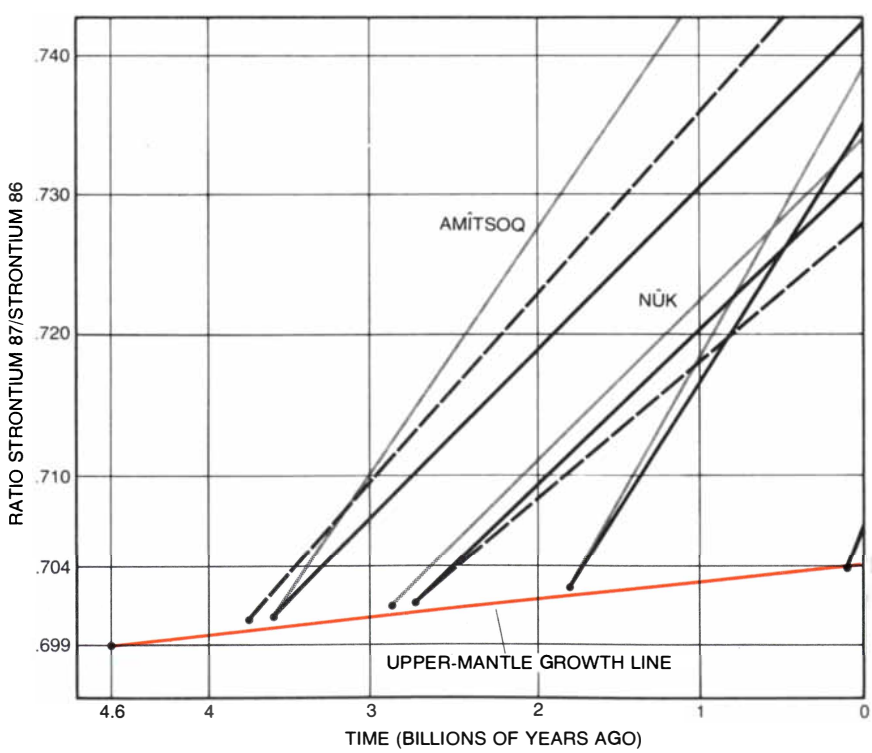
We simply do not know enough to tell what happened in this early period, but it is clear that the search for older rocks could be most rewarding. No clear evidence has yet been found for early meteorite impacts in ancient rocks, and the suggestion of some workers that the later greenstone belts are the terrestrial equivalent of the great lava basins of the lunar maria is at this stage too improbable to contemplate seriously. Neither the geological and temporal relations within the greenstone belts nor the nature and succession of rock types in them fit such a hypothesis. Greenstone belts were almost certainly formed by processes originating within the earth that are probably still operating today, as we shall see.

Let us now return to the question of whether the continental crust has grown with time or was mainly created early in the earth's history—at least 3.8 billion years ago—and then reworked by repeated cycles of melting, sedimentation and metamorphism. Since the density of continental crust is relatively low, it would be impossible for such material to be pushed down and recycled through the mantle. This seems self-evident, yet it is surprising how many geologists regard recycling through the mantle as being an important process in the evolution of the continental crust. The essential unsubductibility and indestructibility of continental crust, once it is formed, is one of the fundamental tenets of global tectonics, as has been emphasized by one of its pioneers, D. P. McKenzie of the University of Cambridge.

Fortunately isotopic methods now exist for deciding whether a given piece of continental crust is old, reworked continental crust or a new addition to the crust derived from the mantle. One of these methods, pioneered in the early 1960's by Gunter Faure and Patrick M. Hurley of the Massachusetts Institute of Technology, is based on the growth in the ratio of strontium 87 to strontium 86 through geologic time. Strontium 87 is produced by the radioactive decay of rubidium 87, which has a half-life of 50 billion years. Strontium 86, on the other hand, is not produced by radioactive decay, so that it remains constant in amount.

The growth in the ratio of strontium 87 to strontium 86 over a given period of time is directly proportional to the ratio of rubidium to strontium in a given sample; the relation is usually expressed as rubidium 87/strontium 86 [*see bottom illustration on page 99*]. In the earth's upper mantle the strontium 87/strontium 86 ratio has changed from a value of .699 at 4.6 billion years ago (based on the known initial ratio in meteorites and on the moon) to a value of close to .704, which is the average for the upper-mantle source of the basaltic lavas that erupt in modern oceans. If we assume that the growth rate from .699 to .704 was approximately linear through geologic time, we can calculate that the rubidium/strontium ratio in the upper mantle is .03. The average rubidium/strontium ratio of continental crust is about .25 (although it varies greatly according to rock type and depth in the crust). It follows that the strontium 87/strontium 86 ratio must increase at a much higher rate in the continental crust than it does in the upper mantle [*see top illustration on opposite page*].

A typical example of the application of the method is provided by rocks from western Greenland. It is hotly debated whether the Nûk gneisses, which make up a large part of western Greenland, are reworked 3.75-billion-year-old Amîtsoq gneisses or whether they represent massive new additions to the continental crust about 2.85 billion years ago (the age of many samples of Nûk gneiss measured by Pankhurst and me). The strontium 87/strontium 86 ratio of the Amîtsoq gneisses 3.75 billion years ago was



STRONTIUM 87/STRONTIUM 86 GROWTH LINES are shown for a variety of granite-gneiss terrains from Greenland (*gray lines*), North and South America (*black lines*) and the Rhodesian shield of Africa (*broken lines*). The slope of each growth line is proportional to the rubidium/strontium ratio of each rock sample. Since initial strontium 87/strontium 86 ratios all converge close to growth line for upper mantle, it is clear that younger gneisses in each area were not made by reworking of old continental crust. Shortest line at right shows initial ratio and growth curve for a granodiorite batholith on coast of Chile only 100 million years old.

.701. The average rubidium/strontium ratio of the Amîtsoq gneisses today is about .3, which is close to the average value for normal continental crust. It can be calculated that by 2.85 billion years ago the average strontium 87/strontium 86 ratio of the Amîtsoq gneisses was about .715 [*see illustration on preceding page*]. The remelting and remobilization of large volumes of Amîtsoq gneisses 2.85 billion years ago would therefore have given rise to igneous precursors of the Nûk gneisses with an initial strontium 87/strontium 86 ratio of about .715.

What do we actually find? The initial strontium 87/strontium 86 ratio of the Nûk gneisses is close to .702. It is clear that the younger gneisses are not remobilized and reworked Amîtsoq gneisses. The best explanation for the respective initial strontium 87/strontium 86 ratios of .701 and .702 for the Amîtsoq and Nûk gneisses is that their granodioritic and tonalitic igneous precursors were produced by the chemical differentiation of material that was derived from the upper mantle by some process of continental accretion close to the measured time of 3.75 and 2.85 billion years ago.

Enough strontium-isotope data have now been published on rocks from Europe, Greenland, Africa, North America and Australia that are between 2.6 and 2.8 billion years old to show that granite-gneiss terrains, greenstone belts and their associated cross-cutting igneous rocks were all formed within a time interval of only about 100 or 200 million years, mainly by the partial melting and chemical differentiation of the upper mantle or of materials derived from it. The reworking of much older continental crust has played a very minor part, although it is isotopically detectable in some areas. Identical conclusions can be drawn from the scantier lead-isotope data on the same rocks.

Isotopic work on rocks 1.7 to 1.9 billion years old from Greenland, North America and elsewhere shows that the entire process occurred again at that time. Low initial strontium 87/strontium 86 ratios in many rocks of this age demonstrate that continental accretion again predominated over the reworking of older crust. With the passage of geologic time to the present, strontium and lead isotope ratios become more complex to interpret. They do indicate an increasing tendency for older continental crust to contribute to the material giving rise to younger continental crust, but continental accretion still predominates wherever oceanic crust is subducted under a continent.

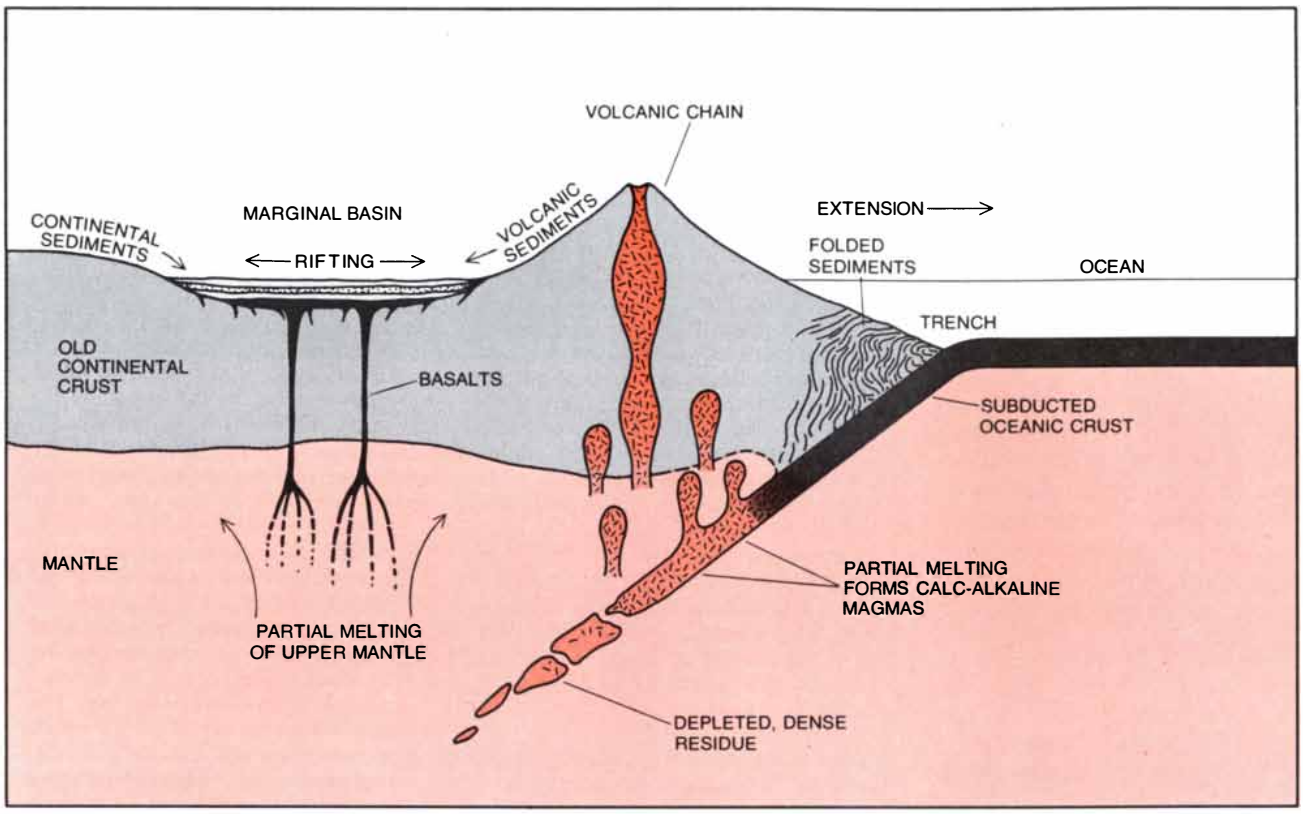
Many geologists have recently come to recognize, in contrast to earlier views, that the rocks of Archean greenstone belts and granite-gneiss terrains show great chemical, petrological and structural similarities to rocks found at those continental margins where oceanic crust is being subducted and partially consumed under an overriding continent. Such areas are known as destructive plate margins. More specifically, Archean granite-gneiss terrains are very similar in composition to the great calc-alkaline batholiths of, say, the west coast of North America and South America. The erosion of those particular batholiths, however, has not yet cut deep enough to expose regions of less siliceous gneisses, depleted in incompatible elements, that probably form the underlying equivalents of the batholiths. Such gneisses would be more strictly comparable to the Archean ones. It must be said we do not really know that the North and South American batholiths continue downward into gneisses of the same age, but the heat-flow studies of Arthur H. Lachenbruch of the U.S. Geological Survey on the Sierra Nevada batholith showed that there is a decrease of heat production downward because of a sharp decrease with depth in the concentration of the geochemically incompatible radioactive elements. In any case, I believe, in agreement with a suggestion made by B. F. Windley of the University of Leicester and Joseph V. Smith of the University of Chicago, that Archean granite-gneisses are the ancient analogues of modern granodioritic and tonalitic batholiths located along destructive plate margins.

The chemistry and petrology of volcanic lavas in Archean greenstone belts have affinities both with lavas erupted from volcanoes situated above subduction zones at or near modern continental margins and with lavas erupted at mid-ocean ridges. Volcanic island arcs (such as the islands of Japan) and volcanoes on continental margins, both of which are at destructive plate margins, represent a geological and chemical environment that is strikingly different from that at mid-ocean ridges, which form constructive plate margins. Yet greenstone-belt volcanic rocks exhibit compositional affinities with both. In order to explain this apparent contradiction John Tarney of the University of Birmingham and Ian W. D. Dalziel and M. J. de Wit of Columbia University, as well as Kevin C. Burke, John F. Dewey and W. S. F. Kidd of the State University of New York at Albany, have recently proposed that Archean greenstone belts are closely comparable to what are called marginal basins, which can be found above subduction zones at or near continental margins.

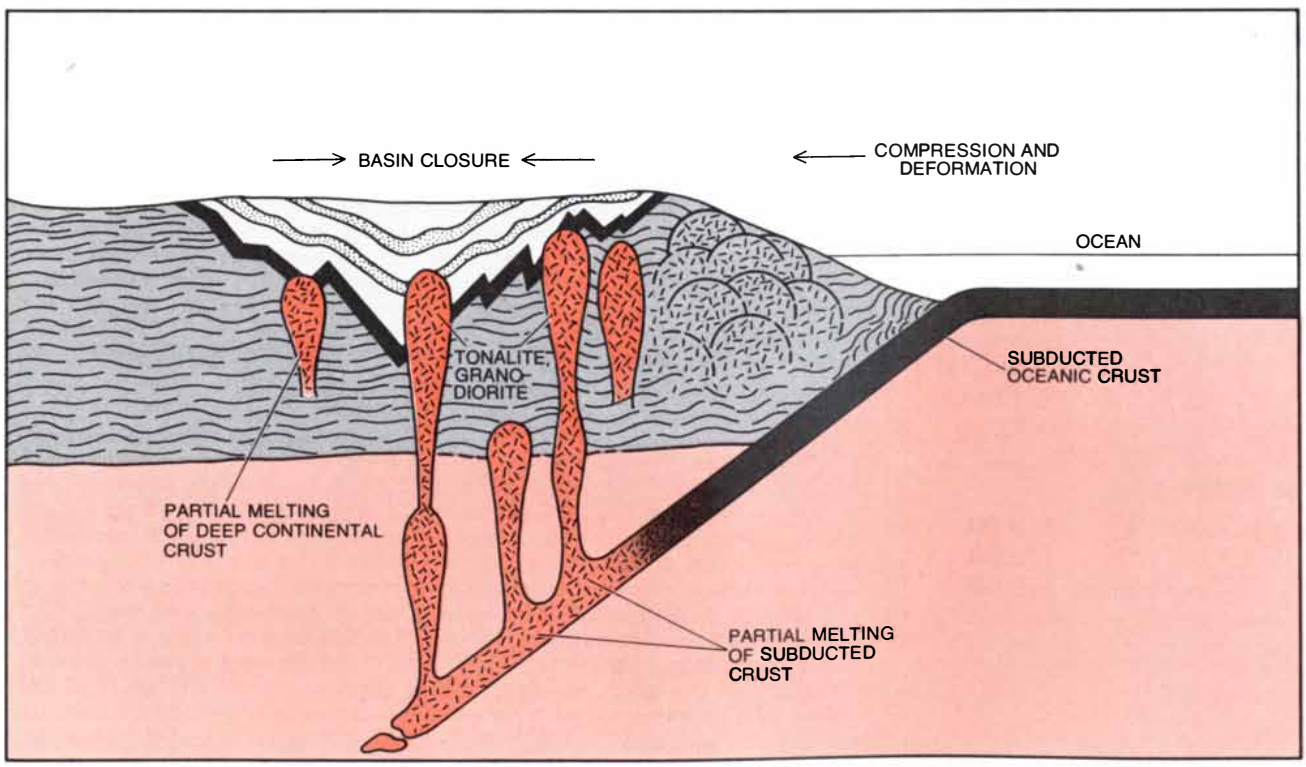
Such basins are true spreading centers, which can cause the overlying continental crust to thin out, rift and even rupture. Nobody yet understands the full reason for their existence, but it may have something to do with the episodic nature of the subduction process. That process may stop for a while, giving the subcontinental mantle above the subduction zone a chance to set up a convection regime, which then causes rifting in the overlying crust. In an extreme case a small piece of continent may split off from the rest, creating a "micro-ocean" in between. A few million years later subduction may resume as before, and the marginal basin will close up again, causing considerable deformation and upheaval in the volcanic and sedimentary rocks of the marginal basin and in the surrounding crustal rocks.

Where does all this leave us? The earliest great continent-forming event on the earth for which we have proof took place about 3.8 billion years ago. There is no justification, however, for thinking this was necessarily the first such event. There is no evidence that rocks of this age represent primordial continental crust. Most of the rock types closely resemble those found in later geological epochs and therefore probably originated by similar processes. The area of the earth's surface occupied by continental crust 3.5 to 3.8 billion years ago is not yet known, but I suspect it was not more than 5 to 10 percent of the present continental area. The average thickness of this ancient continental crust may have approached 20 to 30 kilometers, which is not greatly different from the thickness of continental crust today. The ancient continental crust was already as strong and rigid as modern continental crust. One line of evidence for this conclusion is the presence of numerous vertical fissures filled with basaltic lava, for example the Ameralik dikes, which were injected well before 2.9 billion years ago into a cold, brittle continental crust of Amîtsoq gneisses.

The next great continent-forming event took place between 2.9 and 2.6 billion years ago, and it created perhaps as much as 50 to 60 percent of the area of the present continents. The average thickness of the crust was virtually the same as it is today. There is growing evidence from age measurements that similar worldwide continent-forming events occurred roughly 1.9 to 1.7 billion years ago, again 1.1 to .9 billion years ago and most recently in the past 600 million years. The last major event coincides with the period in which plate tectonics as we now know it led to the breakup and dispersal of the continents, in other words to continental drift. Before that there may have been just the one immense supercontinent known as Pangaea, or possibly two supercontinents: Laurasia to the north and Gondwanaland to the south. There is geological evidence from certain ancient linear structures, akin to the roots of mountain belts, that different parts of the supercontinent underwent differential movement of a few hundred kilometers with respect to one another. Nevertheless, as



MARGINAL BASIN often opens behind volcanic island arcs where oceanic crust is being subducted under a continental margin. Some of the subducted crust melts, forming a calc-alkaline magma. Most of the magma solidifies near the base of the continental crust, but some of it overflows as lava from volcanoes. As subduction ceases a basin opens up behind the volcanic arc and becomes filled with lavas and sediments. One of the best-known marginal basins, the Rocas Verdes complex, is found at western margin of southern tip of Chile.



IF SUBDUCTION RESUMES, the marginal basin will be reclosed with intense compression and deformation of both the old and the new continental crust. The newly formed continental crust becomes chemically differentiated and compositionally layered. In the process it is firmly welded to the old crust. Farther inland the partial melting of subducted oceanic crust causes new magmas to invade deformed rocks of the closed marginal basin. This model would plausibly lead to the type of Archean terrains illustrated at the bottom of page 96.

long as the fairly coherent supercontinent could move in relation to its adjacent ocean floor, the ocean floor could have been subducted, partially melted and chemically differentiated to manufacture typical continental crust at the leading edge of the protocontinent, a process analogous to continental accretion in present-day island arcs and at continental margins.

Unlike some other geologists, I believe both continental crust and oceanic crust have been thick and rigid enough for plate-tectonic processes of one kind or another to have been going on since at least 3.8 billion years ago. This would imply that the nature of the continent-forming process was not greatly different from what we see operating today. There is strong evidence that geological provinces of ancient age in Precambrian-shield areas are not randomly arranged in space; progressively younger areas are adjacent to each other, as would be expected from the hypothesis of progressive continental accretion. The spatial arrangement of ancient geological provinces is a major field for current research.

The reason for the apparently episodic nature of the continent-forming process is not known. Maybe it is only an illusion due to lack of sufficient data, but I do not think so. The same groupings of ages crop up repeatedly on all the continents. During these periods the earth seems to leap into violent action. Then it subsides into relative inactivity, gradually gathering internal strength for the next period of activity. I suspect that it is all connected with the periodic buildup and release of the heat generated within the earth by radioactive decay and by periodic changes in the convection pattern within the mantle (as has been suggested by S. K. Runcorn of the University of Newcastle upon Tyne and by McKenzie). When all the physical parameters are known, it may be possible to enter them into the equations of the new branch of mathematics named catastrophe theory, which describes and predicts episodic, discontinuous phenomena, in other words things that change by fits and starts.

To summarize, I consider that the principal factors that have governed the evolution of the continental crust are as follows. First, the irreversible chemical differentiation of the upper mantle, beginning at least 3.8 billion years ago, has resulted in the predominance of continental growth over continental recycling or reworking. Ancient accretion processes were probably analogous to modern ones, involving some form of plate tectonics and partial melting of the upper mantle, often in admixtures with subducted oceanic crust. The very first continental crust, if it was older than 3.8 billion years, was probably made the same way. It is quite easy to imagine a form of plate tectonics involving only primordial basaltic crust. In any case it can be calculated that the earth's upper mantle is large enough to have manufactured all the continental crust through geologic time without itself having undergone any measurable overall change in its chemical composition.

Second, age measurements clearly demonstrate the relatively short time scale of a continental accretion episode. The immense volume of new volcanic and plutonic igneous material is differentiated from upper mantle or subducted oceanic crust within an interval of 200 million years at the most. At the same time the newly produced igneous material itself undergoes chemical differentiation and metamorphic recrystallization deep within the crust, which results in a downward gradient from ordinary igneous rocks into metamorphic gneisses progressively depleted in the incompatible elements. On this view the time-honored distinction between what is igneous and what is metamorphic begins to break down because of the continuous gradation between the two. The gneisses within newly accreted continental segments are the deep-seated, high-pressure and high-temperature equivalents of the higher-level igneous rocks.

Third, we must take into account the permanence of continental crust because of its relatively low density. Continental igneous, sedimentary and metamorphic rocks cannot be pushed back down subduction zones into the upper mantle to a sufficiently great depth to form significant amounts of new igneous rocks by melting. Instead continental rocks are heavily deformed and pushed upward along great thrust faults within the mountain belts at continental margins. Substantial isostatic uplift of the entire crustal segment by thousands of meters occurs when the pressure eases off, just as a cork pushed under water will pop up when a finger is removed.

The above views are certainly not shared by all geologists. I have been much influenced by my own experience in the isotopic dating of ancient rocks, and above all by the writings of many earlier workers, of whom I should like to single out Hurley, A. E. Ringwood of the Australian National University and W. S. Fyfe of the University of Western Ontario. I am aware of the progressive blurring in this article between Archean and recent geological events. The reason is that the similarities over the past 3.8 billion years appear to me to be greater than the differences, although it would be unrealistic to adhere too rigidly to Hutton's doctrine of uniformitarianism. Perhaps if the evidence for a few really major and distinct episodes of continental accretion becomes much stronger, geologists should seriously consider a modified doctrine of episodic uniformitarianism.

Exciting opportunity to see the world. Inquire at your MG Dealer.

MGB is about as close as you're likely to get, without wings, to the exhilaration of flying.

The wide-open sports car is more than a feeling. It's a living legend of quick, strong, dauntless performance. Our first MG won its first competition in 1929, and today MGB holds the Class E Championship in SCCA racing, as it has for five of the last six years.

For 1977, the legend is advanced with even better handling: we've improved the gearing of our rack and pinion steering to reduce steering effort, decreased steering wheel diameter for quicker response and added anti-roll bars front and rear for increased handling stability.

We have also endowed the 1977 MGB with a new, zip-down rear window and new cockpit including a redesigned dashboard for easier read-out of vital engine functions.

See more of your world than any boxed-in car can show you. For your nearest MG dealer, call these numbers toll-free: (800) 447-4700, in Illinois, (800) 322-4400.

MGB: the wide open sports car.
British Leyland Motors Inc., Leonia, N.J. 07605.

Flashlight Fishes

These marine species light up the water with a large organ under each eye containing living luminous bacteria. They use the light to see by, to communicate, to lure prey and to confuse predators

by John E. McCosker

The production and utilization of light by animals, particularly by fishes, has interested biochemists, behaviorists and observers of nature for centuries. Perhaps the most spectacular example of this phenomenon of bioluminescence is found among the "flashlight" fishes of the family Anomalopidae. Whereas most bioluminescent animals normally use their light for only one purpose—be it communication, luring prey, avoiding predators or improving visibility—flashlight fishes employ their light for all these purposes.

The name Anomalopidae is derived from the Greek for "abnormal eye." The name appropriately describes the appearance of the four species in the family: *Anomalops katoptron, Photoblepharon palpebratus, Kryptophanaron alfredi* and *Kryptophanaron harveyi.* In all these species the fishes have a specialized organ below each eye that is filled with light-emitting bacteria, which collectively generate an illumination that is about as intense as the light from a weak flashlight. The fish can obscure the light in various ways, so that in effect they turn it on and off.

These luminous creatures have been shrouded in mystery since the first specimen was observed in the 18th century by the Dutch naturalist Peter Boddaert. The reason little was known about them was that few were found, and the reason few were found is that they are small and reclusive creatures that tend to be active only on dark nights and in fairly deep water. In recent years a good deal more has been learned about the flashlight fishes, and at the Steinhart Aquarium in San Francisco (a division of the California Academy of Sciences) I have been fortunate enough to be able to maintain a living collection of several of them. As a result my colleagues and I have had an opportunity to observe them and to examine their behavior experimentally. Recent advances in bacteriology, underwater observation and the transportation and maintenance of living fish have brought a synthesis of skills that has resulted in a rapidly increasing body of information about the flashlight fishes.

The common name "flashlight fish" evolved from an article in *Science* in 1975 titled "Light for All Reasons: Versatility in the Behavioral Repertoire of the Flashlight Fish." James G. Morin of the University of California at Los Angeles and his five coauthors concluded the article: "Thus the bioluminescent behavioral repertoire of *Photoblepharon* is extensive and varied. It includes many different offensive, defensive and communicative activities, and is especially unusual because only a single type of light organ is involved. The multiplicity of functions suggest that the organ is like a flashlight, whose owner can exercise options in its use."

My first encounter with flashlight fish was incidental to an expedition mounted by the California Academy of Sciences in 1974 and 1975 to collect the rare coelacanth fish at Grande Comore Island in the Indian Ocean. Although my formal training at the Scripps Institution of Oceanography was in ichthyology, I had never seen a living or preserved anomalopid in the large fish collection at Scripps or in the enormous collections at Stanford University and at the California Academy of Sciences. Through the writings of the late E. Newton Harvey of Princeton University, an authority on bioluminescence, I was familiar with the strange and intensely bioluminescent fishes of the genera *Photoblepharon* and *Anomalops,* and I was therefore intrigued when Francis Debuissy, a French veterinarian and scuba diver stationed at Grande Comore, described a luminous fish called *le petit Peugeot* by the few French divers who had seen it. They were the intrepid men who ventured into the tropical Comoran waters at the time of the new moon. The fish were seen only on the darkest nights and were observed abundantly only at depths greater than 30 meters, where they swam above the reef with their light organs exposed and were indeed reminiscent of a small, dark automobile on a country road at midnight.

Debuissy said rather wryly that he had only seen the fish, being unable to capture any of them because his battery light was too weak to immobilize them. (The flashlight fishes, like many other animals, tend to be transfixed by a bright light.) He felt sure that our newer and brighter diving lights would suffice, particularly when the moon was in its darkest phase. Subsequent nights proved him to be correct. I returned to San Francisco in March, 1975, with numerous living and preserved specimens of *Photoblepharon.*

In addition to the work that my colleagues and I have done, the number of investigations of flashlight fishes has increased significantly since the Arab-Israeli conflict of 1967. The connection between these seemingly unrelated events was explained to me by Morin, who in turn heard it from an Israeli ichthyologist with whom he had studied the population of flashlight fish in the Red Sea near the Heinz Steinitz Marine Biological Laboratory, which is on the Gulf of Eilat. During midnight patrols along the coastline of the Sinai Peninsula after the Six-Day War, Israeli soldiers had observed a faint green glowing mass beyond the coral reef. The soldiers, naturally assuming that they had encountered a team of enemy frogmen, responded by discharging explosives in the glowing shoals. To their surprise the result was a beach littered with the bodies of small, dark fish whose heads continued to blaze with a pair of green, glowing patches.

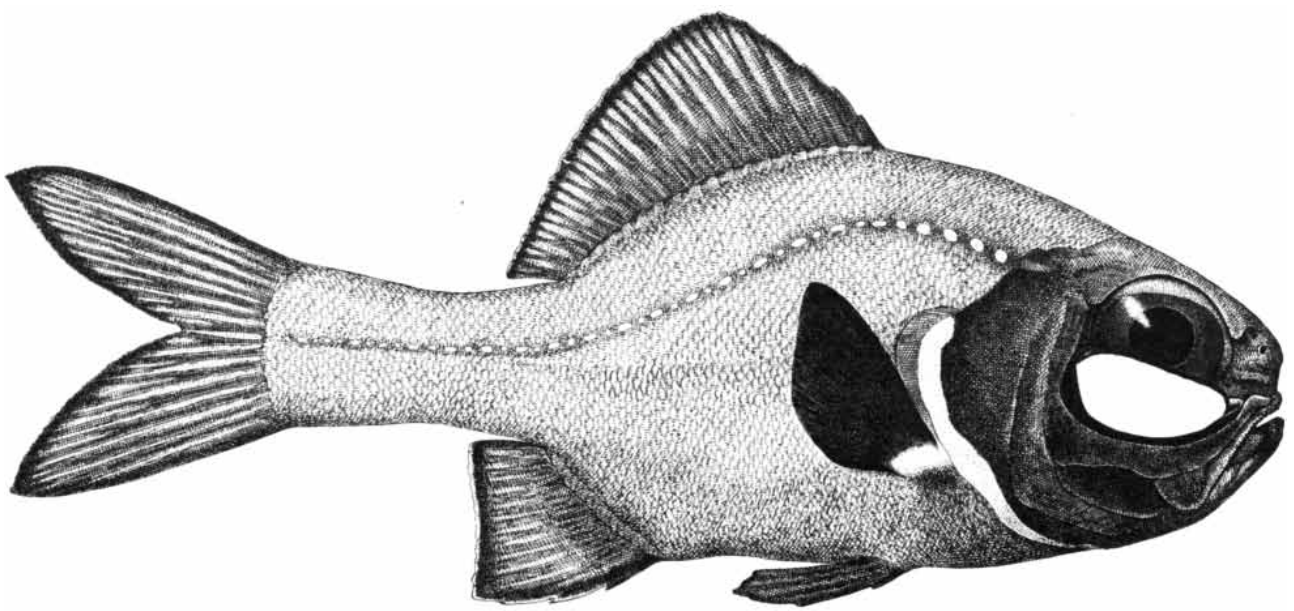
Since 1967 several groups of American and Israeli workers have returned to the Gulf of Eilat to study these little-known fish. They have discovered that during the daytime the flashlight fish live in the dark caves and recesses of the deep reef. On nights with little or no moonlight they venture out of the caves either alone or in small groups to forage along the bottom for small crustaceans

and other plankton of the reef's edge. In the Red Sea they move into shallow water in compact aggregations, which give the appearance of a large green superorganism.

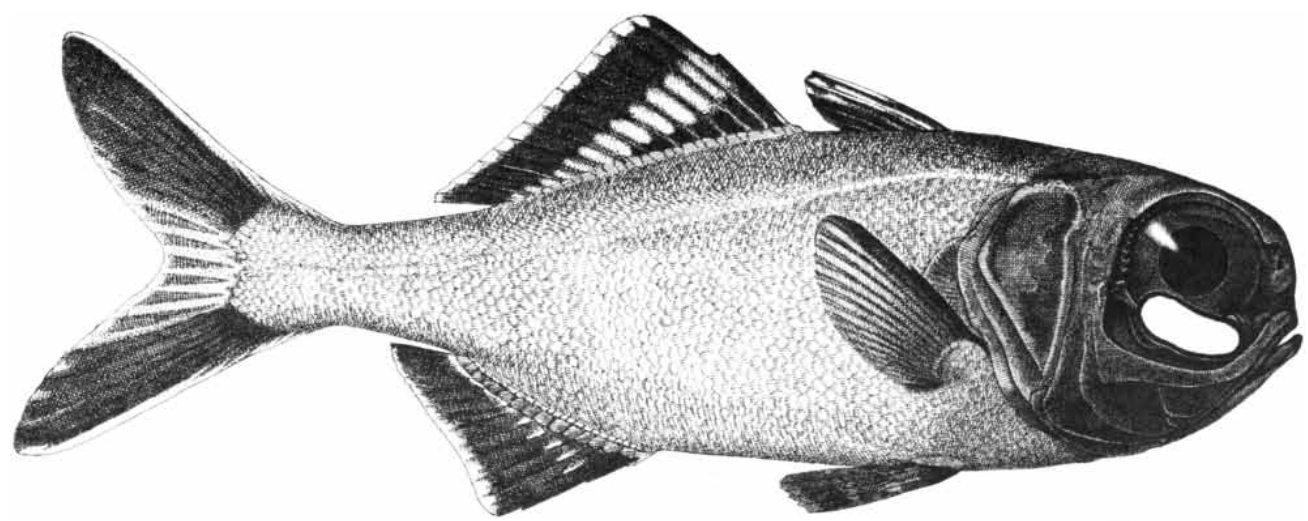
The first person to give a scientific name to the flashlight fishes was Boddaert. In 1781 he named a specimen from Indonesia *Sparus palpebratus,* meaning "the porgy with an eyelid." Boddaert proposed that the function of the large, unusual organs below the eyes was to shield the eyes from injury caused by the coral branches among which the fish lived. In 1803 the French naturalist Count Bernhard Lacépède suggested that the organ served to protect sensitive tissue against sunlight.

It was not until 1900 that the Dutch ichthyologist A. G. Vorderman recorded his observation of light being produced by the living fish. Taxonomists with little comprehension of this odd creature's characteristics subsequently classified it in various genera and families. Eventually the Dutch biologist Max Weber established the genus *Photoblepharon* (meaning eyelid light), in which one species of flashlight fish now resides. The recent discovery of a population of the species in the Red Sea and my discovery of its presence in the Indian Ocean indicate that it is distributed from Indonesia to the Red Sea. It is likely that as night-diving biologists survey the intermediate areas more carefully the fish will be discovered to be widely distributed.

A second species from the Celebes Sea was named in 1856 by the Dutch ichthyologist Pieter Bleeker *Heterophthalmus* (meaning different eye) *katoptron* (meaning mirror, presumably referring to the luminous organ). The species was later placed in the new genus *Anomalops* because *Heterophthalmus* had already been adopted as the generic name



FLASHLIGHT FISH of the species *Photoblepharon palpebratus* is portrayed at about twice the actual size of the mature fish. The light organ under the eye is in the open position. The fish is not only small and dark but also reclusive, living in caves and recesses in deep water.



PORTRAIT OF ANOMALOPS KATOPTRON shows its anatomical differences from *Photoblepharon palpebratus.* The scale of the two drawings is the same. Specimens of *Anomalops* have been collected in Indonesia, in the Solomon Islands and in Japanese waters.

PHOTOBLEPHARON PALPEBRATUS were photographed at night at a depth of about 30 meters in the Indian Ocean by David C. Powell of the Steinhart Aquarium in San Francisco. The fish are active only on dark nights and in deep water, so that information about their habits has been difficult to obtain. Now the aquarium has several living specimens to observe.

ANOMALOPS KATOPTRON is the only other species of flashlight fish that biologists have observed alive. This species differs from *Photoblepharon* in the mechanism of obscuring its light: instead of raising a curtain, it rotates the light organ downward into a pocket. *Anomalops* also are more gregarious, forming schools of some 200 fish that feed near the surface at night.

PINECONE FISH (CLEIDOPUS GLORIAMARIS), also known as the port-and-starboard-light fish, is not a flashlight fish but has a somewhat similar light organ. The organ is fixed in position. Its orange surface acts as a filter to transform the bluish light produced by the organ to a blue-green. Studies at the Scripps Institution of Oceanography indicate that the light organs of each fish are colonized by a different clone of bacteria, as may be true of flashlight fishes.

for a beetle. This gregarious species behaves quite differently from *Photoblepharon* in that it forms schools of as many as 200 individuals feeding on plankton near the surface at night. *Anomalops* is like *Photoblepharon* in being reclusive during periods of daylight or bright moonlight, but its habitat during those times is not known. The species has been collected from several places in Indonesia and the Solomon Islands and is also known from five specimens taken in Japanese waters.

It is somewhat enigmatic that there appear to be two species of anomalopid fishes in the New World, each known from a single dead specimen. A Caribbean form, named *Kryptophanaron* (meaning hidden lantern) *alfredi* (in honor of Alfred Mitchell), was discovered floating on the surface off the coast of Jamaica in 1907 by Ulric Dahlgren of Princeton University. The American workers Charles F. Silvester and Henry W. Fowler described the specimen, which they subsequently lost, as a new genus and species. No other specimens of *K. alfredi* have been found.

In 1972 a small, dark fish with glowing patches under each eye was captured by a shrimp trawler in the relatively shallow water of the Gulf of California. The Mexican captain presented the specimen to W. Linn Montgomery, a graduate student in ichthyology at the University of California at Los Angeles, with the statement that he had not previously seen that species in 35 years of trawling for shrimp. Richard H. Rosenblatt of the Scripps Institution and Montgomery determined that the specimen represented a new species, which they named *Kryptophanaron harveyi* in honor of E. Newton Harvey.

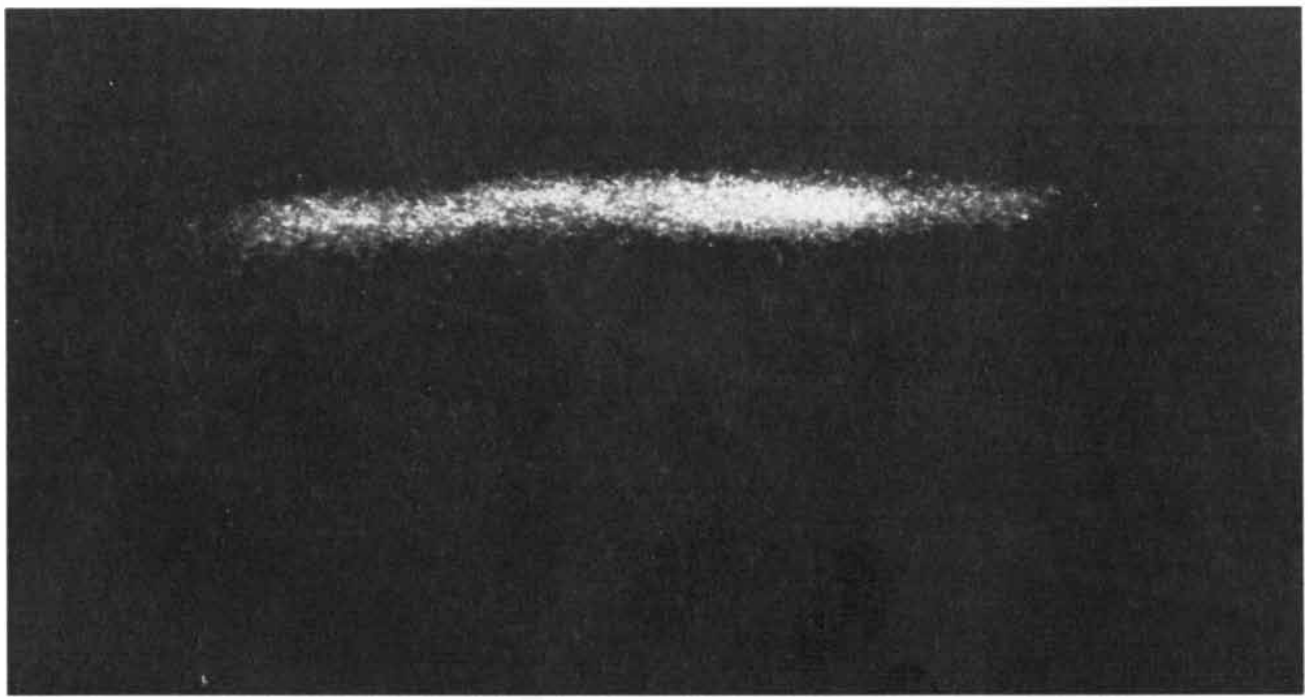
The two forms from the New World are quite similar, suggesting that they have a common ancestry predating the formation of the Central American land bridge. That event, which is believed to have occurred between one and three million years ago, separated the aquatic populations of the Caribbean and the eastern Pacific and provided a basis for their subsequent differentiation into separate species.

When one considers the numerous collections of fishes that biologists and commercial fishermen have made in the Caribbean and the Gulf of California, it is puzzling that *Kryptophanaron alfredi* and *K. harveyi* are known only from single specimens. The rarity of these fishes must be attributable to the habitat they presumably prefer, namely reefs that are below the depths where most scuba divers go and rocky areas that are relatively inaccessible to collection with nets. I might add on the basis of personal experience that for sane biologists deep diving in tropical seas on dark nights with one's diving light turned off is rarely practiced and never enjoyed.

Anomalopids are distinctive in several ways. Their light organ produces what is perhaps the most intense light known to come from a multicellular luminescent organism. An anomalopid fish apparently employs the light not only for attracting prey, confusing predators and communicating with other members of the species but also as a flashlight to see what is in front of it.

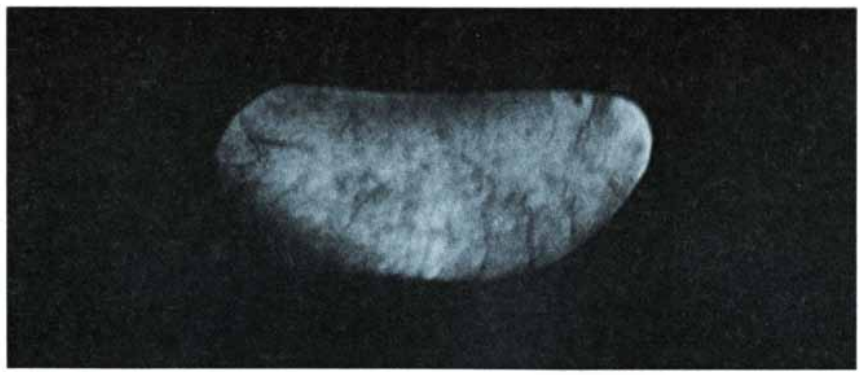
The organ is cream-colored on its outer surface and black on its inner and upper surfaces. (If these surfaces were not black, the light would certainly blind the fish itself.) It is easy to experiment with the organ because it can be removed surgically without difficulty and will continue to glow for eight hours or more after removal.

The glow of the light organ in *Photoblepharon* and *Anomalops* comes from



SCHOOLING FLASHLIGHT FISH were photographed in the Red Sea at night by the light of their own luminous organs. This school of *Photoblepharon palpebratus* contained about 30 fish. Since each individual turns the light off frequently in a blinking pattern by raising a curtain of skin over the organ containing the light-emitting bacteria, it is unlikely that all the lights were "on" in this group of fish.

DAYLIGHT VIEW of the place in the Red Sea where the school of flashlight fish was photographed shows that the far edge of the coral reef (about 12 meters from the camera) was the area where the fish congregated at night to feed on plankton in the shallow water above the reef. In other parts of the world *Photoblepharon* are seldom found abundantly except at depths of 30 meters (100 feet) or more.



LIGHT ORGAN of a flashlight fish was removed surgically and photographed in its own light by means of a time exposure. The bacteria in the organ continue to emit their blue-green light for eight hours or more after the organ has been removed from the fish. This organ, which in the fish has a black inner surface, is from a living specimen of *Photoblepharon palpebratus*.

bacteria that are symbiotic with these fishes. For the bacteria the light is a by-product of metabolism, as heat is a by-product of metabolism in warm-blooded animals. Enormous numbers of these bacteria (some 10 billion per milliliter of fluid in the organ) are packed into special nutritive compartments within the fishes' light organ.

In the Comores my colleague Michael D. Lagios and I had hoped to culture the bacteria in various nutritional mediums. Our thought was that it would be prudent to return to the Steinhart Aquarium with bacteria in culture as well as in the living fish; if the lights "went out," we could reinfect the organs and turn the lights back on. We were dismayed to find that we were unable to culture the bacteria. Later we learned that both Harvey and the Japanese workers Yata Haneda and F. I. Tsuji had failed in similar attempts. Haneda and Tsuji concluded that they were dealing not with normal bacteria but with prokaryotic (lacking a nucleus) cell-like organisms they termed "bacterioids."

A possible explanation for this perplexing failure of efforts to culture the bacteria has been advanced by Kenneth Nealson of the Scripps Institution. His hypothesis is that the light-generating microbe of *Photoblepharon* is so specialized that it is an obligate symbiont: it cannot survive outside its host.

Nealson proposed that this bacterium, unlike more independent bacteria, can only partially metabolize the glucose of its host. Therefore in culture it probably generates toxic concentrations of pyruvic acid, a product of the incomplete oxidation of the sugar. The enzymes the host provides to break down the pyruvic acid are absent in the culture medium, and so the medium becomes too highly charged with pyruvic acid for the bacterium to survive. The phenomenon has been observed in similar symbionts (but not obligate ones). Fortunately the lights of the specimens at the Steinhart Aquarium have remained lit since March, 1975. Indeed, at the time of writing the lights had become brighter rather than weaker as the fish grew accustomed to the conditions in the aquarium.

When I first dissected a specimen of *Photoblepharon* to examine its sexual state and the contents of its stomach, I was surprised to discover a considerable amount of a fatty substance dispersed throughout the coelom: the space between the body wall and the digestive tract. This substance, which has the consistency of lard at room temperature, is probably an energy store. Presumably it provides energy for both the fish and the bacteria during the lunar periods when *Photoblepharon* is less active than it is at other times and is fasting.

The low density of the fatty material would make the fish excessively buoyant if it were not for a compensating reduction in the volume of the gas bladder that is typically found in fishes as a means of controlling buoyancy. The gas bladder in related but nonluminous fishes is much larger. For biologists the reduced gas bladder of *Photoblepharon* is a happy evolutionary circumstance, since it means that specimens can be brought up from deep water without the otherwise common problem of rupture of the bladder from expanding gas.

Flashlight fishes are also unusual among bioluminescent organisms in that their light is on more than it is off. To turn off its light *Photoblepharon* simultaneously raises a black curtain over each light organ, completely blocking out the light. The rate of blinking varies with the water temperature and the conditions of the fish's environment. When live brine shrimp (a food for *Photoblepharon*) are added to an aquarium tank containing several of the fish, a rapid blinking sequence ensues, suggesting that the fish are somehow communicating the information to one another.

In using its light organ to deal with predators *Photoblepharon* relies on a sudden flash of the light that so startles the predator that *Photoblepharon* gains time to escape. Moreover, our studies of the living fish in our aquarium have disclosed a most interesting adaptation that the species presumably employs to confuse predators. Each time a swimming *Photoblepharon* changes direction, it turns off its light. Shortly afterward it turns the light on again, but from the viewpoint of a predator the position at which the light of the prey reappears is unpredictable.

This behavior is continuous. Presumably it makes *Photoblepharon* difficult for a predator to track. Morin has observed the behavior in a natural setting, where he has recorded an average frequency of 75 blinks per minute with an average duration of 160 milliseconds for each blink. His term for the behavior is "the blink and run."

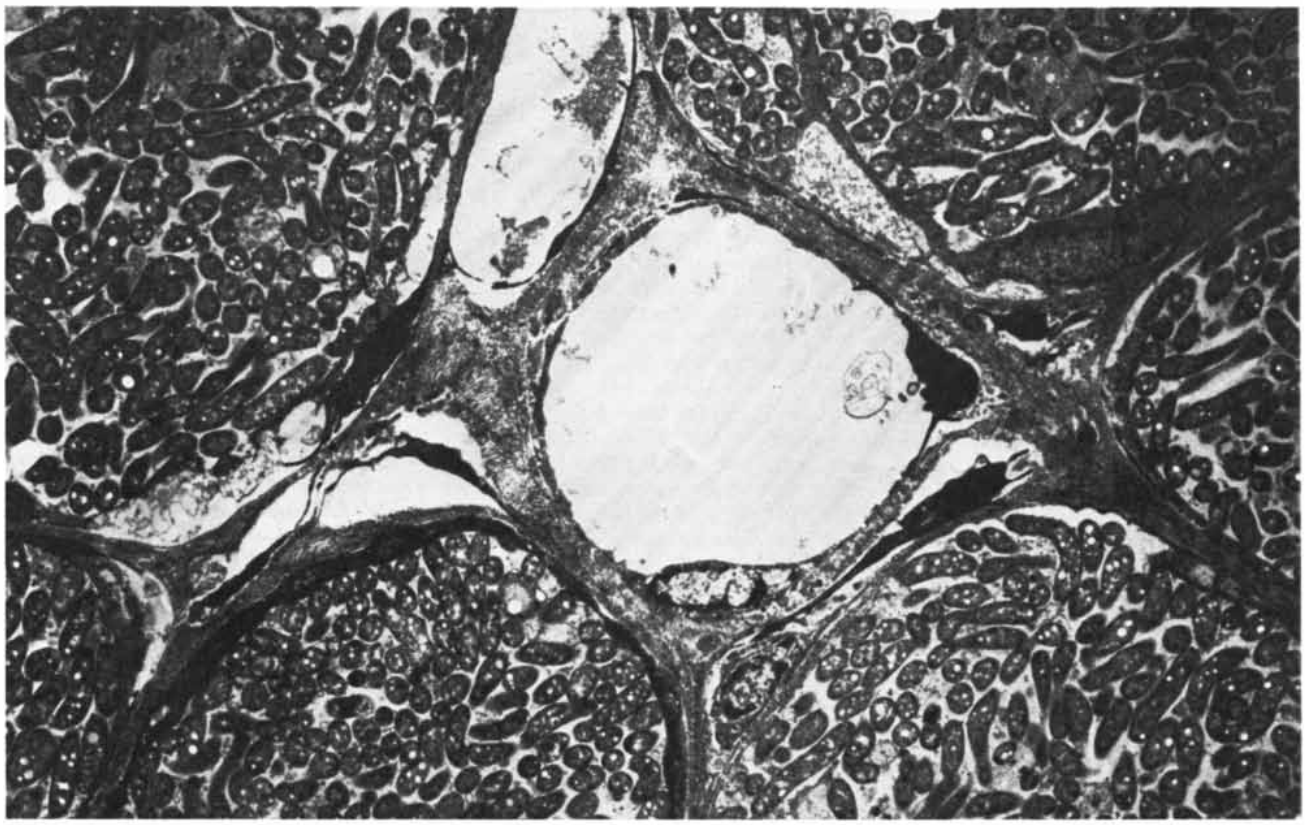
With photometric techniques Morin and his colleagues have identified certain basic patterns of the bioluminescent activity of *Photoblepharon*. The most obvious pattern is an infrequent blinking. (Most other bioluminescent organisms flash rather than blink.) This pattern is typical of undisturbed *Photoblepharon* at night. They blink an average of 2.9 times per minute, and each blink has an average duration of 260 milliseconds.

A second pattern suggests that a more or less daily rhythm exists in the spontaneous blinking frequency of fish kept in continuous darkness. The blinking rate is much higher during the daytime hours than it is during the night, and the duration is somewhat longer. The average frequency is 37 blinks per minute and the average duration of the blinks is 800 milliseconds.

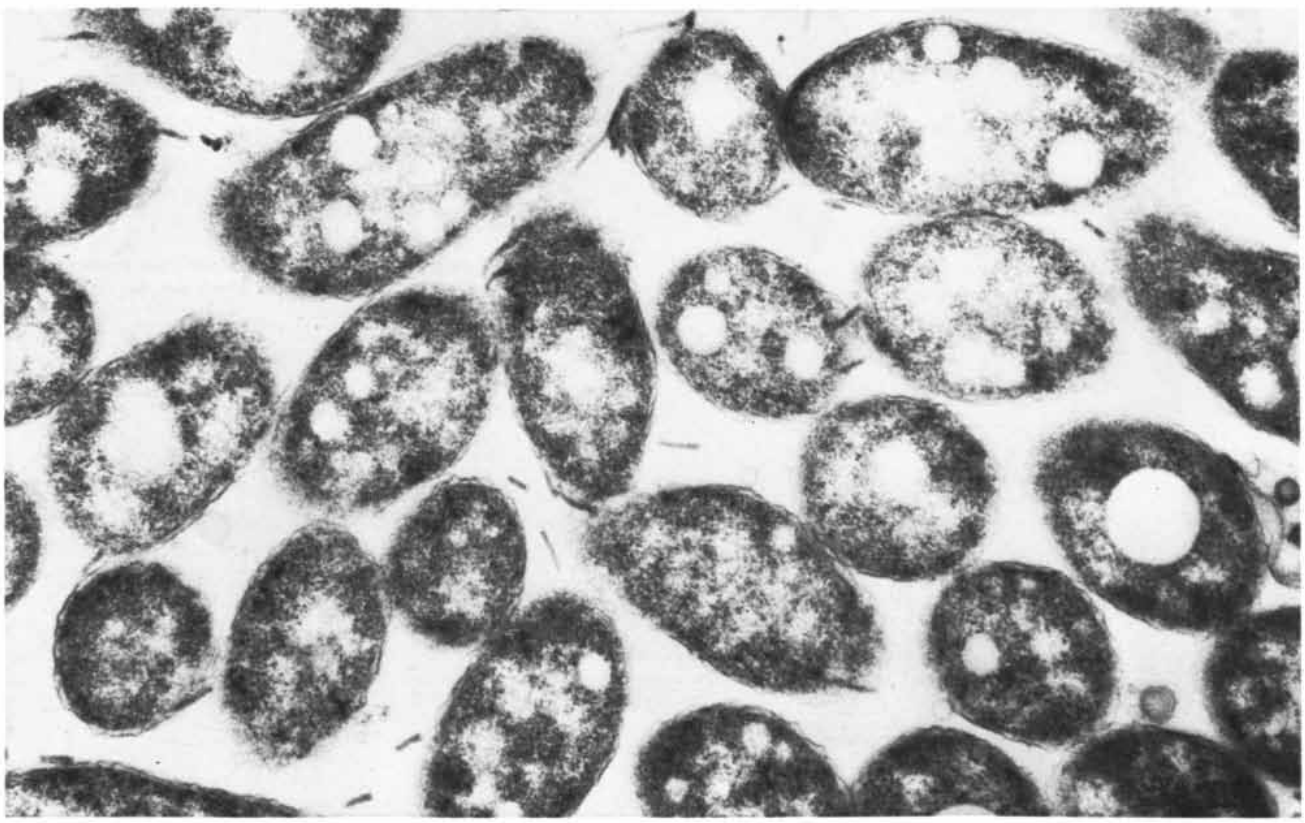
Several specimens of *Anomalops* that arrived at the Steinhart Aquarium in a weakened condition, with their lights extinguished, provided an opportunity to test the hypothesis that flashlight fishes use their light for seeing. When live brine shrimp were put in the tank occupied by the specimens, the lightless fish largely failed to discover the shrimp. The shrimp were eaten immediately when a light approximately equal in intensity to the natural light of the fish was turned on in the room.

We tried without success to rekindle the natural light of these fish. Our first attempt was to reinfect the *Anomalops* organ with a culture of bacteria from glowing specimens of *Photoblepharon*. It did no good to put light and dark fish in the same tank, even though tests of the water showed that many photobacteria were present. Attempts by Edward E. Miller of the Steinhart Aquarium staff to transfer bacteria from *Photoblepharon* to *Anomalops* with a hypodermic needle were also unsuccessful.

The failure of all these efforts suggests that the two species provide different environments for the bacteria. The answer



LIGHT-EMITTING BACTERIA in a specimen of *Photoblepharon palpebratus* appear in an electron micrograph made by Michael D. Lagios of Children's Hospital in San Francisco. The enlargement is about 5,700 diameters. The bacteria are the elliptical and round structures inside the compartments surrounding the central vessel, which provides nourishment to the bacteria from the blood of the fish.



CLOSER VIEW of the light-emitting bacteria shows them at an enlargement of 28,500 diameters. The slender, dark projections are flagella; their function in this species is not clear. Inner circles are artifacts resulting from evaporation of granules by the electron beam.

may be more complicated, however, as Nealson and Edward Ruby discovered at the Scripps Institution in work with the bioluminescent pinecone fish. Their preliminary findings indicate that each pinecone fish is colonized by a different clone of bacteria and so is in effect a biological island.

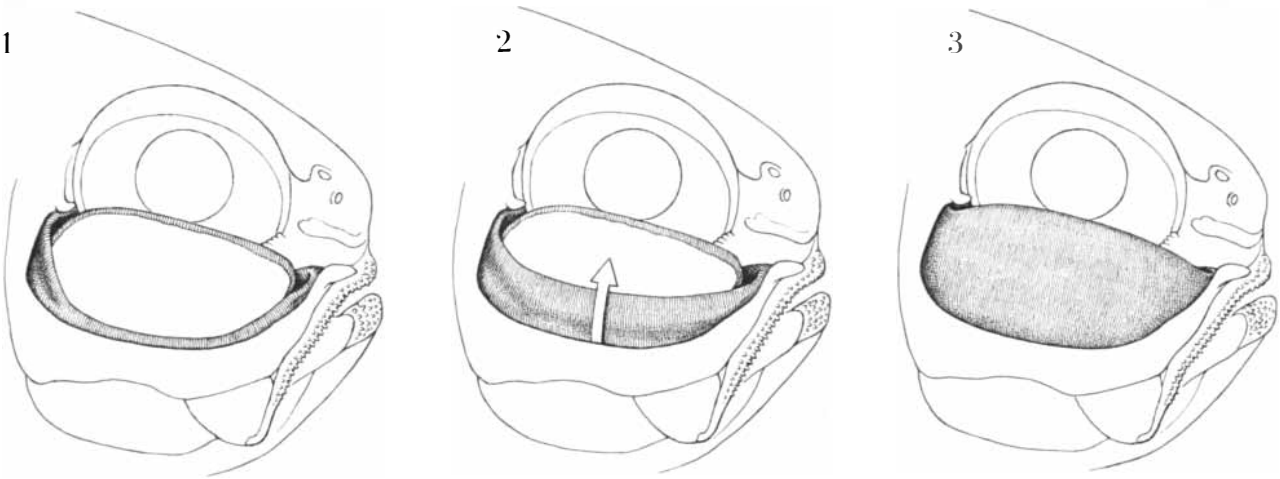
Man has taken advantage of the fact that the flashlight fish lures prey with its light organ. Although the fish captures prey smaller than itself by attracting the prey with its lights, the fishermen of the Banda Islands in Indonesia learned that larger fishes are also attracted to the light. Harvey noted in 1922 that Banda fishermen removed the organ from the fish and attached it to their lines above the hook to act as a lure. The organ remained luminous for many hours as the symbiotic bacteria continued to glow. Other fishermen in Indonesia have learned to take advantage of these alluring properties without harming the flashlight fish. By suspending below their canoe a perforated length of bamboo enclosing a dozen or more living *Photoblepharon* they can fish each night with reusable lures.

Although the light organ is essentially the same in all anomalopid fishes, the way in which it is operated differs considerably. *Anomalops* obscures its light by rotating the light organ downward into a darkened pocket. (The organ is hinged at the front.) *Photoblepharon* extinguishes its light by raising a black curtain over the organ. The curtain is much like an eyelid.

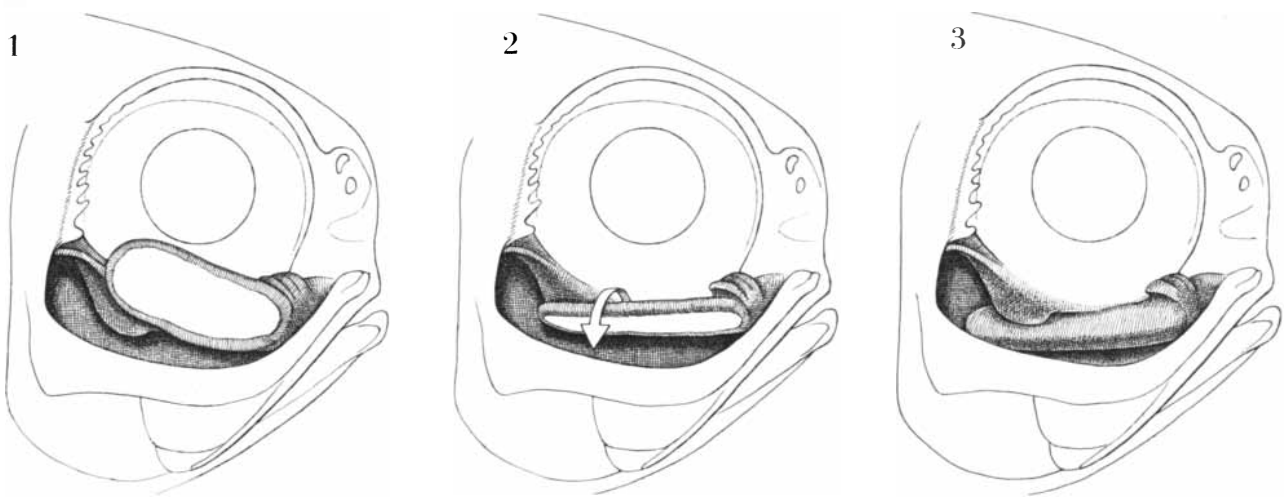
The reason for these differences may be indicated by the recently collected specimen of *Kryptophanaron.* Rosenblatt and Montgomery have noted that the light organ of the fish can be rotated but that a membrane is also associated with the organ. They propose that *Kryptophanaron* controls its light by both rotation and a shutter mechanism, with the rotational movement serving to interrupt the light for fairly long periods and the shutter mechanism accomplishing rapid blinking. Rosenblatt and Montgomery believe both mechanisms may have been present in the ancestor of *Kryptophanaron* and the common ancestor of *Photoblepharon* and *Anomalops.*

If this supposition is correct, the membrane of *Kryptophanaron* or its ancestor provides the genesis for the shutter mechanism employed by *Photoblepharon,* a species that has retained the presumed ancestral habit of staying near the bottom. *Anomalops* has evolved in such a way that it leaves deeper water to feed on plankton near the shore at night. Behavioral accommodations accompanying this evolutionary change include the formation of large schools and a continuous rapid blinking. Both adaptations make it difficult for a predator to single out a lone prey.

Rosenblatt and Montgomery have speculated as follows about the change. "In clear water, near the surface, away



OBSCURATION OF LIGHT ORGAN by *Photoblepharon palpebratus* **is depicted. The mechanism resembles an eyelid, except that in its normal position (*1*) it is folded below the light organ and that the fish raises it (*2 and 3*) in order to black out the light briefly.**



DIFFERENT MECHANISM serves to obscure the light of *Anomalops katoptron*. Its light organ is hinged at the front by a muscle (*1*). The fish employs the muscle (*2 and 3*) to rotate the organ downward into a pouch. Both species of fish normally blink several times per minute.

We go further than anyone to please our passengers.

In terms of distance: We go to 89 far-flung cities in 61 countries on 6 continents.

In terms of time: There are parts of the world to which only Pan Am's long-range 747 SP can take you non-stop. (Tokyo from Los Angeles and New York. Auckland from San Francisco. San Francisco from Sydney. And Bahrain from New York.) Thereby cutting hours off your trip.

In terms of comfort and service: We have the world's largest fleet of big, roomy 747s. Many of which have our exclusive upper-deck dining room in first class.

We offer both first class and economy passengers a choice of meals, a choice of movies, and the choice of conversing in a language other than English. (All our flight attendants are bi-lingual.)

We have one of the most sophisticated reservations systems around. And people experienced in helping sophisticated and unsophisticated travelers all over the world.

In overall terms: We're an airline geared to bringing Americans to the rest of the world as easily and comfortably as possible.

And the rest of the world to America.

PAN AM
America's airline to the world.

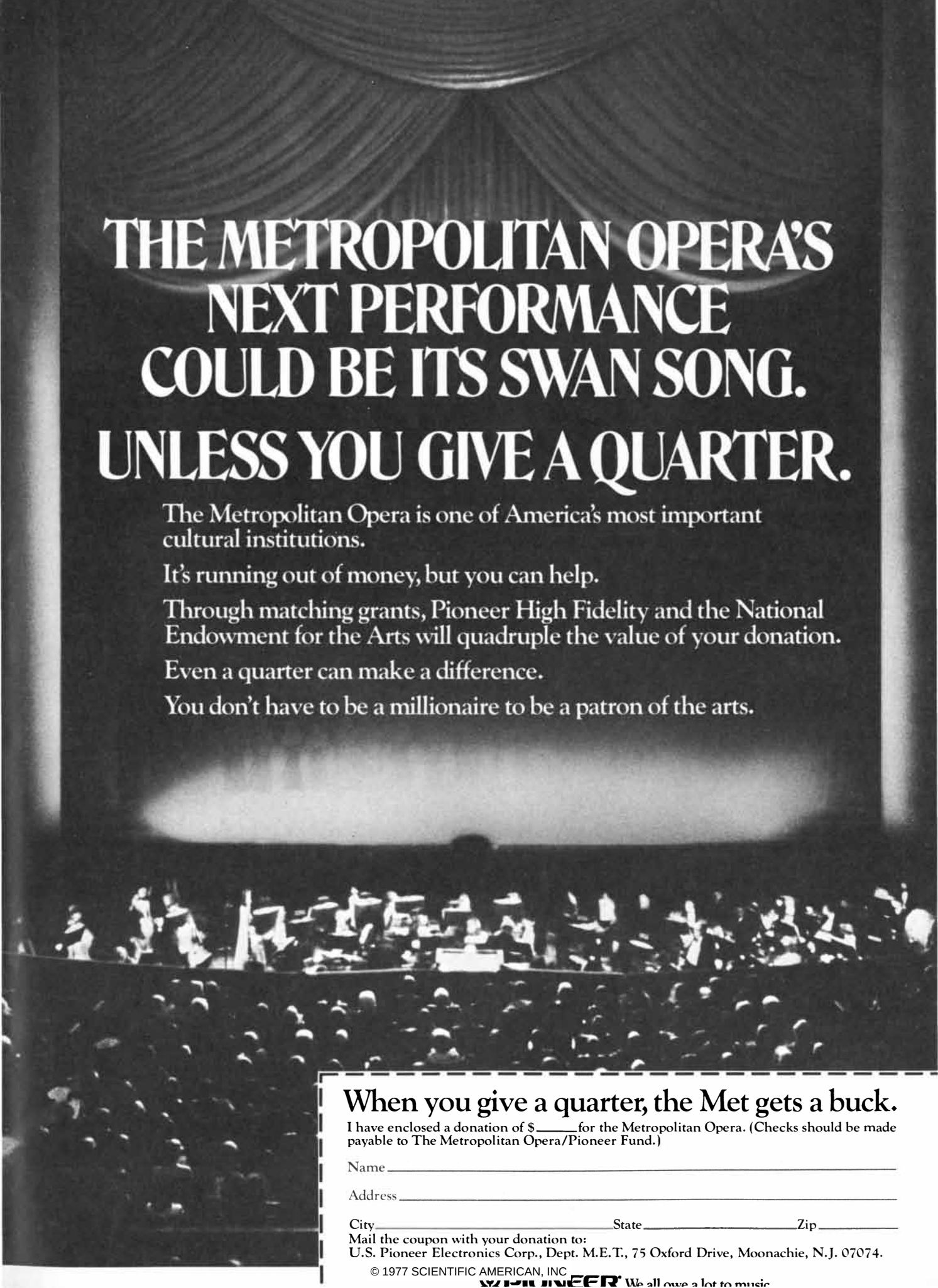
See your travel agent.

from obstructions of the reef, a predator would be able to see the luminous organ, and begin its rush, from a considerable distance. A rapid blink and short dash [as performed by *Photoblepharon*] would require only a small course correction on the part of the predator when the light came on. However, in *Anomalops* the light is occluded for a relatively long period, and all members of the school are blinking. This mechanism should be much more effective in lessening the opportunity of a predator to fixate visually on a single individual. Intermittent flashing in *Anomalops,* then, lessens exposure to predators caused by use of the light organ in feeding and maintaining the school."

Many interesting questions about the anomalopids remain to be explored. What is the behavior of *Kryptophanaron,* the fish that has yet to be observed alive? How do the bacteria in the light organs compare in the two species of *Kryptophanaron* and with *Photoblepharon* and *Anomalops?* Have the symbiotic bacteria evolved at a rate different from that of their hosts? How do larval flashlight fishes acquire their bacteria? Is the photobacterium carried within the fertilized egg, remaining dormant until the time when bioluminescence becomes important to the survival of the juvenile? The answers to these questions will require a combination of persistence and luck in observing the fishes in their natural environments and in experimenting with them in the aquarium.

Finally, it is appropriate to describe here a far-reaching experiment that has been suggested to me by Nealson and by J. Woodland Hastings of Harvard University for a beneficial application of our research on these rather esoteric fishes. Nealson and Hastings are impressed by the remarkable purity of the colonies of bacteria in the light organs of flashlight fishes, that is, by the absence of competition from other bacteria. They suggest that it testifies to the existence in the fish of an extremely effective immune system of the type that in other animals wards off foreign organisms invading the body.

Nealson and Hastings suggest that a genetic engineer might be able to replace certain unnecessary genetic information in the bacterium with information that controls operations useful to the human organism. An example would be an enzyme system that would trigger the production of insulin in a person who is diabetic. By linking such a system to a photobacterium the geneticist would have a visual measure of the purity of his culture. When the light is on, one can assume that the culture is free of contamination. I marvel at the suggestion. During my 50-meter dives in the Indian Ocean on moonless nights, I never suspected that the flashing lights I was following might lead to such an end.


THE METROPOLITAN OPERA'S NEXT PERFORMANCE COULD BE ITS SWAN SONG.
UNLESS YOU GIVE A QUARTER.
The Metropolitan Opera is one of America's most important cultural institutions.
It's running out of money, but you can help.
Through matching grants, Pioneer High Fidelity and the National Endowment for the Arts will quadruple the value of your donation.
Even a quarter can make a difference.
You don't have to be a millionaire to be a patron of the arts.
When you give a quarter, the Met gets a buck.
I have enclosed a donation of $_____ for the Metropolitan Opera. (Checks should be made payable to The Metropolitan Opera/Pioneer Fund.)
Name
Address
City State Zip
Mail the coupon with your donation to:
U.S. Pioneer Electronics Corp., Dept. M.E.T., 75 Oxford Drive, Moonachie, N.J. 07074.
© 1977 SCIENTIFIC AMERICAN, INC
PIONEER We all owe a lot to music.

The Earliest Maya

Archaeological excavations in Belize in Central America have pushed back the origins of the Maya to 2500 B.C. The buildings and pottery uncovered clearly foreshadow the splendor of the Classic Maya period

by Norman Hammond

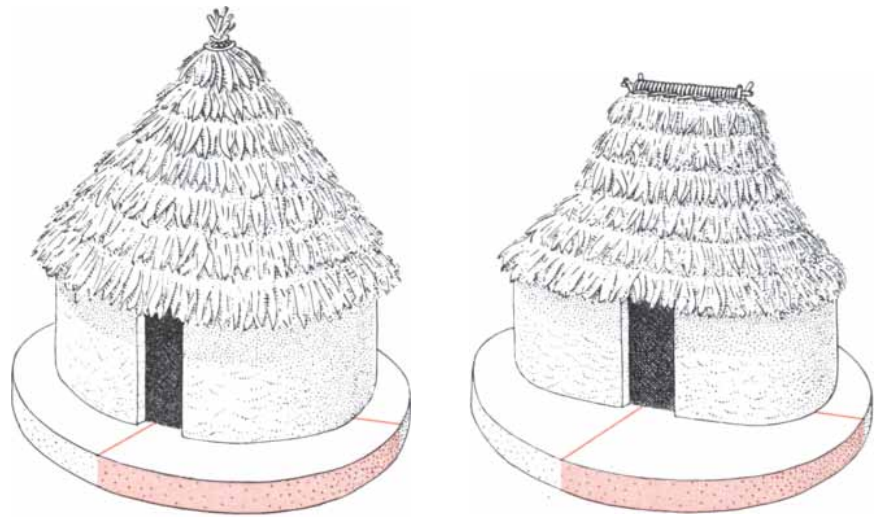
The collapse of Maya civilization, culturally the most advanced of any in the pre-Columbian New World, has inspired almost as many explanations as there are students of American prehistory. An unanswered question of equal importance is how Maya civilization first arose. Until recently that question has received relatively little attention, but its cogency is now greatly increased. Work over the past two seasons on the eastern margin of the Maya area in Belize (formerly British Honduras), often regarded as a backwater, has pushed the beginnings of the Maya Formative (or Preclassic) period back by more than 1,500 years, from about 900 B.C. to perhaps as long ago as 2600 B.C. The new findings place this Early Formative Maya culture among the oldest settled societies in Mesoamerica or, for that matter, in the entire New World.

The term Mesoamerica is often mistakenly thought to be synonymous with Central America: the region extending from southern Mexico to Panama. The term is actually much narrower. Prehistorians define it as the culturally unified area that in pre-Columbian times embraced southern Mexico (including Yucatán), Guatemala, Belize and the western parts of Honduras and El Salvador. The last and politically the most developed of the Mesoamerican civilizations was the Aztec, which the Spanish conquistadors overthrew in 1521. The Aztec capital, Tenochtitlán, was situated where Mexico City stands today, and Aztec political power was centered on the high plateau of Mexico.

Not all Mesoamerican civilizations had this highland focus. The Olmec, one of the earliest of the complex societies in the region, built major ceremonial centers on the low-lying coastal plain of the Gulf of Mexico; examples are San Lorenzo and La Venta. At the same time the Olmec zone of cultural influence and Olmec trade extended into much of the high plateau.

To the east of both the Aztec and the Olmec area in Mexico lies the peninsula of Yucatán and, south of the peninsula, the northern lowlands of Guatemala (the Petén) and Belize. In this southern area during the first millennium of the Christian Era what are regarded as the outstanding characteristics of Maya civilization emerged. This was the start of the Maya Classic period. Extending from about A.D. 250 to 900, the Classic period witnessed the development of mathematics, nontelescopic astronomy and calendrical calculations more advanced than any in other parts of the New World. The data were expressed in a hieroglyphic script utilizing more than 800 characters, many of which still defy decipherment. Paralleling these purely intellectual achievements the Classic Maya civilization gave employment to a school of vase painters as talented as those of Classical Greece and to architects whose great temple pyramids and sacred precincts still amaze the visitor of today.

In common with a number of other students of Mesoamerican prehistory I have been concerned in recent years with the factors underlying the rise of Maya civilization. My own work has taken the form of a series of field studies examining the demographic and economic aspects of the Maya Formative period. The geographical focus of our project, established jointly by the British Museum and the Centre of Latin American Studies at the University of Cam-



MAYA ARCHITECTURE OF THE EARLY FORMATIVE PERIOD is shown in reconstructions based on the partial excavation at Cuello, a site in northern Belize, of two plaster-covered earth platforms that were the foundations for timbered superstructures. The two small superstructures, one circular and one oblong, are alternative conceptions of the timber-and-thatch building that occupied the older of the two platforms, which was probably circular in plan and some six meters in diameter. The platform was built directly on an old soil surface that included burned wood suitable for carbon-14 analysis; the date of construction appears to fall

bridge in 1973, was inspired by the late Sir Eric Thompson, who pointed out the archaeological importance of northern Belize. Physiographically the region is a continuation of the lowlands. To the south and west the Petén, the Classic Maya heartland in northern Guatemala, forms a rain-forest zone with numerous rivers and lakes. To the north the Yucatán peninsula forms an arid zone: a karst landscape of sinkholes, caverns and underground streams. Northern Belize lies within the rain-forest zone but borders on the arid zone. Two of its major valleys are those of the Rio Hondo and the New River. Following the two rivers upstream, one moves south and west toward the Petén heartland [*see illustration on page 119*].

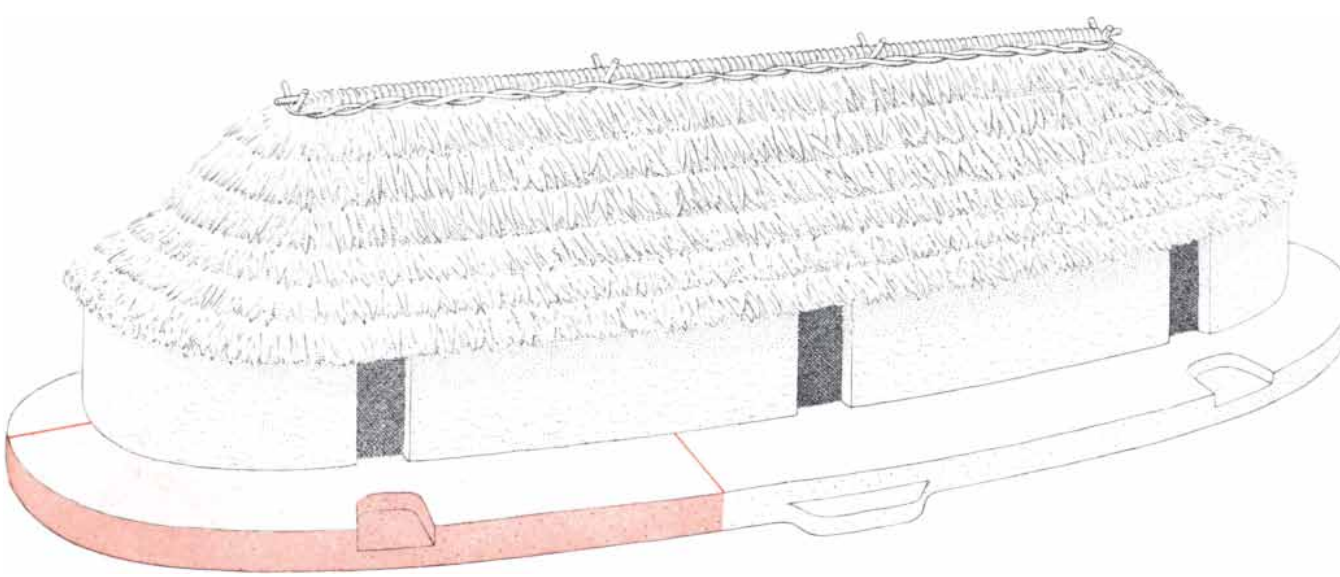
A second reason for selecting the area is that decades of intermittent exploration (mainly by Thomas Gann, a physician and amateur archaelogist, between 1896 and 1936) have uncovered a number of major and minor archaeological sites that evidently were occupied in Late Formative times, a period extending from about 300 B.C. to A.D. 250. Finally, the people of Belize are greatly interested in their country's past, and the government encourages archaeological research. As a result we have enjoyed the friendliest cooperation not only with the government, through the Archaeological Commissioner, Joseph O. Palacio, but also with such representatives of the private sector as Belize Sugar Industries, Maya Airways and G. A. Roe Insurance Services, all of which are generous sponsors of our work.

A line drawn along the 18th parallel defines the southern boundary of our research area in northern Belize; the eastern boundary is the Caribbean, and the western and northern boundaries are the Mexican border along the Rio Hondo and across Chetumal Bay. The entire research area covers some 3,500 square kilometers. During our first field season, in 1973, we concentrated on locating as many archaeological sites as possible. These we classified in terms of size and complexity: scattered, informal residential clusters; formal clusters grouped around a central plaza, and ceremonial precincts, minor or major, surrounded by residential areas of varying extent. We also established a rough regional chronology based on the pottery uncovered in test excavations at several sites. The pottery chronology was achieved by comparing styles and style changes with those already established for pottery from elsewhere in the Maya region, either discovered in association with dated inscriptions (which first appear in about A.D. 250) or dated by means of carbon-14 analysis.

All together we plotted some 60 sites, most of them last occupied during the Classic period and most located in the higher and drier western part of the region: areas of raised ground between the Rio Hondo and the New River and between the New River and Freshwater Creek. Three sites included ceremonial precincts ranging in size from medium to large, and one of them, Nohmul, was surrounded by more than 20 square kilometers of residential settlements. Each of the three ceremonial precincts included a large elevated acropolis, building foundations surrounding large plazas, several tall temple pyramids and at least one parallel-sided court where the sacred ball game had been played.

All three of the major sites, Nohmul, Aventura and El Pozito, are located along the same stretch of high ground, the ridge between the Rio Hondo and the New River. It is easy to visualize the three as the capitals of separate Classic Maya principalities. The lesser sites that surround them are comparable to the towns and villages around the cathedral cities of medieval Europe. One of these lesser sites, located almost exactly midway between Nohmul and El Pozito, stands on land owned by the Cuello family, who hospitably gave us permission to investigate it.

We named the site after its owners. Examination in 1974 of a partially destroyed mound at the site revealed pottery of an unfamiliar type that we had also encountered in the lowest levels of our test excavations at Nohmul and elsewhere. At those sites the unfamiliar pottery was associated with recognizable wares of the Maya Middle Formative period (from 900 to 300 B.C.). The Cuello sherds, however, had no such associations. Was the unfamiliar ware perhaps even older and indigenous? Or had



between 2500 and 2400 B.C., making the platform the oldest architectural endeavor known in Mesoamerica. That a superstructure of some kind once stood on the platform is evident from numerous postholes (*see illustration on page 130*). The larger and more recent platform (*above*) was apparently oblong, some five meters wide and more than 10 meters long, with rounded ends. At the point where one end begins to curve a niche was built into the side of the platform; evidence that the plaster lining of the niche had been renewed suggests that it served as a step for mounting the platform from the patio in front. The platform was probably built between 1700 and 1500 B.C. Considerations of symmetry suggest the second niche shown here, and the small size of both niches suggests the centrally located additional step. The restoration, however, is conjectural; indeed, only those parts of both platforms shown in color have been excavated.

AGES OF EARLY POTTERY IN THE NEW WORLD vary from site to site over a range of 1,200 years. In Mesoamerica (*colored rectangle*) pottery from the Tehuacán valley is estimated to be 4,800 years old and pottery from Puerto Marquez 5,200 years old. Pottery superior in quality to both, recently found in Belize, includes some that is 4,600 years old. One site in the U.S., Stallings Island, has yielded pottery nearly that old, but the most ancient New World pottery now known is from sites in South America: the pottery from Real Alto averages 5,000 years in age, and that from Valdivia and Puerto Hormiga respectively averages 5,600 and 5,800 years.

MAYA LOWLANDS consist of three parts of Mesoamerica: the Yucatán peninsula of Mexico, the Petén region of Guatemala and Belize. Classic Maya civilization reached its apogee in about A.D. 700 in the tropical forest of the Petén; Tikal was a major Classic ceremonial center. Until recently the earliest lowland pottery known was from sites on the Rio Pasión. Ascribed to the Maya Middle Formative period, the wares are dated at about 900 B.C. The earliest lowland pottery now known is that of the Swasey ceramic complex, unearthed at Cuello. Carbon-14 dates for Swasey wares extend from 1250 to 2600 B.C. Swasey pottery, unearthed at four other sites in northern Belize, has also been found at two sites in Yucatán: Becan and Mani.

it been made elsewhere in the first millennium B.C. and reached the sites in northern Belize as an import?

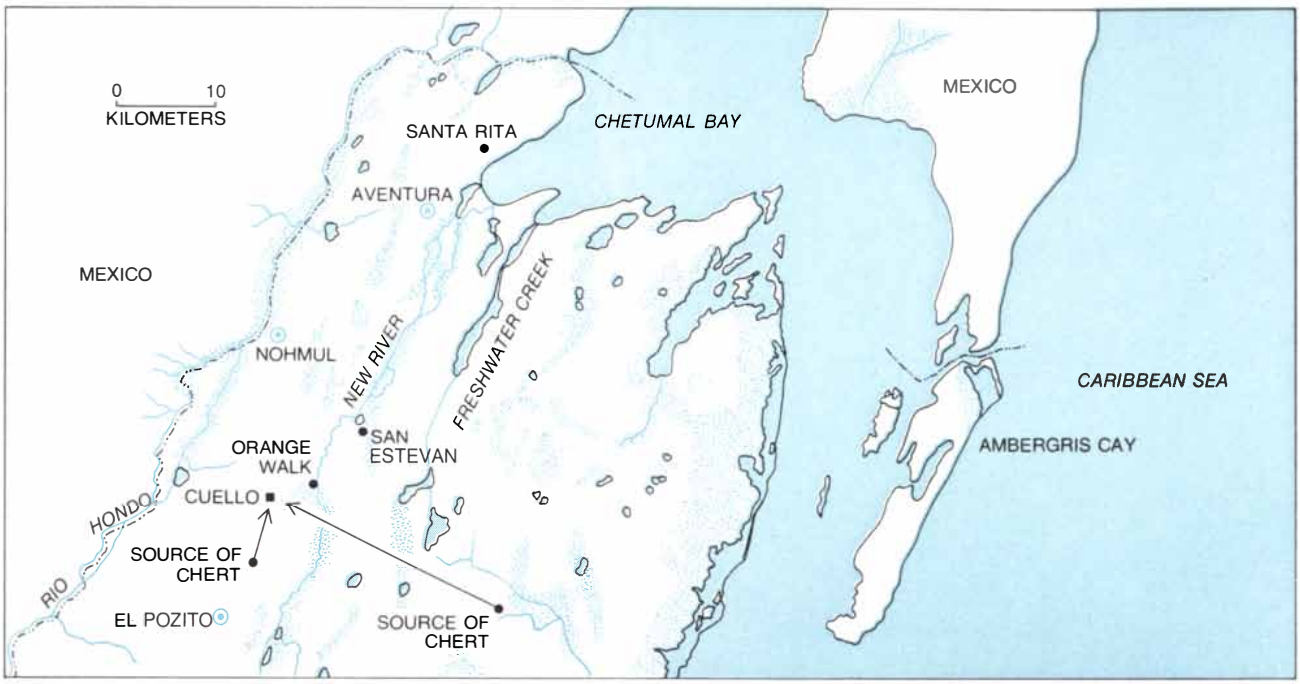
The question could be answered only by excavation. Moreover, the location of Cuello midway between two major ceremonial sites suggested other possible benefits that might accrue from further investigation. What was the nature of the contacts between Nohmul and El Pozito in the Classic period? If the major centers were indeed the capitals of principalities, where was the frontier between them? With questions of this kind in mind we decided to carry out a small-scale excavation at Cuello during our next field season, in 1975.

Cuello lies five kilometers west of Orange Walk Town, capital of the district of Orange Walk. The most obvious part of the uncleared site is a small ceremonial precinct of the Classic period consisting of two linked plazas, each with a small temple pyramid. No stone superstructures survive anywhere at the site, and it is probable that not only the temples but also the many residences and other structures were built of perishable materials: timber frames with palm-thatch roofs.

To the south of the ceremonial precinct we located a series of large platforms. One of them, about four meters high and 80 meters long, particularly attracted our attention because on it was a small temple pyramid, about eight meters high. The pyramid was simply too small for such a large platform. This architectural discontinuity suggested that the temple was a late addition, perhaps built after the platform had been abandoned for some time. On the eroded sides of the temple pyramid we found potsherds of the Classic period. Perhaps, like the unfamiliar pottery we had found in 1973, the platform belonged to the earlier Formative period. Certainly its four meters of material should contain traces of a long period of growth; such had proved to be the case with the large platform at the North Acropolis in Tikal, the great Maya ceremonial center in the Petén.

Work began early in 1975 on the selected structure, designated Platform 34 on the Cuello site map, under the supervision of Duncan Pring, then a graduate student at the University of London and the ceramic specialist for the project. It soon became clear that our guess was right; the pyramid had been built long after the platform. A layer of debris had accumulated on the plaster surface of the platform, evidence that the structure had been out of service for some time, before the pyramid was built directly on top of the debris.

We cut down into the platform, exposing a succession of well-preserved plaster floors. Between the successive layers of plaster we found thin deposits



MAYA SITES IN NORTHERN BELIZE include three large enough to rank as regional ceremonial centers (*color*): Aventura, El Pozito and Nohmul. All are situated on the higher ground between two streams, the Rio Hondo and the New River, that empty into Chetumal Bay. The Cuello site is midway between El Pozito and Nohmul, about five kilometers west of Orange Walk Town. The inhabitants of the site early in the Maya Formative period brought the colored chert they preferred as a raw material for edged tools from two nearby sources, seven and 27 kilometers away. To collect the marine shells they made into ornaments required a minimum round trip of 100 kilometers.

of debris and an occasional thicker layer containing potsherds, animal bones, snail shells and small quantities of burned wood. Two meters down the potsherds included some of the unfamiliar type, this time in association with the remains of pottery identifiable as products of the Middle Formative period. Some 70 centimeters below this the potsherds were all of the unfamiliar type; this was the case thereafter until the excavation reached bedrock at a depth of about four meters.

The finding settled one of our hypothetical questions. The examples of the unfamiliar ware that we had encountered in the early levels of Belize sites elsewhere were not imports. The presence of the ware signified that each site had an even longer history than had been supposed. Until the time of our probes the earliest-known kinds of pottery in the Maya lowlands had been products of the Middle Formative period unearthed at two sites in the Petén: Seibal and Altar de Sacrificios on the Río Pasión. There Gordon R. Willey of Harvard University had turned up pottery of an early phase of the Middle Formative associated with material that yielded carbon-14 dates equivalent to slightly later than 900 B.C. Nowhere in the lowlands had any pottery been unearthed that unquestionably belonged to the Early Formative period (then estimated to run from 1500 to 900 B.C.).

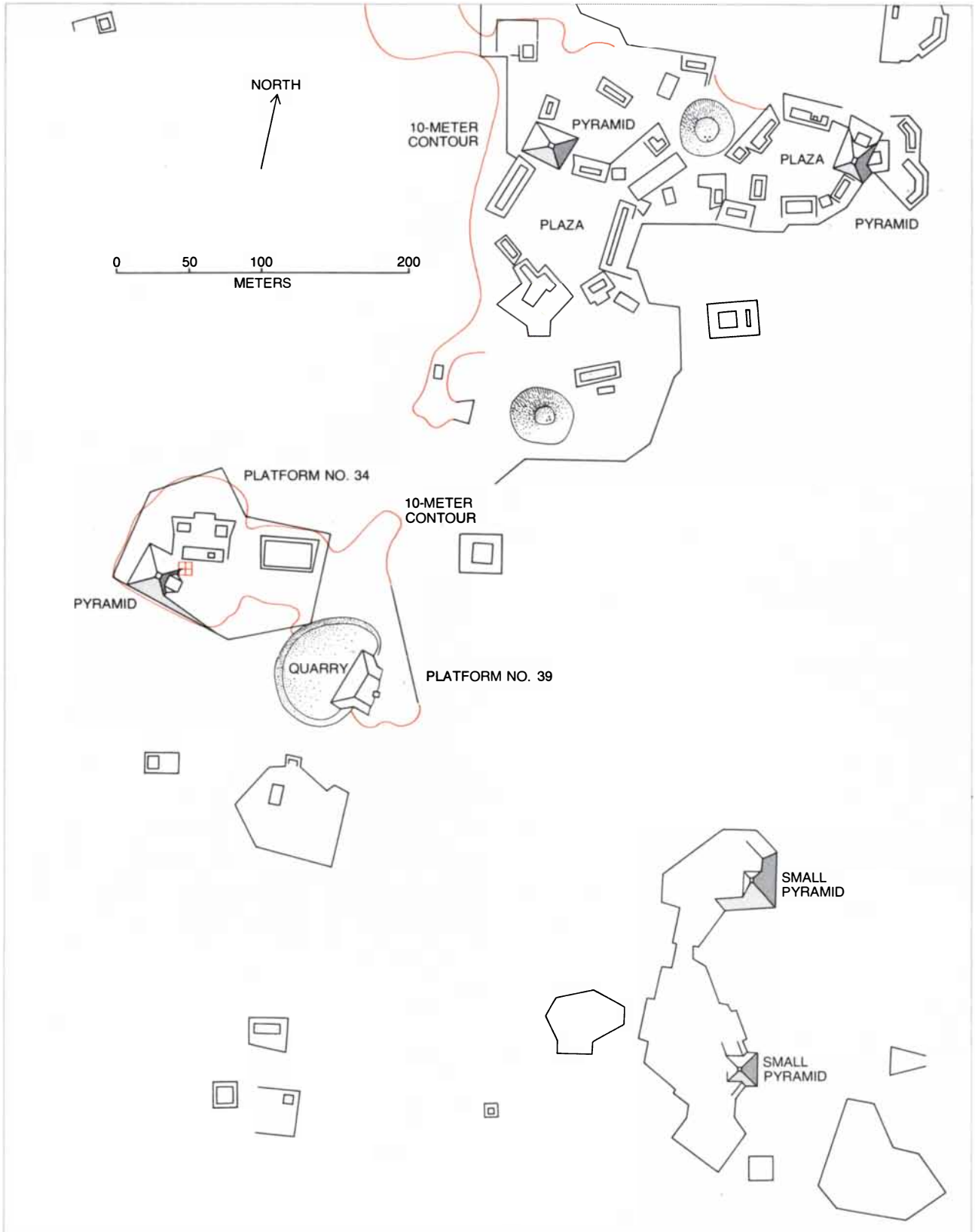
All the ceramic complexes of northern Belize are named after local rivers or lagoons, and so we designated our unknown pottery the Swasey ceramic complex after a nearby tributary of the New River. Obviously the sherds from the lower levels in the Cuello platform were older than the pottery from Seibal and Altar de Sacrificios. The question was, how much older? In the traditional view of early influences in Mesoamerica the Olmec civilization, centered on the Gulf Coast of Mexico west of the Maya area, is considered a probable source of the stimuli affecting the earliest aspects of Maya culture. The period of Olmec influence began around 1300 B.C. Was the Swasey complex early enough to have predated possible Olmec contacts and thus to have inaugurated a cultural tradition independent of Olmec influences?

This remained the question foremost in our minds during the summer and fall of 1975 as two laboratories, one at the University of Cambridge and the other at the University of California at Los Angeles, undertook to make carbon-14 determinations on samples of burned wood from the platform excavation at Cuello. Here two points should be made about carbon-14 dates. First, the dates I am citing are "calibrated." That is, the variations in the regularity of carbon-14 readings, revealed by the analysis of samples of known age from the long-lived bristlecone pine, have been eliminated. Hence the dates represent true calendar years, A.D. or B.C., rather than carbon-14 laboratory years.

Second, there is a degree of statistical uncertainty in carbon-14 dating, as is apparent in the laboratory notation. Take as an example the notation 1000 B.C. $\pm$100. The 1000 B.C. date, usually a mean figure that combines the results of two or more tests of the sample material, is known as the "central" figure; the $\pm$100 indicates odds of roughly two to one in favor of the specimen's age falling somewhere between 1100 and 900 B.C. Statistically this range on each side of the central figure constitutes one standard deviation. If one extends the range to two standard deviations, which in this example would be from 1200 to 800 B.C., the likelihood that the age of the specimen will fall somewhere between these extremes is increased from 68 percent to 97 percent. Extensions of this kind appear in the illustration on page 121. When two such extensions overlap, the two carbon-14 dates are said to be statistically inseparable.

The first carbon-14 date to come through was from Roy Switsur and Alan Ward at Cambridge. It was for a specimen of wood found in a midden deposit representing the transition between the earlier, or Swasey, phase at Cuello and its successor Middle Formative phase, Lopez Mamom. The date of this level proved to be about 1250 B.C., or more than three centuries earlier than the pottery from Willey's Middle Formative sites in the Petén.

The date suggested not only that the



NORTHERN BELIZE SITE, Cuello, is outlined in part; the plan is taken from a survey still in progress. In Classic Maya times Cuello was a minor ceremonial center, its principal focus being two adjacent plazas flanked by pyramids (*top*). Earlier, during the Formative period, the focus was to the south and west and included a massive platform identified by the mappers of the site as Platform No. 34. It was built in the Late Formative period, as was a second, partially destroyed platform, No. 39, situated some 180 meters to the east of No. 34. The area excavated in 1976 (*color*), overlapping a test trench dug during the previous season, bordered on a small pyramid on Platform No. 34 that was not added to the structure until Classic times. Age of the platforms and pyramids at lower right is not yet known.

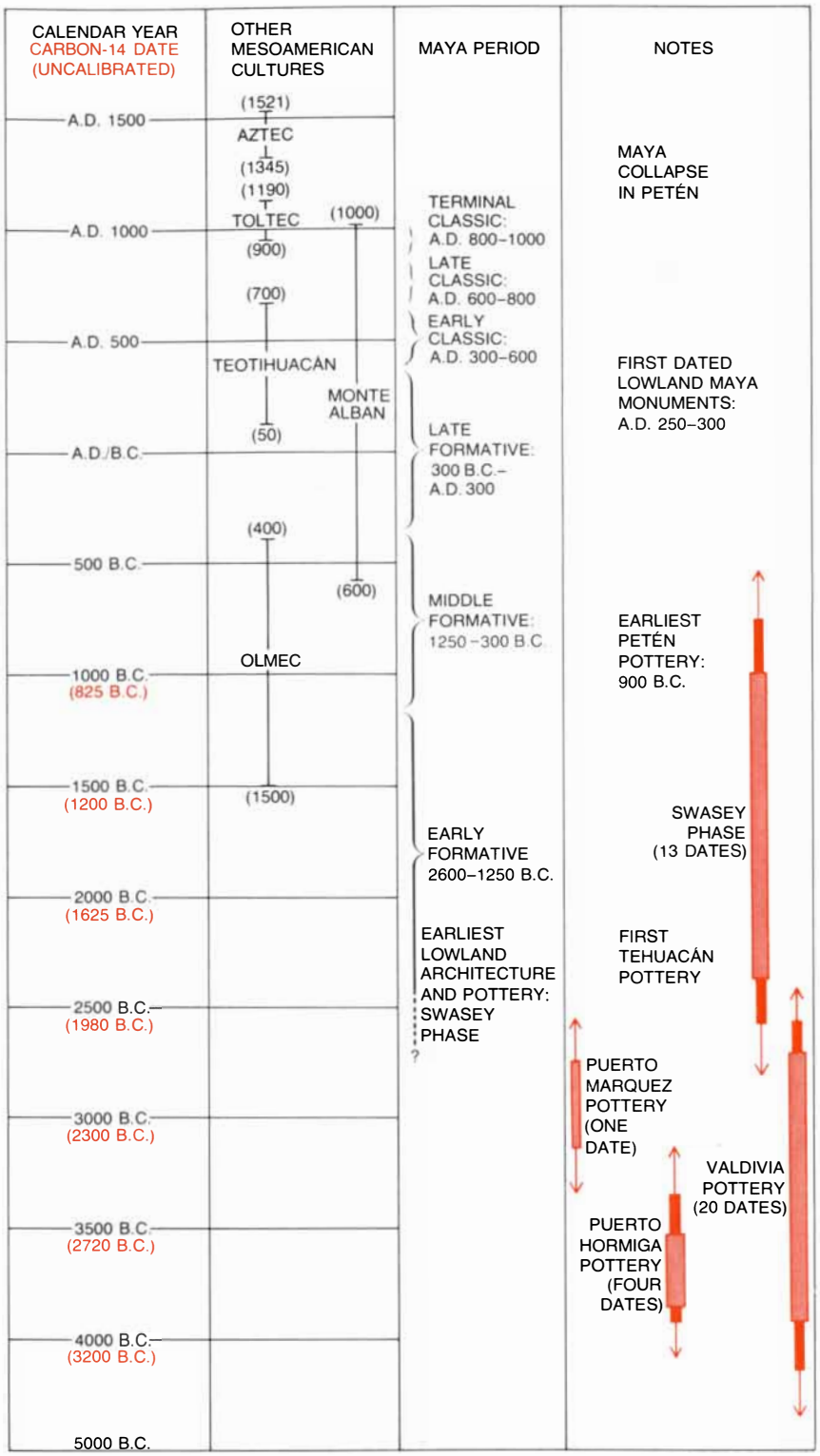
Middle Formative period began earlier than had been supposed but also that the Swasey complex could well represent a hitherto unknown lowland Maya Early Formative period. Moreover, the Swasey upper levels were at least contemporaneous with the Olmec. Two additional dates from Cambridge, one for a Middle Formative layer of the platform and one for a Late Formative layer, reinforced our confidence in the antiquity of the transition deposit.

In November we received the carbon-14 determinations from Rainer Berger at U.C.L.A. A sample from a level immediately below the transition deposit yielded an almost identical reading; a sample from a somewhat lower level indicated a date perhaps two centuries earlier. The agreement between the findings of the two laboratories gave us confidence in the most surprising determination of all. Wood from the lowest midden uncovered in the probe of the Cuello platform, located just above bedrock, was assigned a carbon-14 date by the U.C.L.A. laboratory that ranged from 2450 to 2750 B.C. and thus had a central value equivalent to 2600 B.C. In effect one season of work at Cuello had pushed back the antiquity of the Maya by a full millennium and the prehistory of the lowlands by more than 1,600 years. Moreover, the establishment of such an early date for the possible inception of the Maya Formative period had effectively removed the Olmec civilization from further consideration as the initial stimulus of Maya culture and even suggested the possibility that Maya culture acted as an influence on the emergent Olmec society.

The great antiquity of the Swasey pottery complex had even wider implications. Up to the time of our discovery the earliest-known examples of a ceramic tradition in Mesoamerica had been pottery from two areas in Mexico west of the Maya zone: the Purrón ware of the Tehuacán valley and the Pox pottery of Puerto Marquez on the Pacific coast. Broadly speaking, the Tehuacán pottery (about 2800 B.C.) appears to be coeval with Swasey ware, and the Puerto Marquez pottery, although a good deal earlier (about 3200 B.C.), is statistically inseparable from Swasey because the dates overlap when they are extended by two standard deviations.

Stylistically, however, the Swasey ware is much more sophisticated than either the Purrón or the Pox. In contrast to their limited repertory the Swasey ware has a wide range of forms, finishes and decorations. Could these lowland products of the Early Formative period in the Maya area represent the starting point of a pottery tradition that later expanded over much of Mesoamerica?

Our single very early carbon-14 date from Cuello was clearly in urgent need of confirmation. Furthermore, the nar-



MAYA CHRONOLOGY over a span of 3,500 years is seen against the perspective of other contemporary and later Mesoamerican civilizations' rise and decline. Dates in black (*left*) show calendar years; dates in color are the equivalent carbon-14 years; these have not been calibrated to eliminate their inconsistency, which increases with samples of increasing age. The range of sample age determinations is indicated in four instances by arrows of varying thickness (*color*): for the Swasey ceramic complex at Cuello, for the Mexican Pacific site Puerto Marquez, where pottery slightly earlier than Swasey has been found, and for two South American sites with even earlier pottery. Heads of arrows show maximum ranges of dates, extended from the central carbon-14 reading by ±20 percent (two standard deviations). Wider line measures a one-deviation extension. Where several dates are known the widest line (*light color*) shows the dates' central range. The lowland Formative period was formerly thought to begin in about 900 B.C.

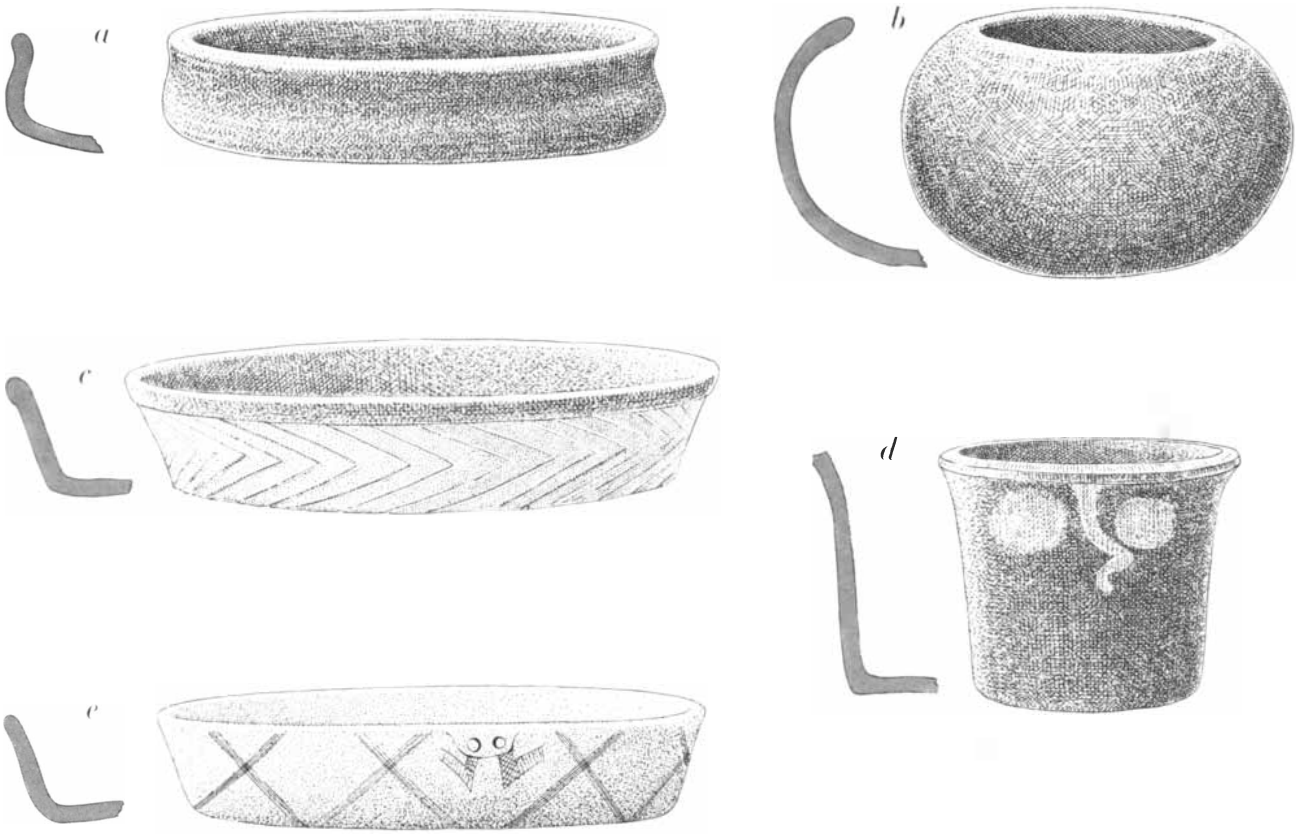
row shaft cut into Platform 34 had provided virtually no information about the economy or the cultural repertory of the Swasey-phase inhabitants of the site, a population that seemed to have a good claim to being the earliest Maya. We therefore decided to put in a short season of further excavations at Cuello in March and April of last year.

That excavation season was supervised by Sara Donaghey of the York Archaeological Trust, which also supplied our group with a drafter, Sheena Howarth, and a conservator, Jim Spriggs. A prime necessity was an accurate map of the site, and work on it had been undertaken in the 1975 season by Michael Walton, an English architect, and Basilio Ah, a Mopan Maya from southern Belize. The two had previously mapped the Maya ceremonial center of Lubaantún [see "The Planning of a Maya Ceremonial Center," by Norman Hammond; SCIENTIFIC AMERICAN, May, 1972]. Ah continued the mapping in 1976, working with another architect, Frederick Johnson of Honolulu. In its present state the map shows the ceremonial precinct at Cuello, surrounded by a scatter of residential compounds [*see illustration on page 120*]. Three massive platforms, spaced several hundred meters apart, lie south of the later center of Cuello along a line running a little south of east. The westernmost of the three is Platform 34, which reached its present dimensions in Late Formative times. The final construction work on the central structure, Platform 39, also took place in the Late Formative period; the age of the easternmost platform has not yet been determined.

The 1976 excavation was confined to a 10-by-10-meter square on Platform 34, its four sides facing the four cardinal directions. Two quadrants of the square were excavated: a five-meter square at the northeast corner and another at the southwest corner. The western side of the southwest quadrant incorporated the 1975 shaft, so that we knew roughly what old floors and midden layers to expect at what depths in the new parts of the excavation. By digging diagonally opposed quadrants we also had the benefit of exposing continuous 10-meter vertical sections through the platform, one running from north to south and the other from east to west.

The 1975 shaft provided excellent guidance during the first two weeks of the season, a time devoted to peeling away successive plaster floor surfaces and screening accumulations of debris. As we passed a floor at a depth of about a meter, however, we encountered two features unlike any we had unearthed in 1975. The first was a layer of rubble, burned plaster and earth, evidence that some structure or structures even older than the platform had been deliberately destroyed. The second, to the south and east of the first, was the surface of a massive rubble dump. Its rough lumps of limestone and chert filled these two sides of the excavation.

Further digging revealed that the layer of rubble, burned plaster and earth covered the remains of two structures that had stood just beyond the rubble dump. We then turned to the task of removing the one-meter layer of rubble. It soon became clear that the rubble had been used to fill up a sunken patio and that the two structures had once stood on individual platforms on the north and west sides of the patio. (We reached the limits of our grid before the south and east sides of the patio were exposed.) The plaster floor of the patio was found to be in a good state of preservation under the rubble.



EXAMPLES OF SWASEY WARES include two of the most abundant variety, Consejo Red (*a, b*). The pots have a cream underslip and a red surface slip. The second shallow dish (*c*), incised with a series of chevrons, is also representative of the Consejo group; its rim and interior are decorated with a red slip. The last two vessels are assigned to other groups of the eight within the Swasey complex. Pot *d* carries a "reserved" design, produced by applying a red slip over an orange underslip. The third dish (*e*), covered inside and out with a buff slip, has been further decorated with fine incisions that form a repeated *X* pattern and frame a false suspension lug. The Swasey-complex groups include 25 ceramic varieties; most of the ones seen here were reconstructed by Louise Christianson from sherd studies.

The 1975 shaft had missed both the west edge of the patio and the rubble fill by about a meter; what we had taken to be earlier floors of the great platform were in fact the interior floors of successive buildings that had stood on the west side of the patio. Further excavation made it plain that in this part of the site the construction of Platform 34 had been preceded by the deliberate razing of the buildings bordering the patio and the filling of the sunken area with rubble.

We were able to place the time of the remodeling toward the end of the Middle Formative period, about 400 B.C. The work had involved a considerable communal effort. Two facts make this clear. First, the limestone available locally at Cuello differs in texture from the limestone used to fill in the sunken patio. Second, there is no chert at all available at Cuello. The nearest source of both fill materials is at least two kilometers from the site.

Platform 34, like the other two great platforms, is obviously a ceremonial structure rather than a residential one. The communal aspect of its construction would thus also seem to involve ceremonial behavior. But what about the structures that had been burned and buried earlier? Were they temples or perhaps residences for a social elite? Or had they been some ordinary cluster of dwellings, razed to make way for the great platform? The evidently ritual nature of the demolition and covering up suggests that the razed structures had been ceremonial ones.

The more impressive of the two buildings had stood on the north side of the patio. A stairway led from the patio up to an open terrace at the front. The front wall of the building, which has been so far only partly excavated, was constructed of small, rounded limestone boulders, laid in courses and covered with a facing of plaster. The doorframe was made out of stiff, perhaps pounded, earth with a core of rubble to give it added strength. The terrace, of similar earth-and-rubble construction, was covered with plaster. Except in the staircase area the plaster finish ran down the face of the terrace and blended into the plaster floor of the patio.

As we dug down, exposing the floor of the building, we found a human burial sealed under the threshold. Such graves are not uncommonly associated with Maya structures and have come to be known as foundation burials. The skeleton was that of a young male, lying on its right side with its head pointing west. In the grave was a small, plain pottery jar and a string of beads made out of mollusk shells and jade.

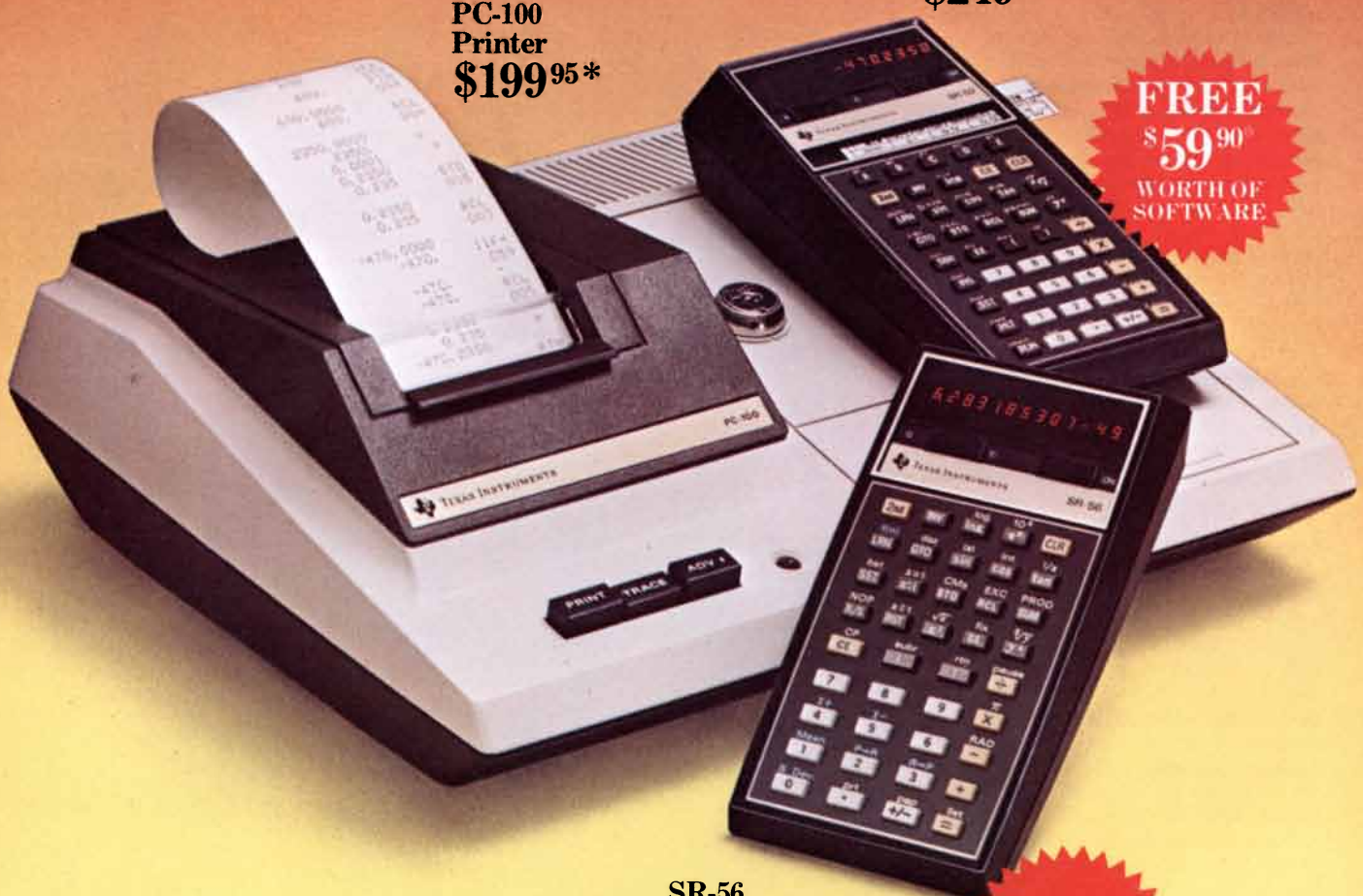
The second building, on the west side of the patio, had been razed level with the patio floor. Only the outline of its entrance stairway and front wall could

Texas Instruments gives you *more* good reasons to buy a programmable calculator now.

SR-52
Card programmable
$249.95*

PC-100
Printer
$199.95*

FREE
$59.90
WORTH OF SOFTWARE



SR-56
Super slide rule
Key programmable
$109.95*

$10
REBATE

New low prices on SR-52, PC-100.

(Now you can get an SR-52 with PC-100 Printer for less than $450.)

Free software with your SR-52.

$10 rebate on SR-56 purchase.

SR-52. The card programmable calculator with computer-like power. 224 program steps and 22 data memories let you solve complex problems. Fast. With less chance of error. You can load complete programs from small magnetic cards in just two seconds. Process and record data or perform complicated calculations automatically. Select a program from the SR-52 Basic Library, which includes 22 prerecorded program cards. There's also a wide variety of optional prerecorded libraries—or you can record your own programs.

Programming is easy with TI's exclusive AOS (Algebraic Operating System). You enter expressions—both numbers and functions—from left-to-right, just as you would write them out on paper.

Now you have the opportunity to own an SR-52 at a lower price than ever before. And, if you buy your SR-52 anytime between January 20, 1977 and March 31, 1977, Texas Instruments will send you two software libraries—a $59.90 value—free.† Choose from Mathematics, Statistics, Finance, or Electrical Engineering. See details in coupon below.

SR-56. A powerful, super slide rule calculator that's key programmable, too. 74 preprogrammed functions and operations easily handle tough scientific problems. 100 programming steps plus 10 memories allow you to obtain results once requiring the assistance of a computer. Yet no programmable is easier to master. There's no special entry sequence to learn with TI's unique AOS (Algebraic Operating System). AOS allows easy left-to-right entry of expressions, both numbers and functions. Build your own programs and use them again and again. Or select one from the SR-56 Applications Library. 56 different programs in mathematics, statistics, finance, electrical engineering, and more are contained in the 192-page Library. Select a program, follow the listing, and you can begin immediately to solve your own problems.

Buy an SR-56 anytime between January 1, 1977 and March 31, 1977 and Texas Instruments will send you a $10 rebate.† See details in coupon below.

PC-100. TI's PC-100 Printer will turn your SR-52 or SR-56 into a quiet, high-speed printing calculator. Hard copy printout provides intermediate results, answers, or will print an entire program listing at the press of a key. In addition, a simple program allows you to print the calculator's entire data memory content.

Your SR-52 or SR-56 plugs directly into the PC-100. Turn the key and you are ready to print. A quiet, reliable thermal printer operates at 3 lines per second, printing 10 significant digits, exponents, decimal points, symbols, and signs.

At its new, low price, the PC-100 is now an even greater value. See it today at your Texas Instruments retailer and learn how inexpensively you can add new versatility to your TI programmable calculator.

†Offers good in United States and Canada only. Offers void where prohibited by law.

*U.S. suggested retail price, may vary elsewhere.

TEXAS INSTRUMENTS
INCORPORATED

SR-52 FREE Software. Expires March 31. M

Indicate choice of 2 software libraries below and (1) return this completed coupon along with (2) your completed SR-52 serialized customer information card (packed in box) and (3) a dated copy of proof of your purchase, verifying purchase between Jan. 20 and March 31, 1977, to:

SR-52 Software Offer
P. O. Box 1210
Richardson, Texas 75080

Name _____
Address _____
City _____ State _____ Zip _____
SR-52 Serial No. _____ (from back of calculator)
☐ Math ☐ Statistics ☐ Finance ☐ EE
Texas Instruments reserves the right to substitute software libraries of equal value based upon availability. Please allow 30 days for delivery.

SA1B

SR-56 $10 Rebate. Expires March 31. M

Texas Instruments will rebate $10.00 of your original SR-56 purchase price when you (1) return this completed coupon, (2) along with your completed SR-56 customer information card (packed in box), and (3) a dated copy of proof of your purchase, verifying purchase between Jan. 1 and March 31, 1977, to:

SR-56 Rebate Offer
P. O. Box 1210
Richardson, Texas 75080

Name _____
Address _____
City _____ State _____ Zip _____
SR-56 Serial No. _____ (from back of calculator)
Please allow 30 days for rebate.

SA1B

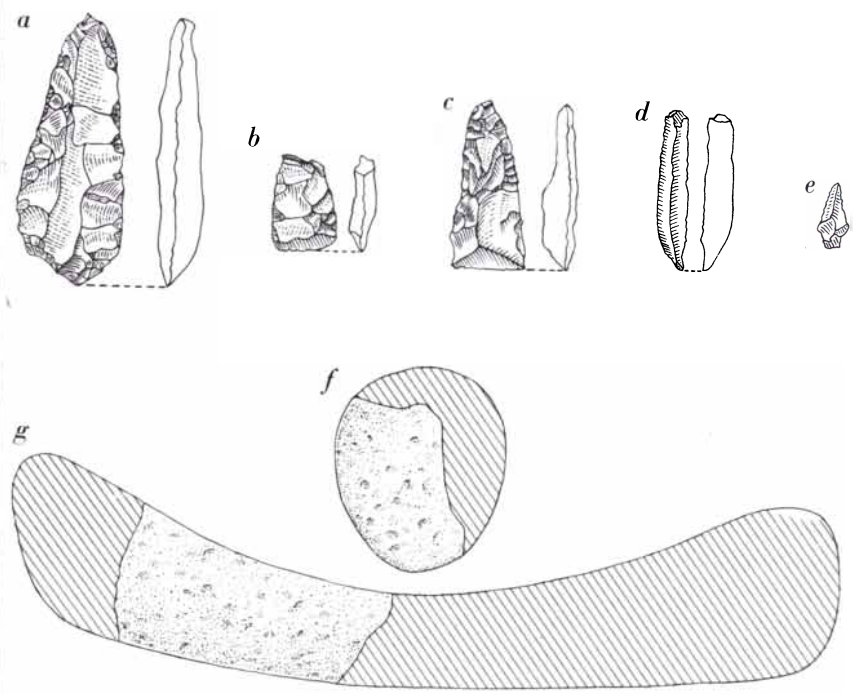
65573

be seen, and the plaster patio floor near it showed signs of intense fire. As we extended our excavation of the area we found that the west edge of the patio at the time of the razing (about 400 B.C.) had in preceding periods been shifted somewhat to the east. Preserved under the Middle Formative buildings and the patio floor were the remains of structures representing three successive periods of construction during the Swasey phase; they constitute the earliest Maya architecture known.

The most recent of the three Swasey buildings is represented by a structure with a poorly preserved floor. Its plaster surface had covered an earlier hole that had once supported the butt of a large timber upright. Remnants of the upright timber, still present in the hole, yielded a carbon-14 reading some three centuries earlier than the readings for samples from this construction period located elsewhere. This suggests that the upright had been a quite mature tree when it was cut down and set in place. On balance the carbon-14 determinations suggest a date of from 1700 to 1500 B.C. for this final Swasey construction period.

The structure was evidently a low platform; its straight east façade had been at least 10 meters long and perhaps substantially longer. At its south end it swept around to the west in a curve, and to judge from what we can see of this feature the platform had been at least five meters wide. Where the curve begins a niche was set into the edge of the platform near the top. Indications that the plaster lining of the niche had been renewed suggest that it was used as a step by those climbing from the sunken patio to the top of the platform. A similar niche may exist at the unexcavated north end of the platform. The small size of the niche we uncovered seems also to argue for the existence of a central stair. The top of the platform had supported a timber superstructure, as was evident from a series of postholes, but no certain plan for the timber structure could be determined.

The structure that represents the next Swasey architectural phase was buried under the last one. It fronts on the same line, but its façade was curved rather than straight. A modest earth-filled platform, it was a mere 30 centimeters high. From what we can see we estimate that it was some seven meters long and four to five meters wide. It had a plaster top and a plaster facing, reinforced along its upper edge with a line of rough stones probably intended to minimize erosion. Like the platform that succeeded it, it had supported timber-framed superstructures; the evidence of the postholes indicates that two such structures had occupied the platform in succession. Both had rounded sides rather than



STONE IMPLEMENTS UNEARTHED AT CUELLO were made chiefly by flaking chert (*top*) or by grinding sandstone (*bottom*). The large tool (*a*) is 15 centimeters long; it is typical of the general-utility choppers of the Formative period. Next (*b, c*) are small axe-adzes that may have served for finer work. The long parallel-sided blade (*d*) is much like later Maya work in obsidian. The small point (*e*) probably was used as a drill or punch rather than as a projectile. Fragments of pinkish sandstone (*f, g*) imported from the Maya Mountains 150 kilometers away are shown in restoration here: a mano and metate, the Mesoamerican corn mill.

straight ones, and the earlier of the two was the larger one. Carbon-14 determinations from numerous samples, some contemporaneous with the platform and some from a succeeding layer, suggest a date for the platform between 2300 and 2000 B.C. Both platforms fronted a plaster-floored patio that covered essentially the same area as the Middle Formative patio had.

The oldest of the three Swasey structures came to light at the juncture of our two five-by-five-meter quadrants and had evidently been built at a time before the Early Formative pattern of constructing platforms around a patio perimeter had become established. So far we have uncovered what appears to be part of a circular platform with a three-meter radius. If our preliminary interpretation is correct, the surface area of the platform would have exceeded 28 square meters.

Like the later Swasey structures the platform was built of earth and had a plaster surface; it supported a timber superstructure and rested directly on a substrate of long-buried soil. The soil is mixed with quantities of trash and other debris of occupation apparently derived from dwelling sites that stood outside our area of excavation. Burned wood from this buried soil has yielded a carbon-14 date that falls between 2500 and 2400 B.C. This would place the date of construction of the earliest platform at Cuello substantially more than 4,000 years ago, making it the earliest example of architecture known in Mesoamerica and one of the earliest in the New World.

The three successive examples of Early Formative architecture unearthed at Cuello may be said, if we take the liberty of rounding the dates, to be roughly 4,400, 4,100 and 3,600 years old. The existence of architectural traditions typical of Classic Maya dwellings, such as plastered floors and platforms with timber-framed superstructures, in the lowlands that long ago is indicative of a developmental period for Maya culture of far greater duration than has been supposed.

Our excavations last year provided significant data on the economy of the lowland Maya during the Early Formative period. For example, we found five more human burials, all associated with the Swasey structures. One was a child four or five years old, three were adult males and the fifth was a young adult female. The biological anthropologist for the project, Frank P. Saul of the Medical College of Ohio, concludes that all four adults show abnormally advanced tooth wear, suggesting the presence of an abrasive substance in their daily diet.

The abrasive could have been either of two abrasives that are found in the diet of the Maya today. One is lime, which is in the diet because the Maya steep the kernels of maize in slaked lime before boiling them; the process softens the hard coating of the kernel and also releases certain amino acids in the maize that would otherwise be unassimilable. The other abrasive is grit derived from the stone roller (mano) and milling table (metate) that are still used today to crush the maize kernels into corn meal. Fragments of these grinding stones, which can also be used to grind seeds other than maize, have been unearthed in the earliest of the kitchen middens at Cuello. They were made out of two distinctive kinds of sandstone that were not of local origin.

For the present it is only an assumption that maize was cultivated at Cuello in the Early Formative period. We recovered a large sample of carbonized plant remains at the site, but they are still being analyzed by Barbara Pickersgill of the University of Reading and their identity is not yet known. It is possible that the lowland economy at this period included such root crops as manioc and sweet potato, but proof is unlikely to be forthcoming because identifiable remains of these plants seldom survive prolonged burial in a tropical lowland soil.

Hunting amplified the diet of the early Maya at Cuello. The bones of both the white-tailed deer (*Odocoileus virginianus*) and the agouti (*Dasyprocta* sp.), unearthed from Swasey-phase middens, have been identified by Elizabeth S. Wing of the Florida State Museum. Snails are also represented; the shells of five edible species, the swamp-dwelling *Pomacea flagellata* in particular, have been identified by Lawrence Feldman of the University of Missouri at Columbia. Between 40 and 60 percent of the snail shells were found in deposits of kitchen refuse.

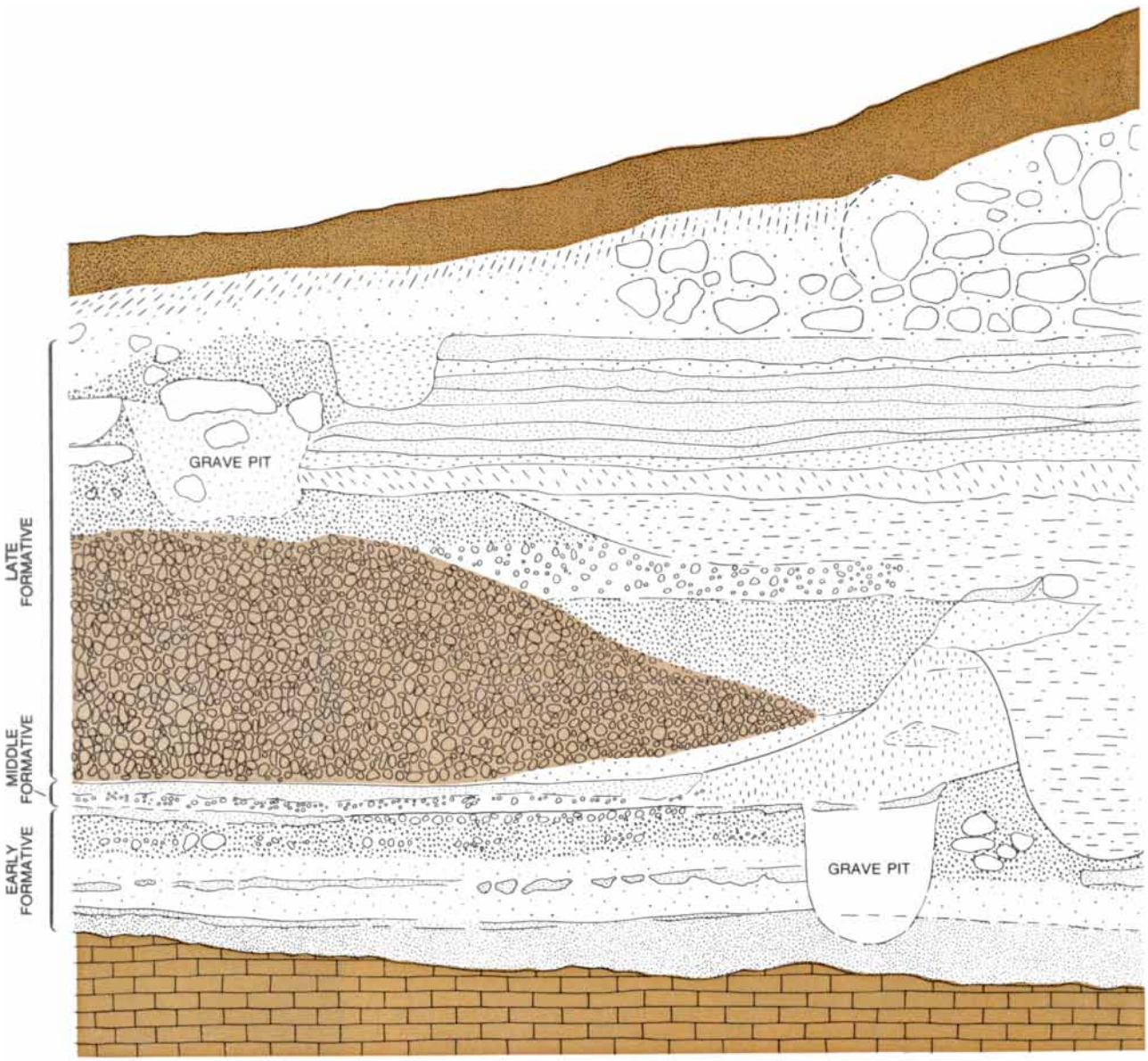
Among the snails are a number of species, seldom if ever eaten, that are common to forest, freshwater and marine environments. Entirely unrepresented in the Swasey material, however, is one snail subspecies (*Neocyclotus dysoni cookei*) that is characteristic of areas that have been burned over and are rich in leaf humus. This suggests that if the Swasey-phase Maya of Cuello did cultivate maize, they did not practice milpa agriculture: planting corn in a field that is prepared by felling and burning the natural brush. Quite the opposite evidence is found in the Middle Formative phase that followed: shells of the swamp snail *Pomacea* decrease in number until they account for only 16 percent of the total, whereas the proportion of the milpa-dwelling *Neocyclotus* rises from zero to 55 percent.

Feldman interprets this change in the snail sample as evidence that the Maya

of the succeeding phase were draining swampy land, thereby diminishing the area suitable for *Pomacea* and bringing the drained land under cultivation. Support for his view comes from the findings of a survey group under the direction of Alfred Siemens of the University of British Columbia and Dennis E. Puleston of the University of Minnesota. This group, working just west of us along the Rio Hondo, on the boundary between Mexico and Belize, has mapped a series of raised-field complexes. The fields were formed in riverside swamps by digging drainage canals and using the spoil from the canals to construct platforms that stand above water level. A wood post retrieved from the bank of one of the Rio Hondo canals has yielded a carbon-14 date that falls at about the end of the Swasey phase at Cuello, or about the time of the decline of the swamp snails.

What was grown on these platforms? Perhaps corn, perhaps root crops, perhaps even a "cash" crop such as cacao. Certainly the construction and maintenance of the raised-field complexes would have required some degree of communal cooperation not inconsistent with a structured society and an elite class. As for the need for a cash crop, the presence of imported materials at Cuello suggests that the early Maya there had reason to produce something with which to barter. (In later centuries the cacao bean was a widely accepted form of currency throughout Mesoamerica.)

The foundation burial on the north



CROSS SECTION of the southwestern five-meter square excavated at Cuello shows the southern exposure of successive strata, from a layer of soil (*color, top*) that accumulated after the site was deserted in about A.D. 900 to bedrock (*color, bottom*) some five meters below the surface. Rock-filled area at right, just below the soil, is the stair platform of the Classic-period pyramid that was built on top of a platform of the Late Formative period in about A.D. 600; the stippled layers under the Classic pyramid represent the successive renewals of the plaster floors that formed the surface of the Late Formative platform. A wood sample from the third of these floors yielded a carbon-14 date equivalent to 200 B.C. The massive rubble fill (*light color, left*), which was not exposed by a 1975 test trench, rests on the plaster floor of a patio built in Middle Formative times. This floor and the one under it end abruptly at the right, evidence that structures facing the patio were razed in about 425 B.C. to clear the way for the Late Formative construction. Samples of wood from the stratum between the two floors date to about 800 B.C. The lowest plaster floor of all, covering an ancient soil that rests on bedrock, is 4,500 years old.

Unique full-function 8-digit wrist calculator... available only as a kit.

A wrist calculator is the ultimate in common-sense portable calculating power. Even a pocket calculator goes where your pocket goes—take your jacket off, and you're lost!

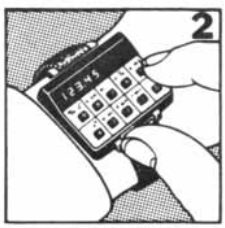
But a wrist-calculator is only worth having if it offers a genuinely comprehensive range of functions, with a full-size 8-digit display.

This one does. What's more, because it *is* a kit, supplied *direct* from the manufacturer, it costs only a very reasonable $19.95. And for that, you get not only a high-calibre calculator, but the fascination of building it yourself.

How to make 10 keys dothe work of 27

The Sinclair Instrument wrist calculator offers the full range of arithmetic functions. It uses normal algebraic logic ('enter it as you write it'). But in addition, it offers a % key; plus the convenience functions $\sqrt{x}$, $1/x$, x^2; plus a full *5-function* memory.

All this, from just 10 keys! The secret? An ingenious, simple three-position switch. It works like this.



1. The switch in its normal, central position. With the switch centered, numbers—which make up the vast majority of key-strokes—are tapped in the normal way.

2. Hold the switch to the left to use the functions to the left above the keys. . .

3. and hold it to the right to use the functions to the right above the keys.

The display uses 8 full-size red LED digits, and the calculator runs on readily-available hearing-aid batteries to give weeks of normal use.

Assembling the Sinclair Instrument wrist calculator

The wrist calculator kit comes to you complete and ready for assembly. All you need is a reasonable degree of skill with a fine-point soldering iron. It takes about three hours to assemble. If anything goes wrong, Sinclair Instrument will replace any damaged components *free*: we want you to enjoy assembling the kit, and to end up with a valuable and useful calculator.

Actual Size

Contents

Case and display window.
Strap.
Printed circuit board.
Switches.
Special direct-drive chip (no interface chip needed).
Display.
Batteries.

Everything is packaged in a neat plastic box, and accompanied by full instructions. The only thing you need is a fine-point soldering iron.

All components are fully guaranteed, and any which are damaged during assembly will be replaced free.

The wrist-calculator kit is available only direct from Sinclair Instrument. Take advantage of this 10 day money-back offer.

Send the coupon today.

Sinclair Instruments Inc.
6 Commercial Street
Hicksville, N.Y. 11801

To: Sinclair Instruments Inc.
6 Commercial Street
Hicksville, NY 11801

Please send me

_____ Sinclair wrist calculator kit(s) at $19.95 $_____

Sales tax (NY residents) _____

Shipping & handling ($2.50 per unit) _____

Enclosed is check/MO (payable to Sinclair Instruments Inc.) TOTAL $_____

Name _____

Address _____

City _____

State _____ Zip _____

I understand that you will refund the purchase price of the kit in full if I return it in saleable conditiion within 10 days of receipt.

SA3

side of the patio under Platform 34 and two of the five earlier graves included beads of jade and shell among the burial offerings. In addition one of the Swasey graves contained a lump of hematite, the hard iron-ore pigment used in powdered form for pottery decoration and body painting. The shells and the hematite could have been obtained within the Maya lowlands and perhaps even in northern Belize, but the jade could not. The nearest known source of the distinctive green gemstone is some 350 kilometers away, in the Motagua valley on the margin of the Guatemala highlands. The presence of jade beads in these early graves is proof of the existence of an extensive exchange network in this part of Mesoamerica more than 3,500 years ago.

Other materials utilized by the Swasey-phase inhabitants of Cuello may also have been obtained through exchange, but the sources were closer. The Maya Mountains of Belize, some 150 kilometers south of Cuello, were evidently the source of the pink-hued kinds of sandstone used to make manos and metates. Richard Wilk of the University of Arizona traced the brightly colored chert the inhabitants favored as a raw material for edged tools to two nearby sources; one, a notable chert workshop at Colha, is 27 kilometers from Cuello and the other, at Richmond Hill, is seven kilometers away. The shells of marine mollusks found at the site must have been transported over a minimum distance of 50 kilometers.

The high quality of the early pottery at Cuello calls for a brief description of the Swasey ceramic inventory. All in all, the variety of vessel forms and surface finishes is considerable, and the workmanship is consistently expert. In type the pots range from rough, unslipped pieces—"earthenware" in modern terminology—to thinner-walled pieces with smooth and glossy surfaces, probably comparable in prestige value and in function to today's porcelains. As the ceramics specialist with the project, Pring has surveyed the Cuello findings from the Early Formative period through the Late Formative, and he sees in them a record of a single, continuous process of development.

The earliest entity in the sequence, the Swasey ceramic complex, can be subdivided into some 25 varieties on the basis of combinations of vessel shape and surface appearance [*see illustration on page 122*]. The commonest are plain, smooth-slipped red bowls: Ramgoat Red and Consejo Red. Ramgoat Red bowls, the earlier of the two varieties, have only a single layer of red slip on their surface. The Consejo Red bowls, which gradually replace the Ramgoat Red as the commonest ceramics, have a cream-colored underslip under the red surface, which gives them a glossier, lighter and more consistent tone.

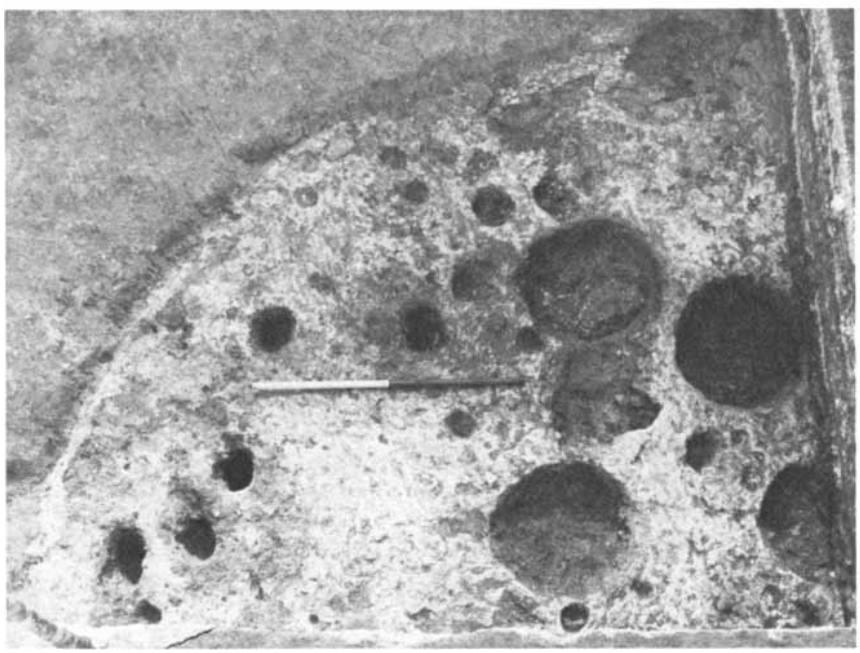
Other notable Swasey-complex varieties include vessels with two-tone surfaces—red on cream, black on red and red on orange—and vessels with surface colors other than red: cream, black, orange, brown and buff. Still others have incised surface decorations: bold chevrons or multiple incisions forming an *X* pattern. A few have animal heads modeled in the round on the rim or on the wall; so far Pring has recognized a frog or toad, a monkey and a turtle.

One particularly striking type of vessel is a long-necked bottle with burnished decorations on an unslipped gray surface. None of the bottles found at Cuello are intact, but the overall character of the material is quite similar to that of ceramics found at Mani in northern Yucatán in 1942 by the late George Brainerd. That was in the days before carbon-14 dating, but Brainerd nonetheless assigned the pottery to the period of about 1500 B.C. His accuracy must be accounted an inspired guess.

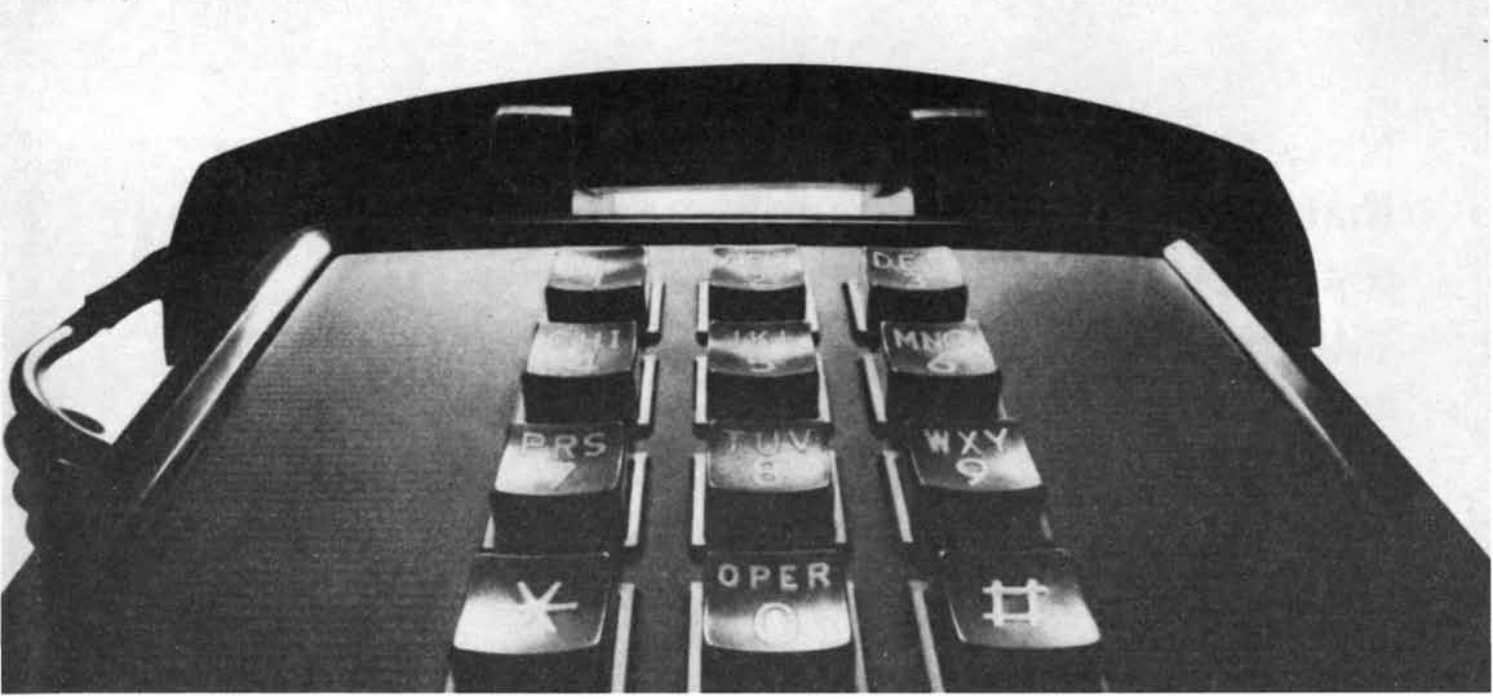
The overall impression conveyed by the Swasey ceramic complex is one of liveliness and variety in both color and decoration. The vessels are indisputably the product of a mature technology rather than an emerging one. Yet these are the earliest ceramics from the Maya lowlands and among the earliest in all Mesoamerica. Where did this technology evolve?

The earliest Swasey-phase remains known are some 4,500 years old. Any answer to the pottery puzzle must therefore be sought in areas where ceramics were known before that time. One possible answer is that the precursor or precursors of the Swasey ceramic complex are to be found in the Maya lowlands, perhaps at Cuello itself or perhaps at other equally early sites. In addition to being uncovered at Mani, well to the north, Swasey-complex pottery has been found at four adjacent sites in Belize: Nohmul, El Pozito, Santa Rita and San Estevan. A few sherds have also come from Becan, a site located almost in the center of the Yucatán peninsula. At none of these other sites has the age of the Swasey-complex pottery been established by carbon-14 dating. A recent carbon-14 date, however, is available from the central Belize site of Barton Ramie, more than 100 kilometers south of Cuello. The date is about 1500 B.C., which falls within the range of the Swasey carbon-14 dates at Cuello. If one also accepts the probability that the earliest Maya occupation of Mani, some 270 kilometers northeast of Cuello, was contemporary with the earliest-known Swasey phase, then a swath of the lowlands some 400 kilometers in length from north to south (from Mani to Barton Ramie), would seem to have been settled by the Maya of the Formative period, no matter how sparsely, at least 600 years earlier than was formerly believed.

What are the prospects that further work at Cuello will uncover still earlier horizons of occupation? Three carbon-



EARLIEST STRUCTURE AT CUELLO, a plaster-covered platform, is seen from above after partial excavation. It was constructed some 4,500 years ago. Its curved edge suggests that the platform was circular in overall shape; two possible reconstructions of the timber structure that stood on it are shown on page 116. The postholes in the plaster surface offer little guidance on the shape of the structure. Large holes are ovens dug into the platform at a slightly later date.



Behind the buttons, there's a "brain."

Electronic intelligence has come to the phone system.

It's part of our new electronic switching machines, working round-the-clock to give you better phone service.

And it's another innovation in telecommunications from Bell Labs and Western Electric.

Intelligent Systems.

Conventional electromechanical switching couldn't have kept up with the enormous increase in telephone calling. Only electronic switching can do that. So we've developed a family of electronic machines that are handling calls in rural, suburban and metropolitan areas.

And our new long-distance electronic system handles 550,000 calls an hour—four times as many as before.

So your Bell telephone company can continue to respond to your needs better and, at the same time, keep your costs reasonable.

Since electronic switching machines run on instructions that can be easily changed, their operations and services can be changed without rewiring or adding equipment.

This flexibility makes many new services practical. Like having your calls forwarded to another number or being signalled the instant a busy line becomes free. The instructions are just programmed into the system's memory.

Electronic switching systems can do more because they're much faster than electromechanical systems. So even though they're handling many more calls, electronic systems have time left over for other tasks.

They continuously check themselves for trouble, diagnose any they find, signal for help, and route calls around a trouble spot.

But trouble is rare with electronic switching.

Today's systems are working round-the-clock with less than 12 minutes "downtime" a year on the average. That makes them 99.9998% reliable.

Our Biggest Job Ever.

The development of electronic switching was the biggest project ever undertaken by Bell Labs and Western Electric in our efforts to keep your phone bill down.

It required an enormous investment in research and development.

Building it took even more work and money.

But now there are 1000 electronic switching offices serving communities of every size as well as handling long distance calls. And a new office is being installed about every working day.

The System Works.

It took the Bell System to develop electronic switching.

Because only the Bell System has Bell Laboratories and Western Electric working together to make your telephone system the most reliable and least expensive in the world.

Which is to say, we had the brawn to put a "brain" behind your phone.

One of a series of messages to keep you informed of how telecommunications technology is changing our world—and the part Bell Labs, Western Electric and your Bell Telephone Company are playing in it.

Western Electric / Bell Laboratories

Now, Dr. Linus Pauling presents new findings that support his arguments even more forcefully than before:
Vitamin C has significant protective activity against many viral diseases, including influenza.

LINUS PAULING

1948 PRESIDENTIAL MEDAL FOR MERIT
1954 NOBEL PRIZE IN CHEMISTRY
1962 NOBEL PEACE PRIZE
1975 NATIONAL MEDAL OF SCIENCE

Vitamin C the Common Cold and the Flu

"In 1976-1977, when another swine-flu epidemic is expected, it is especially important that everyone know that he can protect himself to a considerable extent against the disease and its consequences by the use of this important nutrient, vitamin C"

Available now at bookstores everywhere.

hardcover: $8.95, softcover: $3.45

W. H. FREEMAN AND COMPANY
660 Market Street, San Francisco, CA 94104
58 Kings Road, Reading, England RG1 3AA

14 determinations made at U.C.L.A. late last year provide an ambiguous hint. Samples of burned wood, from layers at the site that have already yielded a sequence of carbon-14 dates firmly linked to the Cuello stratigraphic succession, give considerably earlier readings. The ages of the samples range from early in the fourth millennium B.C. to the middle of the millennium, or from about 6,000 to 5,500 years ago. It is clear from the archaeological context that the much older wood was trash, swept up for construction fill more than a millennium after it had been burned. The fire that burned the wood could have been either a forest fire due to natural causes or the result of human activity in the Cuello area. We calculate the chance that the site was occupied more than 4,750 years ago at about 50–50.

There are alternatives to the possibility of an early evolution of Mesoamerican ceramics centered in the Maya lowlands. One is that ceramics are known to have been made in central and Pacific-coast Mexico and even in the southeastern U.S. some 4,500 years ago. None of the ceramics from these areas, however, have either the variety or the sheer panache of the Swasey ceramic complex.

Pottery 4,500 years old is also known from Monagrillo in Panama, and in northwestern South America pottery has been unearthed that is at least 6,000 years old. The principal South American sites are Puerto Hormiga on the coast of Colombia and Valdivia in southern Ecuador, but it seems likely that during the millennium before the oldest examples of the Swasey ceramic complex first appeared pottery was being made in South America all the way from the Gulf of Guayaquil to the Gulf of Venezuela, both along the coast and in the highlands of the Andes. As for possible relations between this early New World ceramic focus and Mesoamerica, Donald Lathrap of the University of Illinois maintains that there are close resemblances between 4,500-year-old pottery from Real Alto, a site in southern Ecuador that he excavated recently, and the Swasey-complex pottery from Cuello.

For the moment it seems prudent to leave the question open. The origins of Maya pottery may have been local or they may have been exotic. If it was the latter, then a South American stimulus seems more likely than any other now in evidence. What can be said unequivocally is that in the Swasey phase at Cuello we see not only in the pottery but also in the architecture and in the use of stone artifacts the beginning of the cultural tradition that is known as Classic Maya nearly 3,000 years later. Where its creators came from, and when, are questions we hope one day to answer.

MATHEMATICAL GAMES

Cornering a queen leads unexpectedly into corners of the theory of numbers

by Martin Gardner

An analysis of a simple two-person game can lead into fascinating corners of number theory. We begin this month with a charming, little-known game played on a chessboard with a single queen. Before we are through we shall have examined a remarkable pair of number sequences that are intimately connected with the golden ratio and generalized Fibonacci sequences.

The game, which has no traditional name, was invented about 17 years ago by Rufus P. Isaacs, a mathematician at Johns Hopkins University. It is described briefly (without reference to chess) in Chapter 6 of the 1962 English translation of *The Theory of Graphs and Its Applications,* a book in French by Claude Berge. Let us call the game "Corner the Lady."

Player *A* puts the queen on any cell in the top row or in the column farthest to the right of the board; the cells appear in gray in the illustration on this page. The queen moves in the usual way but only west, south or southwest. Player *B* moves first, and then the players alternate moves. The player who gets the queen to the starred cell at the lower left corner is the winner.

No draw is possible, so that *A* or *B* is sure to win if both sides play rationally. It is easy to program an HP-97 printing calculator or the HP-67 pocket calculator to play a perfect game. Indeed, a magnetic card supplied with Hewlett-Packard's recently published book *HP-67/HP-97 Games Pac 1* provides just such a program.

Isaacs constructed a winning strategy for cornering the queen on boards of unbounded size by starting at the starred cell and working backward. If the queen is in the row, column or diagonal containing the star, the person who has the move can win at once. Mark these cells with three straight lines as is shown in part *a* of the top illustration on page 136. It is clear that the two cells shown in color are "safe," in the sense that if you occupy either one, your opponent is forced to move to a cell that enables you to win on the next move.

Part *b* of the illustration shows the next step of our recursive analysis. Add six more lines to mark all the rows, columns and diagonals containing the two previously discovered safe cells. This procedure allows us to color two more safe cells as shown. If you occupy either one, your opponent is forced to move, so that on your next move you can either win at once or move to the pair of safe cells nearer the star.

Repeating this procedure, as is shown in part *c* of the illustration, completes the analysis of the chessboard by finding a third pair of safe cells. It is now clear that Player *A* can always win by placing the queen on the colored cell in either the top row or the column farthest to the right. His strategy thereafter is simply to move to a safe cell, which he can always do. If *A* fails to place the queen on a safe cell, *B* can always win by the same strategy. Note that winning moves are not necessarily unique. There are times when the player with the win has two choices; one may delay the win, the other may hasten it.

Our recursive analysis extends to rectangular matrixes of any size or shape. In the bottom illustration on page 136 a square with 25 squares on a side is shown with all the safe cells colored. Note that they are paired symmetrically with respect to the main diagonal and lie almost on two lines that fan outward to infinity. Their locations along those lines seem to be curiously irregular. Are there formulas by which we can calculate their positions nonrecursively?

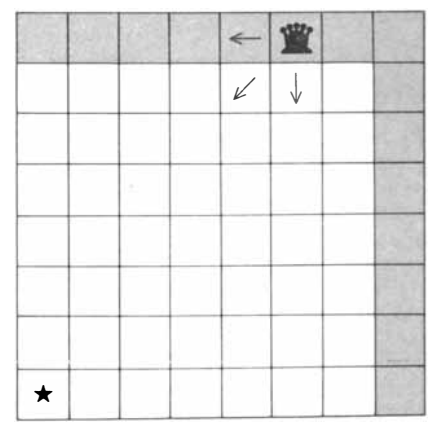
Before answering let us turn to an old counter take-away game said to have been played in China under the name *tsyan-shidzi,* which means "choosing stones." The game was reinvented by the Dutch mathematician W. A. Wythoff, who published an analysis of it in 1907. In Western mathematics it is known as "Wythoff's nim."

The game is played with two piles of counters, each pile containing an arbitrary number of counters. As in nim, a move consists in taking any number of counters from either pile. At least one counter must be taken. If a player wishes, he may remove an entire pile. A player may take from both piles (which he may not in nim), provided that he takes the same number of counters from each pile. The player who takes the last counter wins. If both piles have the same number of counters, the next player wins at once by taking both piles. For that reason the game is trivial if it starts with equal piles.

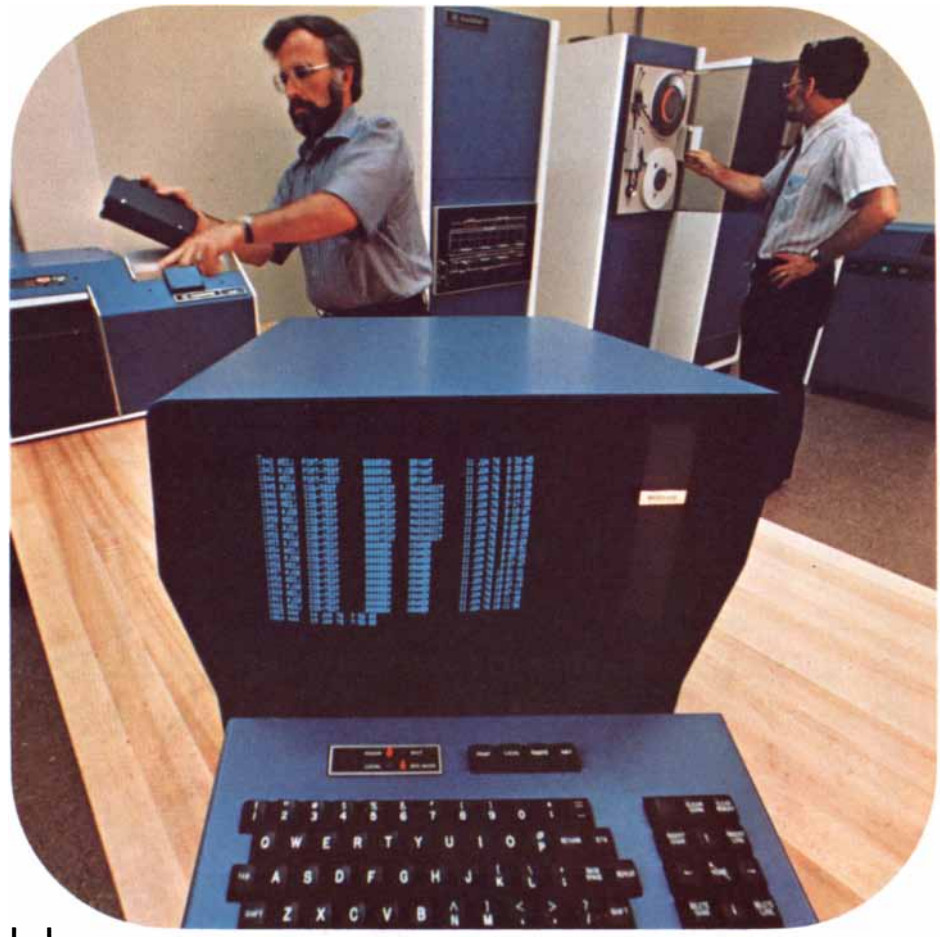
We are ready for our first surprise. Wythoff's nim is isomorphic with the queen-cornering game! When Isaacs invented the game, he did not know about Wythoff's nim, and he was amazed to learn later that his game had been solved as early as 1907. The isomorphism is easy to see. As is shown in the illustration of the 25-square figure, we number the 25 columns along the x coordinate axis, starting with 0; the rows along the y coordinate axis are numbered the same way. Each cell can now be given an x/y number. These numbers correspond to the number of counters in piles x and y. When the queen moves west, pile x is diminished. When the queen moves south, pile y is diminished. When it moves diagonally southwest, both piles are diminished by the same amount. Moving the queen to cell 0/0 is equivalent to reducing both piles to 0.

The strategy of winning Wythoff's nim is to reduce the piles to a number pair that corresponds to the number pair of a safe cell in the queen game. If the starting pile numbers are safe, the first player loses. He is certain to leave an unsafe pair of piles, which his opponent can always reduce to a safe pair on his next move. If the game begins with unsafe numbers, the first player can always win by reducing the piles to a safe pair and continuing to play to safe pairs.

The order of the two numbers in a safe pair is not important. This condition corresponds to the symmetry of any two cells on the chessboard with respect to the main diagonal: they have the same coordinate numbers, one pair being the reverse order of the other. Let us take the safe pairs in sequence, starting with the pair nearest 0/0, and arrange



The cornering game of Rufus P. Isaacs



Harris S-110 Computer System

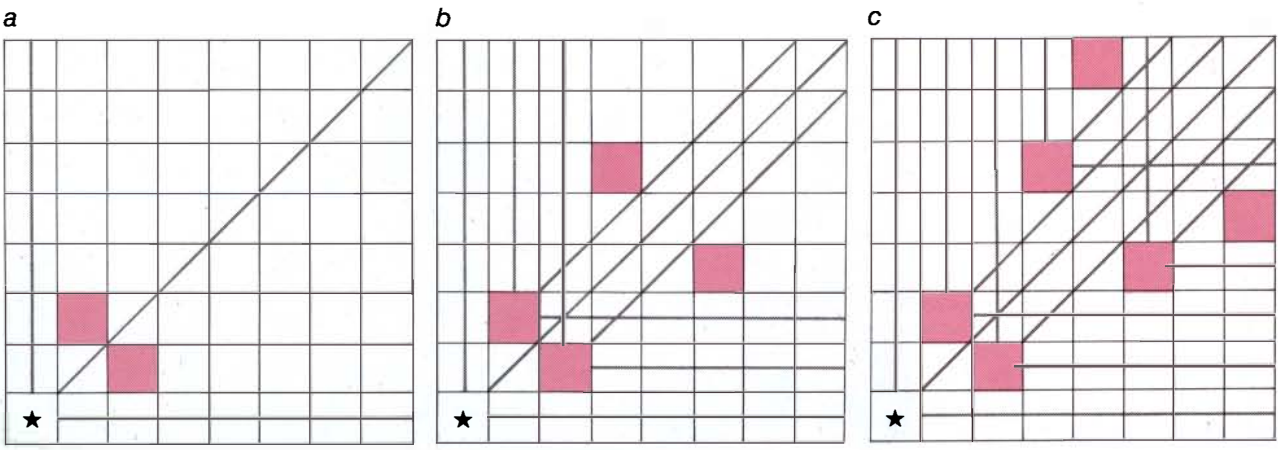
Delivering Computer Power.

Access to big computers can be a problem, and minicomputers have limited power. Harris medium-scale computer systems put number crunching power at the fingertips of the people who need it. Economically.

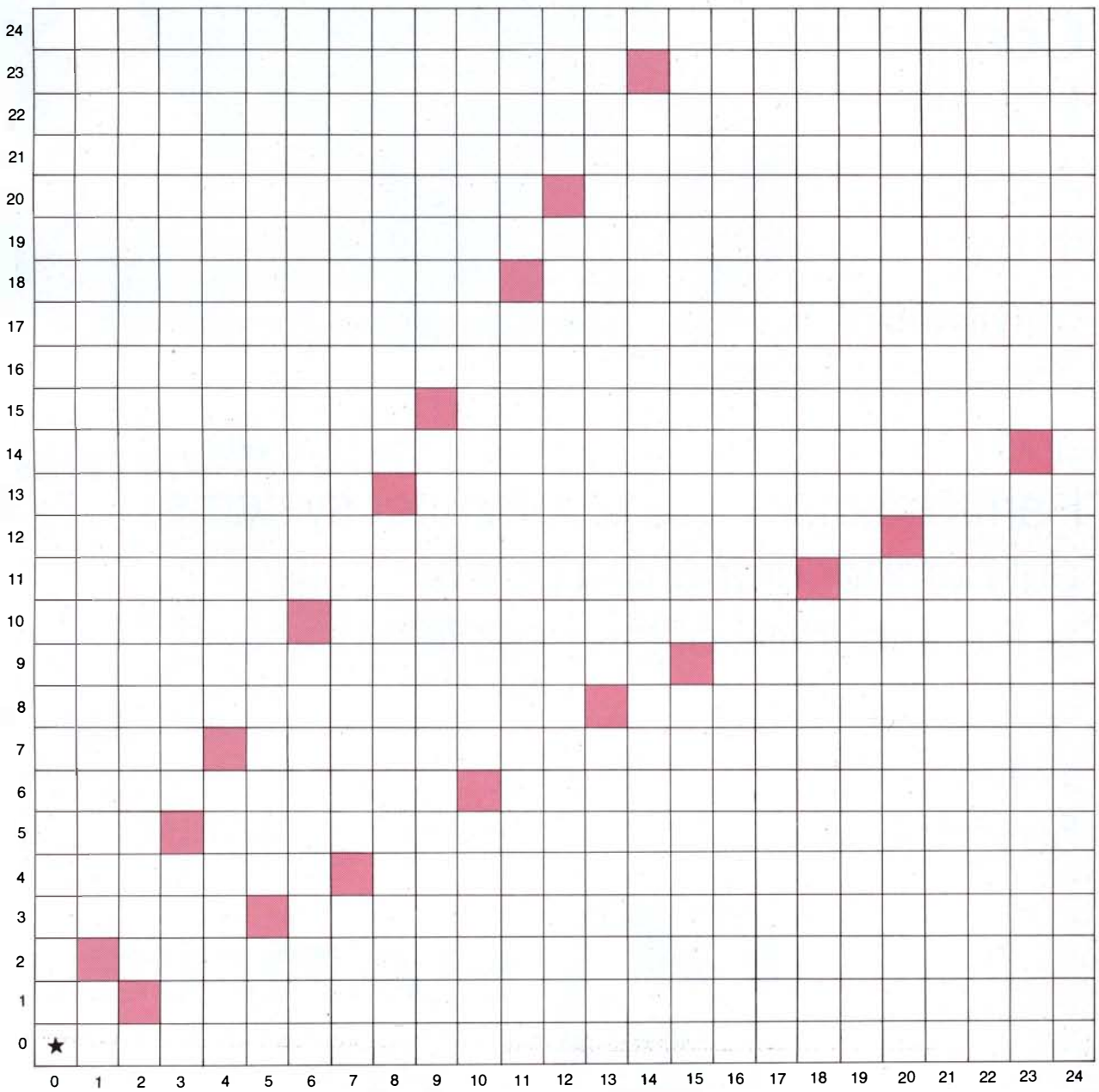
Write: Harris Corporation,
55 Public Square, Cleveland, Ohio 44113.

HARRIS

COMMUNICATIONS AND INFORMATION HANDLING



A recursive analysis of "Corner the Lady"



The first nine pairs of safe cells

them in a row with each smaller number above its partner, as in the top illustration on this page. Above the pairs write their "position numbers." The top numbers of the safe pairs form a sequence we shall call A. The bottom numbers form a sequence we shall call B.

These two sequences, each one strictly increasing, have so many remarkable properties that dozens of technical papers have been written about them. Note that each B number is the sum of its A number and its position number. If we add an A number to its B number, the sum is an A number that appears in the A sequence at a position number equal to B. (An example is $8 + 13 = 21$. The 13th number of the A sequence is 21.)

We have seen how the two sequences are obtained geometrically by drawing lines on the chessboard and coloring cells according to a recursive algorithm. Can we generate the sequences by a recursive algorithm that is purely numerical?

We can. Start with 1 as the top number of the first safe pair. Add this to its position number to obtain 2 as the bottom number. The top number of the next pair is the smallest positive integer not previously used. It is 3. Below it goes 5, the sum of 3 and its position number. For the top of the third pair write again the smallest positive integer not yet used. It is 4. Below it goes 7, the sum of 4 and 3. Continuing in this way will generate series A and B.

There is a bonus. We have discovered one of the most unusual properties of the safe pairs. It is obvious from our procedure that every positive integer must appear once and only once somewhere in the two sequences.

Is there a way to generate the two sequences nonrecursively? Yes. Wythoff was the first to discover that the numbers in sequence A are simply multiples of the golden ratio rounded down to integers! (He wrote that he pulled the discovery "out of a hat.")

The golden ratio, as most readers of this column are aware, is one of the most famous of all irrational numbers. Like pi it has a way of appearing in unlikely places. Ancient Greek mathematicians called it the "extreme and mean ratio" for the following reason. Divide a line segment into parts A and B in such a way that the ratio of length A to length B is the same as the ratio of the entire line to A. You have divided the line into a golden ratio. Because this has been widely thought to be the most pleasing way to divide a line, the golden ratio has provoked a bulky literature (much of it crankish) about the use of the ratio in art and architecture.

We can calculate the golden ratio by assigning a length of 1 to line segment B. Our method of dividing the line is expressed by $(A + 1)/A = A/1$, a simple quadratic equation that produces for A a positive value of $(1 + \sqrt{5})/2 =$ 1.61803398..., the golden ratio. Its reciprocal is .61803398.... It is the only positive number that becomes its own reciprocal when 1 is taken from it. In Britain the golden ratio is usually signified by the Greek letter τ (tau). I shall follow the American practice of calling it ϕ (phi).

POSITION (n)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A. $[n\phi]$	1	3	4	6	8	9	11	12	14	16	17	19	21	22	24
B. $[n\phi^2]$	2	5	7	10	13	15	18	20	23	26	28	31	34	36	39

The first 15 safe pairs in W. A. Wythoff's nim

The numbers in sequence A are given by the formula $[n\phi]$, where n is the position number and the brackets signify discarding the fractional part. B numbers can be obtained by adding A numbers to their position numbers, but it turns out that they are rounded-down multiples of the square of phi. The formula for sequence B, therefore, is $[n\phi^2]$. The fact that every positive integer appears once and only once among the safe pairs can be expressed by the following remarkable theorem: The set of integers that lie between successive multiples of phi and between successive multiples of phi squared is precisely the set of natural numbers.

Two sequences of increasing positive integers that together contain every positive integer just once are called "complementary." Phi is not the only irrational number that generates such sequences, although it is the only one that gives the safe pairs of Wythoff's nim. In 1926 Sam Beatty, a Canadian mathematician, published his astounding discovery that any positive irrational number generates complementary sequences.

Let k be the irrational number. Sequence A consists of multiples of k, rounded down, or $[nk]$, where n is the position number and the brackets indicate discarding the fraction. Sequence B consists of rounded-down multiples of $k/(k-1)$, or $[nk/(k-1)]$. Complementary sequences produced in this way are called Beatty sequences. If k is phi, the second formula gives rounded-down multiples of $1.618+/.618+ = 2.618+$, which, owing to the whimsical nature of phi, is the square of phi. Readers might like to convince themselves that Beatty's formulas do indeed produce complementary sequences by letting $k = \sqrt{2}$, pi, e or any other irrational, and that rational values for k fail to produce such sequences.

Whenever the golden ratio appears, it is a good bet that Fibonacci numbers lurk nearby. The Fibonacci sequence is 1, 1, 2, 3, 5, 8, 13, 21, 34..., in which each number after the first two is the sum of the two preceding numbers. A general Fibonacci sequence is defined in the same way, except that it can begin with any pair of numbers. A property of every Fibonacci sequence is that the ratio of adjacent terms gets closer and closer to phi, approaching the golden ratio as a limit.

If we partition the primary Fibonacci sequence into pairs, 1/2, 3/5, 8/13, 21/34..., it can be shown that every Fibonacci pair is a safe pair in Wythoff's nim. The first such pair not in this sequence is 4/7. If we start another Fibonacci sequence with 4/7, however, and partition it 4/7, 11/18, 29/47..., all these pairs are also safe in Wythoff's nim. Indeed, these pairs belong to a Fibonacci sequence of what are called Lucas numbers that begins 2, 1, 3, 4, 7, 11....

Imagine that we go through the infinite sequence of safe pairs (in the manner of Eratosthenes' sieve for sifting out primes) and cross out the infinite set of all safe pairs that are pairs in the Fibonacci sequence. The smallest pair that is not crossed out is 4/7. We can now cross out a second infinite set of safe pairs, starting with 4/7, that are pairs in the Lucas sequence. An infinite number of safe pairs, of which the lowest is now 6/10, remain. This pair too begins another infinite Fibonacci sequence, all of whose pairs are safe. The process continues forever. Robert Silber, a mathematician at North Carolina State University, calls a safe pair "primitive" if it is the first safe pair that generates a Fibonacci sequence. He proves that there are an infinite number of primitive safe pairs. Since every positive integer appears exactly once among the safe pairs, Silber concludes that there is an infinite sequence of Fibonacci sequences that exactly covers the set of natural numbers.

Take the primitive pairs 1/2, 4/7, 6/10, 9/15... in order and write down their position numbers, 1, 3, 4, 6.... Does this sequence look familiar? As Silber shows, it is none other than sequence A. In other words, a safe pair is primitive if and only if its position number is a number in sequence A.

Suppose you are playing Wythoff's game with a very large number of counters or on a chessboard of enormous size. What is the best way to determine whether a position is safe or unsafe, and how do you play perfectly if you have the win?

You can, of course, use the phi formu-

...	10	9	8	7	6	5	4	3	2	1	
...	55	34	21	13	8	5	3	2	1	1	
				1	0	0	1	0	1	0	= 17

Fibonacci notation for 17

Leaner, cleaner and 10 inches trimmer in size—the new 1977 Thunderbird still has the things that make Thunderbird famous—the special luxury, the features, the appointments. They're all there.

Every inch a Thunderbird. From the sumptuous velour upholstery and the luxurious comfort of its optional split bench seat (shown above) to the handsome design and fine craftsmanship of its instrument panel ... Thunderbird truly lives up to its name.

With all that, it's almost unbelievable that this Thunderbird is sticker priced dramatically less than last year's Thunderbird. So see it at your local Ford Dealer and let a test drive convince you.

FORD

When America needs a better idea. Ford puts it on wheels.

THUNDERBIRD

FORD DIVISION Ford

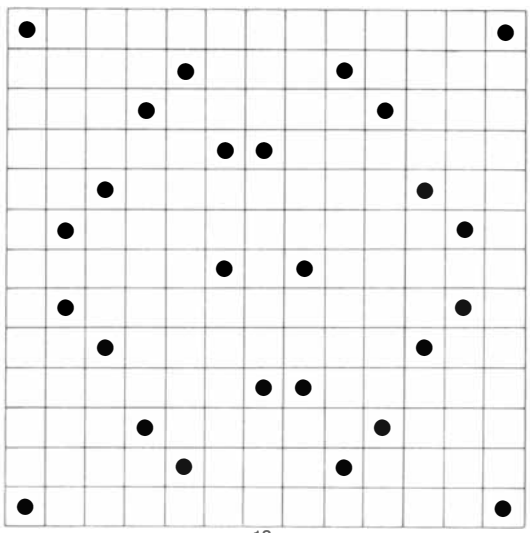
las to write out a sufficiently large chart of safe pairs, but this is hard to do without a calculator. Is there a simpler way comparable to the technique of playing perfect nim by writing the pile numbers in binary notation? Yes, there is, but it uses a more eccentric type of number representation called Fibonacci notation that has been intensively studied by Silber and his colleague Ralph Gellar and also by other mathematicians such as Leonard Carlitz of Duke University.

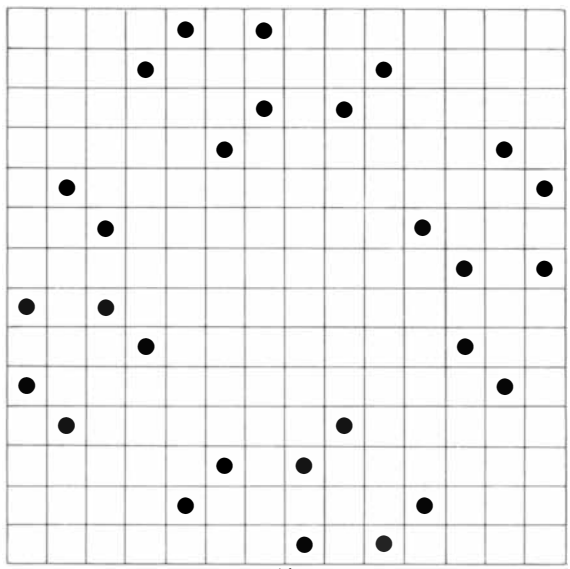
Write the Fibonacci sequence from right to left as is shown in the bottom illustration on page 137. Above it number the positions from right to left. With the aid of this chart we can express any positive integer in a unique way as the sum of Fibonacci numbers. Suppose we want to write 17 in Fibonacci notation. Find the largest Fibonacci number that is not greater than 17 (it is 13) and put a 1 below it. When we move to the right, we find the next number that, added to 13, gives a sum that does not exceed 17. It is 3, and so a 1 goes below 3. When we move to the right again, the next number that gets a 1 is the 1 in the second position. The unused Fibonacci numbers get 0's.

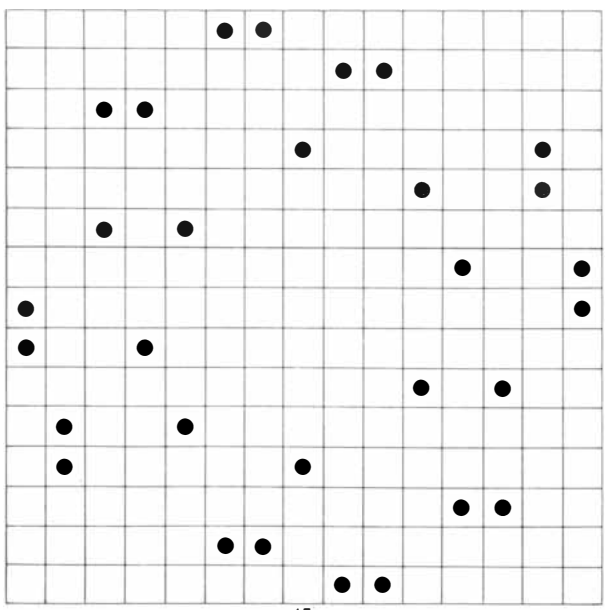
The result is 1001010, a unique representation of 17. To translate it back to decimal notation sum the Fibonacci numbers indicated by the positions of the 1's: $13 + 3 + 1 = 17$. The 1 farthest to the right in the Fibonacci sequence is never used, so that all numbers in Fibonacci notation end in 0. It is also easy to see there are never two adjacent 1's. If there were, they would have a sum equal to the next Fibonacci number on the left, and our rules would give that number a 1 and give 0's to the original pair of adjacent 1's.

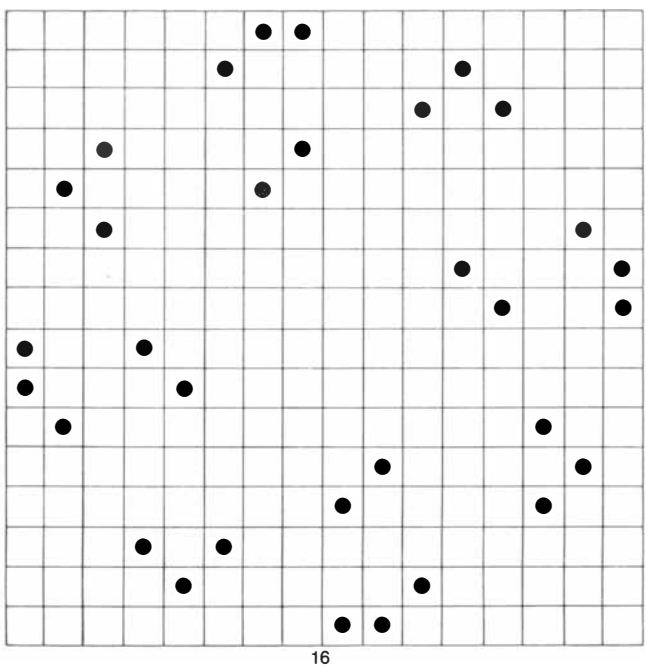
In Fibonacci notation the sum of a safe pair is the *B* number with 0 appended. From this it follows that the Fibonacci sequence is obtained by starting with 10 and adding 0's: 10, 100, 1000, 10000.... The same procedure gives any Fibonacci sequence generated by a primitive pair. For example, the Lucas sequence starting with 4/7 is 1010. 10100, 101000, 1010000....

Every *A* number in Fibonacci notation has the 1 farthest to the right at an even position from the right. Every *B* number is obtained by adding 0 to the right of its *A* partner. Therefore every *B* number has the 1 farthest to the right in


13


14


15


16

New solutions for the maximum no-three-in-line problem

an odd position. Since every counting number is either an A number or a B number, we have a simple way of deciding whether a given position in Wythoff's nim is safe or unsafe. Write the two numbers in Fibonacci notation. If the smaller one is an A number, and if adding 0 produces the other number, the position is safe, otherwise it is unsafe.

An example of the method is $8/13 = 100000/1000000$. The 1 in 100000 is at position 6, an even position, so that 100000 is an A number. Adding 0 produces $1000000 = 13$, the partner of 8. We know that 8/13 is safe. If it is your turn, your opponent has the win. If you think he cannot play perfectly, make a small random move and hope that he soon will make a mistake.

If the pair is unsafe and it is your turn, how can you determine the safe position to which you must play? There are three cases to consider. In each case call the unsafe pair x/y, with x the smaller number, and write both numbers in Fibonacci notation.

In the first case x is a B number. Move to reduce y to the number equal to the number obtained by deleting the right-hand digit of x. For example, $x/y = 10/15 = 100100/1000100$. Since 100100 has the 1 farthest to the right at an odd position, it is a B number. Delete its last digit to obtain $10010 = 6$. The safe numbers you must produce (by removing from the larger pile) are 10 and 6. On a chessboard this corresponds to an orthogonal queen move.

In the second case x is an A number, but y exceeds the number obtained by appending 0 to x. Move to reduce the value of y to that number, for example $x/y = 9/20 = 100010/1010100$. Because x's 1 farthest to the right is in an even position it is an A number. Appending 0 produces $1000100 = 15$. This is less than 20. Therefore the safe pair to play to is 9/15. On the chessboard this too is an orthogonal queen move.

If the numbers do not conform to cases 1 and 2, do the following:

1. Find the positive difference between x and y.

2. Subtract 1, express the result in Fibonacci notation and change the last digit to 1.

3. Append 0 to get one number. Append two 0's to get a second number. These two numbers are the safe pair you seek, even though the resulting Fibonacci numbers may be "noncanonical" in having consecutive 1's.

An example of the third case is $x/y = 24/32 = 10001000/10101000$. The first and second cases do not apply. The difference between 24 and 32 is 8. Subtracting 1 leaves 7. In Fibonacci notation 7 is 10100. Changing the last digit to 1 produces 10101. Appending 0 and 00 yields the safe pair $101010/1010100 = 12/20$. This result is reached by taking 12 from both piles. It corresponds to a diagonal queen move.

It is impossible to go into the whys of Silber's bizarre strategy. Interested readers will find the proofs in Silber's paper, "Wythoff's Nim and Fibonacci Representation," in the February issue of *The Fibonacci Quarterly*. (The periodical is available at $15 per year from the Fibonacci Association, Leonard Klosinski, University of Santa Clara, Santa Clara, Calif. 95053.)

Neither can I go into the ways in which Wythoff's game has been generalized, but a word or two should be added about the game's misère form: the last person to play loses. As T. H. O'Beirne makes clear in *Puzzles and Paradoxes* (Oxford, 1965), misère Wythoff's nim, like misère nim, requires only a trivial alteration of the chart of safe pairs. Remove the first pair, 1/2, and substitute 0/1 and 2/2. The misère strategy is exactly like the standard strategy except that at the end you may have to play to 2/2 or 0/1 instead of 1/2.

Let us modify Wythoff's nim as follows. A player may take any positive number of counters from either pile, or he may take one counter from one pile and two counters from the other. Can the reader determine the chessboard model and the winning strategy before it is given here next month?

Here are the answers to last month's questions:

The scrambled poem is constructed from the following lines of Oscar Wilde's "The Ballad of Reading Gaol."

The vilest deeds like poison weeds
Bloom well in prison air:
It is only what is good in Man
That wastes and withers there.

The secret of Mary Youngquist's poem is that the last letter of every word is the initial letter of the next word. The property even extends to the poet's name at the end.

The letters etaoin shrdlu can be rearranged to spell South Ireland.

The unsolved maximum no-three-in-line problem of October was to prove or disprove that on a chessboard of side n it is always possible to place $2n$ counters so that no three are in "line." Lines may have any orientation. I gave solutions for boards through order 12. Several readers have extended solutions through order 16.

Richard Byfield, Richard Jacobson, Anne De Lamper and Robert Van Clampitt each found by hand an order-13 solution. Michael Meierruth wrote a nonexhaustive computer program that turned out 29 solutions of order 13 and one of order 15. Later he found a solution for order 14 and four solutions for order 16. Eric Jamin found by hand four order-14 solutions and one for order 16. The illustration on the preceding page shows sample patterns for 13, 14 and 16 and the only known pattern for 15.

What owning an original print says about you.

Anyone with a dollar or two can own one of millions of reproductions of fine art.
Very few people can own one of a limited edition of a fine original print, even if the price is quite modest.
Which immediately sets you apart.
When you own an original print, you've acknowledged the need for quality, beauty, and individuality in your surroundings.

You've made a commitment to art, and backed it up by investing in your own good taste.

And you own something of real aesthetic and monetary value.

(Unlike reproductions, which almost never grow more valuable, original prints by recognized artists tend to appreciate, sometimes considerably.)

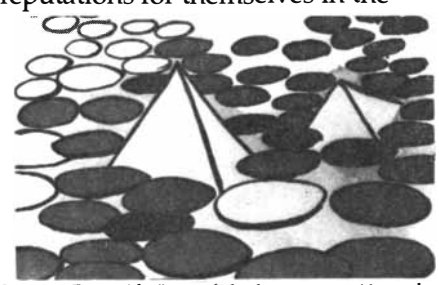
At the Original Print Collectors Group, we offer original etchings, engravings, lithographs, woodcuts, and silkscreen prints, by artists whose works hang in the world's great museums and galleries.

These include prints by Calder, Chagall, Dali, Miro, and Picasso, as well as artists who are just beginning to make reputations for themselves in the world of art.

Gatja Rothe used the method of mezzotint engraving, almost a lost art, to create the striking "Moondance". Signed limited edition of 160.

Each print we offer is a true, museum-quality original.

Each print is made from an image created directly on the plate by the artist himself.



"Circle of Pyramids," one of the late master Alexander Calder's last original lithographs. Signed limited edition of 50.

Each is hand-printed, signed and numbered by the artist, custom-framed, and accompanied by a money-back guarantee and a certificate of authenticity.

Wayne Cooper's original lithograph "Rainbow," a timeless record of rural America. Signed limited edition of 300.

And each print has been chosen not only for its beauty, but for its investment potential.

(Over the past decade, works by the artists we offer have, on the average, appreciated more than 30% *each year*. Perhaps that's why *Business Week*, *Money*, and UPI have quoted us as an authority on investing in art.)

In the etching "Clearing in the Woods," Alvin Sandler's masterful impressionism captures the immediacy of a moment in nature. Signed limited edition of 99.

To enhance our members' enjoyment of their original prints, we keep them informed with frequent newsletters, which tell how to recognize original prints; the different methods by which original prints are made; and what trends and which artists are causing the most excitement among knowledgeable collectors and dealers around the world.

Owning an original print says that you have a sincere interest in art appreciation *and* in art that appreciates.

And the best way to develop that interest is to find out more about the Original Print Collectors Group. Send in the coupon. There's no obligation.

Original print collectors group, Ltd.
120 East 56th Street, Dept. SA6, New York, NY10022

☐ Please send me your free color brochure, latest newsletter, and membership information.

Name_____

Address_____

City_____ State_____ Zip_____

BOOKS

A flotilla of picture books, map symbolism, muscular exercise and the Pantheon of Rome

by Philip Morrison

WEEDS IN WINTER, written and illustrated by Lauren Brown. W. W. Norton & Company, Inc. ($8.95). BERGGASSE 19: SIGMUND FREUD'S HOME AND OFFICES, VIENNA 1938, THE PHOTOGRAPHS OF EDMUND ENGELMAN, with an introduction by Peter Gay and captions by Rita Ransohoff. Basic Books, Inc. ($25). THE WORLD YOU NEVER SEE: UNDERWATER LIFE, written and photographed by Peter Parks. Rand McNally & Company ($9.95). BIRDS OF THE WEST COAST, paintings, drawings and text by J. F. Lansdowne. Houghton Mifflin Company ($40). Printed books mainly bring text, but of course their pages can carry images as well. Johannes Gutenberg himself knew it; his aim was to mechanize not only the work of scribes but also the art of the illuminator and the miniaturist. Here are four volumes as disparate in medium as they are in topic but all distinguished by their visual displays.

Weeds in Winter is a book as modest as the quiet winter countryside from New York to Boston, which is its center: spare and subdued but rewarding close inspection. Those roadsides and fields are not empty. The same herbaceous plants, the pinks and the grasses and the worts that delight and intrigue the wanderer by summer, are there still, "dead, woody tissue, various shades of brown or gray, sometimes with fruits, sometimes just as stalks; and many of them are spectacularly beautiful." The gifted young author and artist has made a field guide for winter that is clear enough for the wildflower beginner, wisely concentrating on the most common and most conspicuous, just what one most easily notices.

The plants included, about 135 species, can be found from the Atlantic to Minnesota and from lower Canada to Virginia. The careful key and the lively comments are clear and instructive. Page after page bears filigreed inky drawings of the plants—hardhack and tickseed, loosestrife and cinquefoil—transformed to a purity of structure that delights the eye yet is tinged by a philosopher's melancholy. Thoreau saw it the same way: "More obvious and interesting frequently than in summer even, as if their beauty was not ripe till then... decent weeds at least, which widowed nature wears." Here is the reed grass *Phragmites*, the "plant of industrial America" whose 20-foot bamboolike stems crowd dying salt marshes such as the New Jersey flats, no longer under the tides; bergamot, its heads crowded with flower tubes that crush to release a richly minty smell; teasel, its bracts curving elegantly around a head like a spiny egg. (Here the author's common sense transcends her sources; she makes it plain that this fragile teasel head could not have been adequate to the heavy task of fulling newly woven cloth in the strict sense—thickening and softening it—as some authorities appear to say. It was, however, used for raising the nap of woven woolens, which is part of fulling in the broad sense: the cloth-finishing process. Teasel was still grown for this purpose as recently as 20 years ago, as it has been for a long time. We can read in *Piers Plowman:* "Cloth that cometh from the wevying is noght comly to were/Tyl it is fulled under fote, or in fullying stokkes / Wasshen wel with water and with tasles cracched.") Just to turn these delicately figured pages is to evoke the long shadows and the complex brown vistas of our countryside.

The black-and-white photographs of *Berggasse 19* record a scene not from nature's winter but from a winter of human society. In the dreary spring of 1938 Edmund Engelman (now a New York photographic consulting engineer)

Freud's consulting room, shown in* Berggasse 19: Sigmund Freud's Home and Offices, *was filled with antiquities

CHOOSE ANY ONE OF THESE VALUABLE COMPLETE WORKS

as a trial member of the BOOK-OF-THE-MONTH CLUB®

The Suggested Trial: You simply agree to buy four Club choices within a year at substantial savings on most books you choose

CARL SANDBURG's *Pulitzer Prize-Winning* ABRAHAM LINCOLN

The entire six-volume set

More than a portrait of a great American, this monumental 6-volume biography is a tapestry of America in crisis and resolution. Winner of a Pulitzer Prize for history...a work of timeless beauty and humanity.

for only $18.50

PUBLISHER'S LIST PRICE: $120

THE STORY OF CIVILIZATION

by WILL and ARIEL DURANT

Eleven matched volumes

This multi-volume history of all mankind—ancient and modern, Occidental and Oriental, is one of the great literary achievements of the twentieth century.

for only $24.75

PUBLISHER'S LIST PRICE: $244.45

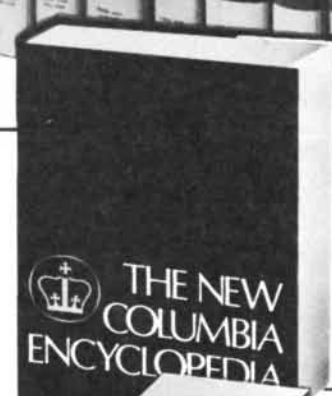
THE NEW COLUMBIA ENCYCLOPEDIA

New fourth edition . . . completely revised after a lapse of twelve years

Cited in earlier editions as the best single-volume reference work, this latest illustrated edition has over 7000 new articles from world affairs to sports, many useful features: list of Nobel Prize winners...tables of rulers of many countries...Third World persons, events, art forms and places.

for only $15

PUBLISHER'S LIST PRICE: $79.50



THE COMPACT EDITION of the OXFORD ENGLISH DICTIONARY

Contains the complete contents of the 13-volume *Oxford English Dictionary*, generally considered the final arbiter of the meaning, origin, history and usage of words in the English language. Bausch & Lomb magnifying glass included.

for only $17.50

PUBLISHER'S LIST PRICE: $95

IF YOU are not aware of the many clear advangages, both economic and intellectual, of shopping in America's Bookstore,® the dramatic opportunity outlined here is well worth your attention. Until now, high retail prices have put the magnificent works shown here beyond the reach of many reading families. However, as a trial member of the Book-of-the-Month Club, you can choose any one of these works at an extraordinary saving—as much as $219.70

This bounty for booklovers is not a one-time affair but is a demonstration of a continuing benefit. Not only will you be able to keep abreast of the important new books you are eager to read, but if you continue your membership you will enjoy similar savings on expensive library sets and high-priced reference works. It's a wonderful way to build a well-rounded library and save money, too.

BOOK-OF-THE-MONTH CLUB, INC.
Camp Hill, Pennsylvania 17012

was a dispossessed former owner of a photography shop in Vienna, waiting to emigrate. He had been a skilled devotee of photography from childhood, and a mutual friend arranged for him to enter the Freud apartment inconspicuously (the Gestapo was watching) in order to document its look for posterity, beyond the years of the Third Reich. The Freud family had been given permission to seek refuge in London; they were to flee within 10 days. The photographer had to be discreet: flash or floodlights would not do, and the 82-year-old Freud could not be disturbed; the photographer had to work his way through the rooms attentive to Freud's well-established routine in order not to encounter the preoccupied old man.

On the third day the two nonetheless met each other, young photographer face to face with old seer. The mutual friend smoothed out the encounter, and Freud looked at the collection of prints, already growing, that would record the Freud home. He smiled broadly and said: "Ich danke Ihnen herzlich. Das wird für mich viel bedeuten." (Idiomatically, "I am most grateful. It means a lot to me.") We too are grateful; the photographs mean much to us also. There are 54 plates, including portraits of the entire family, ranging from the street and the façade to the silent ranks of figurines from Rome, China and Egypt that crowded the desk where Freud worked. There is an intimate gloss for each plate and a perceptive and graceful introduction by Peter Gay, a distinguished American historian of German culture.

It was in this place, notable mainly for the embarrassment of objects it held, that Freud once "playfully wondered ... whether some day there might be a 'marble tablet'" explaining that here "the Secret of Dreams was revealed to Dr. Sigm. Freud." The photographs cannot of course reveal or conceal, but they do give signs, as did the Delphic oracle and the writer and thinker who collected those crowded pictures, rugs and sculptures. The depths of the human unconscious, its tension of clear reason and dark fear, can hardly be more sharply present than they are in the third photograph: in the street in front of Berggasse 19, between two everyday shop doors, we see across Freud's lintel a bright Nazi banner, the dread *Hakenkreuz,* which we know as black on red.

In *The World You Never See* true bright colors appear: close to 300 color photographs, together with a substantial personal text and full captions. The author is a zoologist-microscopist-filmmaker with Oxford Scientific Films, justly held to be about the best team now working in the rich domain of closeup wildlife cinematography. Most of the pictures are macro- or micrographs, the world seen at magnifications from a few diameters to a few hundred. We get some idea of what it takes to make such photographs in a couple of pictures of the special equipment Peter Parks has developed. His dark-field optical benches are thickets of lamp tubes and adjusting racks. He has taken them to Jamaica and Bermuda, to the shore near Melbourne and to the wetlands of Wisconsin, and as a valuable by-product of what was primarily television motion-picture production he has arranged this striking book. In it the watery world is sampled widely, from the familiar *Daphnia* and ostracods in a drop of pond water to the life of a trout stream and on to the shores of beaches around the world, surface plankton, the rich middle depths, the deep waters and the diatoms and radiolarians that can appear in seawater under high magnification.

The anecdotal quality of the readable text, which is never very technical, makes this a book more for browsing than for reference. The images are genuinely beautiful. A few animals, such as the ghost crab pirouetting on the Bermuda sand and the sleeping perch, are from our hand-sized world. The remainder, as the title asserts, are from a world we never see. Although many of them are known to readers from drawings and photographs, quite a few are first shown here, and few can be seen so beautifully in any other collection. The pelagic snail *Janthina* is familiar to collectors for its delicate violet shell; here its extraordinary life (it floats on a bubble raft on the high seas until by chance it collides with the small jellyfish *Velella,* on which it preys) is documented in photographs as colorful as they are strange. The eye of a needle helps to establish many a shot. Clumps of sargasso weed are disclosed as so many floating oceanic meadows, grazed by barnacles and prowled by incredibly well-hidden tiny crabs and a little clambering angler fish, whose fin rays act like clasping fingers that clench and release the fronds. The book is a real bargain at its price; it is eye-catching, exploiting interference colors and the dramatic dark-field technique for visual excitement. It ought to be an inducer and a complement to more systematic reading, as in the books of Alister Hardy, who adds a one-page foreword in honest praise.

Birds of the West Coast exhibits carefully reproduced paintings in full natural color. They represent wild birds mostly at life size; the birds appear in extraordinary detail, lifelike and attentive, floating against a white world empty save for the firm rest on which the creature stands, a twig or a rail or a square foot of beach. There are scores of color plates, four of which are fold-outs double the 10-by-14-inch page size of the book (which is to be complemented by a second similar volume).

The artist is a young Canadian painter of birds who has for 15 years grown in achievement within the subculture of birders. The foreword is by a famous birder, S. Dillon Ripley of the Smithsonian Institution. Ripley sees in this book the evidence that the artist's "bird representation has come of age." The birds are portrayed "with a superb romantic power. They are as impressive in their way as Audubon's birds were a century and a half ago." No new continent is on display here, but it is apparent even to a cool onlooker that these paintings reach a standard beyond the steady competence of the many admirably illustrated books that stream out for this audience. Lansdowne knows birds well; he has lived on Vancouver Island, where he could see the snowy owl, the bald eagle, the tufted puffin and the common little chickadee. These he presents in splendor, along with 50 more. The bald eagle still lives in that region in "relatively undisturbed abundance." Once on a brilliant frosty winter's day he saw a "gathering of eagles.... A hundred or so had come to glut themselves on the thousands of dead salmon washed down from the spawning grounds."

THE NATURE OF MAPS: ESSAYS TOWARD UNDERSTANDING MAPS AND MAPPING, by Arthur H. Robinson and Barbara Bartz Petchenik. The University of Chicago Press ($8.95). All too uncommon is the aim of this small, serious book: applied philosophy. Here is the thought of two practitioners of a specialized science, cartography, engaged in self-conscious examination of the meaning of what they do. The senior author is professor of cartography at the University of Wisconsin, the junior a working cartographic editor with a graduate degree. Their readable arguments ought to encourage many another worker in a specialized discipline to make a start toward a general theory, as they here courageously begin their long road toward a general theory of cartography. Such metascience is too important and too difficult to be left to the philosophers.

The book begins with a sense of shock. The very literature through which they had hoped to come to understand their own domain was founded on a "universal metaphor": the map itself. How to gain understanding when the map you do not understand is the means of explanation? If all theories seem to be like maps, then a theory of maps is surely a reflexive and convoluted goal. The authors proceed to simplify the now too heavy task. It will be enough to consider tangible, visible maps, not metaphorical representations of a part of physical space. Space is too impersonal a word for them; they prefer the term "milieu" to connote the relation of the cartographer to the region he describes. For neither mapper nor percipient can the map

Model CRF-320 Horizon Master Portable Shortwave Radio

Leave it to Sony to bring the whole world right in under your thatched roof.

With this amazing new radio you could tune in everywhere from Aberdeen to Zanzibar. Relax with a coconut cocktail and the BBC. Spend Sundays with the NFL.

The radio has 32 worldwide bands, and includes the entire shortwave spectrum. You'd even know if it's snowing on Kilimanjaro.

There are three different power systems: AC, batteries, and optional 12V car battery.

Inputs for tape recorder playback and monitoring.

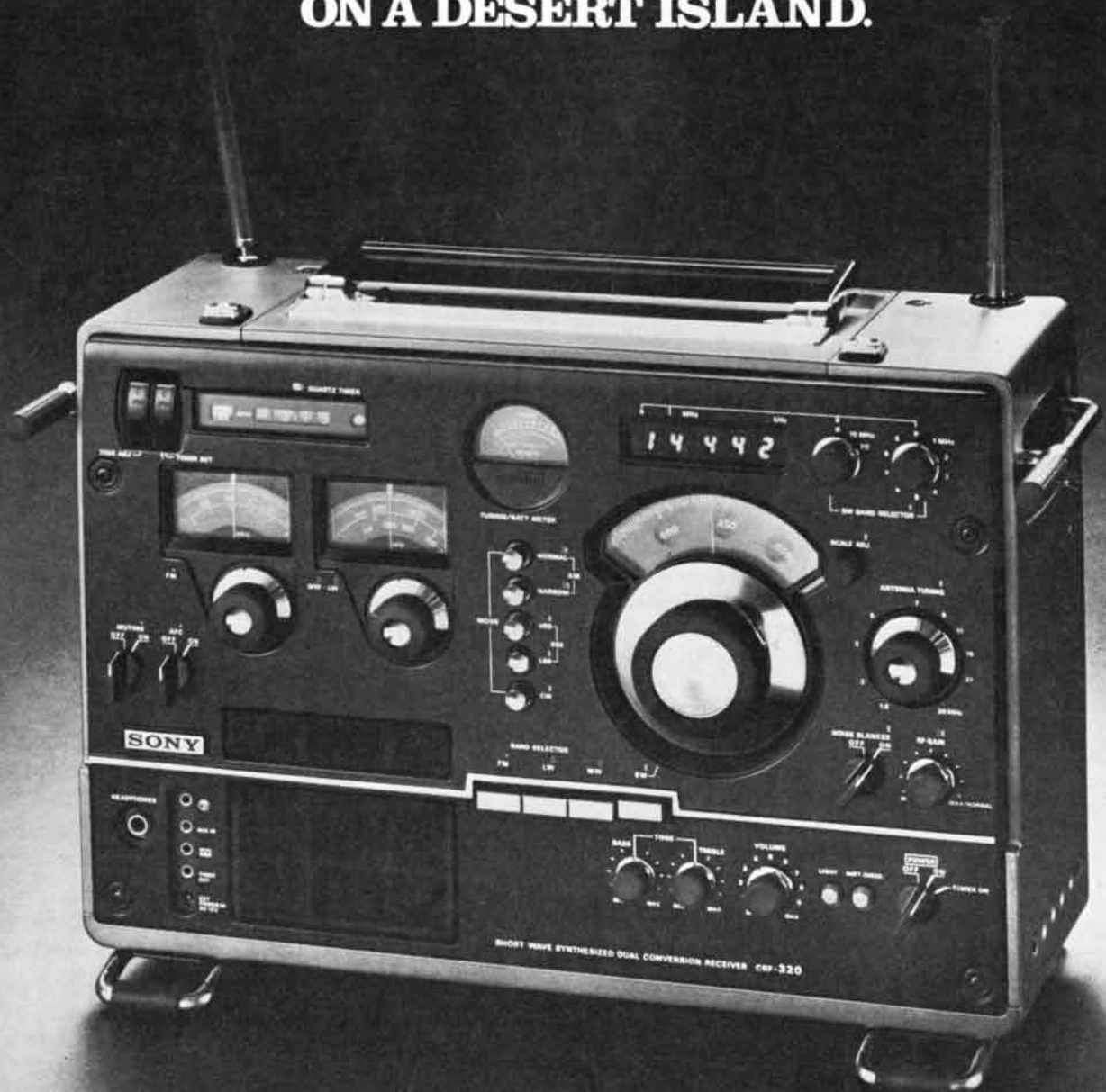
A crystal-controlled dual conversion synthesizer with LED frequency display for ultra-precise tuning. Which simply means you don't have to fiddle the dials to find France.

And a quartz-crystal digital clock accurate to within 10 seconds a month.

It's also equipped with an automatic Timer Set which activates the radio and tape recorder output. So, conceivably, you could set the dials to a different frequency each night, and wake up to a new country each morning.

Isn't it wonderful? Even on a desert island, Sony can make your world complete. Well, almost.

THE SECOND THING YOU'D WANT WITH YOU ON A DESERT ISLAND.



"IT'S A SONY."®

be a mere passive modeling of some portion of the world. Seeing is itself active, a comparison with internal models, an ordering never freely given by what is without but actively constructed within by mapper and map reader alike.

Robinson and Petchenik follow the leads provided in plenty by the philosophers and psychologists of communication today. They wander from Claude Shannon and R. V. L. Hartley to Michael Polanyi and Jean Piaget, from Susanne Langer to Karl Lashley, without much light. Information theory holds little for them; indeed, they touch on the central idea of dimensionality, but they do not emphasize the problem that lies in the long, long tapes of the universal Turing machine. Information theory can reduce every structure to a linear tape that is its formal equivalent. That is only a clumsy and tedious substitute for the flashing clarity we can gain from multidimensions, which offer the viewer free choice among the dense cluster of potential messages in every map. The issue turns on one question: the concept of space itself. Here the authors rely on Piagetian experiments and the personal interpretations of Polanyi. These may be too light a framework for the task (although they are surely of value). On them Robinson and Petchenik rest their most explicit result: a start at the classification of maps.

They take two dimensions, one the logical mode of like and unlike, the other the geometrical one of neighboring and separated. A general map of one area shows "everything" that lies together, and an atlas of such maps represents separation. But a thematic map, say a map of rainfall, collects like properties in space and separates out rainfall in distinct places. Thus are established the functional classes of maps. Another cut establishes the scale classes: one can map a small area topographically, map "everything" in some area or map only the orange groves. A fourth box includes long-view maps, say of Europe, along with thematic maps of oranges around the world. Here in some disarray Robinson and Petchenik close, touching their central conclusion: "The concept of spatial relatedness . . . which indeed is the reason for the very existence of cartography, is a quality without which it is difficult or impossible for the human mind to apprehend anything."

The essays are interesting, but they are only a start. One needs more mathematics. Plainly there is a duality in our apprehensions; might it be seen as the contrast between algebra and geometry? Symbols and words can be combined to make any relation we choose; "mimetic" arrangements in a space of more dimensions appeal to us as being less arbitrary, closer to the world. Yet there arises between the two modes a profound isomorphism. Maybe we de-

tect two styles of two brain hemispheres, but coequal rather than subordinate. There is plenty we do not know.

Two small nuggets: It was Edmund Halley who may have made the first thematic maps, those of winds and of magnetic declination. And the popular step to abstraction today is the cartogram, those maps that show, say, the states drawn with their area scaled to their population. That powerful blend of symbol and spatial relationship sets knotty problems for any logician fallen into reality.

BIOMECHANICS AND ENERGETICS OF MUSCULAR EXERCISE, by Rodolfo Margaria. Oxford University Press ($21). In his own Milan Institute of Human Physiology, Professor Margaria has for a decade tried to elucidate the physiology of human exercise, of man walking, running, hopping strangely across the dusty face of the moon. He began work on these problems with "passion" and "interest" just after the heroic days of human physiology, nearly 50 years back, when Bloomsbury novelists laughed at long rides on the anchored bicycle as a parable of the strangeness and artificiality of science.

This summarizing book treats old and new alike, putting the biochemistry of muscle work and the cunning of our limbs and motions to the stern assay of hard muscular exercise. Here are graphs of the energy changes, step by step, of an Olympic athlete at the start of a sprint. At the peak of his output during this time of maximum acceleration he is releasing about 1.8 kilowatts of power, which is 15 or 20 times the basal metabolic power spent at rest. To measure that all you need to do is ask your subject to run upstairs two steps at a time while you take fast motion pictures. The vertical velocity will yield values quite comparable to the classical method "introduced by Fenn in 1932 with brilliant results." For that one uses a platform sensitive to three components of acceleration, all graphically recorded and properly integrated over time.

Such a platform measures only external work, work that moves the center of gravity of the body. Internal work may be spent against internal friction, or in contracting muscles without movement; these cannot be measured directly. Limb movement that does not move the center of gravity can be measured from motion pictures. It is not too hard to measure the energy cost of running, say by collecting expired air from a treadmill runner. From these data the efficiency can be found; about a third of the work is internal. The efficiency comes out very high, however: 40 percent. That contradicts measurements of efficiency carried out under slow walking conditions, where many possible losses could not be important. It seems the result is misleading. Indeed, the conclusion is hard to avoid: The energy output is overestimated.

What could be wrong? Plainly more than gravitational potential energy and kinetic energy is involved. There is in fact a third invisible energy, which smooths out the curve peaks. The mechanical work done by the muscles is stored in part within the muscles themselves, elastically, and is returned at the right part of the cycle of motion. "The term 'elastic' recurs very frequently in the terminology of . . . athletes, trainers, or the public in general to describe the movement and the performance of athletes." Their instinct is correct: the stored elastic energy is nearly equal to the energy expended in contraction. All the scaling and speed limits confirm this inference. The loss of energy in muscular viscosity, given much importance in the 1920's, is probably not significant at the speeds of runners. A simple model for running is the forward progress of an elastic bouncing ball. (Walking, on the other hand, is more like the end-over-end rolling of an egg.)

It is gravity that fixes our oscillating rhythm of locomotion. Analysis shows that a man could not run on the moon (even if the surface were hard) faster than 12 kilometers per hour, a third of the earthly limit. The energy cost would be low, but the runner would be frustrated. In fact, a jumping technique would be much more suitable.

The first half of the book connects human performance with the mechanisms of muscle-energy release at the biochemical level. It now seems that fast direct-energy release by the splitting of high-energy phosphate bonds supplies our initial power, up to half of the overall oxygen debt. Only later does lactic acid production by the breakdown of the stored starchlike glycogen take part. Finally, of course, the oxygen mortgage must be paid, but that is done over a longer period of time. The studies are confirmed by direct measurements on isolated mouse muscle and by inference from such evidence as the inability of the final sprint in a long race to reach the same speed as a sprint starting from rest. Extra oxygen is of no avail to runners between events, although it does pay to store up glycogen by eating foods rich in carbohydrates for a couple of days before a contest. The mass of muscle determines the amount of high-energy phosphate bonds available, since the composition of muscle cannot change much. Hard training alone builds muscle mass. Perhaps covert injections—aerated blood, glycogen?—will become winning maneuvers.

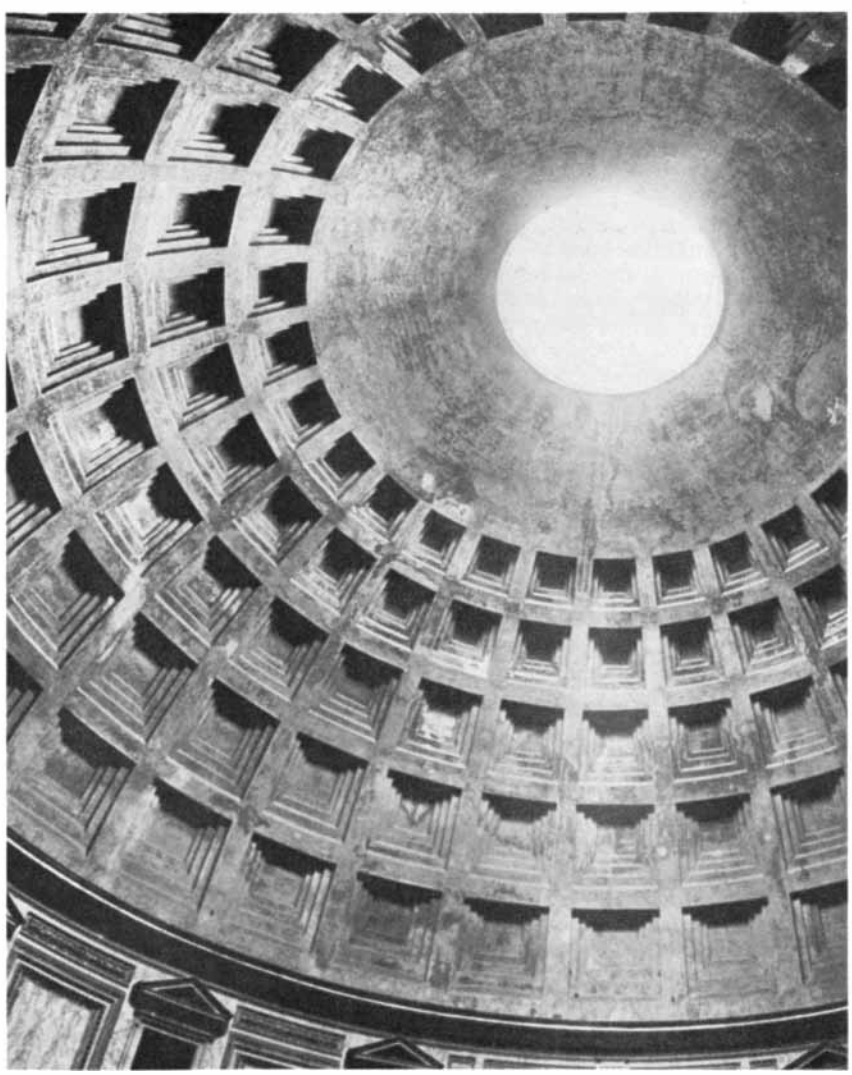
The generally meticulous publishers have nodded over this fine little book. The editing has not helped readers past Professor Margaria's nonstandard usage with regard to units. Let the read-

er remember that "kilogram metres per second" translates into about 10 times as many watts. There are a number of other notational and cross-reference puzzles as well.

THE PANTHEON: DESIGN, MEANING, AND PROGENY, by William L. MacDonald. Harvard University Press ($12.50). Hear Shelley on this "icon of Rome's claims," perhaps the most influential building in the world: "It is as it were the visible image of the universe; in the perfection of its proportions, as when you regard the unmeasured dome of Heaven, the idea of magnitude is swallowed up & lost. It is open to the sky, & its wide dome is lighted by the ever changing illumination of the air. The clouds of noon fly over it and at night the keen stars are seen thro the azure darkness hanging immoveably, or driving after the driven moon among the clouds." Its magnitude is not slight: the dome is larger than any the Renaissance and Baroque masters built after it in stone, wood or iron, whether Brunelleschi in Florence, Michelangelo in Rome, Wren in London or Mansart in Paris.

We do not know its architect, who must have been a skilled professional, but it was the Emperor Hadrian's conception, and his personality molded its boldly original form. We date it securely by the brick stamps found on a number of the large tile-shaped bricks that face many walls, stamped by the brickmakers long ago for inventory or taxation. It was begun after A.D. 117 and dedicated between A.D. 126 and 128 under that cultivated and many-faceted ruler. His name is not on it; indeed, the large inscription he placed there recalls Marcus Agrippa, the eminent minister of Augustus who had built a sanctuary on the same site a century earlier.

The building stands today after almost two millenniums as sound as it was built. There are three parts: the formal columned entrance porch of stone, an intermediate block and the rotunda, a great cylinder topped by a hemispherical dome. Nine-tenths of the structure of the intermediate block and of the rotunda is of concrete. The cylinder was poured in layers between brick walls until the dome was reached; then the pouring continued over "an immense hemispherical form, supported by a forest of timbers" and covered by the negative molds for the powerful coffered ceiling. That concrete hemisphere is graded in the weight of the aggregate in regular layers, the uppermost part of the dome being of concrete made out of pumice. Concrete suited imperial Rome, where gangs of cheap labor could build on the carefully sequenced and organized work of a smaller group of master craftsmen. The long effort of skilled individuals needed for a classical temple in cut marble was costly and anachronistic. At the spring of the dome, to be sure, there is a series of steplike rings of masonry that bind and weight its lower part, ring buttresses, so to speak.



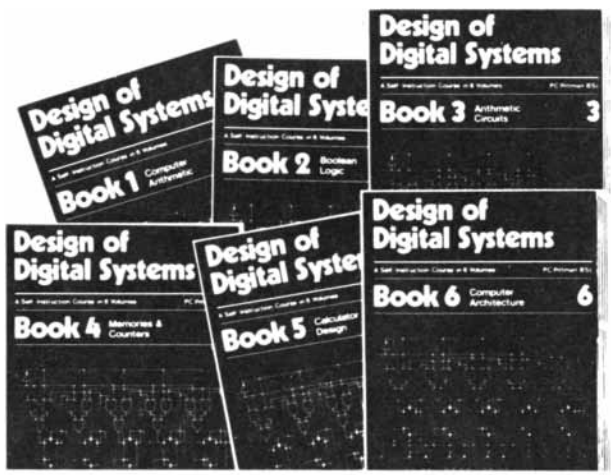
The Pantheon of Rome, from **The Pantheon.** *At the top is the oculus*

Most of this learned and much-illustrated book is architectural and symbolic rather than structural. What the author, professor of art history at Smith College, seeks is an understanding of the meaning of the building as a clue to its enduring and pervasive influence over the whole of the West. We have little documentation from Rome itself, but that little helps. It is pretty sure that the domed rotunda, with the temple front, represented the order of the world invoking the order of the heavens. A 30-foot opening the oculus, admits the heavenly lights 150 feet above the geometrically paved floor. The Pantheon was sacred to all the gods, but surely chief among them were the seven planetary deities in the seven now empty niches. The shape of the building is manifestly unified; for example, the cylinder is just as high as the dome, and its radius and height match. "The theme, of course, was unity—the unity of gods and state, of people and state, and the unity of the perpetual existence and function of the state with the never-ending revolutions of the planetary clockwork . . . , a cosmology expressed in architecture."

The Pantheon was never copied, but buildings were made afresh in its image for churches and noblemen, kings and commonwealths. It has stood in Rome, always accessible to the West, all those years, a symbol that offers the archetype of the "stability of the firmament itself." Palladio built after it, as did Thomas Jefferson, and we know it in the capitols of our republic. Not one other dome displays the full triumph that was Hadrian's, the "long cylinder of light" that moves through the still architecture each day, uniting the fabric of human hands with the ceaseless order of the heavenly round, "neither sacred nor secular, but a place of man and nature," partaking of unchanging law, circular, shining and perpetual.

Understanding Digital Electronics

New teach-yourself courses



Design of Digital Systems is written for the engineer seeking to learn more about digital electronics. Its six volumes — each 11-1/2" x 8-1/4" are packed with information, diagrams and questions designed to lead you step-by-step through number systems and Boolean algebra to memories, counters and simple arithmetic circuits, and finally to a complete understanding of the design and operation of calculators and computers.

The contents of Design of Digital Systems include:

Book 1 Octal, hexadecimal and binary number systems; conversion between number systems; representation of negative numbers; complementary systems; binary multiplication and division.
Book 2 OR and AND functions; logic gates; NOT, exclusive-OR, NAND, NOR and exclusive-NOR functions; multiple input gates; truth tables; De Morgans Laws; canonical forms; logic conventions; Karnaugh mapping; three-state and wired logic.
Book 3 Half adders and full adders; subtractors; serial and parallel adders; processors and arithmetic logic units (ALUs); multiplication and division systems.
Book 4 Flip flops; shift registers; asynchronous and synchronous counters; ring, Johnson and exclusive-OR feedback counters; random access memories (RAMs) and read only memories (ROMs).
Book 5 Structure of calculators; keyboard encoding; decoding display data; register systems; control unit; program ROM; address decoding; instruction sets; instruction decoding; control program structure.
Book 6 Central processing unit (CPU); memory organization; character representation; program storage; address modes; input/output systems; program interrupts; interrupt priorities; programming; assemblers; computers; executive programs; operating systems and time sharing.

Digital Computer Logic and Electronics is designed for the beginner. No mathematical knowledge other than simple arithmetic is assumed, though the student should have an aptitude for logical thought. It consists of four volumes — each 11-1/2" x 8-1/4" — and serves as an introduction to the subject of digital electronics. Everyone can learn from it — designer, executive, scientist, student, engineer.

Contents include: Binary, octal and decimal number systems; conversion between number systems; AND, OR, NOR and NAND gates and inverters; Boolean algebra and truth tables; De Morgans Laws; design of logic circuits using NOR gates; R-S and J-K flip flops; binary counters, shift registers and half adders.

In the years ahead the products of digital electronics technology will play an important part in your life. Calculators and digital watches are already commonplace. Tomorrow a digital display could show your automobile speed and gas consumption; you could be calling people by entering their name into a telephone which would automatically look up their number and dial it for you.

These courses were written by experts in electronics and learning systems so that you could teach yourself the theory and application of digital logic. Learning by self-instruction has the advantages of being faster and more thorough than classroom learning. You work at your own pace and must respond by answering questions on each new piece of information before proceeding.

After completing these courses you will have broadened your career prospects and increased your fundamental understanding of the rapidly changing technological world around you.

The six volumes of Design of Digital Systems cost only: **$19.88**

And the four volumes of Digital Computer Logic and Electronics cost only: **$14.88**

But if you buy both courses, the total cost is only: **$29.90**

a saving of over: **$5.00**

SEVEN-DAY MONEY-BACK GUARANTEE: If you are not satisfied with your Cambridge course, return it within 7 days for a full refund.

To order your books, complete the order form below and send it together with your check or money order to GFN Industries, Inc., 6 Commercial Street, Hicksville, N.Y. 11801.

To: GFN INDUSTRIES, INC.
6 COMMERCIAL STREET, HICKSVILLE, NY 11801

Please send me:

___ Sets of Design of Digital Systems $19.88 _____
___ Sets of Digital Computer Logic & Electronics $14.88 _____
___ Sets of both courses $29.90 _____
Sales tax (N.Y. residents) _____
Shipping and handling $2.50 per set _____
Enclosed is check/mo (payable to GFN Industries, Inc.)

Total $_____

Name _____

Address _____

City/State/Zip _____

Prices include overseas surface mail postage. SA 3B

BIBLIOGRAPHY

Readers interested in further reading on the subjects covered by articles in this issue may find the lists below helpful.

SUPERPHÉNIX: A FULL-SCALE BREEDER REACTOR

FAST BREEDER REACTORS. Glenn T. Seaborg and Justin L. Bloom in *Scientific American,* Vol. 223, No. 5, pages 13–21; November, 1970.

THE NECESSITY OF FISSION POWER. H. A. Bethe in *Scientific American,* Vol. 234, No. 1, pages 21–31; January, 1976.

THE BREEDER: WHEN AND WHY. *EPRI Journal,* Vol. 1, No. 2; March, 1976.

WAVES IN THE SOLAR WIND

THE SOLAR WIND. E. N. Parker in *Scientific American,* Vol. 210, No. 4, pages 66–76; April, 1964.

INTRODUCTION TO THE SOLAR WIND. John C. Brandt. W. H. Freeman and Company, 1970.

CORONAL EXPANSION AND SOLAR WIND. A. J. Hundhausen. Springer-Verlag, 1972.

SOLAR WIND STREAM EVOLUTION AT LARGE HELIOCENTRIC DISTANCES: EXPERIMENTAL DEMONSTRATION AND THE TEST OF A MODEL. J. T. Gosling, A. J. Hundhausen and S. J. Bame in *Journal of Geophysical Research,* Vol. 81, No. 13, pages 2111–2122; May 1, 1976.

OPIATE RECEPTORS AND INTERNAL OPIATES

OPIATE RECEPTOR: DEMONSTRATION IN NERVOUS TISSUE. Candace B. Pert and Solomon H. Snyder in *Science,* Vol. 179, No. 4077, pages 1011–1014; March 9, 1973.

PROPERTIES OF OPIATE-RECEPTOR BINDING IN RAT BRAIN. Candace B. Pert and Solomon H. Snyder in *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 70, No. 8, pages 2243–2247; August, 1973.

REGIONAL DISTRIBUTION OF OPIATE RECEPTOR BINDING IN MONKEY AND HUMAN BRAIN. Michael J. Kuhar, Candace B. Pert and Solomon H. Snyder in *Nature,* Vol. 245, No. 5426, pages 447–450; October 26, 1973.

OPIATE RECEPTOR MECHANISMS. Edited by Solomon H. Snyder and Steven Matthysse with Huda Akil, Ian Creese, Peter A. Mansky and Gavril W. Pasternak in *Neurosciences Research Program Bulletin,* Vol. 13, No. 1; February, 1975.

THE OPIATE RECEPTOR: A MODEL EXPLAINING STRUCTURE-ACTIVITY RELATIONSHIPS OF OPIATE AGONISTS AND ANTAGONISTS. Andrew P. Feinberg, Ian Creese and Solomon H. Snyder in *Proceedings of the National Academy of Sciences of the United States of America,* Vol. 73, No. 11, pages 4215–4219; November, 1976.

BIOLOGICAL NITROGEN FIXATION

THE BIOLOGY OF NITROGEN FIXATION. Edited by A. Quispel. North Holland Publishing Company, 1974.

SURVEY OF NITROGENASE AND ITS EPR PROPERTIES: A COMPREHENSIVE TREATISE. R. H. Burris and W. H. Orme-Johnson in *Microbial Iron Metabolism,* edited by J. B. Neilands. Academic Press, 1974.

NITROGEN FIXATION IN BACTERIA AND HIGHER PLANTS. R. C. Burns and R. W. F. Hardy. Springer-Verlag, 1975.

REGULATION AND GENETICS OF BACTERIAL NITROGEN FIXATION. Winston J. Brill in *Annual Review of Microbiology,* Vol. 29, pages 109–129; 1975.

THE ACOUSTICS OF THE SINGING VOICE

SINGING: THE MECHANISM AND THE TECHNIC. William Vannard. Carl Fischer, Inc., 1967.

TOWARDS AN INTEGRATED PHYSIOLOGIC-ACOUSTIC THEORY OF VOCAL REGISTERS. John Large in *NATS Bulletin,* Vol. 28, No. 3, pages 18–25, 30–36; February/March, 1972.

ARTICULATORY INTERPRETATION OF THE "SINGING FORMANT." Johan Sundberg in *The Journal of the Acoustical Society of America,* Vol. 55, No. 4, pages 838–844; April, 1974.

FUNDAMENTALS OF MUSICAL ACOUSTICS. Arthur H. Benade. Oxford University Press, 1976.

THE OLDEST ROCKS AND THE GROWTH OF CONTINENTS

CONTINENTS ADRIFT: READINGS FROM SCIENTIFIC AMERICAN. With introductions by J. Tuzo Wilson. W. H. Freeman and Company, 1972.

THE EARTH'S AGE AND GEOCHRONOLOGY. Derek York and Ronald M. Farquhar. Pergamon Press, 1972.

A DISCUSSION ON THE EVOLUTION OF THE PRECAMBRIAN CRUST. Organized by J. Sutton and B. F. Windley in *Philosophical Transactions of the Royal Society of London: Series A,* Vol. 273, No. 1235, pages 315–581; February 8, 1973.

PLANET EARTH: READINGS FROM SCIENTIFIC AMERICAN. With introductions by Frank Press and Raymond Siever. W. H. Freeman and Company, 1974.

THE EARLY HISTORY OF THE EARTH. Edited by Brian F. Windley. John Wiley & Sons, 1976.

FLASHLIGHT FISHES

BIOLUMINESCENCE. E. Newton Harvey. Academic Press Inc., 1952.

LES PETITS PEUGEOTS OF GRANDE COMORE. John E. McCosker and Michael D. Lagios in *Pacific Discovery,* Vol. 28, No. 5, pages 1–6; September–October, 1975.

LIGHT FOR ALL REASONS: VERSATILITY IN THE BEHAVIORAL REPERTOIRE OF THE FLASHLIGHT FISH. James G. Morin, Anne Harrington, Kenneth Nealson, Niel Krieger, Thomas O. Baldwin and J. W. Hastings in *Science,* Vol. 190, No. 4209, pages 74–76; October 3, 1975.

KRYPTOPHANERON HARVEYI, A NEW ANOMALOPID FISH FROM THE EASTERN TROPICAL PACIFIC, AND THE EVOLUTION OF THE ANOMALOPIDAE. Richard H. Rosenblatt and W. Linn Montgomery in *Copeia,* No. 3, pages 510–515; August 20, 1976.

THE EARLIEST MAYA

THE RISE AND FALL OF MAYA CIVILIZATION. John Eric Sidney Thompson. University of Oklahoma Press, 1966.

THE LOST CIVILIZATION: THE STORY OF THE CLASSIC MAYA. T. Patrick Culbert. Harper & Row, 1974.

PRECLASSIC TO POSTCLASSIC IN NORTHERN BELIZE. Norman Hammond in *Antiquity,* Vol. 48, No. 191, pages 177–189; September, 1974.

RADIOCARBON CHRONOLOGY FOR EARLY MAYA OCCUPATION AT CUELLO, BELIZE. Norman Hammond, Duncan Pring, Rainer Berger, V. R. Switsur and A. P. Ward in *Nature,* Vol. 260, No. 5552, pages 579–581; April 15, 1976.

THE ORIGINS OF MAYA CIVILIZATION. Edited by Richard E. W. Adams. University of New Mexico Press, in press.

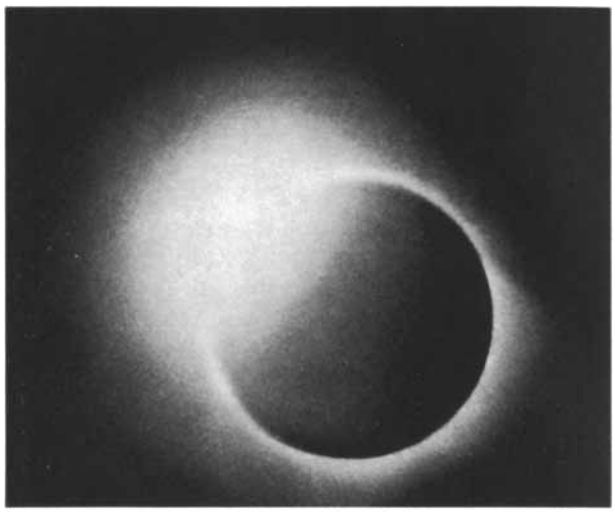
MATHEMATICAL GAMES

THE GOLDEN SECTION, PHYLLOTAXIS, AND WYTHOFF'S GAME. H. S. M. Coxeter in *Scripta Mathematica,* Vol. 19, Nos. 2–3, pages 135–143; June–September, 1953.

A FIBONACCI PROPERTY OF WYTHOFF PAIRS. Robert Silber in *The Fibonacci Quarterly,* Vol. 14, No. 4, pages 380–384; November, 1976.

WYTHOFF'S NIM AND FIBONACCI REPRESENTATIONS. Robert Silber in *The Fibonacci Quarterly,* Vol. 15, No. 1, pages 85–88; February, 1977.

SITMAR'S SOLAR ECLIPSE CRUISES. THE ECLIPSE IS ONLY PART OF THE SCIENTIFIC ADVENTURE.



On October 12, 1977, one of nature's rare spectacles occurs—a total eclipse of the sun.

On that date, Sitmar's two Liberian-registered luxury vessels rendezvous 1200 miles southwest of Los Angeles to provide an unequaled vantage point for viewing the eclipse.

But Sitmar offers far more than just a spectacular view.

On board both ships will be Sitmar's "Science at Sea" program. It features distinguished experts lecturing in fields such as Astronomy, Archaeology, Navigation, Oceanography, Meteorology and Photography. Passengers may even earn professional academic credits in Astronomy through UCLA Extension.

All this in addition to the usual Sitmar amenities. Visits to popular foreign ports. Lavish continental cuisine. European service from a 500-man Italian crew. And 25,400 square feet of open deck space.

On October 1, the T.S.S. Fairwind departs from Fort Lauderdale, Florida, transits the Panama Canal, sails to the eclipse viewing site, then terminates in Los Angeles (17 days in all). En route you visit exciting South American and Mexican ports. The T.S.S. Fairsea departs Los Angeles on October 8 on an 11-day cruise, which also visits Mexico.

FLY FREE FROM 127 CITIES

You'll even fly free roundtrip from any of 127 cities throughout the U.S., Canada and Mexico aboard regularly scheduled airlines for the Fairwind cruise with Sitmar's incomparable Air/Sea program. Or you can enjoy major air/sea savings for the Fairsea cruise.

And you'll enjoy stopover privileges with Sitmar's "Cruise Plus" feature.

For more information see your travel agent.

"SCIENCE AT SEA" LECTURERS

FAIRWIND CRUISE

Dr. Edward M. Brooks, Professor, Geology and Geophysics, Boston College;
Dr. Mark Chartrand III, Chairman, American Museum, Hayden Planetarium, New York;
Dr. Frank D. Drake, Director, National Astronomy & Ionosphere Center; Professor of Astronomy, Cornell University;
Dr. Carl Sagan, Professor of Astronomy; Director, Laboratory for Planetary Studies, Cornell University;
Mr. Walter S. Sullivan, Science Editor, New York Times

FAIRSEA CRUISE

Dr. George Abell, Chairman, Dept. of Astronomy, UCLA;
Dr. Joseph Chamberlain, Program Coordinator; Director, Adler Planetarium in Chicago;
Dr. J. Allen Hynek, Professor, Dept. of Astronomy, Northwestern University
Dr. James C. Sadler, Assoc. Professor, Dept. of Meteorology, University of Hawaii.

Plus Other Distinguished Guest Lecturers.

Sitmar Cruises

The Caribbean and South America. Mexico. Canada and Alaska

For a copy of our Solar Eclipse brochure, please write to Sitmar Cruises, 1801 South Hill Street, Los Angeles, CA 90015. Please include the name and address of your travel agent.

Merci, Monsieur You sure got

Look where the science of transportation is now.

Today, thanks to the inventiveness of people like Monsieur Cugnot, transportation is a science. A science that moves people and products from where they are to where they want to be. Efficiently. Economically.

Rockwell International is very much a part of that science with products that range from undersea navigation systems to the world's first reusable spacecraft . . . and many other forms of transportation in between.

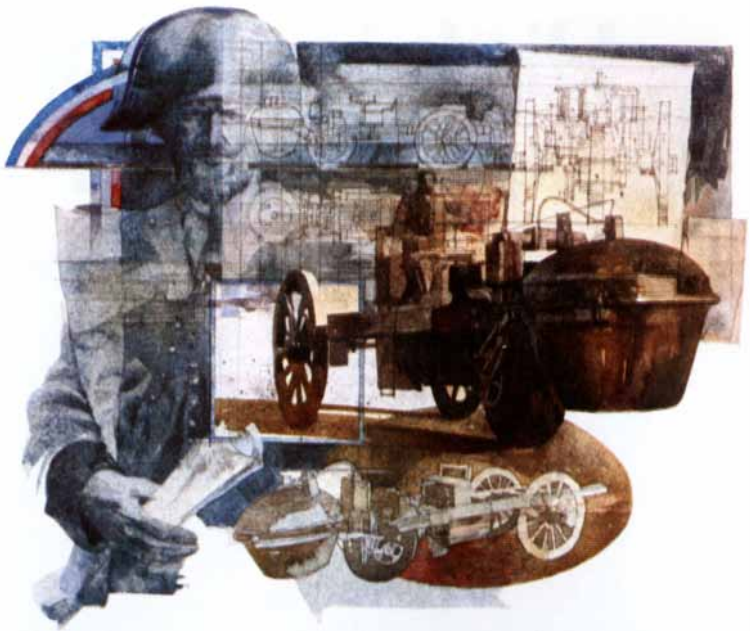
The science of transportation is just one of the sciences of Rockwell, one of the ways we're putting science to work for man.

Further information, including our Annual Report, is available by writing: Rockwell International, Department 815S, 600 Grant Street, Pittsburgh, PA 15219.

Over half of all on-highway heavy-duty trucks and tractors ride on Rockwell axles and stop with Rockwell brakes. Other Rockwell components for America's vital highway transportation system include: suspension springs, U-joints and reinforced plastic assemblies.

Rockwell builds driveline components and heavy-duty brakes for off-highway equipment used in mining, forestry, construction, agriculture and many other fields.

Nicolas Joseph Cugnot, of France, built the world's first self-propelled road vehicle. When? Over 200 years ago. It travelled at about three miles per hour with stops at 15-minute intervals to build up steam. Monsieur Cugnot is credited with another first: While motoring through the streets of Paris, he took a sharp turn too speedily and became the world's first automobile accident statistic. His vehicle was impounded, never to be driven on public streets again. It has been preserved in the Conservatoire des Arts et Metiers in Paris.



Cugnot. things moving.

All 300 cars of the Washington Metropolitan Area Transit Authority will ride on undercarriages from the broad Rockwell line of components for the mass transit industry.

The Space Shuttle, which we're building for NASA, will become the space transportation system for the 1980s and beyond. Its payloads will help put space to work for man.

High-speed hydrofoil boats, such as this, have accumulated over 50-million passenger miles, powered by Powerjet® propulsors developed from Rockwell's rocket engine pump technology.

Our Rockwell Turbo Commander® 690B can take busy executives into untapped markets — fast. It outperforms every other existing business propjet in its class.

Rockwell International

...where science gets down to business

High technology is our basic strength. We make it work.

Mobile Power

Power. Combining a gas turbine and all necessary controls, this truck-mounted power unit from our Power Systems Division can be driven almost anywhere—and can produce enough electricity for a city of 25,000 people.

Pilot pleaser

Systems. Built to the rugged requirements of the army of the future, our Sikorsky Aircraft Division has developed the U.S. Army's troop carrying helicopter of the 1980's and 1990's.

People mover

Industrial. At the Charles de Gaulle International Airport in Paris, our Otis Elevator subsidiary helps take the walking out of flying. Moving walkways carry up to 10,000 passengers an hour from curbside to seven satellite buildings.

Results

Crowning five years of achievement, our 1976 sales and earnings were at record highs. Since 1972, sales have increased by 155% and earnings by 211%.

UNITED TECHNOLOGIES®

Pratt & Whitney Aircraft Group • Otis Group • Essex Group • Sikorsky Aircraft • Hamilton Standard Power Systems Division • Norden • Chemical Systems Division • United Technologies Research Center

United Technologies Corporation, Hartford, Conn. 06101.

Tune in Sundays to MEET THE PRESS on NBC-TV—sponsored by United Technologies Corp.