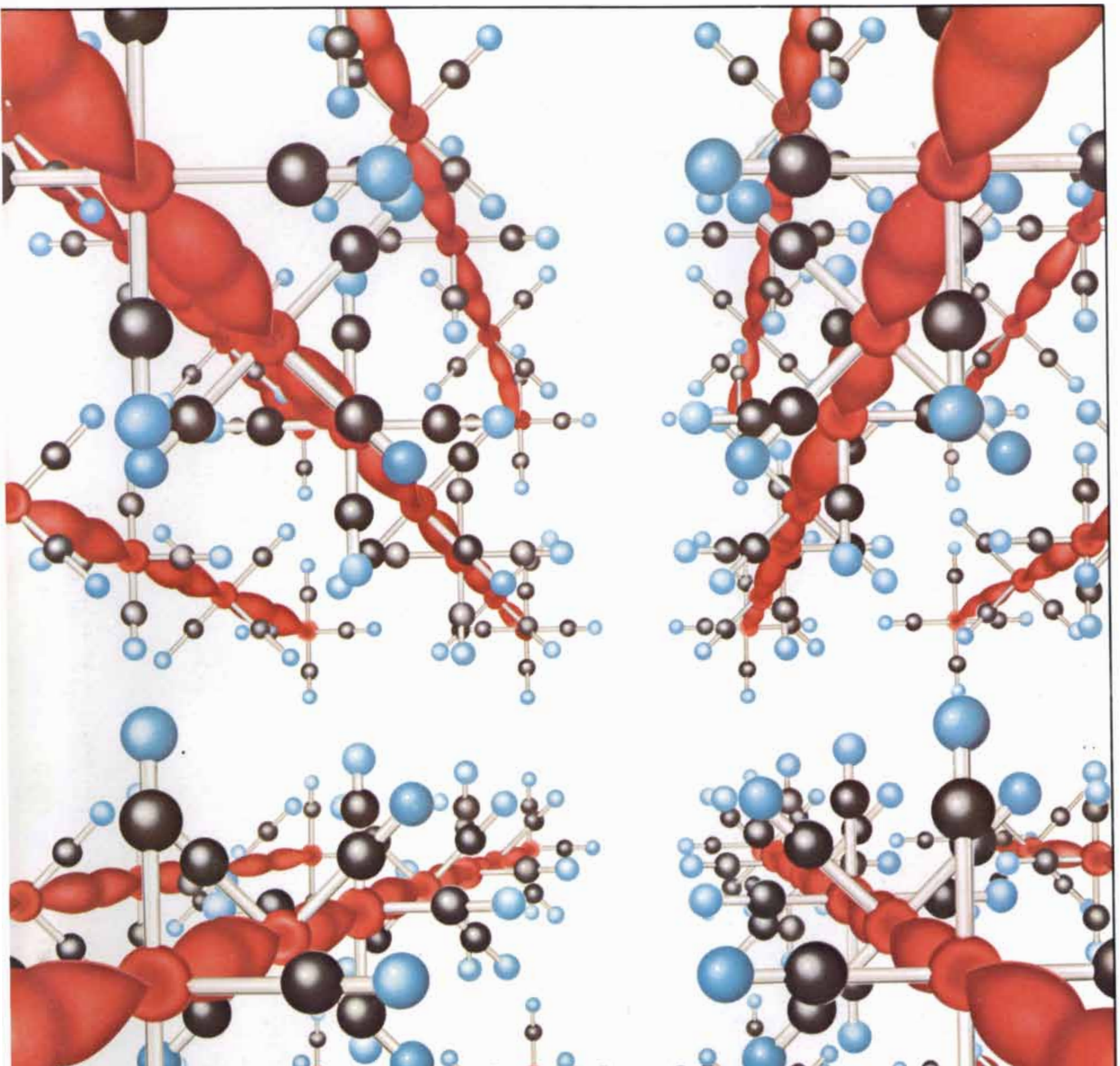
SCIENTIFIC AMERICAN



LINEAR-CHAIN CONDUCTORS

$2.00

October 1979


GM
Eldorado 1980
BY CADILLAC...


12
CHIVAS REGAL
FOUNDED 1801
BLENDED SCOTCH WHISKY
PRODUCE OF SCOTLAND
750 ML
(25.4 FL. OZ.)
100% SCOTCH WHISKIES
BLENDED & BOTTLED IN SCOTLAND BY
CHIVAS BROTHERS LTD.
Of course you can afford it. Think of all the money you've saved drinking ordinary Scotch.
12 YEARS OLD WORLDWIDE • BLENDED SCOTCH WHISKY • 86 PROOF • GENERAL WINE & SPIRITS CO., NEW YORK, N.Y.

Established 1845 SCIENTIFIC AMERICAN October 1979 Volume 241 Number 4

ARTICLES

DEPARTMENTS

SCIENTIFIC AMERICAN (ISSN 0036-8733), PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. SECOND-CLASS POSTAGE PAID AT NEW YORK, N.Y., AND AT ADDITIONAL MAILING OFFICES. AUTHORIZED AS SECOND-CLASS MAIL BY THE POST OFFICE DEPARTMENT, OTTAWA, CANADA, AND FOR PAYMENT OF POSTAGE IN CASH. SUBSCRIPTION RATE: $18 PER YEAR, U.S., ITS POSSESSIONS AND CANADA; $22 PER YEAR, ALL OTHER COUNTRIES. POSTMASTER: SEND ADDRESS CHANGES TO SCIENTIFIC AMERICAN, 415 MADISON AVENUE, NEW YORK, N.Y. 10017.

executive health

the report that briefs you on what to watch

Volume XV, Number 11 • August, 1979 • Pickfair Bldg., Rancho Santa Fe, Calif. 92067 • Area 714:756-2600

Leonard Hayflick, Ph.D.:

A NEW THEORY ON THE CAUSE OF HEART DISEASE AND STROKES . . .

Vitamin B_6 may be the key.

What new research has revealed that may downgrade cholesterol as the main predisposing factor . . . and why.

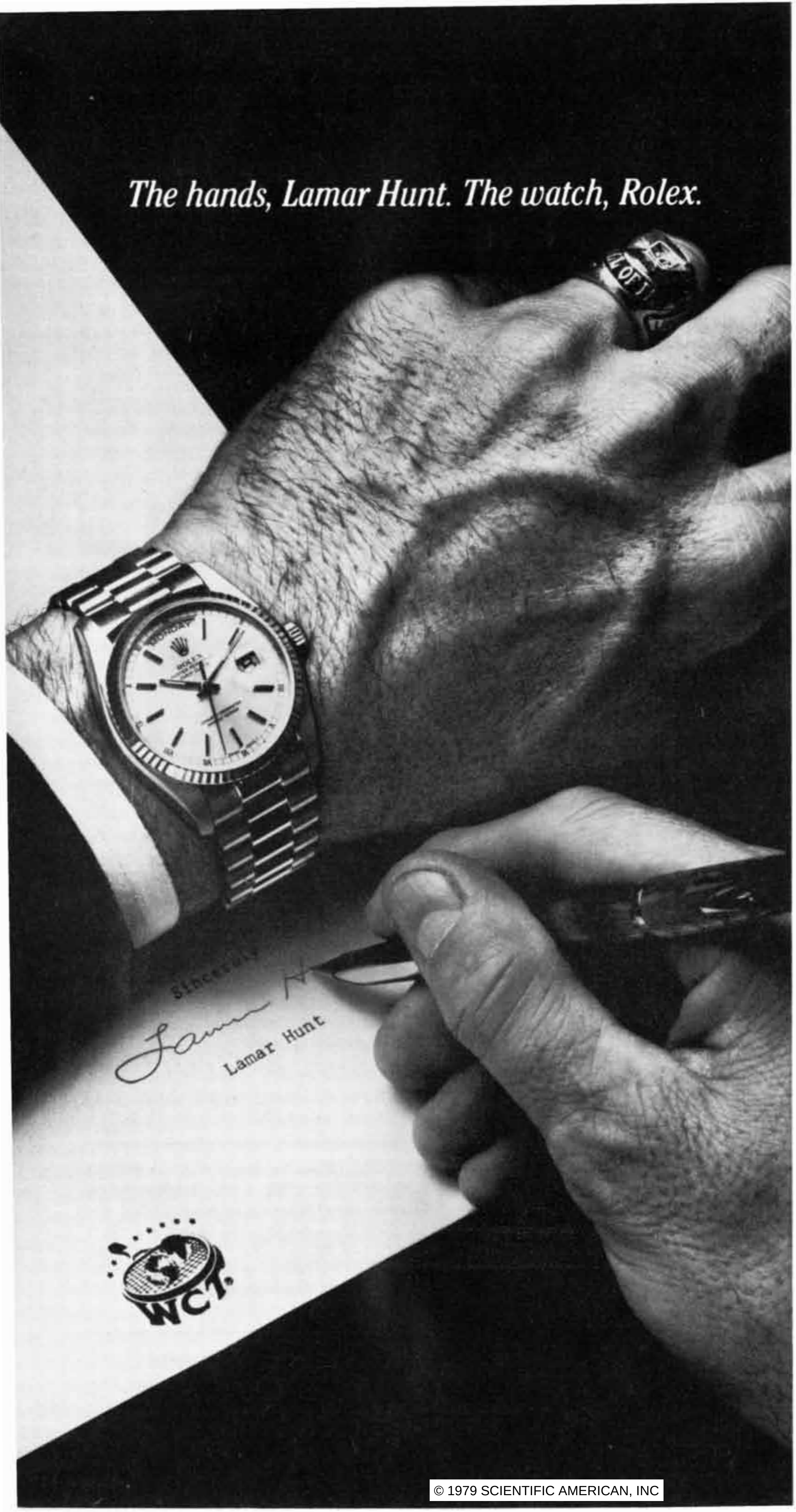
PUBLISHER'S NOTE: *Dr. Leonard Hayflick, a member of our Editorial Board, is Senior Research Cell Biologist at Children's Hospital Medical Center, Oakland, California. He is a leading authority on the biology of aging and has won several major awards for his research in this field. In 1961 he discovered a new micro-organism that causes a common form of human pneumonia. He is the author of 150 scientific research papers and books and is presently the editor of several scientific journals.*

—Richard Stanton

Nearly one million Americans will die this year from vascular diseases, that is, those diseases of the arteries that cause heart attacks and strokes. What is more, about 28 million people are already afflicted by vascular diseases. For comparison, this year, a third of a million people will die of cancer, and one hundred thousand from accidents.

It may not be too surprising to learn that vascular disease is the major killer. After all, the heart beats about 100,000 times a day while pumping 4300 gallons of blood through vessels that, if laid end to end, would circle the earth twice! There is certainly sufficient opportunity for something to go wrong.

What is going wrong is that the old causes of death in our youth, mainly infectious diseases, have been largely conquered and we are now living long enough to encounter vascular diseases. We have been advised to protect ourselves by avoiding smoking, sedentary habits, hypertension and high cholesterol diets. But none of these no-no's alone, or together, guarantee a healthy vascular system. Something may be wrong with this "conventional wisdom" and an important new theory



ROLEX

The Rolex crown.
For the innovators of our time.

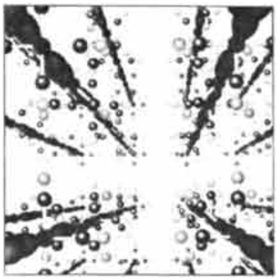
Lamar Hunt.
Founder of the American Football League.
Co-founder of World Championship Tennis.
Member of the Pro Football Hall Of Fame.
A dynamic sports entrepreneur.

Rolex.
Dynamic innovator in watchmaking.

Mr. Hunt's Rolex is the Rolex Day-Date Oyster Perpetual Superlative Chronometer, self-winding, handcrafted from a solid block of 18kt. yellow gold, pressure proof to 165 feet, with matching hidden clasp President bracelet. Also available in 18kt. white gold and platinum.

For brochure, write Rolex Watch U.S.A., Inc., Dept. 116, Rolex Building, 665 Fifth Avenue, New York, New York 10022.

Geneva · Bombay · Brussels · Buenos Aires · Caracas · Cologne · Dublin · Hong Kong · Johannesburg · London · Madrid · Mexico City · Milan · New York · Paris · Sao Paulo · Singapore · Sydney · Tokyo · Toronto



THE COVER

The illustration on the cover is an interior view of a crystalline material that conducts electricity efficiently only along one axis (see "Linear-Chain Conductors," page 52). The material is a salt of tetracyanoplatinate. The long, columnar structures that recede from view are stacks of tetracyanoplatinate molecules. Carbon atoms (*gray*) and nitrogen atoms (*blue*) are shown schematically as balls, but the platinum atom (*red*) at the center of each molecule is represented by an orbital, which defines the distribution of electrons in the atom. Because the orbitals extend well above and below the plane of the molecule, orbitals of adjacent molecules in the same column overlap. The overlapping gives the crystal a comparatively high conductivity along the axis parallel to the columns; the conductivity measured perpendicular to this axis is smaller by a factor of 100,000. The crystal also incorporates potassium and bromine atoms and water molecules, but here they have been omitted for clarity.

THE ILLUSTRATIONS

Cover illustration by George V. Kelvin

Page	Source
19–26	Ilil Arbel
36	*The Rise of Surgery: From Empiric Craft to Scientific Discipline.* © 1978, University of Minnesota Press
44–50	Andrew Tomko
51	Photo Researchers, Inc.
53	George V. Kelvin
54	Xerox Corporation
55–61	George V. Kelvin
63	Daniel Rugar, Stanford University
64–67	Dan Todd
68	Abdullah Atalar, Stanford University
69–70	Calvin F. Quate, Stanford University
83	Joseph E. McDade, Center for Disease Control, U.S. Public Health Service (*top left*); William B. Cherry, Center for Disease Control (*top right*); Francis W. Chandler, Center for Disease Control (*bottom*)
85	Francis W. Chandler
86–98	Bunji Tagawa
103	Kenneth R. Miller, Harvard University
104–105	Jerome Kuhl
106	Nam-Hai Chua, Rockefeller University
107	Kenneth R. Miller, Harvard University (*top*); Jerome Kuhl (*bottom*)
108	Jerome Kuhl
109–112	Kenneth R. Miller
113	Kenneth R. Miller (*top and bottom left*); L. Andrew Staehelin, University of Colorado (*bottom right*)
115	Dallas L. Peck and Thomas L. Wright, U.S. Geological Survey; Robert W. Decker, Hawaiian Volcano Observatory
116–117	Dallas L. Peck and Thomas L. Wright, U.S. Geological Survey; Robert W. Decker, Hawaiian Volcano Observatory (*left*); Walken Graphics (*right*)
119–120	Walken Graphics
124	Dallas L. Peck and Thomas L. Wright, U.S. Geological Survey; Robert W. Decker, Hawaiian Volcano Observatory
126–128	Walken Graphics
131	Eric Buffetaut, French National Center of Scientific Research
132–144	Patricia J. Wynne
147	Scala / Editorial Photocolor Archives
151–166	Ilil Arbel
173	James Young
174–184	Michael Goodman

There is only one real pioneer It's Sony.

1957: The world's first pocket transistor radio.

In 1954, a fledgling Japanese tape recorder manufacturer visited America to investigate a new device called the transistor.

At first, things were less than encouraging. "Transistors are only good for hearing aids," they were told. "And besides, they can't be mass produced."

Undeterred, the Japanese representatives returned to Tokyo.

1954: The first

Thirty-six months later, the world saw its first pocket transistor radio.

Followed by the world's first all-transistor FM radio.

And, partiall sign of their continuing dedication to audio, the Tokyo Telecommunications Engineering Corporation adapted the Latin word for sound—"sonus"—and changed its name to Sony.

In the years that have followed, Sony has never faltered in its dedication to technological innovation. And we'd be loathe to estimate how often our advances have ended up on the circuit boards and front panels of our competitors' equipment as "technological breakthroughs."

But enough of the past.

The hi-fi components featured here stand as eloquent proof that Sony—the company that virtually founded the era of transistorized high fidelity—is still at its very forefront.

A few Sony Audio firsts:

1949: Obtained patent on the basic magnetic tape-recording system.
1952: Developed stereo broadcasting in Japan.
1954: Introduced condenser microphone.
1955: First consumer stereo tape recorder in Japan.
1959: Invented "Tunnel Diode"; basis of all high-speed, low-distortion semiconductors.
1965: First all-silicon solid state amplifier.
1966: The first servo-controlled turntable. Forerunner of quartz-locked turntables.
1968: First electronic end of record sensor.
1969: First digital-synthesized FM tuner.
1969: Invented the ferrite tape head.
1973: Invented the V-FET: Opened era of high-speed transistors.
1973: First to manufacture ferrichrome tape.
1973: Dr. Esaki wins Nobel Prize in Physics for "Tunnel Diode."
1975: First turntable with carbon-fiber tone arm.
1977: The world's first consumer digital audio processor.
1977: First consumer amplifier with pulse power supply.
1978: Patented liquid crystal recording meters.

The V5 receiver: To this day, only Sony offers Sony quality.

Unlike hi-fi receivers designed to impress you with a facade of magic buttons and switches, Sony receivers are designed to impress you with rich sound.

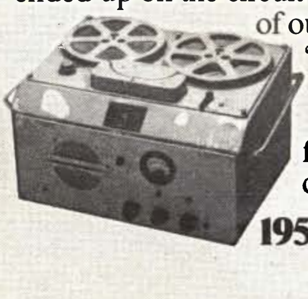
Case in point: the V5.

In technical terms, the V5 delivers 85 watts per channel at 8 ohms from 20 to 20,000 hertz with no more than 0.07% total harmonic distortı

In human terms, this means the receiver can reproduce every note of music any instrument can play with no audible distortion. And of speakers without straining.

But that's only the beginning.

Instead of using the mundane power transformers found in competitors' products, the V5 utilizes more expensive toroidal core transformers that provide richer bass.



1950: Japan's first tape recorder, the "Type G."

1979: The V5 receiver: Designed for people who appreciate value as much as they appreciate sound.

in high fidelity.

1979: The new TA-F40 integrated amplifier and ST-J60 digital synthesized FM tuner. Separate components that sound as sophisticated as they look.

Instead of cutting corners by using a flimsy pressboard bottom, we've cut interference by encasing the *entire* receiver in metal.

And for better FM reception, instead of using the standard three- or four-gang variable-tuning capacitor, we've opted for a higher quality five-gang model.

All of which explains why if you pay a few dollars less for one of our competitors' receivers, it's probably because you're getting less receiver.

The new Sony cassette decks: The state of the art, from the people who invented it.

1979: The new TC-K65. Sony remains one of the only hi-fi companies to produce our own tape transports, motors, meters, heads—even the tape itself.

Since we introduced tape recording to Japan in 1950, Sony has sold millions of tape decks.

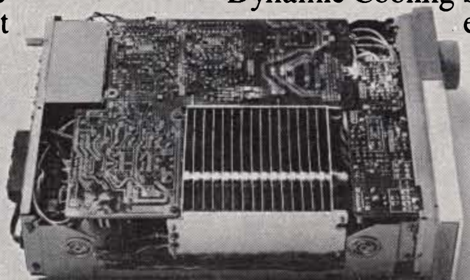
A quick look at our new TC-K65 cassette deck will explain why.

Like all two-motor cassette decks, the TC-K65 is designed for low wow and flutter. Unlike others, however, we feature "brushless and slotless" motors that reduce this problem to the point of being inaudible.

Instead of using just any tape head material, the TC-K65 features Sony "Sendust and Ferrite" heads that combine wide response with extreme durability.

Instead of using an ordinary metering system, we've developed a 16-segment LED meter whose life expectancy far exceeds the fancy blue fluorescent models other companies are currently touting.

And there's also a "Random Music Sensor" for preprogramming tapes, settings for metal tape, remote control and timer capabilities, and the kind of high-quality D.C. tape head amplifier you'll find in almost no one else's tape decks.



1979: The Sony "ThermoDynamic Cooling System." Until now, only available in satellites.

But you really haven't heard anything yet.

Unfortunately, we don't have enough space here to tell you the complete Sony hi-fi story.

Like the way a recent dealer survey rated our turntables #1 in value and performance.

Or the way our new separate tuners and amplifiers (not to mention micro components) utilize highly advanced light-weight pulse power supplies whose levels of distortion are virtually unmeasurable.

Or how they use a NASA developed "Thermo-Dynamic Cooling System" that eliminates heat, excess wire and the distortion and interference that normally accompany them.

If you'd like to hear more about the complete line of Sony hi-fi components (or if you need the name of your nearest dealer) write to Sony, P.O.Box CN 04050, Trenton, New Jersey 08650.

In the meantime, if somebody makes noise about innovations in high fidelity, think of the biggest pioneer in audio. And remember Sony.

SONY AUDIO

We've never put our name on anything that wasn't the best.

LETTERS

Sirs:

One of the procedures for measuring the radius of the earth described in your department "The Amateur Scientist" for May was discussed by none other than Albert Einstein. In a letter to a friend dated February 4, 1912 (*Albert Einstein et Michele Besso, Correspondance,* Hermann, Paris, 1972), he considered how an approximate value, higher than the actual one resulting from atmospheric refraction, can be obtained with the formula $R = S^2/2h$, where R is the radius of the earth, S is the radius of the visible horizon and h is the observer's elevation. Einstein also mentioned the relation $R = 2h/\alpha^2$, where angle α is the dip of the horizon. This formula is equivalent to the one used by Joseph L. Gerver in "The Amateur Scientist."

According to the Michelin guide to Paris, on a clear day it is possible to see to a distance of 67 kilometers from the third level of the Eiffel Tower, which is at a height of 274 meters. The application of Einstein's first formula gives a radius for the earth of 8,200 kilometers, which is about 30 percent above the actual value.

C. D. GALLES

University of Paris
Paris

Sirs:

We wish to point out that "The Year without a Summer," by Henry Stommel and Elizabeth Stommel [SCIENTIFIC AMERICAN, June] suggests a correlation between volcanic eruptions and climate based on very limited data. This topic has been dealt with in the journal *Weatherwise* by Helmut E. Landsberg and J. M. Albert, who suggested that the summer of 1816 was not outside the realm of a long-return-period cold spell and was by no means global. In fact, the "year without a summer" was unusually warm in many places.

We do not deny that volcanic dust affects climate, but the mechanisms involved are far more complex than the Stommel article suggests. Let us point out several inadequacies in the argument that the cold summer of 1816 was due entirely to the eruption of the volcano Tambora in 1815. Little is known about the eruption other than contemporary reports. The "climatologists [who] rank the eruption as the greatest producer of atmospheric dust between 1600 and the present" are doing so by the somewhat circular reasoning that since 1816 was cold, its unusual weather must have been caused by a large amount of dust, hence the eruption was huge. Another "large" eruption that, on the basis of climate data, was said to have produced a large dust veil and widespread cooling in the Northern Hemisphere was the 1835 eruption of the volcano Cosigüina in Nicaragua. As part of an ongoing climate-research program at the Goddard Institute for Space Studies of the National Aeronautics and Space Administration one of us has begun a volcanological investigation of this volcano. It now appears that the records are exaggerated; the eruption was large and clearly had a great impact on a local scale, but it was not of the type that would be capable of injecting large volumes of gas and dust into the *stratosphere,* which is essential if a widespread dust veil is to be formed. It would appear now that the post-1815 cooling had other causes.

To return briefly to Tambora, no modern volcanological study has been made of it. We have just undertaken one. Two of the only pieces of data published in this century are of interest: one is that of Petrochevsky, who visited the volcano in 1947 and commented on the lack of young ash deposits resulting from the great 1815 eruption. The other is that the 1815 "pumice" is, again according to Petrochevsky, a peculiar type of basalt. This is not at all the kind of magma usually associated with large explosive eruptions. Obviously there is a need for more data.

The temperature curves presented by the authors and others (for example the curve for New Haven from 1790 to 1860) show one conspicuous feature: the temperatures were falling consistently from 1811–12 to reach a minimum in 1816, and they were rising after 1816. Surely this suggests a control other than a volcanic injection of aerosol into the stratosphere in 1815. Such a curve can also be "made" to dip before an actual low is reached by the "running average" technique of compilation. The authors do not state whether this record is a running average or actual yearly temperatures.

This decrease in temperature before 1815 is also shown in the data derived from Clifford Mass and Stephen H. Schneider. What the authors neglect to point out was that Mass and Schneider also state that the 1810–20 interval was a period with the lowest number of observed sunspots in the past 250 years. Solar activity might also be an important control on climate. The generally colder period worldwide in the 1810–20 interval may well be correlated with reduced solar activity.

STEPHEN SELF

Department of Earth Sciences
Dartmouth College
Hanover, N.H.

MICHAEL R. RAMPINO

Goddard Institute
for Space Studies
New York

Sirs:

The New Haven temperatures are means for individual months of June according to A. Loomis' "The Meteorology of New Haven" published in Volume 1 (1866) of *Transactions of the Connecticut Academy of Sciences.* A recent graph showing the exceptional dust-veil index attributed to Mount Tambora's 1815 eruption is in Alan Robock's "Internally and Externally Caused Climate Change" (*Journal of the Atmospheric Sciences,* Vol. 35, No. 6, pages 1111–1122; June, 1978).

Climate is the sum of many causes. In a chapter of a forthcoming book on the cold summer we attempt to discuss possible causes other than volcanic dust, including sunspots. We also take up the important paper by Landsberg and Albert. In the limited space of an article aimed chiefly at describing the social impact of the cold summer perhaps there was not enough discussion of differing opinions. We agree that the case for Mount Tambora is not conclusive.

HENRY STOMMEL

Woods Hole Oceanographic
Institution
Woods Hole, Mass.


JOHN
JAMESON
EST'D
1780
A BLEND
LIMITED
DUBLIN

Attention to detail. It's part of our character.

A calligraphic brush stroke that is just so. To us, it is the difference between penmanship and art. This attitude of attentiveness aboard

Japan Air Lines is the difference between transportation and transport.
It may also be manifest in a meal tray arranged so artfully that dinner looks almost too good to eat; the exquisite detailing of a kimono; or warm sake, served with classic grace, in a delicate little cup.
Seichi. Attention to detail.
Because we never forget how important you are.

The way we are is the way we fly.

JAL

JAPAN AIR LINES

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

OCTOBER, 1929: "Rightly or wrongly, many Americans still consider that they voluntarily relinquished naval supremacy in 1921, and that meanwhile Great Britain has built up her strength in cruisers until she has recovered actual supremacy. No lasting concord will exist unless it is founded on parity. Neither country will brook second place. Developments on the naval question have now been coming so fast that it is difficult for a monthly magazine to keep abreast of them. Shortly before this issue went to press Prime Minister MacDonald, as a gesture of goodwill toward the United States, halted construction on two British cruisers and slowed down other naval work. President Hoover answered this *beau geste* by suspending construction on three American cruisers on this year's program. It is expected that these acts will have a direct influence on the final agreement for parity."

"Three years ago the United States lagged behind the world in nearly every department of flying. Two years ago a young American airmail pilot linked Roosevelt Field with Le Bourget in his lonely and unexampled flight across the Atlantic and not only leaped into world fame overnight but also lighted a blaze of enthusiasm for aviation that burns more brightly day by day. Today the United States dominates the sky. Approximately 4,500 airplanes were manufactured last year in the United States. Estimates for 1929 put the figure at 9,000. Airplane exports have tripled in a year. Air-transport operators flew some 8,000,000 miles in the first six months of 1929 as against 10,510,000 for all of 1928. The total length of the airway network is now estimated at 30,000 miles as compared with 16,667 at the close of the year. Forty thousand passengers were carried in the first six months of this year and but 35,000 in the 12 months preceding."

"A striking similarity between nervous action and the behavior of iron wire in a bath of nitric acid has been discovered and studied by Ralph S. Lillie, professor of physiology at the University of Chicago. In a demonstration of the experiment a wire of pure iron, one to five centimeters long, is immersed in a bath of nitric acid of 60 to 80 per cent. A colorless film immediately forms over the wire. If the wire is scratched at one end, a wave travels rapidly along it. What happens, according to Dr. Lillie, is that a sort of battery is formed. The film, which is only one molecule thick, has an electric charge negative with respect to that of the wire. When the film is scratched, a current is set up that dissolves the film next to the bare spot. This continues until the bare spot, which is seen as a wave, has passed the length of the wire. The film forms again after the wave has passed."

"A career of remarkable achievement ended on August 3 with the death in Washington of Emile Berliner, inventor of the first disk-record talking machine and of the telephone transmitter from which has evolved the radio microphone of today. An immigrant from Germany in 1870, Berliner sold glue, painted backgrounds on photographic enlargements, was a traveling salesman and in 1877 became a store clerk in Washington. His inventive genius soon manifested itself when, after experimenting in his spare hours, he invented the loose-contact transmitter that placed the telephone on a commercial basis three years after its invention by Bell. Then came the gramophone and Berliner's invention of the present method of duplicating disk records. He worked on a helicopter before the Wright brothers' success with the airplane, traced the cause of high infant mortality to raw milk and invented the first light-weight internal-combustion engine for aircraft."

SCIENTIFIC AMERICAN

OCTOBER, 1879: "When Mr. Edison said he was about to produce a practical economical electric light for general use, the entire business world took it for granted that it was forthcoming. Gas stocks tumbled both in this country and in Europe, and people waited for the coming light. Scientists having had experience in this direction shook their heads but for the most part suspended judgment. Delay on the part of Mr. Edison comforted the holders of gas stocks, and confirmed the scientists in their belief that he had undertaken not only a task of great magnitude and difficulty but also one that would require more time and means than are available to most experimenters. But Mr. Edison, flushed by his earlier victories, is undaunted and determined to yield to no obstacle. Unfortunately, however, the daily papers have from time to time printed articles about progress in electric lighting, which, because of their extravagance and inaccuracy, have placed Mr. Edison in an extremely embarrassing position. At present little is said by Mr. Edison concerning the electric light, except that he considers it an assured success and that he is perfecting the details of his electric lighting system as rapidly as possible."

"One cannot take a ferry in New York without being amazed at the movement of breadstuffs visible on all sides. On the Hudson River Railroad and all the other iron thoroughfares converging on the city long trains of grain cars are almost constantly in sight, while on the river vast rafts of grain-laden canal boats more than rival the railway trains in carrying capacity. During the week ending September 6 the exports of breadstuffs from New York included 113,224 barrels of flour, 2,519,409 bushels of wheat, 914,623 bushels of corn, 2,996 bushels of oats and 103,701 bushels of rye. The storage capacity of the port is about 12,000,000 bushels, but the present active demand for grain for foreign shipment, due to the general deficiency of European crops, prevents any large accumulation here. Indeed, the bulk of shipping devoted to the transportation of grain from New York to foreign ports is at this season something unprecedented in the history of the world."

"Our English contemporary *Journal of the Society of Arts* lists the estimated value of the production of coal-tar colors in 1878: Germany, £2,000,000; England, £450,000; France £350,000; Switzerland, £350,000. There are now in England six coal-tar color works, in Germany no fewer than 17, in France five and in Switzerland four. There are also three works in Germany and three in France that manufacture aniline in enormous quantities for the production of coal-tar colors. The total production of artificial madder, or alizarine, alone is estimated at 9,500 tons. Calculating its selling price at £150 per ton, the annual value amounts to no less than £1,425,000. As a dye alizarine is now not more than one-third of the average price of madder in 1859–68. Consequently in the United Kingdom, where the annual value of madder imported was £1,000,000, the annual saving is very great. Such is the wonderful growth of this industry, which dates only from 1856."

"A contract has been made by Alvan Clark & Sons of Cambridgeport, Mass., with the Russian government, relative to the great objective for the Imperial Observatory at Pulkowa. The proposed glass is to be the largest in the world. The contract provides that the definition of the glass shall not be inferior to that of the telescope in the Naval Observatory in Washington. The objective in Washington is 26 inches in diameter; the proposed glass is to be from $31\frac{1}{2}$ to 32 inches in diameter, with a clear aperture of 30 inches. When finished, the glass will be mounted in Hamburg. The cost of the glass alone will be $32,000."

THE AUTHORS

GENE E. LIKENS, RICHARD F. WRIGHT, JAMES N. GALLOWAY and THOMAS J. BUTLER ("Acid Rain") share an interest in the influence of human activities on biogeochemical processes. Likens is professor of ecology at Cornell University. He obtained his Ph.D. in zoology from the University of Wisconsin in 1962 and taught at Dartmouth College before joining the Cornell faculty in 1969. Likens has been codirector of the Hubbard Brook Ecosystem Study since 1962. Wright is on the staff of the Norwegian Institute for Water Research in Oslo, where he has worked on a major study of the effects of acid precipitation on forests and fishes. He received his bachelor's degree in chemistry at Dartmouth, his master's degree in geology at Yale University, a second master's in geochemistry at the California Institute of Technology and his Ph.D. in ecology from the University of Minnesota, the last in 1974. Wright spent the 1977–78 academic year at the University of Bern working on sedimentation processes in large lakes. Galloway is assistant professor of environmental sciences at the University of Virginia. He obtained his Ph.D. in chemistry from the University of California at San Diego in 1972 and then worked as a postdoctoral associate with Likens at Cornell. Galloway joined the Virginia faculty in 1976. Butler is a research aide in the section of ecology and systematics at Cornell and an energy analyst at Cornell's Center for Environmental Research.

ARTHUR J. EPSTEIN and JOEL S. MILLER ("Linear-Chain Conductors") have been actively involved in the study of highly conducting molecular materials for nearly a decade. Epstein is a physicist at the Xerox Corporation's Webster Research Center in Rochester, N.Y. He did his undergraduate work at the Polytechnic Institute of Brooklyn and went on to obtain his Ph.D. in physics from the University of Pennsylvania in 1971. After working for a year at the Mitre Corporation, Epstein joined the Xerox research staff in 1972. Miller is a chemist with the Occidental Research Corporation in Irvine, Calif. He did his undergraduate work at Wayne State University and received his Ph.D. in inorganic chemistry from the University of California at Los Angeles in 1971. After postgraduate work at Stanford University he went to Xerox to study the chemistry of linear-chain molecular materials. He moved to Occidental Research this year.

CALVIN F. QUATE ("The Acoustic Microscope") is professor of electrical engineering and applied physics and chairman of the department of applied physics at Stanford University. He obtained his bachelor's degree in electrical engineering at the University of Utah and his Ph.D. from Stanford in 1950. In 1949 he joined the technical research staff of Bell Laboratories, where he was later appointed associate director of electronics research. After 10 years at Bell Laboratories he moved to the Sandia Corporation in Albuquerque, N.M., where he subsequently became vice-president and director of research. He went to Stanford in 1961.

DAVID W. FRASER and JOSEPH E. McDADE ("Legionellosis") are on the staff of the Center for Disease Control of the U.S. Public Health Service in Atlanta, Ga. Fraser is chief of the Special Pathogens Branch of the Bacterial Diseases Division in the center's Bureau of Epidemiology. He was graduated from Haverford College and received his M.D. at the Harvard Medical School in 1969. After an internship and residency at the Hospital of the University of Pennsylvania he was Special Projects Associate in epidemiology at the Mayo Graduate School of Medicine in Rochester, Minn. He then returned to the University of Pennsylvania as chief resident in medicine and taught at the medical school. Fraser moved to the Center for Disease Control in 1977. McDade is a research microbiologist in the Leprosy and Rickettsia Branch of the center's Virology Division. After obtaining his Ph.D. in microbiology from the University of Delaware in 1967, he spent two years studying rickettsial infections at the U.S. Army Biological Center at Fort Detrick, Md. He then worked for two years as director of the cell-production department of Microbiological Associates, Inc. From 1971 to 1975 he was a research associate in microbiology at the University of Maryland School of Medicine, spending most of his time in Cairo and Ethiopia on studies of putative extrahuman cycles of epidemic typhus. He went to the Center for Disease Control in September, 1975.

KENNETH R. MILLER ("The Photosynthetic Membrane") is associate professor of biology at Harvard University. He did his undergraduate work at Brown University and received his Ph.D. in biology from the University of Colorado at Boulder in 1974. His research interests include the function of the photosynthetic membrane in the light reaction of photosynthesis, experimental techniques in high-resolution electron microscopy and the biochemistry of membrane-associated large molecules. In his spare time Miller is active in Amateur Athletic Union (A.A.U.) Masters swimming: he is the New England champion in the 30–34 age group in the 50-yard butterfly and the 50-yard backstroke.

DALLAS L. PECK, THOMAS L. WRIGHT and ROBERT W. DECKER ("The Lava Lakes of Kilauea") began their studies of molten "lakes" of lava in 1963, when they were stationed at the Hawaiian Volcano Observatory. Peck is chief geologist of the U.S. Geological Survey. He obtained his bachelor's and master's degrees at the California Institute of Technology and his Ph.D. from Harvard University in 1960. Peck then did studies for the Geological Survey in western Oregon, in the Sierra Nevada granite-batholith area of California and in Hawaii. Wright is a staff geologist with the Survey. He did his undergraduate work at Pomona College and received his Ph.D. from Johns Hopkins University in 1961. He then did studies for the Survey in the Cascade Range of Washington and in Hawaii. Since leaving Hawaii in 1969 Wright has investigated basalts from the moon, the mid-Atlantic Ridge and the Columbia River. Decker is scientist-in-charge of the Hawaiian Volcano Observatory. He did his undergraduate and master's work at the Massachusetts Institute of Technology and obtained his D.Sc. from the Colorado School of Mines in 1953. After teaching and doing volcanological research in Indonesia he joined the faculty of Dartmouth College. While he was at Dartmouth he did field work in Hawaii, Iceland and Central America. Decker was appointed head of the Volcano Observatory early this year.

ERIC BUFFETAUT ("The Evolution of the Crocodilians") is on the research staff of the French National Center of Scientific Research, working in the Laboratory of Vertebrate Paleontology of the University of Paris. A native of Normandy, he became interested in fossils as a child. He studied geology and paleontology at the University of Paris and received his doctorate in 1974 for a thesis on fossil crocodilians from the upper Cretaceous of Niger. Buffetaut's research on the evolutionary and biogeographical history of the crocodilians has involved field work in France, Canada, Tunisia, Mali and Pakistan.

ALLAN CALDER ("Constructive Mathematics") is lecturer in mathematics at Birkbeck College of the University of London. He was educated at the University of London, obtaining his B.Sc. in mathematics in 1967 and his Ph.D. in topology in 1970. He has since taught at a number of universities, including Carleton University in Ottawa, Louisiana State University, the University of Essex and the University of Missouri, and he is now visiting associate professor at New Mexico State University.

MATHEMATICAL GAMES

Some packing problems that cannot be solved by sitting on the suitcase

by Martin Gardner

"It *is* a Square!... Beautiful! Beau-tiful! Equilateral! *And* rectangular!"
—LEWIS CARROLL,
A Tangled Tale, Knot Two

In this department for January I discussed the unsolved general problem of packing n identical circles into squares of minimum area so that the circles do not overlap. Packing n identical squares into larger squares of minimum area presents a similar problem, which except for very low values of n is also unsolved. This month, after summarizing the sparse results pertaining to this difficult task, I shall turn to some other curious questions concerning the packing of squares and then make a brief excursion into the third dimension.

Like the task of packing circles into a square, the task of packing squares into a square can be viewed in two ways. On the one hand, if the outer square is assumed to have a side of length 1, the problem is to determine how large n identical squares can be and still fit into the outer square. On the other hand, if the identical squares are assumed to have a side of 1 (in which case they are called unit squares), the problem is to determine the smallest square into which the n unit squares will fit. The latter procedure is the one that will be considered here. Amazingly, minimal solutions have been found only in those cases where n is the square of an integer or is equal to 2, 3 or 5.

Call the side of the square to be packed with unit squares k. The illustration on the opposite page displays packings for the best, or lowest, values known for k when the number of unit squares n ranges from 1 through 15, with brief comments on each packing. Obviously whenever n is a square number, k is the square root of n, and in the cases where n equals 2, 3 or 5 it is not difficult to prove that the values of k shown in the illustration are minimal. When n is not a square number, the side k must be greater than the square root of n, and it can always be taken to be less than or equal to the lowest integer greater than the square root of n.

When n is not large and equals $a^2 - a$ for some integer a, it is conjectured that the side k of the minimum bounding square is equal to a. Or to put the conjecture geometrically, the unit squares in an $(a - 1)$-by-a array cannot be rearranged to fit into a square of side smaller than a. If the conjecture is true, it follows at once that a square of the same size is also minimal when the number of unit squares is increased up to and including a^2. In order to see this assume that the bounding square can have a side smaller than a and still accommodate one or more unit squares in the empty, or ath, row. Removing the extra square or squares will leave $a^2 - a$ squares inside a square of side smaller than a, which contradicts the conjecture.

The smallest value of n for which the conjecture does not hold is not known. Ronald L. Graham of Bell Laboratories has found that the conjecture is violated when n equals $40^2 - 40$, or 1,560 (that is, when the outer square measures 40 on a side), but he believes the smallest value of n that violates the conjecture is considerably lower, perhaps closer to $10^2 - 10$, or 90. The smallest n for which the best packing requires a tilting of squares at any angle other than 45 degrees is also not known.

These results, as well as those given in the illustration on the opposite page, are taken from the only reference I know for this fascinating problem, the paper "Geometrical Packing and Covering Problems," by the Dutch mathematician F. Göbel. It appears in *Packing and Covering in Combinatorics*, an anthology (in English) edited by A. Schrijver and published in 1979 by the Mathematical Centre, Tweede Boerhaavestraat 49, Amsterdam. When n equals 16, of course, the value of k is 4. Packings for the best values known for k when n equals 17 and 18 are shown at the top in the illustration on page 20. If the conjecture mentioned above is true, k equals 5 when n equals 20 through 25.

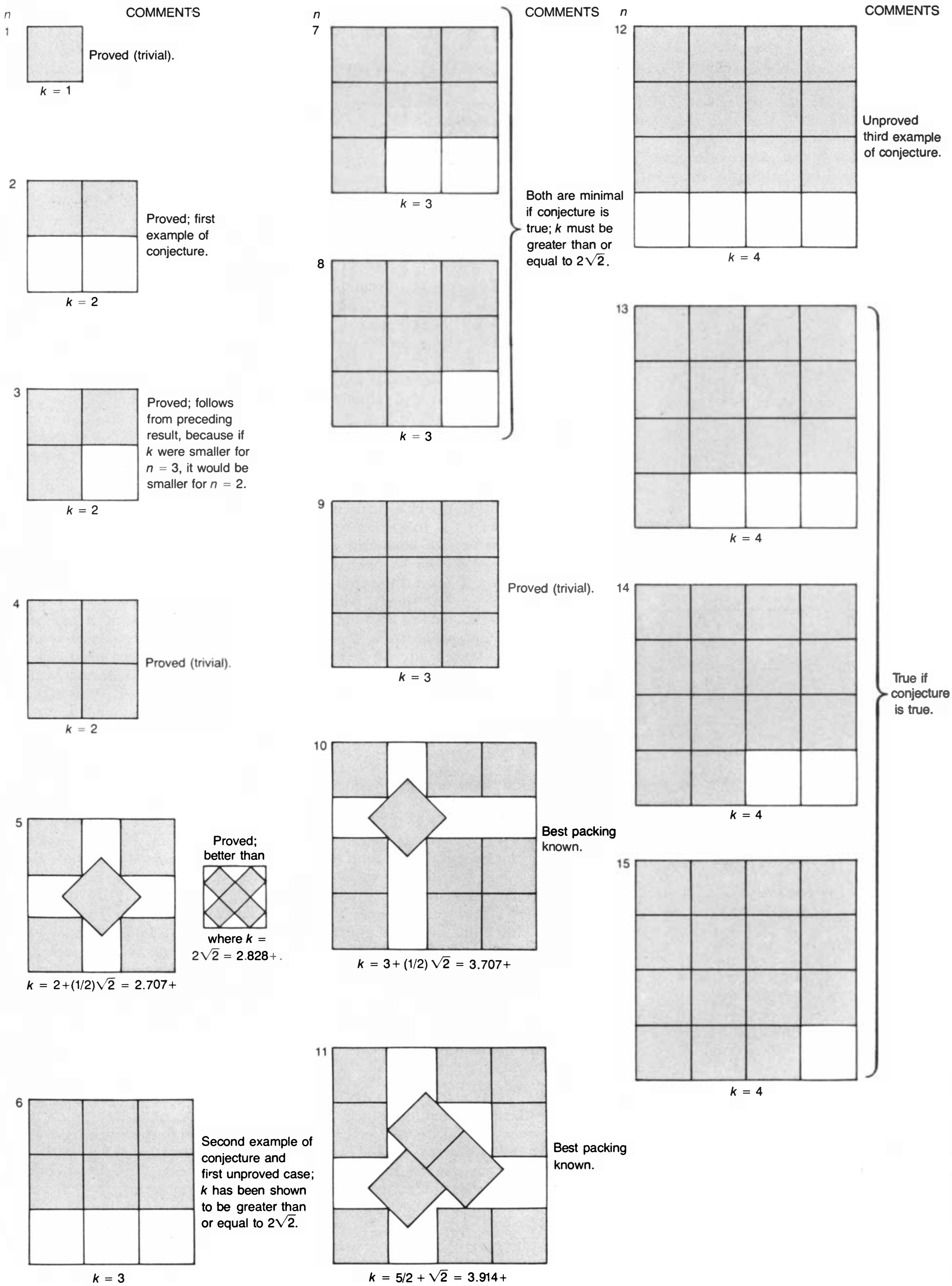
The case $n = 19$ is of special interest. Early this year, after reading in this department about the packing of unit circles into squares, Charles F. Cottingham began to investigate the packing of unit squares into squares. Unaware of the previous work in this area, Cottingham duplicated Göbel's results up through $n = 25$, with one exception. For the case $n = 19$ he obtained a slightly better packing than the one shown at the bottom in the illustration on page 20. Can the reader improve on this pattern before I give Cottingham's next month? Cottingham also improved on Göbel's patterns for certain values of n higher than 25.

As n becomes larger the task of proving that a packing square is minimal (when n is not a square number) becomes increasingly difficult. Even a proof for the case $n = 5$ is not trivial. One way to go about determining the minimum side length of the square that bounds five unit squares is to divide it into four equal square regions. Since five unit squares must go inside, the pigeonhole principle ensures that the centers of at least two of the unit squares must be inside or on the border of at least one of the four square regions. The final step is to show that a square region large enough to accommodate the centers of two unit squares cannot be smaller than $1 + (1/4)\sqrt{2}$ on a side. Hence the side of the minimal bounding square must be twice that length, or $2 + (1/2)\sqrt{2}$.

No general procedure for finding minimum bounding squares is anywhere in sight, but in 1975 Paul Erdös and Graham published a paper in which they proved a remarkable theorem. They showed that as the number of unit squares increases to a sufficiently large number, clever packings can lower the amount of wasted space to an area that never exceeds $k^{7/11}$, or $k^{.636+}$, where k is the side of the bounding square. (The area of wasted space is zero when there is a square number of unit squares.)

Hugh Montgomery later lowered this asymptotic bound slightly, to $k^{(3-\sqrt{3})/2}$, or $k^{.633+}$, and it can probably be lowered further. Graham conjectures that as n approaches infinity the ultimate bound may well be $k^{.5}$, or the square root of k. In a paper published in 1978 Klaus F. Roth and Robert C. Vaughan, two British mathematicians, showed that the bound cannot have a limit below $k^{.5}$. Therefore as the matter now stands in optimal packings of large numbers of unit squares into minimal bounding squares the wasted area will be greater than or equal to $k^{.5}$ and less than or equal to $k^{.633+}$.

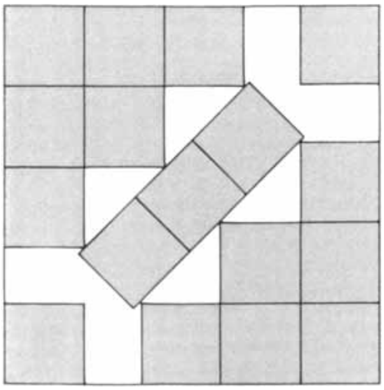
To dramatize how much space can be saved when the number of unit squares is large Graham considers the packing of unit squares into a large square of side 100,000.1. If the squares are conventionally packed in straight rows, $100{,}000^2$ unit squares can be fitted in.



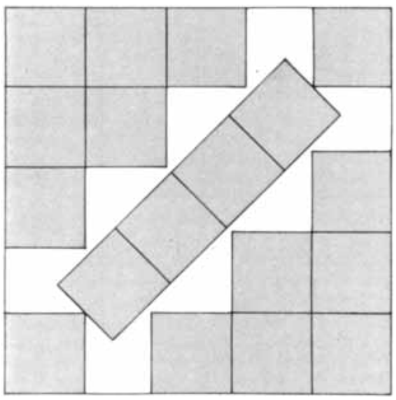
Best results known for packing **n** *unit squares into the smallest square possible when* **n** *ranges from 1 through 15*

This packing leaves an extremely thin, linear border of empty space into which no unit square can fit along two sides of the outer square. If the squares are tilted and packed according to the Erdös-Graham-Montgomery technique, however, more than 6,400 additional unit squares can be fitted in!

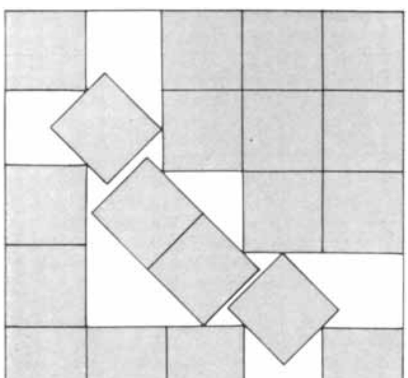
Now consider the sequence of squares whose sides are the consecutive integers 1, 2, 3, 4, 5.... The consecutive areas of these squares form the infinite sequence 1, 4, 9, 16, 25..., and the consecutive partial sums of this sequence are 1, 5, 14, 30, 55.... Does the sequence of sums include a square number? Yes, but only one: 4,900. In the mid-1960's this number, which is the square of 70 and the sum of the first 24 square numbers, suggested the following problem: Can the 70-by-70 square be packed with the first 24 squares whose sides are consecutive integers? The answer to the question is no, but there are packings of squares from the set that omit the 7-by-7 square and leave as little as 49 unit squares of wasted space (see my *Mathematical Carnival,* page 149). This packing is thought to be minimal, although I know of no proof of the conjecture. Nor do I know whether a packing with less wasted space can be achieved if the order-70 square is viewed as a cylindrical or toroidal surface.



$n = 17, k = 4 + \sqrt{2}/2 = 4.707+$



$n = 18, k = 2 + 2\sqrt{2} = 4.828+$



$n = 19, k = 4 + (2/3)\sqrt{2} = 4.942+$

Packings for **n** *= 17,* **n** *= 18 and* **n** *= 19*

Will a set of these consecutive squares (starting with 1 or a higher number) tile a rectangle? This question, first asked by Solomon W. Golomb, remains unanswered. If the answer is no, as is suspected, Golomb suggests the following refinement: What pattern of consecutive squares starting with 1 can be packed into a larger square to leave the smallest percentage of wasted space? The same question can be asked of a bounding rectangle. I know of no work that has been done so far on any of these questions.

No square can be packed with squares of sides 1, 2, 3, 4, 5..., and the same is probably true for a rectangle, but is it possible to tile the infinite plane with consecutive squares starting with 1? This unusual problem was posed by Golomb in *Journal of Recreational Mathematics* (Vol. 8, No. 2, pages 138–139; 1975–76). Consecutive squares of sides 1 through 11 can be spiraled around a center as shown in the illustration on page 24, but that tiling leaves an unavoidable hole: the space marked *X.* No algorithm for tiling the entire plane with consecutive squares beginning with any integer has been found, but the task has not been proved impossible.

The upper illustration on page 26 shows one way consecutive squares can be placed to form a simply connected region (an area with no holes) that will cover more than three-quarters of the plane. The southeast quarter is covered by squares with sides in the familiar Fibonacci sequence: 1, 2, 3, 5, 8, 13.... The southwest quarter is covered by a truncated Lucas sequence that begins 7, 11, 18, 29... and a thin strip of overlap from the squares on the right. (The Lucas sequence, which begins 1, 3, is the simplest of the generalized Fibonacci sequences.) The northeast quarter is covered by squares in a truncated Fibonacci sequence that begins 6, 9, 15, 24... and a thin strip of overlap from the squares below. Verner E. Hoggatt, Jr., the editor of *The Fibonacci Quarterly,* has shown that no numbers are duplicated in these three sequences. The squares with the sides that were omitted from the sequences are arranged along the top of the Lucas squares. The black spot in the illustration is the origin point of the Cartesian plane. This pattern of consecutive squares covers more than three-quarters of the plane, of course, but it is not known whether even the quarter plane can be tiled exactly with consecutive squares.

There are many ways to tile the entire plane with squares whose sides are not consecutive but only different. Here is one way, which makes use of the fact that a square can be dissected into as few as 21 different squares. (For a report on this recent discovery see "Science and the Citizen" in *Scientific American* for June, 1978.) The smallest square in any such dissection is necessarily an interior one, and so to tile the plane let a dissected square, with its 21 subsquares, serve as the smallest square in a larger replica of the same pattern. The larger (tiled) square can then be the smallest square in a still larger replica, and so on. By continuing in this way it is possible to completely cover the infinite plane with squares of different sizes.

For more than a decade it has been known that any set of squares with a total area of 1 (the sides need not be rational) can be packed without overlap into a square of area 2 (see *Mathematical Carnival,* Chapter 11). Suppose a rectangle has unit width. What is the minimum length that will allow the rectangle to bound any set of squares with a combined area of 1? A long-conjectured answer of $\sqrt{3}$ was verified by Daniel J. Kleitman and Michael M. Krieger in their paper "Packing Squares in Rectangles I" in *Annals of the New York Academy of Sciences* (Vol. 175, Article 1, pages 253–262; July 31, 1970). Later in "An Optimal Bound for Two Dimensional Bin Packing" (*16th Annual Symposium on Foundations of Computer Science,* IEEE Computer Society, Institute of Electrical and Electronics Engineers, 1975) the same two authors showed that the rectangle of smallest area into which any set of squares with a total area of 1 can be packed has sides of $2/\sqrt{3}$ and $\sqrt{2}$. In each of these results the wasted area is huge—more than 60 percent of the area of the set of squares.

A classic square-packing problem known as Mrs. Perkins' Quilt is named for a puzzle mentioned in one of Henry Ernest Dudeney's books. (John Horton Conway's work on this problem is also described in Chapter 11 of *Mathematical Carnival.*) Here the task is to divide a square with integral side *n* into the smallest number of nonoverlapping subsquares with integral sides so that no space is wasted. In this puzzle the sub-

squares are not required to be different.

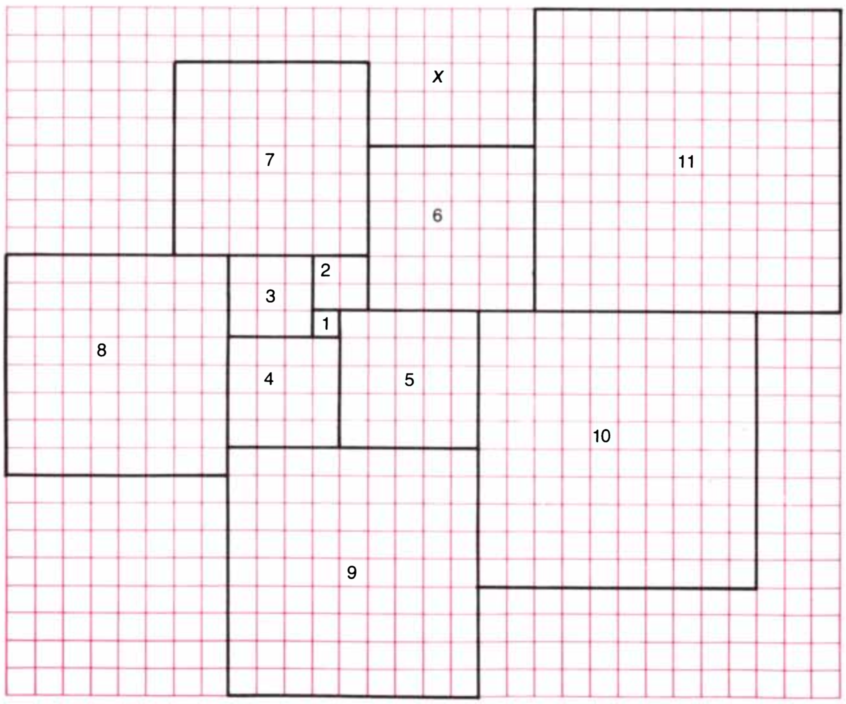
The following question is closely related to Mrs. Perkins' Quilt: What is the largest number of subsquares (allowing duplicates) into which no square can be cut? Or to put it another way, what is the smallest number n of subsquares such that a square can always be cut into n and all higher numbers of subsquares? It is not hard to show that a square cannot be cut into two, three or five subsquares, but it can be cut into six subsquares and all higher numbers of subsquares. (To cut a square into six subsquares, divide a 3-by-3 square into nine unit squares and then mark off one subsquare of side 2 and five of side 1.) Hence the answer to the question phrased the first way is 5, and the answer to the question phrased the second way is 6.

The natural extension of this problem to cubes turned out to be considerably more difficult. It was established that a cube can be subdivided into n subcubes when n equals 1, 8, 15, 20, 22, 27, 29, 34, 36, 38, 39, 41, 43, 45, 46, 48, 49, 50, 51, 52, 53, 55 and all higher numbers. The task was proved to be impossible for all the remaining values of n except one: $n = 54$. For many years the question of whether a cube could be cut into 54 subcubes remained a perplexing problem.

As long as the case $n = 54$ was unsolved it was impossible to answer the following question: What is the largest number of subcubes, not necessarily different, into which a cube cannot be cut? The question became known as the Hadwiger problem because it had been framed by a Swiss mathematician, Hugo Hadwiger of Bern. The answer was thought to be 54, but several years ago a dissection of a cube into 54 subcubes was independently found by two other residents of Bern, Doris Rychener, a flute teacher, and A. Zbinden of the International Business Machines Corporation. The solution was published by Richard K. Guy in his department on research problems in *The American Mathematical Monthly* (Vol. 84, No. 10, page 810; December, 1977). The lower illustration on page 26, based on an illustration provided by Guy, shows 42 cubes of side 1, four of side 2, two of side 3 and six of side 4. The four pieces shown go together in an obvious way to make a cube of side 8.

It is now possible to answer Hadwiger's question. The largest number of cubes into which a cube cannot be cut is 47; or 48 is the smallest number such that a cube can be cut into that number of subcubes and all higher numbers of them. Incidentally, it is not possible to cut a cube into any number of subcubes of different sizes. The proof, one of the most beautiful in combinatorial geometry, can be found in my addendum to William T. Tutte's article "Squaring the Square," reprinted in my *2nd Scientific American Book of Mathematical Puzzles & Diversions.*

Many curious square-packing problems concern only a single square that fits into a region of a specified shape. For example, what is the largest square that can be inscribed in a regular pentagon of



Can the plane be tiled with consecutive squares?

Step up to excellence. Step up to Series E.

Why settle for anything less than excellence. Especially now that it's available at a very affordable price in Series E from Hewlett-Packard—five precision calculators for science, engineering and business starting at just $50.*

EXCELLENCE BY DESIGN. Experience the HP difference in Series E: a new larger display with commas for easy reading; built-in diagnostic systems that help catch errors; RPN logic for ease-of-use and efficiency; a complete documentation system—Owners' Manual and Applications Books; plus a host of HP extras—low battery warning light, rechargeable batteries, positive feedback keys, sturdy impact-resistant case. In sum—excellence by design.

A SERIES E FOR EVERY NEED. Series E gives you more power and features at a lower price than any comparable calculator HP has ever offered.

NEW LOW PRICES:

HP-31E. Scientific. $50.*
HP-32E. Advanced Scientific with Statistics. $70.*
HP-33E. Programmable Scientific. $90.*
HP-37E. Business Management. $75.*
HP-38E. Advanced Financial with Programmability. $120.*

SEE YOUR HP DEALER. For details, send in the coupon or visit your nearest HP dealer. For the name and address, CALL TOLL-FREE 800-648-4711 except from Alaska or Hawaii. In Nevada, call 800-992-5710.

HEWLETT hp PACKARD

Dept. 236M, 1000 N. E. Circle Blvd., Corvallis, OR 97330

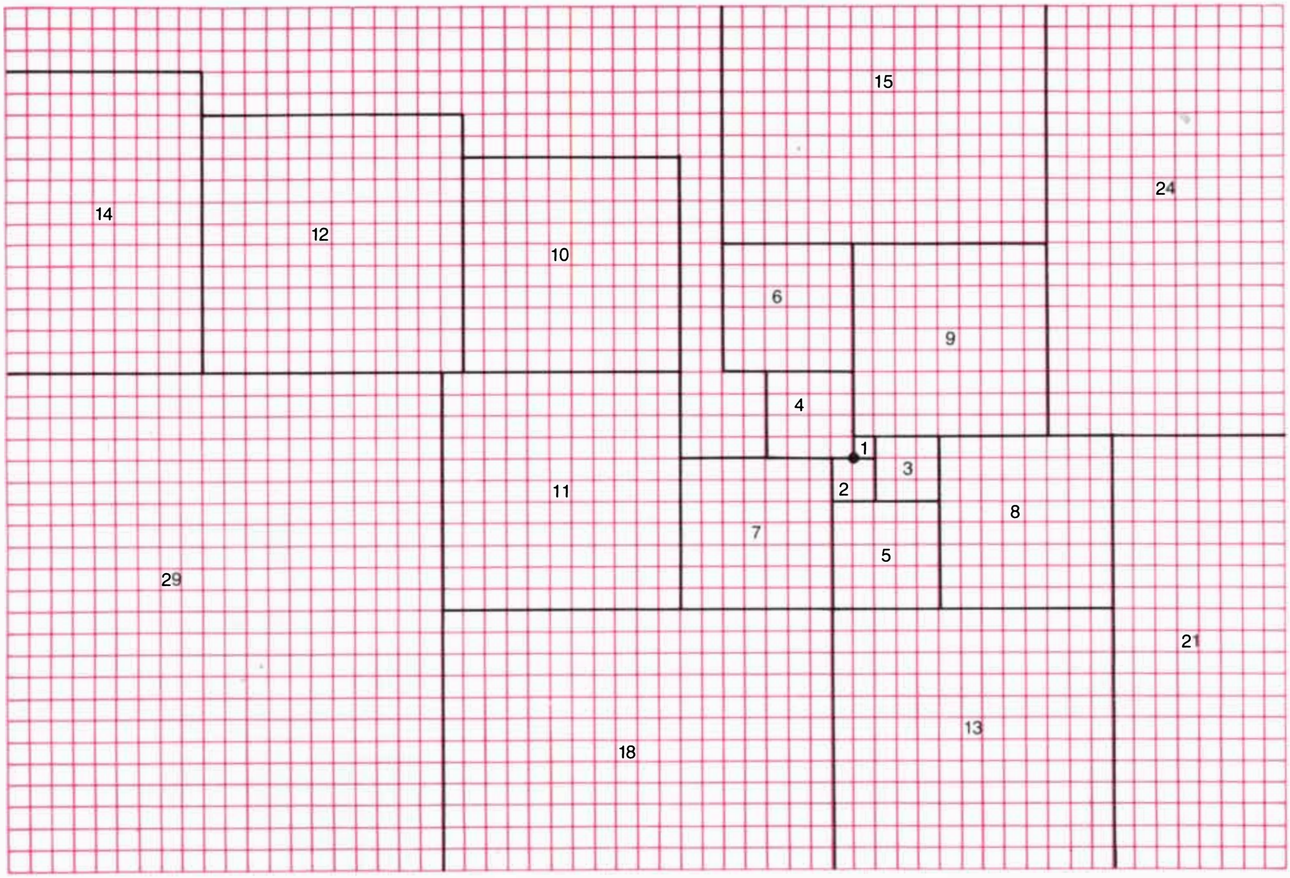
HEWLETT-PACKARD
Dept. 236M
1000 N.E. Circle Blvd.
Corvallis, OR 97330
Please send details on Series E:
_____ Business Calculators
_____ Scientific Calculators
Name_____
Address_____
City_____
State_____ Zip_____

*Suggested retail price excluding applicable state and local taxes—Continental U.S.A., Alaska and Hawaii.
Displays photographed separately to simulate typical appearance. 619/17

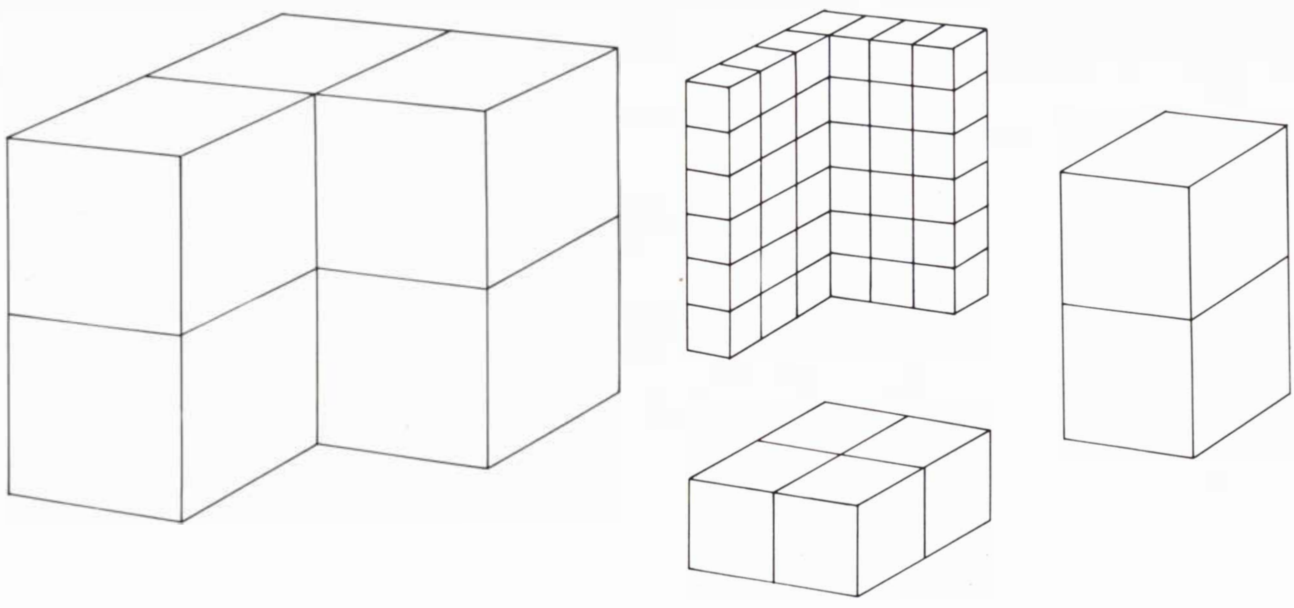
side 1? The solution of this little-known and deceptively simple problem will be given next month.

Last month readers were asked to find a solution in integers of $x^y = y^x$, with x greater than y, other than the familiar 4, 2. The second solution is $-2, -4$.

Another exercise was to prove that it is impossible for the decimal expansions of pi and e to contain all of each other. If it were possible, then after a finite number (x) of decimal digits in pi, e would begin, and after a finite number (y) of decimal digits in e, pi would begin. In that case both pi and e would be repeating decimal fractions, each with a period equal to $x + y$. All repeating decimal fractions are rational, however, and since pi and e are known to be irrational, the assumption must be false.



Covering more than the three-quarter plane with consecutive squares



Doris Rychener's dissection of the 8-by-8-by-8 cube into 54 subcubes

BECAUSE only rare European Oak has the character to gentle and mature our Chardonnay, we are cradling our Chardonnay in these handmade casks until the day the winemaker judges it ready for release. Every step we take, we take with care because THE WINE REMEMBERS

THE WINERY OF
ERNEST & JULIO
GALLO

Expected release: Spring, 1980.

BOOKS

SALT and beyond, the autobiography of a physicist, the evolution of surgery

by Philip Morrison

Arms Control and Salt II, by W. K. H. Panofsky. University of Washington Press ($6.95). World Armaments and Disarmament: SIPRI Yearbook 1979, by the Stockholm International Peace Research Institute. Distributed in the U.S. by Crane, Russak & Company, Inc. ($47.50). The muddy flood of debate over SALT II, with its heavy load of political opportunism and mixed motives, has not managed to silt over the giant shape of the real issues. They loom underneath; in the words of Professor Panofsky, a physicist whose insider's measured but critical thought has long been heard in public debate on nuclear arms, the "stockpiles of nuclear weapons today represent a qualitatively different situation from that ever faced by man. The total destructive power releasable by nuclear weapons now becomes comparable to cosmic forces." In two brief lectures he offers a cogent and informed overview, in his fair-minded and open manner. The lectures make one strong point: The physical reality and the physical uncertainties inherent in the use of the nuclear stockpile are in conflict with the symbolic perceptions of military strength so often held to be the measures of superiority, resolve and status. He treats fully the question of the neutron bomb, which is not a revolutionary tactical weapon but became a major issue touching on resolve and on the risk of nuclear escalation. In a special final section he puts forward a strong argument for SALT II, with a direct and simple response to the major objections. The effort to limit the "truly insane" accumulation of weapons requires that we view clearly the factual physical threat hidden behind the symbolism of political discourse. If we do not, we may lose all "ability to answer the fateful question: When is there enough?" A summary of the numerical content of SALT II ends this booklet, a model of concise and sensible argument.

Much more detailed, quite technical and gloomier, the latest and thickest annual summary volume from the Stockholm International Peace Research Institute (SIPRI) deserves special notice. Invaluable data are presented, such as the strategic-weapons holdings of the powers, world military spending, the arms trade and the chief arms manufacturers worldwide, reconnaissance satellites—all in more detail than can be found elsewhere. There is a fresh study (by an analyst named Owen Wilkes) examining in detail the communications system that can bring the terrible command to the missile submarines deep in the ocean and that also directs the hunter-killer submarines there. A full-page diagram of that system, U.S. Navy version, is a fascinating tangle of redundant links and modalities. Airborne very-low-frequency radio transmitters (one is at this moment in flight over the Atlantic) are not the last-ditch backup for the ground centers and the orbiting satellites. That is the role of the Emergency Rocket Communications System, long-range transmitter relay stations carried as payload by a few missiles out there somewhere among the thousand silos, ready to pop up for an hour or two to speak the word from on high.

The SIPRI study concludes that the combination of the powerful counterforce against land-based Russian missiles that our plans for Missile X represent, taken with the wide lead now held by U.S. submarine and antisubmarine technology, implies that the U.S. nuclear-deterrent stance "will be abandoned in favor of a war-fighting and war-winning strategy." In his scenario Wilkes foresees that all Russian missiles will be vulnerable in a few years to a U.S. first strike from the land, the sea and the air, except perhaps for the long-range missiles on the Russian *Delta*-class submarines in or near their subarctic home ports on the Barents Sea. And true to the worst-case style he cites various U.S. hints of penetration of even that icy sanctuary.

No doubt such surprise schemes lie at the limits of plausible planning. If that, however, is the open concern of an academic student of war in Stockholm, what is the dour view of the responsible Russian analysts working in secret in some target strongpoint? The present balance is growing visibly unstable; American hawks are sure to alert, or to induce, their Russian counterparts. Even with SALT II the sky is unsafe. Without it the decade glows ominously.

Disturbing the Universe, by Freeman Dyson. Harper & Row, Publishers ($12.95). "In September 1947 I enrolled as a graduate student in the physics department of Cornell University... to learn how to do research in physics under the guidance of Hans Bethe.... Two things about him immediately impressed me. First, there was... mud on his shoes. Second, the other students called him Hans. I had never seen anything like that.... In England professors were treated with respect and wore clean shoes." That alert, diffident young man is now professor at the Institute for Advanced Study in Princeton, a man whose originality and versatility of mind and articulate pen have led him to lasting contributions in mathematics, quantum field theory and public policy, and to startlingly advanced engineering designs, from neutron bombs and power reactors to optics and space travel. This reviewer first heard Dyson's name at Cornell a few months before Dyson came to the U.S. It was Hans Bethe who read it out, from a letter of recommendation Bethe was sharing with the other theorists, who had met to find agreement on who should be taken into the busy group. It came from a senior professor at Cambridge, a man who had long been a friend of Bethe's. The letter was terse but clear. It ran more or less like this: "Dear Bethe, I think you should admit Freeman Dyson to study with you. Though he is in his early twenties, he is the finest mathematician in England." We took him, and the choice was wise.

Not a word of his mathematical career, and no formula of any kind, enters this extraordinary and beautiful book. In the form of an autobiography, but a selective one, it is full of estimates of other people, of reflections and motives, of the inward man more than of the outward one. It is written in a swift, ribbon-clear prose, humane and self-revealing, much blood and no little fire in the ink. Everywhere enriched by the work of poets, it stands as the deepest and most readable account of the personal choices, influences and interior guides of a productive scientist yet to see print.

What the reader should do at this point is seek out the book. In our too busy era, however, reviews generally find more readers than the dearer and longer books on which they comment. There follows, therefore, some account of a few of the two dozen brief chapters, of Dyson's life in England and in America, and of "points beyond," a characteristically daring and personal set of extrapolations on the long-range destiny of our kind and on the even larger questions of cosmic meaning.

In the last years of World War II, Dy-

son worked as a civilian analyst in the Operational Research Section of Bomber Command. His study took him at 20 as a visitor to Wyton R.A.F. Base, "as ugly as a wartime military base can be. Endless puddles, barracks, warehouses full of bombs, rusting wreckage" surrounded the runways from which the overloaded Lancasters of 83 Squadron took off night after night into the blackness. They went to Berlin, to shower it with fire, which led to a steady dehousing of wretched civilians, while the many important military targets were well protected by their trained fire-fighting teams.

The air crews paid slowly but heavily. Their long-run chances were small; the special pathfinder crews from Wyton base had about one chance in 11 of fulfilling alive their allocated tour of duty: 60 missions. Dyson knew those odds; he had reckoned them. Not so the crewmen. He intended to find out from them firsthand what crew experience might have to do with plane survival. The records showed that it was coming to mean less and less; veteran crews were shot down like tyros. But serious conversation turned out to be impossible. "What could they say to me, or what could I say to them, across the gulf that separated us?" They too were boys of 20, who would face flaming death 60 times, if they were lucky. "They knew, and I knew that they knew, that I was one of those college-educated kids who found themselves cushy civilian jobs and kept out of harm's way."

Such was Bomber Command; it "might have been invented by some mad sociologist... to exhibit as clearly as possible the evil aspects of science and technology." It was a huge, mindless organization, improving flying machines to make them death traps for the youths who flew them, burning ancient cities and killing their people to little purpose, acting in secrecy mainly to conceal its own "failures and falsehoods" from the high authorities in London and from the men in the squadrons. Its commander-in-chief "never admitted mistakes, and appeared to be as indifferent to the slaughter of his own airmen as he was to the slaughter of German civilians." His scientific advisers were "too timid to challenge any essential element of his policies" they daily measured to be mad, a "crazy game of murder."

A medically discharged young soldier shared Dyson's office at Command headquarters. "He knew less mathematics than I did but more about the real world. When we looked around us at the brutalities and stupidities of the Command, I got depressed and Mike got angry. Anger is creative; depression is useless." What Mike grew angry about was his own inference from the fact that some designer's half-conscious choice had made the Lancaster floor escape hatch of minimum size, inches smaller than that on the older Halifax. When American bombers were shot down in daylight, half the men escaped, to turn up as prisoners of war. The Halifaxes, from which the men dropped by parachute at night, showed a smaller recovery, about 25 percent. The Lancasters, bigger and better night bombers in every other way, returned 15 percent. The too small hatch in a burning aluminum hull going out of control delayed the bailing out in terror of seven men hasty for a last chance at life, impeded at those cold, breathless heights by their bulky flying suits and parachute harness. In the entire command only Mike worried about such numbers; Lancaster crewmen did not know, and they were not encouraged to think, about how small a number of them would succeed in parachuting.

Those missing inches of hatch cost the lives of several thousand men. Mike spent two years forcing the enlargement of the Lancaster hatch. He succeeded, although it was maddeningly slow. His anger had to climb up the rungs of the command and down a long ladder again to the production plants with an unwelcome message of error. Dyson was desolate. From pacifism he had come to serve the war on Hitler, from war had come to bombing and then to indiscriminate city attacks, at least justified by victory. But did such attacks in fact aid victory? No; then at least one could work to save the crews. By the last wartime spring Mike had saved lives but "I had saved none. I had surrendered one moral principle after another, and in the end it was all for nothing."

Postwar America was sunnier. The tale of Dyson with Richard Feynman on the road to Albuquerque is a classic of youthful discovery and insight, as memorable as Werner Heisenberg's happiness waiting for sunrise on the dunes of Heligoland after his day's find of the matrix mechanics. Over the years we encounter Robert Oppenheimer, Edward Teller and Theodore Taylor, the visionary designer of bombs turned herald of the danger of nuclear terrorists. In 1960 Dyson published a startling piece asserting that fission-free nuclear bombs were the wave of the future, that the nuclear-test ban would deny their development and perhaps reduce the U.S. to "the position of the Polish army in 1939, fighting tanks with horses."

Dyson has changed. He does not mince phrases; that article "was a desperate attempt to salvage an untenable position with spurious emotional claptrap..., wrong on at least four counts: wrong technically, wrong militarily, wrong politically and wrong morally." In 1962, now a U.S. citizen, he became a "migrant worker" for the Arms Control and Disarmament Agency. He spent two summers there. This English theoretical physicist, with "an abiding love for the Russian language and an ability to read it with reasonable fluency," studied Khrushchev for the ACDA.

But the best postwar chapter tells of H. G. Wells's Doctor Moreau and J. B. S. Haldane's Daedalus. ("Black is his robe from top to toe, / His flesh is white and warm below, / ... in his eyes a still small flame / ... as he rides / Singing my song of deicides.") Young Haldane's diction was overcharged, but his message, that biological understanding will confer a terrible power, was correct enough. "Man cannot play God and still stay sane.... We still can choose to be masters of our fate." The key is to allow biological research freedom, on condition of no introduction of new species to disturb nature's balance.

Here another real hero appears onstage in the 1960's at Dyson's ACDA. Like Mike at Bomber Command, Matthew Meselson, professor of biology at Harvard, has saved the lives of many who will not know of him. By adroit formulations, tireless energy, directed persistence, indisputable knowledge and the luck to know Henry Kissinger, this firm, quiet, brilliant man has done more than any other person to rid the world of biological weapons, that clearest realization of the dire prophecies of Wells and Haldane. It is a hopeful story, one far too little known. The campaign clearly led to "Nixon's finest hour," our unilateral renunciation of biological warfare, and then on to Russian concurrence, formally signed by 1972. It had taken Meselson and the fates just nine years from his initial visit to the ACDA, when he first read the Army Field Manual 3–10. "Seldom in human history has one man, armed only with the voice of reason, won so complete a victory."

We also read in Dyson's book of the many human tongues seen as so many species, of the implications of self-reproducing automata, of extraterrestrial beings, of individual homesteading space travel on the model of the *Mayflower* or the wagon trains of Brigham Young. Dyson sees us and our counterparts diverse and millionfold among the stars, some worlds "insane, and as crazy as Doctor Moreau's island," and some who will "shit on the morning star" (from Robinson Jeffers). In all of this there is much to learn and something to dispute. The author quite consistently exalts choice and will but perhaps underestimates their inherent thermodynamic costs. He was the first to write of seeking civilizations that had completely tamed their sun's energy, so that from far away they would shine as bright as a sun, but one visible only in the deep infrared appropriate to our kind of ambient temperature. Actually engineering experience appears to suggest that such technological suns could not exist; long before the last few percent of the radiant

SCOTCH SO SMOOTH IT NEVER MAKES WAVES.

All good Scotch begins with the best of intentions and the finest of ingredients. It is in the delicate process of blending that differences are struck.

And in the taste of Cutty Sark there are striking differences.

To make good Scotch, the whiskies comprising the blend are aged in separate oaken casks, then blended together. To make Cutty Sark, the whiskies are also aged separately, then blended together, then returned to cask for up to a year and a half longer. And only <u>then</u> bottled.

This results in a Scots Whisky whose elements are perfectly knit, with a taste that's unusually smooth.

And this is one reason why, whether you prefer Cutty Sark with water, with soda, or neat, it glides over the palate with nary a ripple.

CUTTY SARK

"CUTTY SARK" "CUTTY" THE CUTTY SARK LABEL AND THE CLIPPER SHIP DESIGN ARE REGISTERED TRADEMARKS OF BERRY BROS. & RUDD LTD. LONDON, ENGLAND. 86 PROOF BLENDED SCOTCH WHISKY DISTILLED AND BOTTLED IN SCOTLAND. IMPORTED BY THE BUCKINGHAM CORPORATION, NEW YORK, N.Y.

energy coming from the nearby natural star had been captured the clever devices that would be needed to capture it would have ceased paying for themselves. Is the law of diminishing returns itself not one face of the laws of thermodynamics?

It is not hard to see an engaging bias toward the effectiveness of reason throughout Dyson's marvelous preludes and inventions. One may doubt any argument from design, even the one put here "in its newest form." The world is comprehensible more perhaps because our comprehending minds are indeed of the world—worldly—than because mind plays an essential role outside humankind. But Dyson's eloquence is worth hearing out.

No review should fail to cite at least one of the fine poems that illuminate and explicate this unique work of devotion and honesty. The best, perhaps, ends the high tragedy of Robert Oppenheimer, a man Dyson knew well over the last 20 years of life. The episodes are fascinating: they show Dyson long ambiguous about the quality of the man. In the end what prevails over the enigmatic remarks and the inexplicable frosts in a warm relationship is a late view of Oppenheimer, again as we saw him in the old Los Alamos years, a figure of clear human greatness, who bore with insight, courage and magnanimity his intolerable burden. Thus too was Oppenheimer in his last year of life, and Dyson reports it. Dyson took part in the arrangements for the funeral in Princeton one chill gray February day. Oppenheimer's wife Kitty wanted a poem read. She even knew the poem, one of Oppenheimer's favorites: "The Collar," by George Herbert. But in the end she decided against it; it was too personal for a public occasion, too vulnerable to easy distortion. Dyson reprints it now, since "Kitty is dead too, and Robert... beyond the reach of any further journalistic distortion." For Dyson it now helps to link the innermost nature of Robert Oppenheimer more to King Lear than to Hamlet. A reader might also catch Dyson within its metaphysical eloquence, as a man who sees above all the human condition, held between wisdom and folly, as a worthy enterprise to which we in hope commit all we have. At the risk of rude distortion about half of the lines are these: "I struck the board, and cry'd, 'No more; I will abroad.' / What, shall I ever sigh and pine?... / Sure there was wine / Before my sighs did drie it; there was corn / Before my tears did drown it; / ... there is fruit, / And thou hast hands. / Recover all thy sigh-blown age / On double pleasures; leave thy cold dispute / Of what is fit and not; forsake thy cage,... / Away! take heed; / I will abroad. / Call in thy death's-head there, tie up thy fears; / He that forbears / To suit and serve his need / Deserves his load. / But as I rav'd and grew more fierce and wilde / At every word, / Methought I heard one calling, 'Childe'; / And I reply'd, 'My Lord.'"

TISSUES AND ORGANS: A TEXT-ATLAS OF SCANNING ELECTRON MICROSCOPY, by Richard G. Kessel and Randy H. Kardon. W. H. Freeman and Company ($30). The sharp photograph on the wide page recalls a pile of understuffed powder puffs lying in a drawer among some shaggy tennis balls. It is, of course, no such unlikely collection; it is what the eye can make of a view of mammalian red blood cells and white cells in an arteriole, magnified more than 6,000 times. The view is that of the scanning electron microscope, with its remarkable depth of field and rich textural quality. This atlas offers about 700 such plates, each with a careful explanatory column, mostly showing the tissues of the rat, the guinea pig or the rabbit at a scale appropriate for the fine structure not of cells or membranes but of tissues and organs. Magnification is between 100- and 5,000-fold, so that a page presents a field that is at the widest the better part of two millimeters and at the narrowest less than the breadth of a hair. A pile of wallboard, ends linked by cunning ball-and-socket joining, turns out to be the transparent lens of the eye; a bunch of cables are the tails of some sperm cells in the process of formation; a wall covered with protruding tongues, quite a piece of surrealism, is the villi-covered surface of the central small intestine.

The authors have summoned to their aid here and there transmission electron micrographs at a magnification an order of magnitude higher to display the ultratexture of key structures, the layers and fibers of the living world where the view is dominated not by textural form but by modular pattern. They have also reproduced, and often redrawn, some painstaking modern microanatomical diagrams to explain what is seen in spatial and functional array. A tour through the skin of the rabbit and a set of views of a remarkable microcast of the kidney's vascular network at magnifications up to a few hundred are striking. The microcasts are done with a proprietary polymerizing compound, repeatedly washed, treated with caustic and gently oven-dried, in the end removing all tissue and leaving behind networks of extraordinary intricacy.

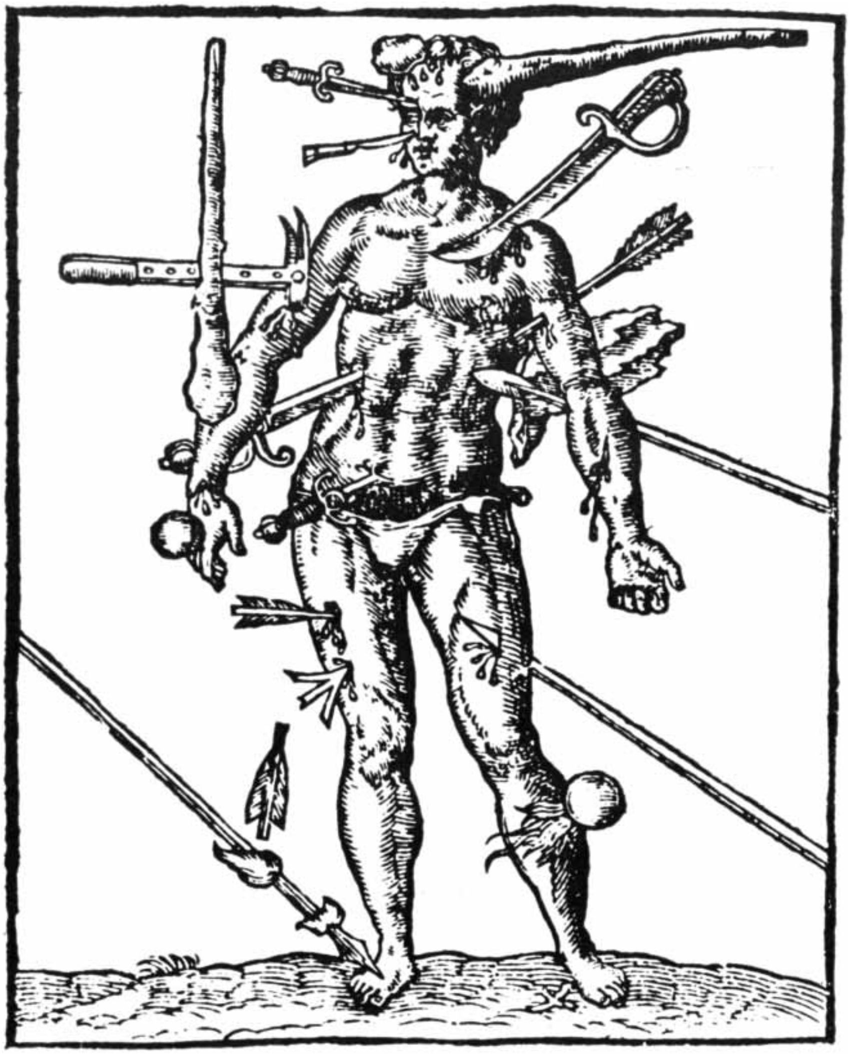
The whole is an exercise in surfaces and in connectivity, the essence of the organs of life, worth any general reader's time and fascinating for students.

THE RISE OF SURGERY: FROM EMPIRIC CRAFT TO SCIENTIFIC DISCIPLINE, by Owen H. Wangensteen and Sarah D. Wangensteen. University of Minnesota

Press ($39.50). In the 1880's a doctor from the provinces visited a friend, a distinguished surgeon newly called to head a famous Berlin clinic. "'What is new in surgery?' Replied Bergmann, 'Today we wash our hands *before* operation!'" The reply was telling. The rise of prophylactic antisepsis with its logical foundation in microbiology turned surgery from a measure of heroic necessity into a therapeutic choice. In 1881 President Garfield was shot in the back. He lingered for a couple of months, finally to die from a fatal hemorrhage "occasioned by sepsis." His wound had been probed 17 different times by "socially clean" hands. "Had a strip of sterile adhesive plaster been applied over the point of entry of the bullet to the right of the president's spine, and all probing of the bullet tract with nonsterile fingers and probes been avoided, Garfield probably would have survived to run for reelection."

From ancient times the knife was used in urgent treatment. The body impaled by an arrow and the live baby in the womb of the dead mother were visible imperatives for surgery. Even blades of flint were used with skill to cut small circular disks of bone from the skull; these operations must have given obvious relief in cases of depressed skull fracture. Survival was a not uncommon outcome.

War is always an epidemic of trauma, so that the military surgeon has had much scope, the more since the coming of the villainous gunpowder. A severe compound fracture of the leg was so grave in prognosis that amputation became a treatment of choice, even across the terrible barrier of pain. The battle surgeons aimed for speed, became knowing in the anatomy of muscle, bloodstream and joint, and devised entire tool shops, the grim instruments often ornately decorated. Firm, open statistics and rigorous accountability were all but unknown. In 1842 the mortality



"WOUND MAN," depicting injuries suffered by combatants, was a feature of military-surgery books from the 15th century. This one in *The Rise of Surgery*, by Owen H. Wangensteen and Sarah D. Wangensteen, is from the works of the 16th-century surgeon Ambroise Paré.

INCREDIBLE.
A POCKET CALCULATOR/STOPWATCH/ALARM CLOCK THAT MAKES MUSIC.

It's as small as a candy bar, and even thinner.

Yet Casio's new Melody-80 combines no fewer than nine separate timekeeping functions, the talents of a sophisticated electronic calculator, and it entertains you with Beethoven. As well as your own immortal musical creations.

And you can be among the first to have one in your pocket, simply by picking up your phone and ordering directly from The Sharper Image.

A 2-alarm alarm clock.

As a start, this is likely to be the most useful home or travel timekeeper you've ever owned.

With ±15 seconds per month quartz accuracy, it displays hours, minutes and every second, plus the day of the week—in crisp, oversized liquid crystal digits. It also gives you a four tone chime on the hour, if desired.

You press a button, and there's the date. Incidentally, its calendar is programmed to make automatic end-of-month adjustments for the next 120 years.

It gives you two independent alarm systems; each of which may be set to signal you AM or PM with either electronic beeps or 20 seconds of classical music. (Alarm 1 plays *Tarantella Napoletana;* alarm 2, several bars of Beethoven's *Für Elise.*)

You can also choose between melody or beep with its additional countdown timer, which can be set to go off at the exact second—from 1 second to 24 hrs.

Operating continuously, the Melody-80 performs for up to 12 months without a battery change (replacements are available from any jeweler or camera store).

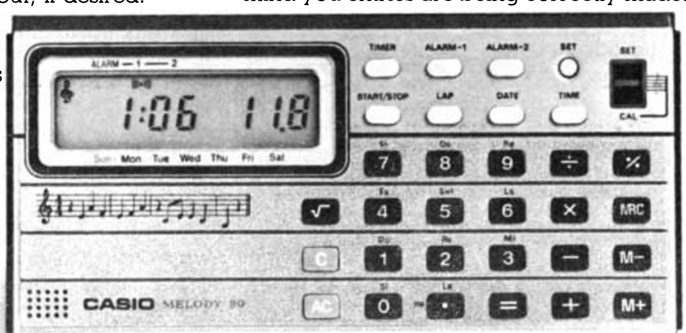
Full function stopwatch plus calculator.

While remarkably compact, no sacrifices have been made for size.

In stopwatch mode, the Melody-80 times events to the nearest 1/10 second, up to 24 indicated hours (and beyond, with automatic startover).

It tracks elapsed times, laps and timeouts, even 1-2 finishes—a feat very few stopwatches can accomplish.

The calculator solves your toughest figuring to 8 digits. Provides every calculator key with its own musical note to remind you entries are being correctly made.



While marvelously compact (4-1/2" × 2-1/4" × 5/16"), the Melody-80 provides an ample amount of space between its keys, reducing the chance of making incorrect entries. Shown approximately 3/4 actual size.

It works with plus and minus numbers, square roots, percentages, discounts, mark-ups and more. Even solves mind boggling calendar calculations, like: "What day of the week falls on January 1, 1982?"

And your calculations don't even interrupt clock or stopwatch functions.

Ready, Maestro?

As you must have figured out by now, the Melody-80 doesn't confine its genius to strictly serious business.

By flicking the upper right-hand switch, you convert its calculator buttons into music keys, a full 11 note scale from A to D.

Entertain your friends with clean, synthesizer style sounds. Play anything from boogie to Bach. Tunes can even be amplified with a microphone and home stereo for surprisingly rich tones.

Ready for an encore? Simply touch the [=] button and hear your creation repeated in a rapid, even rhythm.

Applaud. Or get your money back.

This new everything-in-one instrument is only $49.50. Which includes complete instructions, carrying case, desk-top stand, batteries, one-year manufacturer's warranty and convenient service by mail, if needed. Plus, The Sharper Image's own two-week return privilege.

If it's not the most superb calculator you've ever seen, just send it back as new for a complete and courteous refund, no questions asked. And we really mean it.

ORDER NOW TOLL-FREE.

For fastest delivery, credit card holders may use our toll-free ordering number. Or send check for $49.50 plus $1.50 delivery. Add $2.97 tax in California.

(800) 227-3436

In California (800) 622-0733

THE SHARPER IMAGE™

260 California St., Dept. CM—071
San Francisco, CA 94111
(415) 788-4747

of major amputations in nine leading Paris hospitals was 52 percent, and during the siege of Paris in 1870 not one of 70 amputees survived at the hands of a famous, if ailing, old surgeon in a city stricken by famine.

The "last critical factors looked at by surgeons in considering the nature and causes of infection and failure... were their hands." There certainly were hints. Here and there over time and place a surgeon had phenomenal success. He might ascribe his record to care in "little circumstances," as did the great Alexander Monro of Edinburgh, founder of a surgical dynasty, whose loss of patients under amputation before 1752 was only eight out of 99. Monro was among the first to offer an inclusive report of mortality; to be sure, it was easy to disclose results so benign for the time. He emphasized his concern to limit bleeding, but he had nothing to say about surgical cleanliness.

For two centuries those who "cut for stone" (entered the bladder to remove calculi) were the most discussed surgeons of Europe. Among them were the famed, the charlatans and the showmen, the adept healers and the itinerant barber-surgeons. A century before the rise of antiseptic surgery the uniquely successful Lyons lithotomist Claude Pouteau insisted on surgical cleanliness, in a work with a "very modern ring." Pouteau had as a student recovered from a severe infection that followed the prick of a finger as he was working on a cadaver. A half-dozen such farseeing men were the forerunners of it all.

The ability to perceive the facts, to set experience before learned argument and common belief, was rare. The insight and courage of Ignaz Semmelweis in the wards of Vienna mark him as hero of this tale, the effective discoverer of nailbrush and surgical scrub. His big book of 1860 did not immediately carry the day among the German professors of obstetrics, but by 1867 (with Semmelweis two years dead) even his critics had adopted his methods. The pioneer was not alone, of course. Before him went the hospital reformers, the sanitarians and those who insisted on the accountability of the profession, people such as Florence Nightingale and John Simon. After him came the time of controversy and change, particularly with Joseph Lister and his carbolic acid washes and sprays. By the late 1880's the insights of Semmelweis, given compelling power by the experimental microbiology of Louis Pasteur and Robert Koch, ruled in almost every operating room. Thermal treatment came then to replace the earlier chemical antisepsis of all tools and supplies. The surgeons began their steady expansion, now to elective intervention even into the three closed cavities of the body: the abdomen, the thorax, the cranium.

Everyone knows the sequel: anesthesia, X rays, antibiotics, detailed specialization, impressive efforts on the battlefield. The dramatic surgical amphitheater, surely nourished by motives less reasonable than pedagogy, has for the present lost its appeal, save perhaps on television. The long frock coats, "sleeves encrusted with dried blood and pus," have given way to masks, rubber gloves and sterile gowns, gowns that slowly got greener between the wars. The surgeon's role expands, transplanted organ and bypass plumbing standing for his present sovereignty. The profession today pursues a complex art of biological engineering. A mass spectrometer checks the patient's air intake, and a single versatile British surgeon of the day is cited as developer of acrylic-cemented alloy implant hip joints, laminar airflow (to reduce infection by bacteria carried to the operating field by turbulence) and so simple a change as impervious shields under the surgeon's gown (to end bacterial penetration from the sweating body of a hardworking orthopedic surgeon!).

The authors are an experienced couple, the senior author a veteran professor of surgery, teacher, ingenious experimenter and clinical innovator, the second a professional historian of medicine. They have made a fine chunky book, well studded with period illustrations and rich with citations of the extensive literature. The treatment is topical, most of the two dozen chapters spanning a long time period; it is informed everywhere by personal experience and by a simplicity and compassion of viewpoint that is endearing. But statistical and economic matters, so urgent for public health, are hardly treated, and the present state of affairs does not receive detailed attention. It is the rise of the discipline that is at center stage. The big index of names here is a real resource, because it includes the copious footnotes at the end of the book. John Snow of London is celebrated as the first anesthesiologist; we ought to recall that he was also a pioneer epidemiologist, tracing Soho cholera to one Broad Street pump.

CAUSES OF CLIMATE, by John G. Lockwood. John Wiley & Sons, Inc. ($32.95). ICE AGES: SOLVING THE MYSTERY, by John Imbrie and Katherine Palmer Imbrie. Enslow Publishers, Short Hills, N.J. 07078 ($12.95). Climate varies on every time scale: daily as the earth turns, yearly as the axis tilts, by decades as the ocean currents modulate and shift, by millenniums as the ice sheets grow and melt. It is easy to add more effects: men cultivate the land, yearly burning vast areas to clear the stubble. Others flock to enlarge the paved streets of the cities. The volcanoes pour out their quasi-random plumes of dust, mimicked—so far at just comparable scale—by the myriad smokestacks of industry. Oases arise, lakes form and dwindle. Even the oceans grow and die, as the very continents drift across the globe, turning sea to land in geological epochs. Meanwhile the rhythmic pull of moon and planets incessantly modulates the elliptical path and the tilt of the spinning earth, whose axis toplike precesses slowly in space. The swifter ocean of the air follows all these forcing terms, its energy including the heat motion of the air molecules, the store of gravitational potential energy as the parcels rise and sink and the potential energy of the water it always carries, as separate molecules or as droplets. The ruling input from the sun may well vary on scales we know not, because the sun does change its spots. In that input we all live, by crops that annually ripen by the sun, the wind and the rain.

There is one quite general consolation for those who would understand this cacophony. The variance spectrum of climate is not white, pure random noise. It is red, both east and west. At historical frequencies it is very red (periods of more than one millennium), red at intermediate frequencies and nearly white for high frequencies (periods of less than 10 years). The implication is of course that the climate has a memory, long held in the ocean gyres and held longer still on the heights piled with the slow-changing ice. Lockwood, a climatologist of the University of Leeds, offers an interesting summary of all of this—and a great deal more—at expository level, with simple mathematics but many complicated tables and graphs. He describes the physics of the radiative input and how it is shuffled, the surface where heat and water mainly exchange, the general circulation of the air and the water that spreads the surface burdens, and the evolution of it all in time. A chapter offers a prediction of the future, and a few pages of conclusion follow.

The book is long on input, short on output. That is a fair map of our knowledge. Worldwide snow, the reflectivity of the earth mapped, the half-dozen cells of the general average meridional mass motion of the air, the energy and particulate output of volcanoes and agriculture, power plants and forest fires, the runoff and soak-in of rainfall might sample the tables and maps exhibited as the chapters flow. Feedback loops are not few, and a couple of flow diagrams, all storage boxes and flow lines neatly parametrized, are here to show what a quantitative climate model consists of. Quite a few results of such model studies are summarized, on the whole rather uncritically. The book is full of interest, and it is a guide to students, but it is likely to leave the reader sadder and wiser. It is a useful reference at the first level for all who would seriously enter this vital, still quite unsure applied science.

(The hardback is certainly expensive: eight pages per dollar.)

The Imbries present a quite different book. It is a bargain by the standard set above, 17 pages for a dollar. The first half of the text, well furnished with period drawings and maps, consists of a well-written history of the ideas of worldwide glaciation, from Louis Agassiz onward. We read of Joseph Adhémar, James Croll and particularly of Milutin Milankovitch, the three authors who over a century proposed and developed to authenticity the story of the variations in the seasonal insolation of the earth that must arise from the known perturbations of the earth's orbit. The material is fresh and full of interest; the men and their work are far too little known by general readers.

The rest of the book closely chronicles research of today, mainly since the 1960's, when marine geology, particularly the improving chronology of deep-sea cores, opened a new volume in the study of past climate. There are no uncertainties here: the book avers that the cause of the ice ages is fully understood at last. The periods up to 100,000 years show up in a number of core data, just as Milankovitch would have predicted. The geological rarity of ice ages—three long episodes over all earth history—must be explained away, because the perturbations we have with us always. Continental drift can take care of that (in three pages). The absence of reliable quantitative estimates of the effects of insolation change, like the fact that the evidence is based almost entirely on the presence of peaks of about the right periods in the deviation of the power spectrum of the isotope ratio from randomness, do not bother the confident authors. Of course, no one denies the reality of such perturbations in sunlight, but their prior causal nature cannot be established even by such undoubtedly fascinating hints. Maybe even more important forcing functions of that nonlinear system are as aperiodic as continental drift. The future is laid too confidently before us: an unpublished paper shows that the ice began to return about 7,000 years back: the glaciers will be maximally extended again by about A.D. 25,000, we are told. Maybe so.

This little volume is well worth reading. It does appear, however, to be a sort of climatological *samizdat*. The oceanographer and the historian of science who sign the work are respectively father and daughter, and that promised paper on the future is by father and son. The publishing is intimate as well: the editor bears the name of the firm. All this coziness has not given much scope to any referee; there is an advocate's case here, but it may turn out to be right or wrong. Meanwhile the solution so neatly expounded should be taken along with a cooling cube of ice.

What makes the American Express Card indispensable?

I'd just be so limited

without it.

"**Traveling, shopping in London or wherever,** I like being able to buy what I want to buy. And the American Express® Card lets me. Or if I decide to stay longer someplace, I can. Unlike bank cards, no computer will stop me because an embarrassing spending limit has been set in advance. And with my crazy tour schedule, unexpected things do pop up and I simply count on the Card to handle them."

Lena Horne

Vacation time is another time when the Card is indispensable. For plane tickets, car rentals, hotels, restaurants, resorts. Not to mention shopping. And the Card gives you the flexibility to stay longer if the skiing is perfect or go further if you fall in love with the countryside. After all, there's practically no limit to what you can do with the Card.

Here's proof the Card is indispensable. Those are copies of the receipts you signed, and they come back with your monthly bill. They usually do. Making monthly budgeting expense accounts and annual tax time easier to figure and prove. Just one more reason to be a Cardmember. If you aren't already one, call 800-528-8000.

The American Express Card. Don't leave home without it.℠


IMPORTED BY CALVERT DIST. CO. N.Y.C.
PRODUCED IN SCOTLAND
IMPORTED
ONE QUART
100% SCOTCH WHISKIES
PASSPORT SCOTCH
BLENDED SCOTCH WHISKY
WILLIAM LONGMORE & COMPANY, LIMITED KEITH, SCOTLAND
ONE QUART
EST. 1864
80 PROOF
LE GRAND
CAFE DE LA PAIX
Because you enjoy going first class.
In Paris or at home, life's more satisfying when you're enjoying the best. That's Passport.
Enjoyed worldwide because it's made of Scotland's finest whiskies. Ask for Passport—go first class.
Passport Scotch.

Established 1845 SCIENTIFIC AMERICAN October 1979 Volume 241 Number 4

Acid Rain

In recent decades the acidity of rain and snow has increased sharply over wide areas. The principal cause is the release of sulfur and nitrogen oxides by the burning of fossil fuels

by Gene E. Likens, Richard F. Wright, James N. Galloway and Thomas J. Butler

Measurements of the acidity of rain and snow reveal that in parts of the eastern U.S. and of western Europe precipitation has changed from a nearly neutral solution 200 years ago to a dilute solution of sulfuric and nitric acids today. In the most extreme example yet recorded—a storm in Scotland in 1974—the rain was the acidic equivalent of vinegar (*p*H 2.4). The main reason for this trend is the rise in the emission of sulfur and nitrogen oxides to the atmosphere accompanying the rise in the burning of fossil fuels. Partly because acid precipitation has a variety of harmful effects on plant and animal life, efforts have been made to curb the emission of such oxides. Notwithstanding these efforts the level of emissions is expected to rise further under the present policy of turning increasingly to coal as a source of energy. At the very least the prospect suggests that a more substantial effort should be made to determine more precisely the effects of acid precipitation; it might also suggest a more vigorous exploration of ways to reduce the release of both sulfur and nitrogen oxides into the atmosphere.

Acidity in a solution such as rain is synonymous with the presence of hydrogen ions (H^+). A common measure of acidity is *p*H, which is defined as the negative logarithm of the hydrogen-ion concentration. The *p*H scale ranges from 0 to 14, with a value of 7 representing a solution that is neutral, values below 7 indicating greater acidity and values above 7 indicating greater alkalinity. One should bear in mind that the *p*H scale is logarithmic, so that solutions of *p*H 6, 5 and 4 contain respectively one, 10 and 100 microequivalents of acidity (H^+) per liter (abbreviated μeq/1).

What determines the chemistry of rain and snow? The water comes from evaporation and transpiration (water vapor lost by plants) and is essentially distilled, or pure. Once the vapor reaches the atmosphere, however, it condenses on solid particles and soon reaches equilibrium with atmospheric gases. One of the gases is carbon dioxide (CO_2), and as carbon dioxide dissolves in the water carbonic acid (H_2CO_3) forms. Carbonic acid, being a weak acid, dissociates only slightly in distilled water, yielding hydrogen ions and bicarbonate ions (HCO_3^-). At normal concentrations and pressures of carbon dioxide in the atmosphere the *p*H of rain and snow would be 5.6.

Other substances reaching the atmosphere tend to shift the *p*H one way or the other. For example, small amounts of dust and debris are swept from the ground into the atmosphere by the wind. Soil particles are usually slightly basic, or alkaline, in distilled water and release into solution base cations (positive ions) such as calcium, magnesium, potassium and sodium (Ca^{++}, Mg^{++}, K^+ and Na^+), with bicarbonate usually the corresponding anion, or negative ion. Ammonia gas in the atmosphere originates largely from the decay of organic matter. The ammonium ion (NH_4^+) in rain and snow tends to increase the *p*H. In coastal areas sea spray may contribute significantly to the chemistry of precipitation. Here the important ions added are the ones most abundant in seawater, namely the cations Na^+ and Mg^+ (sodium and magnesium), the anion Cl^- (chloride) and to a lesser extent cations of calcium and potassium (Ca^{++}, K^+) and anions of sulfate (SO_4^{--}).

Gases such as sulfur dioxide (SO_2) and hydrogen sulfide (H_2S), which come from volcanoes and from other natural sources, can also alter the chemistry of precipitation locally. Sulfur dioxide and hydrogen sulfide are oxidized and hydrolyzed in the atmosphere to sulfuric acid. Nitrogen oxides are similarly converted into nitric acid. If these acids are present in significant quantities, they can acidify precipitation to below *p*H 5.6.

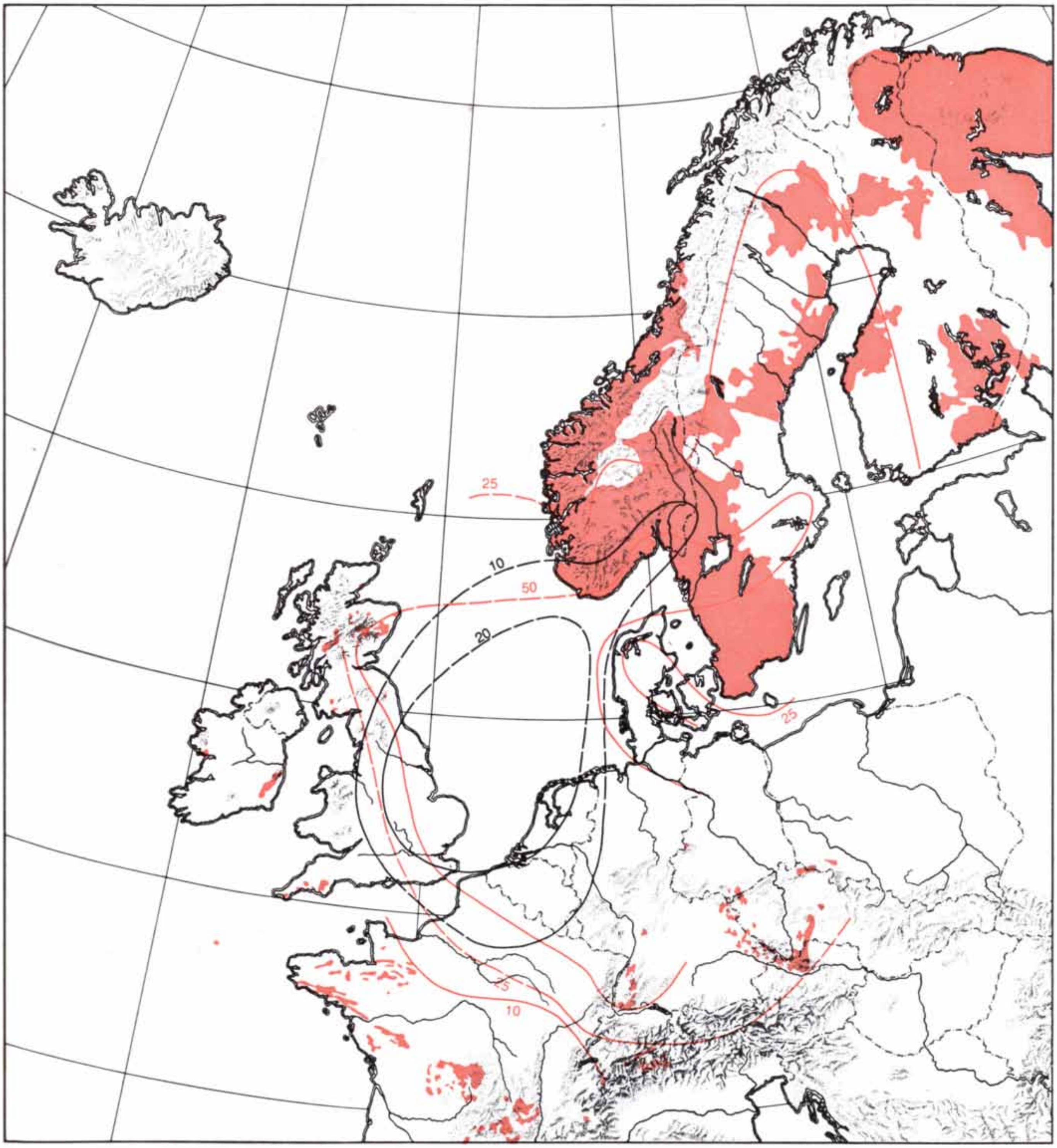
The chemistry of natural precipitation, then, depends on the relative amounts of these various substances in the atmosphere. The precipitation in coastal areas can be generally characterized as very dilute seawater if winds are toward the shore. In areas with high winds and sparse vegetation, such as deserts, the precipitation may contain relatively large amounts of base cations. Precipitation in regions with calcareous soils often contains calcium and bicarbonate, presumably as a result of the incorporation of dust, and the *p*H of rain and snow is usually well above 6.

The *p*H of precipitation that fell before the Industrial Revolution and has been preserved in glaciers and continental ice sheets is generally found to be above 5. Swiss workers in Greenland recently measured the *p*H of ice that had originated as snow some 180 years ago; it had a *p*H ranging from 6 to 7.6.

Human activities have significantly changed this picture on a continental scale and perhaps a global one. Large quantities of sulfur and nitrogen oxides are emitted into the atmosphere by the combustion of fossil fuels and the smelting of sulfide ores, particularly in the heavily industrialized and urbanized North Temperate Zone. These oxides are converted into strong acids (sulfuric and nitric) in the atmosphere and fall to the ground in rain and snow. Strong acids dissociate completely in dilute aqueous solutions and lower the *p*H to less than 5.6. The definition of acid precipitation is rain and snow at a *p*H below that value.

An irony in this matter is that the trend toward building taller stacks to relieve local pollution problems has turned the local problems into regional ones. In 1955 only two stacks in the U.S. were higher than about 180 meters; now essentially all stacks being built are taller than that. By 1975 at least 15 stacks taller than 300 meters had been built in the world. In the United Kingdom tall stacks were an integral part of measures employed to reduce the severity of smog in London after episodes of life-threatening pollution in the 1950's.

The copper-nickel smelter in Sudbury, Ont., has a superstack more than 400 meters tall. Some 1 percent of the total annual emissions of sulfur throughout the world (from both natural sour-

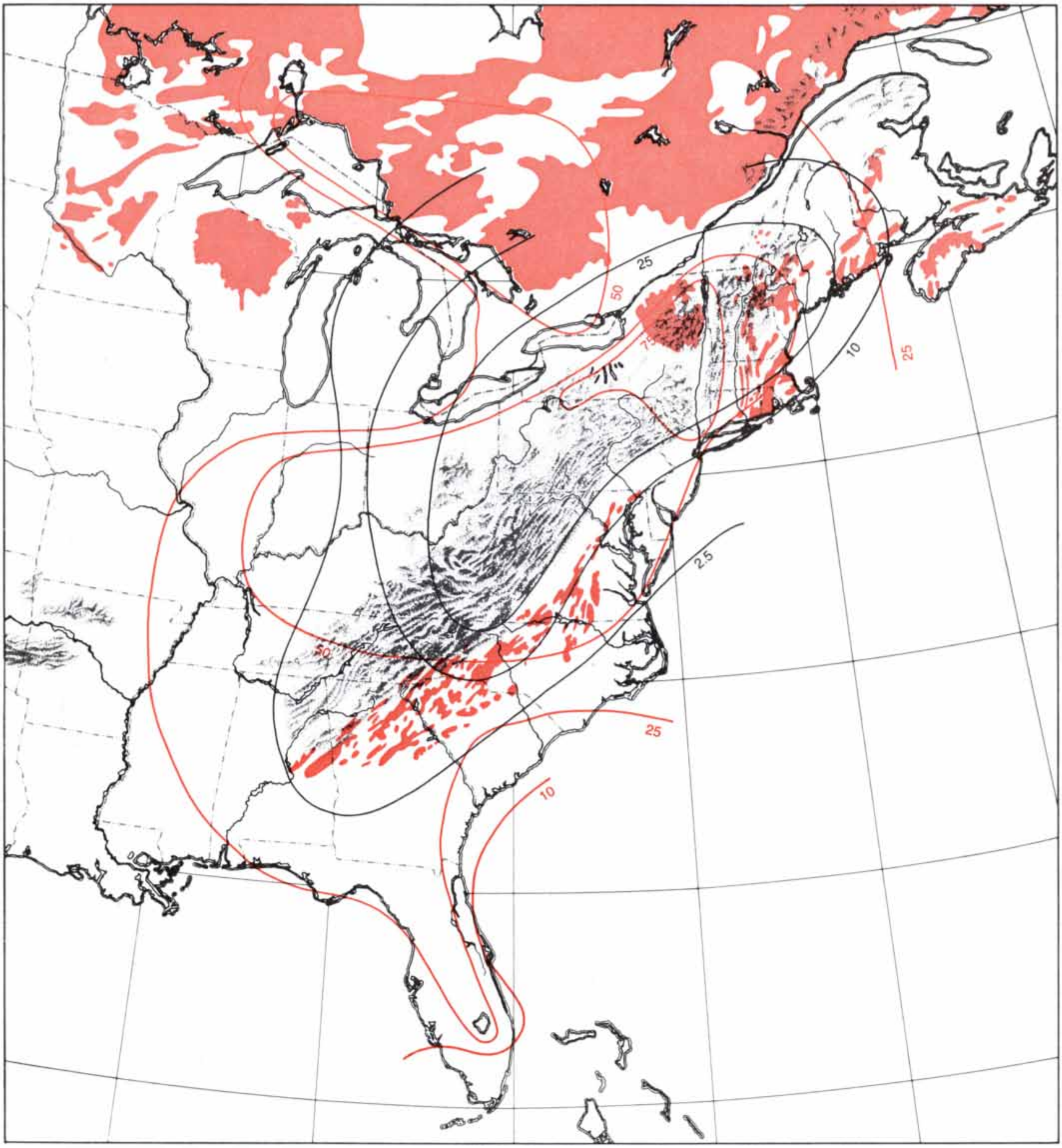


ACID PRECIPITATION in northwestern Europe is compared for 1956 (*black*) and 1974 (*color*). Acidity (the concentration of hydrogen ions in rain and snow) is shown by the numbers on the curving isolines; the larger the number, the higher the concentration of acid. The two sets of isoline numbers are based on volume-weighted, average concentrations of hydrogen ions in terms of microequivalents per liter of precipitation. The colored areas of the map indicate areas that are more susceptible to acid rain and snow because of the geology of their bedrock. In general they are areas with highly siliceous types of bedrock, among them granite, quartzite and quartz sandstone.

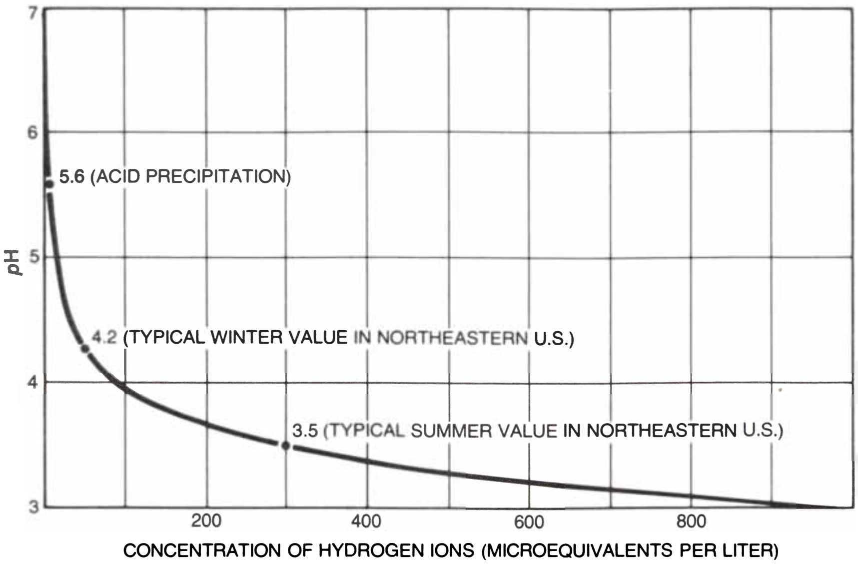
ces and human activity) come from this one smelter complex. During the past decade the annual emissions of sulfur from Sudbury have about equaled the amount thought to have been emitted by all the volcanoes of the world.

It is beyond argument that such pollutants are transported over long distances through the atmosphere. Studies of the deposition of sulfur, lead and other materials on the Greenland ice cap and in other remote places demonstrate such mechanisms of dispersal. What is a matter of sometimes heated discussion, particularly when it is alleged that pollution is crossing a political boundary, is the proportion of pollutants that can be attributed to remote sources.

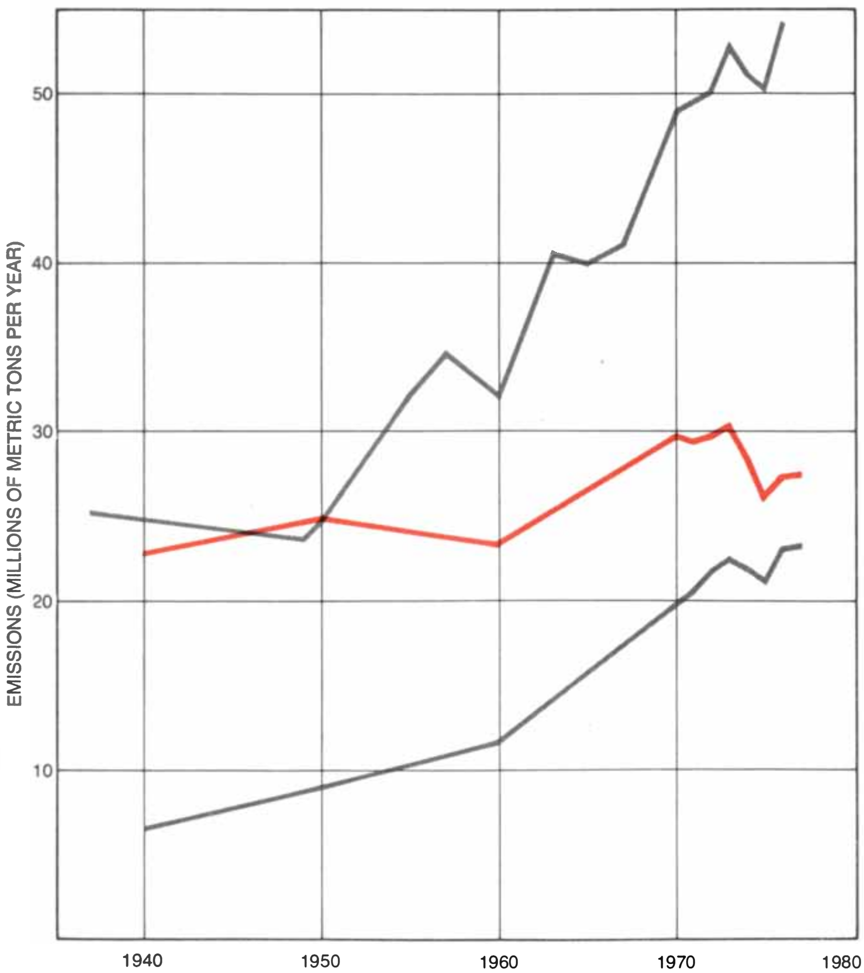
On an annual basis rain and snow over large regions of the world are now from five to 30 times more acid than the low-



SPREAD OF ACID RAIN in eastern North America is depicted for 1955–56 (*black*) and 1975–76 (*color*) with data for the isolines based on the concentration of hydrogen ions in microequivalents per liter. The quality of the data for North America is variable and the sampling methods employed for the two periods represented on this map were somewhat different, but the general pattern and trend are clear. Areas sensitive to acid rain and snow are also shown. Similar areas are found farther north in Canada and in the Rocky Mountains. They have a low ability to neutralize acids. Values higher than 2.5 microequivalents per liter (*p*H 5.6) are defined as acid precipitation.



ACID CONCENTRATION is measured by *p*H, which is the negative logarithm of the hydrogen-ion concentration, and more directly by the concentration of hydrogen ions expressed in microequivalents per liter. The relation between the two measures is depicted here for aqueous solutions. Since the *p*H scale is logarithmic, *p*H 6 is 10 times and *p*H 5 is 100 times more acidic than *p*H 7. Ten times more hydrogen ions are required to shift from *p*H 5 to 4 than from 6 to 5.



RISE IN EMISSIONS of sulfur and nitrogen oxides is charted. The curves reflect increases due to human activity, namely the burning of fossil fuels and the smelting of ores. From bottom to top the curves respectively trace the rise of nitrogen oxides in the U.S., sulfur dioxides in the U.S. and sulfur dioxides in Europe, excluding the U.S.S.R. Sulfur and nitrogen may be converted into strong acids in precipitation. The increase of acids parallels the rise in emissions.

est value expected (*p*H 5.6) for unpolluted atmospheres. The rain of individual storms can be from several hundred to several thousand times more acid than expected. Concern over the environmental effects of acid precipitation has led to major research efforts in Sweden, Norway, the U.S. and Canada.

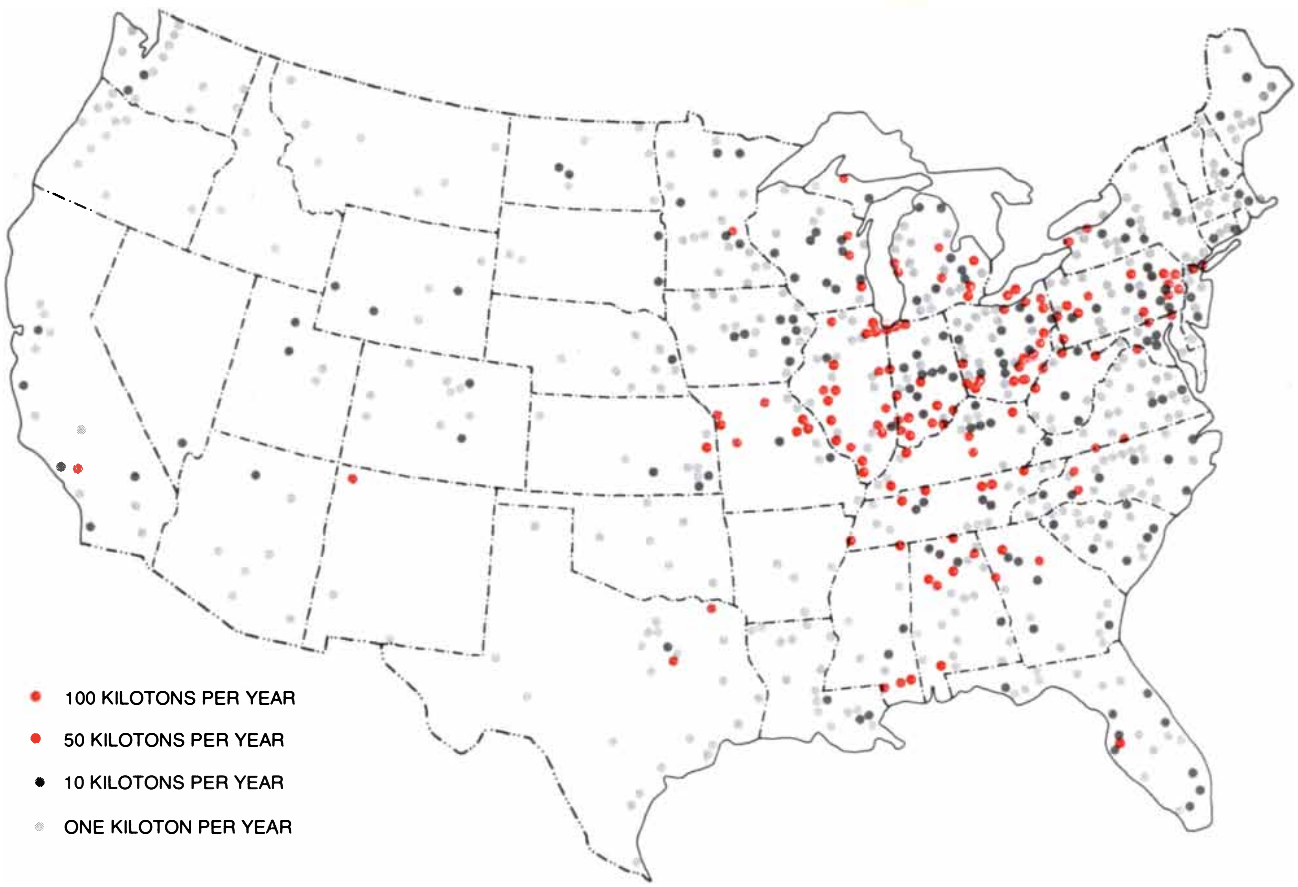
Acid precipitation has been known for many decades in the vicinity of large cities and industrial plants such as smelters, but the phenomenon is now much more widespread. In large areas of the eastern U.S., southeastern Canada and western Europe the annual average *p*H of precipitation ranges from about 4 to 4.5. Since the data are fuller for Europe than for North America, we shall discuss Europe first.

The discovery that precipitation in large regions of Europe had become decidedly acidic followed the establishment in the 1950's of the European Atmospheric Chemistry Network (also known as the IMI Network, for the International Meteorological Institute in Stockholm). At times the network encompassed up to 175 stations in northern and western Europe, from which monthly samples of bulk precipitation (a method of collecting with sampling apparatus that is always open) were analyzed for the concentration of 11 major ions. Although many of the stations are no longer in operation, the data provide a record for the *p*H of precipitation on a regional scale in Europe more than 20 years ago.

In 1968 Svante Odén of Sweden pointed out that precipitation in Europe had clearly become more acidic since the network was established. Maps of volume-weighted mean annual *p*H in precipitation showed that a central region of high acidity (*p*H 4 to 4.5) had spread from the area of Belgium, the Netherlands and Luxembourg in the late 1950's to most of Germany, northern France, the eastern British Isles and southern Scandinavia by the late 1960's.

The Organization for Economic Cooperation and Development (OECD) recently concluded a three-year study that gives a more detailed picture of the present acidity of precipitation in Europe. The results of the study appear in *Long-Range Transport of Air Pollutants, Measurements and Findings,* a summary of the daily sampling done at 67 stations from 1973 to 1975 as part of the OECD project. The resulting map shows that acid precipitation has continued to spread and today encompasses nearly all of northwestern Europe. (The situation in eastern Europe is unclear for lack of long-term data; clarification may be at hand from data from a new network for precipitation chemistry initiated by the United Nations' Economic Commission for Europe.)

The lowest *p*H value reported for an individual storm in Europe up to now is



PATTERN OF EMISSIONS of sulfur dioxide in the U.S. in 1975 is plotted on the basis of data obtained by Brookhaven National Laboratory. The plots indicate the areas where emissions are concentrated because of the activity of generating stations and industrial plants.

2.4 on April 10, 1974, at Pitlochry, Scotland. In the same month values of 2.7 and 3.5 were reported respectively on the west coast of Norway and at a remote station in Iceland. The lowest annual *p*H yet recorded is 3.78 in 1967 at De Bilt in the Netherlands.

At some stations the increase in the annual concentration of hydrogen ions in precipitation since 1955–56 has been about tenfold. This increase in acidity is matched by increases in the concentrations of sulfate and nitrate over the past 20 years. In northern Europe the excess sulfate concentration (in excess of contributions of sea salt) rose markedly from 1955 to about 1963 and has remained virtually constant since then; overall the rise from 1955 to 1975 was about 2 percent per year. The concentration of nitrate in precipitation has increased quite steadily since 1955 at the rate of about 4 percent per year. Both increases parallel rises in the emission of sulfur and nitrogen oxides from the combustion of fossil fuels.

Emissions of sulfur dioxide in Europe were estimated at about 25 million metric tons of sulfur in 1973. A map showing the regional sources was prepared as part of the OECD study; it revealed regions with particularly high emissions in central Britain, the Ruhr area of West Germany, the coal areas of southern East Germany, southern Poland and Czechoslovakia. It is estimated that human activity accounts for about 90 percent of the sulfur emissions.

It is also estimated that the annual anthropogenic emissions of nitrogen oxides from the 11 OECD countries of Europe totaled about two million metric tons of nitrogen in 1973. The OECD study indicated that these emissions almost doubled from 1959 to 1973. A similar trend might be expected for eastern Europe. In addition biological processes supply oxides of nitrogen to the atmosphere, but the magnitude and significance of these emissions are not known.

The general pattern of the deposition of acid in European rain and snow reflects the long-range transport of air pollutants from the sources of emission. The *p*H of precipitation is generally lowest near the sources and increases with distance. The normal flow of air from southwest to northeast results in an asymmetrical pattern, with acid precipitation reaching far up into northern Scandinavia. Swedish investigators have calculated that more than 70 percent of the sulfur in the atmosphere over southern Sweden is from human activity, and 77 percent of it is believed to originate from sources outside Sweden.

The gradient of acidity is accentuated in mountainous areas that are generally downwind of urban and industrial sources. Mountains enhance precipitation, which continuously removes pollutants from passing masses of air. This situation is found in southernmost Norway, where air masses from the major source areas in the British Isles and central Europe travel many hundreds of kilometers over the North Sea and over land areas of relatively low relief before most of the pollutants are removed in the precipitation over a belt from 100 to 200 kilometers wide along the mountainous Norwegian coast. By the time the air masses have passed over mountainous areas with an elevation of 1,000 meters or so the *p*H of the precipitation has changed from an average of 4.2 at the coast to 4.6 or 4.7, representing a decrease in the concentration of hydrogen ions from about 60 microequivalents per liter to about 20.

Since mountains enhance precipitation, their windward flanks receive large amounts of acid. Data from the OECD project show that the annual deposition of acid in rain and snow in 1974 was highest along the mountain slopes of southernmost Norway, southern Germany and Switzerland. Values ranging from 88 milliequivalents per square meter per year to 120 were observed in those areas.

Dry deposition of sulfur dioxide gas and particulate sulfate also represents the deposition of an equivalent amount of hydrogen ions. Data from the OECD study show maximum values of dry deposition of 10 grams of sulfur per square meter per year in central Europe, or four times the deposition rate by rain and snow, with rates decreasing rapidly with increasing distance from the emission sources to values of about .1 gram in northern Scandinavia. Wet deposition of sulfate in rain and snow begins at three grams of sulfur per square meter per year near the major sources and decreases less rapidly to .5 gram in northern Scandinavia. Dry deposition of sulfur is therefore the dominant mode near the emission sources, whereas wet deposition dominates in remote areas of northern Europe.

The record of changes in precipitation chemistry over North America is much less clear. No monitoring network has been maintained from the 1950's to the present. Even today the monitoring that is done is not sufficient to give a detailed picture.

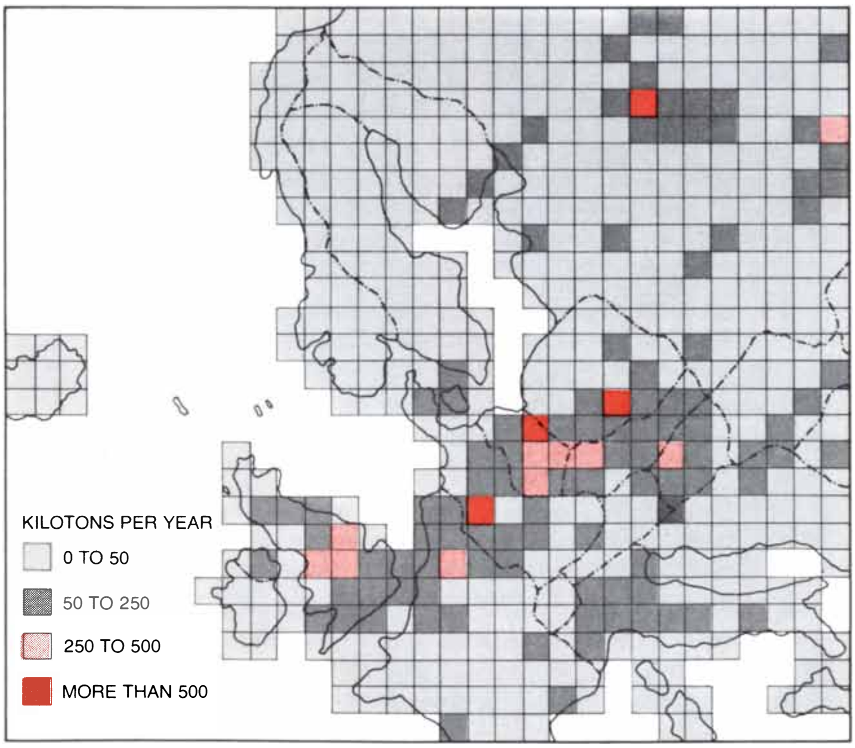
It is known that before about 1930 relatively large amounts of bicarbonate were found in samples of precipitation in Virginia, Tennessee and New York. Although no determinations of *p*H were made, the presence of bicarbonate in the samples shows that they could not have had a *p*H of less than 5.6. The earliest measurement of precipitation *p*H in the U.S. that we know of was made during a rainstorm in Maine in 1939 by Henry G. Houghton of the Massachusetts Institute of Technology. He obtained a value of 5.9.

If accurate chemical data are available and the hydrogen-ion concentration is more than about three microequivalents per liter, reliable *p*H values can be calculated from a detailed chemical analysis of dilute aqueous solutions. Such chemical data were collected for precipitation at 24 sites in the eastern U.S. in 1955–56 by C. E. Junge of the Air Force Cambridge Research Laboratory. The data have been employed to calculate the *p*H distribution at the time. Although the density of sampling points was low, the general pattern is quite clear: a large area of the northeastern U.S. was receiving acid precipitation by 1955.

From 1959 through 1966 the U.S. Public Health Service and then the National Center for Atmospheric Research operated a sampling network for precipitation chemistry throughout the 48 contiguous states of the U.S. During the latter part of the period (1964–66) some 30 stations, 17 of them east of the Mississippi, collected monthly samples of rain and snow with automatic samplers that were open only during precipitation. The majority of the stations west of the Mississippi consistently showed monthly *p*H values above 5.6; most commonly the values were above 6.5. East of the Mississippi the precipitation was much more acidic. Notwithstanding the paucity of the data from this now discontinued network, the general pattern was quite obvious: areas of acid precipitation were widespread in the eastern U.S., with the most acidic rain and snow falling on New England, New York and Pennsylvania.

Even today there is no high-density, synoptic coverage of the chemistry of precipitation in North America. The situation may soon improve, however, because both the U.S. and Canada have started long-term programs for measuring the composition of atmospheric deposition. The Canadian Network for Sampling Precipitation (CANSAP) began operating in 1976 and the U.S. National Atmospheric Deposition Program (NADP) was initiated in 1978.

In an effort to obtain synoptic data on the current *p*H of precipitation in the eastern U.S. we personally communicated with all the individuals, groups and state and Federal agencies we knew to be studying precipitation chemistry. Although the density of sampling points is still low (46 stations east of the Mississippi, 19 more in Canada), it is much higher than it was in 1955–56 or 1964–66. The data reveal a southward and westward extension of the area receiving acid precipitation and an intensification of acidity in the Northeast and Southeast since 1955. (We wish to acknowledge in this context the contributions from Gerald M. Aubertin and J. Currier of data for West Virginia; Patrick L. Brezonik for Gainesville, Fla.; T. Burton for Tallahassee, Fla.; D. Charles for Ray Brook, N.Y.; K. G. Dickson for the Lake Okeechobee area in Florida; M. J. Duever for Corkscrew Swamp in Florida; the Environmental Data and Information Service for Caribou, Me., Wake County and Macon County in North Carolina, Lauderdale County in Mississippi, Salem, Ill., and Atlantic County in New Jersey; Donald F. Gatz for Urbana, Ill.; Bruce L. Haines for Athens, Ga.; Evelyn B. Haines for Sapelo Island, Ga.; G. S. Henderson for Oak Ridge, Tenn.; Jay S. Jacobson for Yonkers, N.Y.; K. Kuntz for Ontario; J. G. McColl and David F. Grigal for Minnesota; J. M. Miller for Washington, D.C.; M. P. Olson for Ontario and Sable Island in Canada; Rosa de Pena for State College, Pa.; C. J. Richardson for Michigan; W. T. Swank for Asheville, N.C.; the U.S. Geological Survey for sites in the Northeast, and Herbert L. Vol-



EMISSION PATTERN of sulfur dioxide for Europe in 1973 is plotted in terms of thousands of tons of sulfur per year. The map is based not on particular "hot spots," as the companion map of the U.S. is, but on a grid that divides Europe into squares about 127 kilometers on a side.

chok for Argonne, Ill., New York, N.Y., Chester, N.J., and Woods Hole, Mass.)

It is worth noting that the most rapid increase in precipitation acidity over the past two decades has been in the Southeast. (The trend is more apparent than it is in the Northeast because a smaller amount of acid is required to cause a change of *p*H from, say, 6 to 5 than from 5 to 4.) In the same period the Southeast has had a large expansion of the urban and industrial activities that give rise to emissions of sulfur and nitrogen.

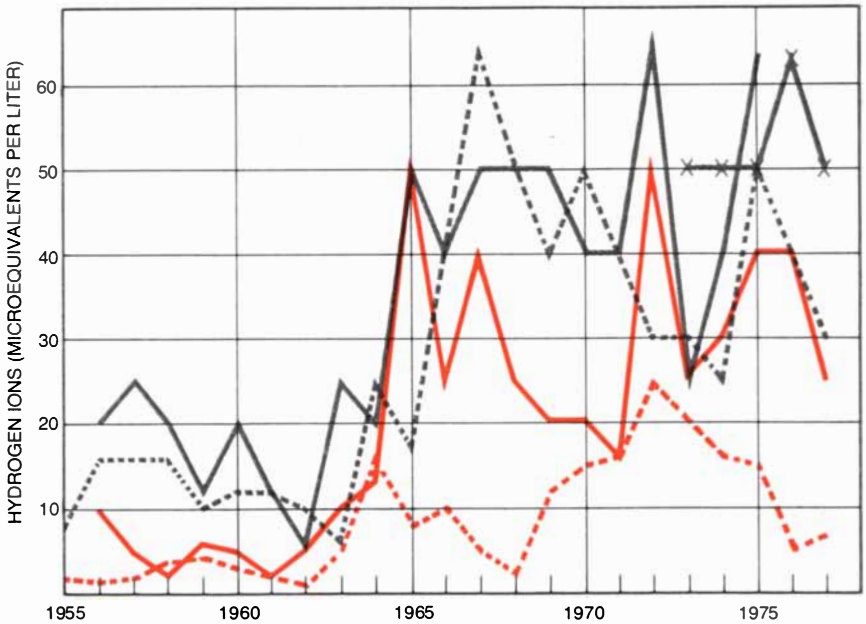
As in Europe, nitrate concentrations in precipitation have increased severalfold in the northeastern U.S. during the past decade or so. Although sulfuric acid is the dominant acid in acid precipitation, nitric acid appears to be contributing about 30 percent (and an increasing proportion) of the hydrogen ions in acid precipitation.

Measurements of the dry deposition of sulfur in the U.S. and Canada are highly scattered. A recent report by the National Academy of Sciences on sulfur oxides suggests that the dry deposition of sulfur is approximately equal to the wet deposition in the eastern U.S.

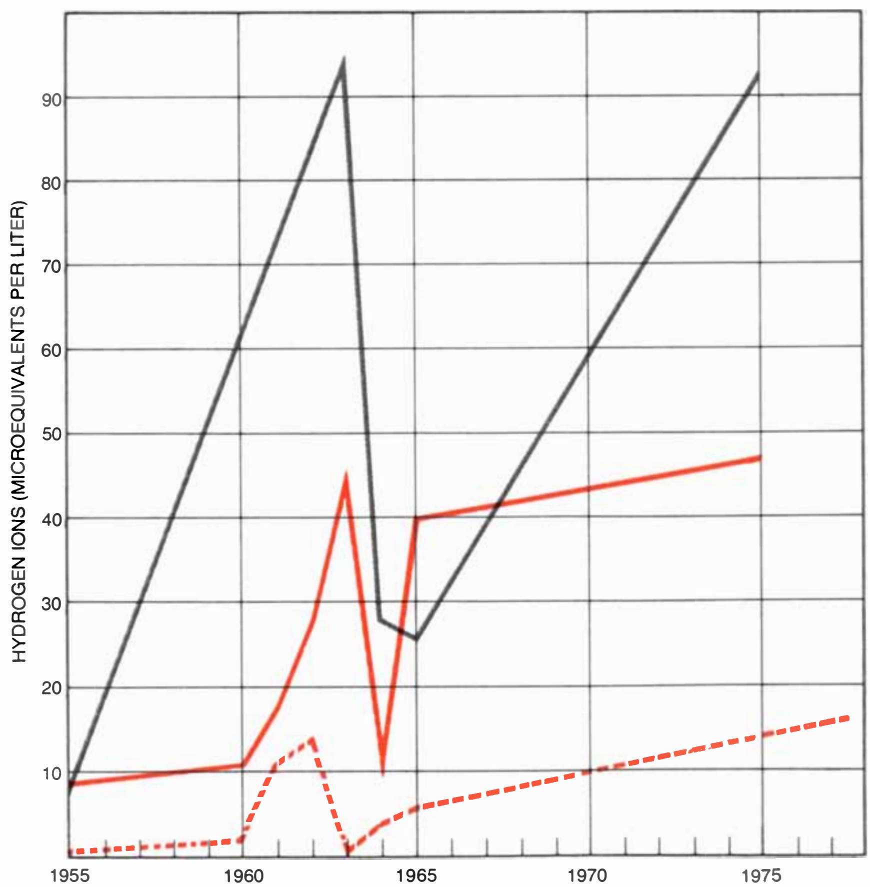
Since there are large sources of emissions in the middle-western U.S. and the general flow of air is from west to east, it is understandable that acid precipitation is a phenomenon of eastern North America. The phenomenon could become national, however, in view of the planned increases in the combustion of fossil fuels in the western regions of the continent.

What effect has the increased acidity had on the environment? The question has not been fully answered, but certain freshwater ecosystems seem to be particularly sensitive to inputs of acid. Such ecosystems are generally in areas underlain by highly siliceous types of bedrock such as granite, some gneisses, quartzite and quartz sandstone. These rock types are highly resistant to dissolution through weathering, with the result that surface waters in such areas contain very low concentrations of ions derived from weathering. Characteristically such soft waters have a low buffering capacity (a low ability to neutralize additions of acids or alkalis). Hence when acid precipitation falls on such an area, the acids are not fully neutralized in the terrestrial watershed or catchment, and so streams and lakes become acidified.

Bedrock geology therefore provides a rough indication of the location of such sensitive areas. Maps of the bedrock, however, do not take into account the mineralogy of overlying, unconsolidated deposits such as glacial till, glaciofluvial materials and marine sand, which may have been transported far from their source and may have a mineralogy



ACIDITY OF RAIN AND SNOW in recent years is charted for Lista in Norway (*solid gray*), Ås in Norway (*broken gray*), Kise in Norway (*solid color*) and Kiruna in Sweden (*broken color*). The acidity is expressed in terms of the annual volume-weighted concentration of hydrogen ions in microequivalents per liter. Most of the data come from samples collected by the European Atmospheric Chemistry Network. The additional curve for Lista from 1973 to 1977 reflects data from the Organization for Economic Cooperation and Development (OECD).



LOCAL TRENDS in acid rain and snow are depicted for Washington, D.C. (*solid gray*), Nantucket, Mass. (*solid color*), and Tampa, Fla. (*broken color*). The 1977 value for Florida is from Bradenton and the 1975–76 value for Massachusetts is from Woods Hole; each place is within about 40 kilometers of the original sampling site. Ranges shown for two periods in Washington represent values from different sampling stations in or adjacent to the metropolitan area.

	UNPOLLUTED INLAND (NORTHERN SWEDEN) 1956 *p*H 5.4	UNPOLLUTED COASTAL (WESTERN IRELAND) 1967 *p*H 5.8	POLLUTED INLAND (SOUTHERN NORWAY) 1974–75 *p*H 4.32	POLLUTED COASTAL (SOUTHERN NORWAY) 1972 *p*H 4.15	POLLUTED INLAND (NEW HAMPSHIRE) 1973–74 *p*H 3.94
HYDROGEN (H^+)	4	1.6	48	71	114
SODIUM (Na^+)	7	930	9	480	6
POTASSIUM (K^+)	3	24	3	15	2
CALCIUM (Ca^{++})	35	76	8	45	10
MAGNESIUM (Mg^{++})	8	216	4	132	32
AMMONIUM (NH_4^+)	0	1	21	48	22
BICARBONATE (HCO_3^-)	6	21	0	0	0
CHLORINE (Cl^-)	10	1,026	11	574	12
SULFATE (SO_4^{--})	29	131	52	164	110
NITRATE (NO_3^-)	0	1	26	54	50
TOTAL CATIONS	57	1,249	93	791	164
TOTAL ANIONS	45	1,179	89	792	172

CHEMISTRY OF PRECIPITATION is portrayed for five areas in terms of the concentration (in microequivalents per liter) of 10 ions. Positive ions are cations, negative ones anions. Hydrogen, sulfate, nitrate and chloride contribute to acidity (as when hydrogen and sulfate combine to form sulfuric acid); the other ions tend to neutralize an aqueous solution. In coastal areas seawater contributes many ions.

quite different from that of the bedrock. Local variations in the chemistry of surface waters then may result from differences in these two geological substrates, in soil type and in vegetation.

Areas of highly siliceous bedrock are widespread on the Precambrian Fennoscandian Shield in Scandinavia, the Canadian Shield, the Rockies, New England, the Adirondacks, the Appalachians and smaller areas elsewhere. Fresh waters acidified to below *p*H 5 as a result of acid precipitation are found in southern Norway, southern Sweden, Britain, the Netherlands, East Germany, Nova Scotia, Ontario, northern New England, the Adirondacks and Florida. One clear effect has been on the fish populations, which have been virtually wiped out or severely reduced in acidified lakes in some of these regions.

For example, in southern Norway the acidification of thousands of freshwater lakes and streams has affected fish populations in an area of 33,000 square kilometers. In addition several other ecological changes have been observed as a result of the acid stress in such lakes. The rate of decomposition of organic matter is slowed, presumably because of changes in the predominance of bacteria and fungi. Bacteria, which are efficient decomposers, generally do worse in an acid environment. Changes occur at all levels of the food web: phytoplankton, zooplankton and fishes, all of which decrease in number of species. Even the behavior of organisms may change because of changes in their environment.

The distribution of acid lakes is congruent with the distribution of acid precipitation. Further evidence of the impact of acid precipitation is that acid fresh waters are normally not found in regions that would be sensitive to acid (being nutrient-poor and poorly buffered) but lie outside the regions of highly acid precipitation. These observations lead to the conclusion that both the amount of acid in precipitation and the geological characteristics of the area are important in determining the sensitivity of fresh waters to acid precipitation.

Working with water-chemistry data from more than 1,000 lakes in Norway, Arne Henriksen of the Norwegian Institute for Water Research has recently developed an empirical relation by which the sensitivity of a given freshwater system to inputs of acid precipitation can be estimated. For example, he predicts that a calcium bicarbonate soft-water lake of the kind found by the thousands in North America and northern Europe with about 80 microequivalents per liter of bicarbonate (corresponding to about 1.7 milligrams of calcium per liter and a *p*H of about 6.5) will lose its bicarbonate buffer to the point where the *p*H drops below 5 and fish are eliminated if the long-term average *p*H of precipitation is below about 4.3. It would take precipitation even less acidic to give rise to similar conditions in lakes with originally lower concentrations of bicarbonate.

Summer rains are appreciably acid in large areas of North America and Europe, particularly the first part of a rainstorm. They are certainly within the range of acidity where laboratory studies indicate that harmful effects could be expected on growing vegetation if it were exposed and sensitive. It is difficult, however, to isolate the effects of one environmental stress from all the others affecting vegetation, and so the overall effect of acid precipitation on plants has to be regarded as ambiguous in the light of the present state of knowledge.

A natural ecosystem is extremely complex, with numerous interactions at all levels of organization. Evaluating the effects of acid precipitation on such a system is difficult and expensive. The social, economic and political aspects of acid precipitation are particularly vexing because air, rain and snow are not confined by political boundaries. For example, Norway is now completing a $10-million study of the effects of acid precipitation on its forests and lakes. In the U.S. a recent report by the National

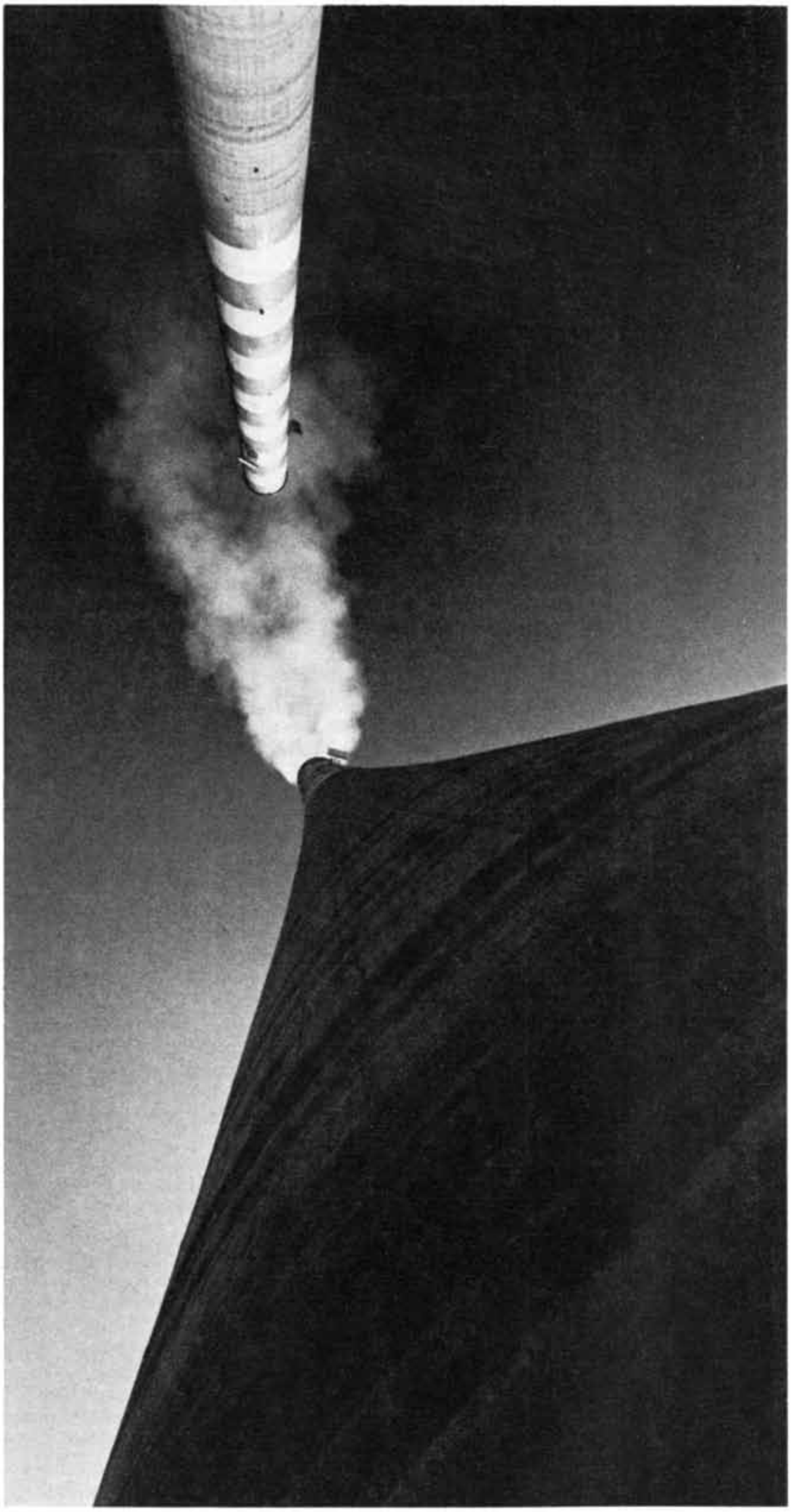
Atmospheric Deposition Program to the President's Council on Environmental Quality (CEQ) concluded that a program of research costing $9 million per year will be necessary to evaluate the phenomenon of acid precipitation and its effects.

It seems unlikely that the combustion of fossil fuels will decline in the industrialized countries over the next 30 years or so. It is more likely to increase because of the increased reliance on coal as a replacement for oil. The question of the broad environmental effects must be faced. In the U.S. the President's Committee on Health and Environmental Effects of Increased Coal Utilization has identified acid precipitation as one of the two major global environmental problems, the other being the increased emissions of carbon dioxide with their potential effects on climate.

In the U.S. the Environmental Protection Agency (EPA) has ruled that from 70 to 90 percent of the gaseous sulfur must be removed from the emissions of all new coal-based power plants. Old plants remain largely uncontrolled. The control in new plants would be achieved by the installation of scrubbers: flue-gas desulfurization units.

Even so, the EPA predicts that emissions of sulfur dioxide by utilities, which are by far the largest single source in the U.S., will rise from 18.6 million metric tons in 1975 to between 20.5 (with conservation and the best available controls) and 23.8 million (under existing regulations) in 1995. Although such projections are dependent on a variety of known and unknown factors and hence are tenuous, with increased emissions of sulfur dioxide (and nitrogen oxides) it can generally be expected that the deposition of hydrogen ions with precipitation will increase.

The installation of scrubbers is expensive, but the environmental costs of inadequately controlled emissions are also large. Many costs are hidden, particularly the ones associated with the synergistic effects of the stresses that acid puts on organisms (for example, increasing their susceptibility to pathogens and predators). The easy solutions, such as burning low-sulfur coal and oil, have already been applied in the control of emissions in some areas. Nevertheless, in some areas the *p*H of rain and snow has already been pushed through the "easy" part of the scale (from 7 to 4). Reductions in acidity will now be more difficult and far costlier to effect. Ultimately the solution to the acid-precipitation problem lies in the decrease in emissions of sulfur and nitrogen oxides to the atmosphere. This objective can be achieved with a combination of improved control technology and the conservation of fuel.



TALL STACKS of the generating station of the Tennessee Valley Authority in Cumberland City, Tenn., reach a height of 305 meters (1,000 feet). They typify a widespread trend to build taller stacks to reduce the concentration of sulfur and nitrogen oxides near the ground, since that is where the Environmental Protection Agency (EPA) and similar state agencies make measurements of such concentrations. By putting the oxides higher into the atmosphere, however, the stacks cause them to be carried farther, turning a local problem into a regional one.

Linear-Chain Conductors

A few materials with a linear or columnar architecture conduct electricity well only along a single axis. The one-dimensional organization of such a material shapes its electronic properties

by Arthur J. Epstein and Joel S. Miller

Experience with common substances teaches one to associate certain properties with general classes of materials. For example, metals such as copper, iron and nickel, as well as alloys of metals, are good conductors of electricity. Ionic materials such as table salt, molecular solids such as sugar and polymeric solids such as Teflon and polystyrene are generally poor conductors. In recent years, however, a number of materials have been created that confound these expectations. They are ionic crystals, molecular solids and polymers that mimic some aspects of the electronic structure of a metal and therefore exhibit certain metal-like properties. In particular they are good conductors of electricity.

The most remarkable property of these materials is their anisotropy: their conductivity is different when it is measured along different directions in the solid. In most cases there is one axis along which the conductivity is high, sometimes approaching that of the elemental metals. Perpendicular to that axis the conductivity can be smaller by a factor of 100,000. This anisotropy reflects the underlying structure of the materials: they are made up of many parallel chains or stacks of molecules, and they conduct well only along the chains or stacks. They are often called quasi-one-dimensional solids or linear-chain materials.

The electrical, magnetic, optical, structural and chemical properties of these materials are all of interest, but we shall discuss mainly their electrical conductivity. In some cases the mechanism of conduction is much like that in an elemental metal, or at least it is metal-like above a threshold temperature. At lower temperatures the electrons in these materials undergo a phase transition and the materials become semiconductors. Other linear-chain solids are semiconductors or insulators at all practical temperatures. Still others have properties that can be described adequately only by more elaborate models of conduction.

Not all the highly conductive linear-chain solids are recent discoveries. In 1842 W. Knop prepared a salt of tetracyanoplatinate, a chemical group in which four cyanide ions (CN^-) surround a platinum ion. The water-soluble salt had the color and sheen of gold or bronze. In 1910 Frank Playfair Burt of University College London synthesized the inorganic polymer poly(sulfurnitride), which consists of alternating sulfur and nitrogen atoms bound together in long chains. The chemical formula is $(SN)_x$, where x stands for some large but indeterminate number. Both of these substances form highly conductive linear-chain solids. In $(SN)_x$ the long polymeric strands line up parallel to one another. The tetracyanoplatinate group has the form of a flat disk, but in the salt prepared by Knop the disks are piled like poker chips in many parallel stacks. This structure was first identified in 1964 by Klaus Krogmann of the University of Stuttgart. Several other salts of tetracyanoplatinate have a similar structure, and we shall refer to all of them by the generic abbreviation TCP.

Although TCP and $(SN)_x$ have been known for many years, their properties were not examined in detail until the 1970's. By then other linear-chain systems had been prepared. A particularly important molecule was synthesized and characterized in 1960 by a group of investigators at E. I. du Pont de Nemours & Company. Its formal name is 7,7,8,8-tetracyano-*p*-quinodimethane, which is usually abbreviated TCNQ. It is another planar molecule, but unlike TCP it is an organic substance, being made up entirely of carbon, nitrogen and hydrogen atoms. Many salts in which TCNQ is combined with other atoms or molecules form linear-chain solids.

Part of the impetus for the study of linear-chain solids was provided in 1964 by W. A. Little of Stanford University. If a linear-chain material could be designed to the right specifications, he suggested, it might exhibit superconductivity not only at very low temperatures, as known superconductors do, but at room temperature. The structure Little proposed had a conducting spine with side groups that would stabilize the superconducting transport of electrons. A room-temperature superconductor would be of great technological importance, and so Little's idea inspired a major effort to synthesize such structures. No room-temperature superconductor has yet been discovered, however, and Little's proposal remains controversial.

In 1973 Alan J. Heeger and Anthony F. Garito and their co-workers at the University of Pennsylvania reported observing very high conductivity in a salt made up of TCNQ and another organic molecule, tetrathiafulvalene, or TTF. At comparatively low temperatures (between 50 and 60 degrees Kelvin) they found that the conductivity of (TTF) (TCNQ) approaches that of copper at room temperature. At first it was thought this highly conductive state might signal the onset of superconductivity, but it turned out otherwise.

What has been found is no less interesting. The columnar structure of the anisotropic conductors is essential to their nature and has a major influence on their properties. The mechanisms of conduction observed in these substances differ from the mechanism that operates in a three-dimensional elemental metal. Certain electronic phase transitions, which convert a metal into a semiconductor, are unique to one-dimensional materials. These phenomena are of considerable interest for what they reveal about the physics and chemistry of the solid state, and the materials themselves may ultimately have technological applications.

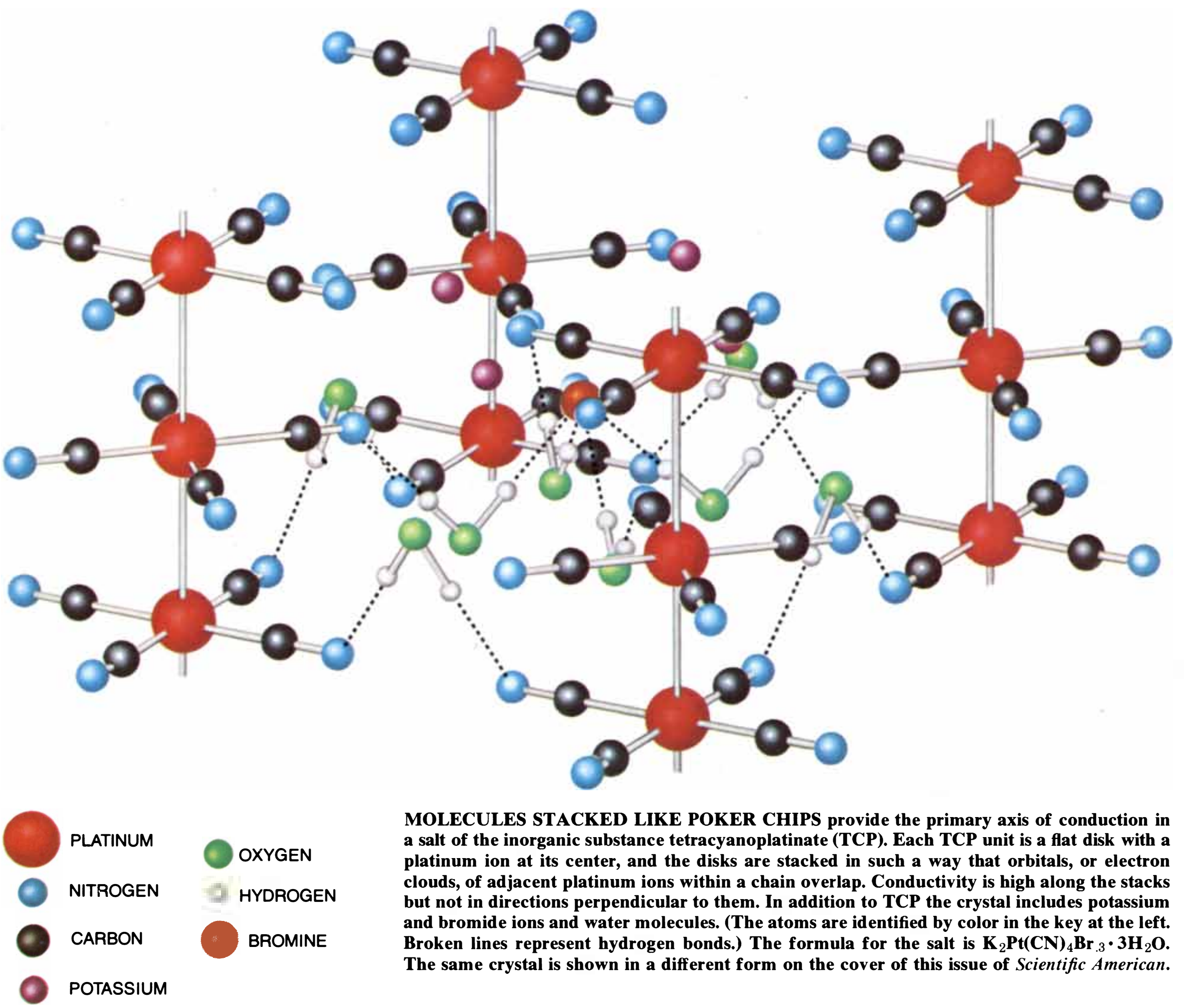
Electrical conductivity is a property of matter that has an enormous range of values. The unit of measurement is the mho per centimeter, which is the reciprocal of the unit of electrical resistivity, the ohm centimeter. ("Mho" was coined by Lord Kelvin as the backward spelling of "ohm," the unit named for Georg Simon Ohm.) The best insulators, such as Teflon and polystyrene, have conductivities of about 10^{-18} mho per centimeter; the best room-temperature conductors, copper and silver, approach 10^6 mhos per centimeter. Several linear-chain materials have conductivities (measured at room temperature along the favorable axis) that range from a hundred to a few thousand mhos per centimeter.

The conductivity of a solid is determined by its electronic structure. The distribution of the electrons in the vicinity of an atom is described by a system of orbitals, each of which has a characteristic shape and size. In a highly conductive solid orbitals of adjacent atoms or molecules overlap, so that electrons can readily move from site to site through the lattice. Linear-chain conductors are materials in which there is extensive overlapping of orbitals only along one axis. In the tetracyanoplatinate systems, for example, the orbitals in question are particular orbitals of the platinum ion, designated d_z2, which extend well above and below the TCP planes. In a salt of TCP the platinum ions are separated by some 2.88 angstrom units (one angstrom is equal to 10^{-8} centimeter), which allows the d_z2 orbitals of platinum ions in adjacent planes to overlap. For the purpose of comparison, the interatomic distance in metallic platinum, an excellent conductor, is 2.79 angstroms.

As a result of the orbital overlap, electrons are delocalized along the TCP columns. The movement of electrons from one column to another is much less efficient. A typical TCP salt, incorporating potassium and bromide ions as well as water, has the formula $K_2Pt(CN)_4Br_{.3} \cdot 3H_2O$. Hansrudi R. Zeller and his colleagues at the Brown-Boveri Research Center have reported that the room-temperature conductivity of this material parallel to the TCP stacks is about 300 mhos per centimeter. Conductivity between stacks is 100,000 times lower.

(TTF)(TCNQ) is made up of two kinds of flat molecules, which form parallel, conducting stacks. The molecules in each stack are tilted with respect to the axis of the stack, and adjacent columns are arranged to form a herringbone pattern. Within both the TTF and the TCNQ stacks orbitals of each molecule overlap those of the molecules above and below it in the same stack. The orbitals involved are associated with the molecule as a whole rather than with any particular atom. There is little contact between the orbitals of molecules in different stacks, and so charge transport perpendicular to the stacks is inefficient. The maximum room-temperature conductivity of (TTF)(TCNQ) parallel to the columns is higher than that of TCP: about 650 mhos per centimeter. The anisotropy of the material is smaller, however; conductivity in the most favorable direction is about 500 times the conductivity in the least favorable direction.

Many polymers are made up of molecules chemically bonded to form long

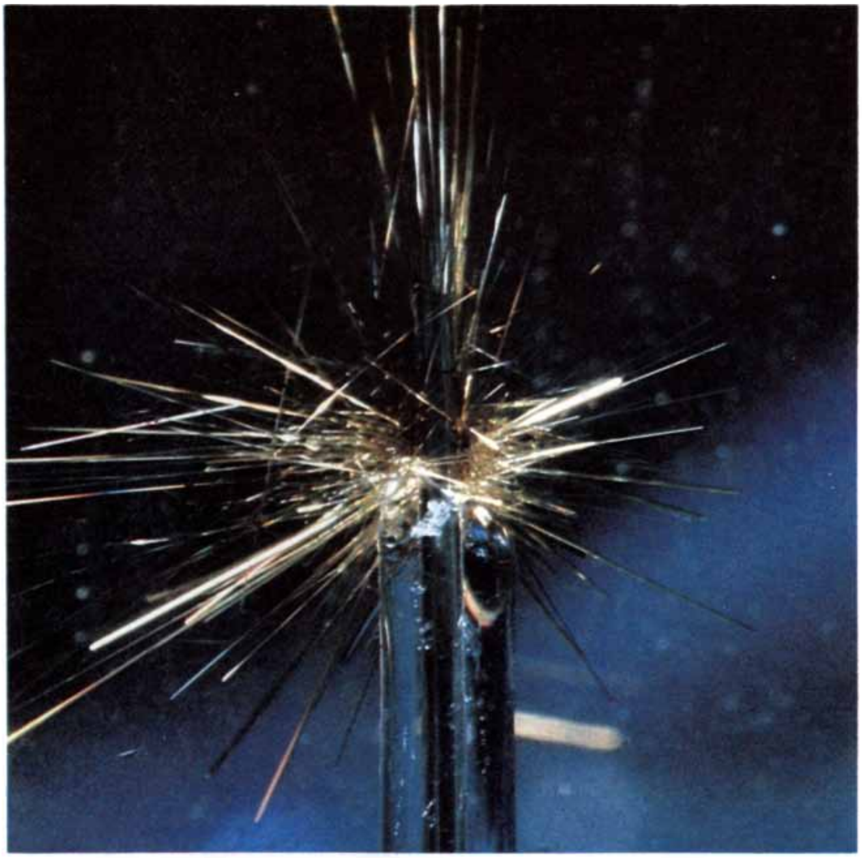


MOLECULES STACKED LIKE POKER CHIPS provide the primary axis of conduction in a salt of the inorganic substance tetracyanoplatinate (TCP). Each TCP unit is a flat disk with a platinum ion at its center, and the disks are stacked in such a way that orbitals, or electron clouds, of adjacent platinum ions within a chain overlap. Conductivity is high along the stacks but not in directions perpendicular to them. In addition to TCP the crystal includes potassium and bromide ions and water molecules. (The atoms are identified by color in the key at the left. Broken lines represent hydrogen bonds.) The formula for the salt is $K_2Pt(CN)_4Br_{.3} \cdot 3H_2O$. The same crystal is shown in a different form on the cover of this issue of *Scientific American*.

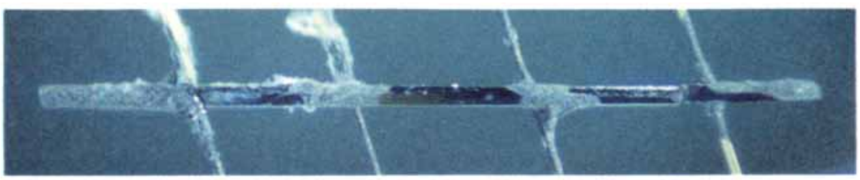
chains and therefore exhibit certain anisotropies. Most polymers, however, are poor conductors of electricity. An exception is $(SN)_x$, which has a room-temperature conductivity of about 2,000 mhos per centimeter. This high conductivity results from the overlapping of orbitals parallel to the polymer strands, creating a pathway for conduction throughout the length of the strand. Conductivity perpendicular to the strands is also comparatively high (that is, the anisotropy is small), indicating that there is also significant overlapping of orbitals between chains. $(SN)_x$ becomes a conventional, three-dimensional superconductor at temperatures below .3 degree K.

An organic polymer that can be given a high conductivity is polyacetylene, which has the formula $(CH)_x$. It is a continuous chain of carbon atoms, usually represented as having alternating single and double bonds connecting them, and with a single hydrogen atom attached to each carbon. By changing the concentration of dopants, or impurities, in the structure the conductivity can be adjusted over a wide range, from about 10^{-9} mho per centimeter to more than 2,000 mhos per centimeter.

One consequence of a linear structure is extreme sensitivity to defects and disorder. In a three-dimensional solid a point defect can reduce conductivity by scattering electrons, but it cannot completely block conduction because the electrons can detour around it. In a linear-chain material, on the other hand, an interruption of the chain can halt conduction entirely in that strand. If a large proportion of the strands are broken at some point along their length, conduction is seriously impaired.



NEEDLELIKE CRYSTALS of a highly conductive linear-chain material were grown electrochemically at the Webster Research Center of the Xerox Corporation. The material is a potassium salt of TCP with the formula $K_{1.75}Pt(CN)_4 \cdot 1.5H_2O$. Its metallic luster is characteristic of many one-dimensional conductors. The crystals are less than a millimeter long. Stacks of TCP disks are parallel to the needle axis, and so conductivity is highest in that direction.



MEASURING THE CONDUCTIVITY of a linear-chain conductor requires attaching four probes to a crystal 1.5 millimeters long. The two outer probes pass a known current through the crystal; the voltage developed across the inner probes is then measured. The voltage is proportional to the electrical resistance of the specimen. The reciprocal of the resistance is the conductance, and the conductivity is found by adjusting for the dimensions of the specimen. The crystal shown is a sample of (NMP)(TCNQ); it is mounted on gold wires .001 inch in diameter.

Because linear systems are highly susceptible to disruption, their intrinsic conductivity can be measured accurately only in crystals that are pure and crystallographically almost perfect. Special techniques have been devised for preparing such crystals. One of these methods, developed at the Webster Research Center of the Xerox Corporation, employs the property to be measured—the conductivity—to grow the crystal. When an electric current is passed through a solution of potassium ions and tetracyanoplatinate ions, crystals of TCP form on the anode. Specimens of the salt $K_{1.75}Pt(CN)_4 \cdot 1.5H_2O$ prepared in this way have a room-temperature conductivity some 100,000 times greater than crystals grown by more conventional means.

Single crystals of linear-chain solids often have a needlelike shape, with the conducting strands or columns parallel to the needle axis. The crystals are generally quite small: they are typically less than a millimeter long and several hundredths of a millimeter thick. In ordinary light many linear-chain conductors have a metallic luster. The inorganic substances that incorporate platinum are generally gold or bronze in color, and so is $(SN)_x$. Films of polyacetylene look like aluminum foil but have the texture and pliability of plastic food wrap. The organic materials, such as (TTF)(TCNQ), are usually black.

In order to explain the mechanism of electrical conduction it is necessary to discuss the state of the electrons in a solid. In an isolated atom each electron occupies a particular orbital, which has a certain energy associated with it. The orbital defines the distribution of the electron's charge in space. Each orbital can be occupied by no more than two electrons. In an atom in its ground state the first two electrons fill the lowest-energy orbital, and any additional electrons must go into higher-energy orbitals. Because the electrons are confined to orbitals they can have only discrete energy levels; all intermediate energies are forbidden. If an electron is to change its energy at all, it must jump from one allowed level to another.

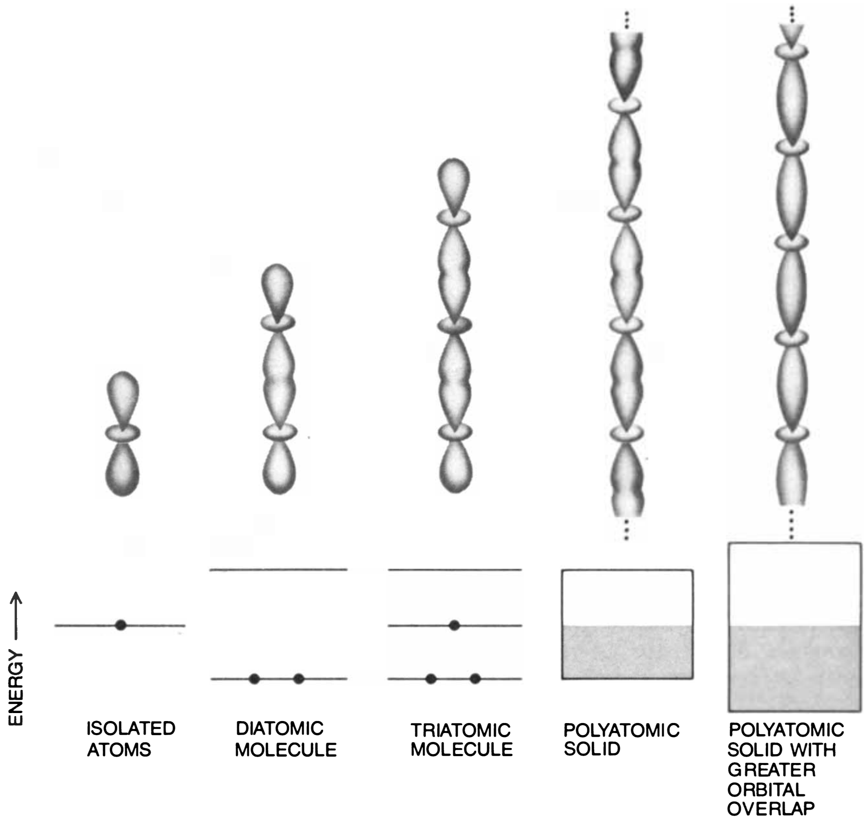
When two identical atoms are brought together closely enough for their orbitals to overlap, each energy level is split to create two new levels, one below the original level and the other above it. The magnitude of the splitting is determined by how extensively the orbitals overlap, or in other words by how closely the two atoms approach each other. It is important to note that the total number of electrons that can be accommodated by the system does not change when the two atoms coalesce to form a molecule. Whereas before there were two independent atomic orbitals of the same energy, each of which could hold two electrons, there are now two molecular orbitals with two different energy levels, and again each level can hold two electrons. If three atoms are brought together, three new energy levels are created, one below the original level, one above it and one between these two levels.

A three-dimensional crystalline solid brings together within the range of orbital overlap on the order of 10^{23} atoms. Each energy level of the original atoms is split into a corresponding number of finely spaced levels. Indeed, the interval between adjacent energy levels is so small that each set of levels can be regarded as a continuous band where an electron can take on any energy between the lower and the upper bound of the band. Since each of the levels that make up the band can hold no more than two electrons, in a solid made up of n atoms each band can be filled with no more than $2n$ electrons. If any more electrons are added, they must go into the lowest unoccupied band, sometimes at a substantial cost in energy.

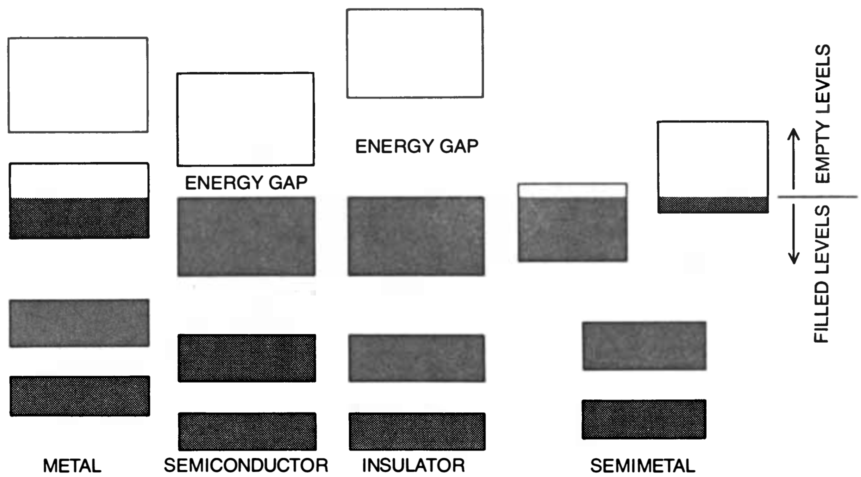
The extent of occupation of the energy bands and the magnitude of the gaps between them determine most of the electrical properties of a solid. The electrons in a band are said to be delocalized: they are free to wander throughout the solid. In order for an electric current to be sustained, however, there must be a net movement of electrons in one direction. If the energy levels in a band are filled continuously from the lowest available level to the highest allowed one, there can be no net motion of the electrons.

It is the presence of a partially filled electron-energy band that gives a metal its most distinctive properties. In a metal the electrons in the highest filled energy levels can readily be promoted to unoccupied levels that are only infinitesimally higher in energy. Thus electrons are always freely available for conduction. A partially filled band is created whenever the orbitals that combine to form the band are not fully occupied. For example, if each orbital is occupied by just one electron (instead of the maximum of two electrons), the corresponding energy band in the solid will have only n electrons rather than the maximum of $2n$. The remaining levels in the band are available for conduction.

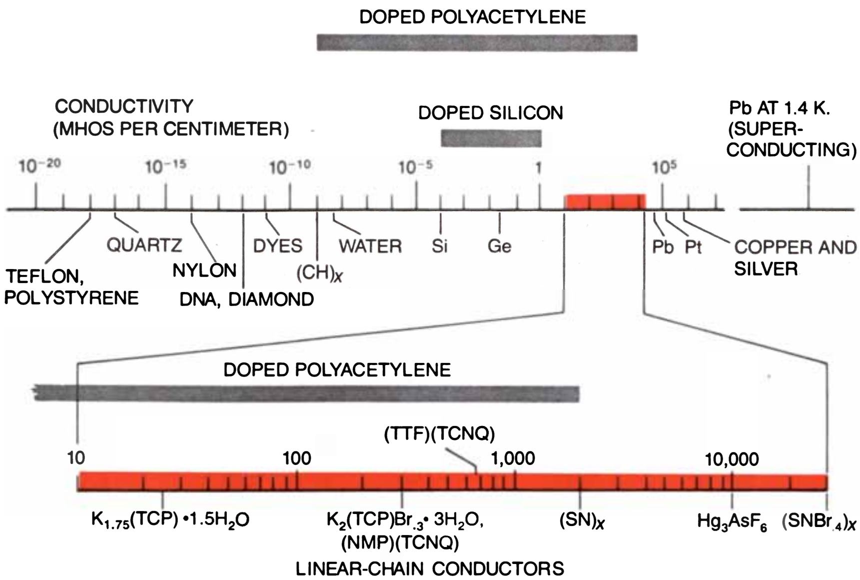
The conductivity of a metal is limited by the scattering of moving electrons, which changes their direction and tends to restore a state in which the net momentum of all the electrons is zero. At high temperature the main cause of scat-



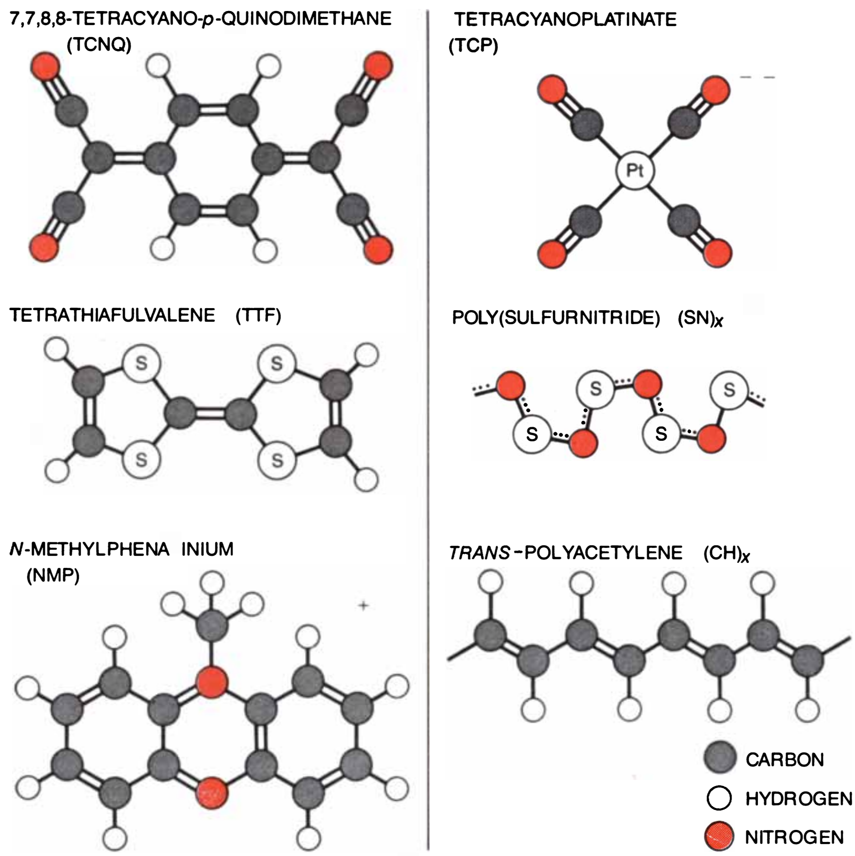
FORMATION OF ENERGY BANDS in a solid results from the overlapping of the orbitals of adjacent atoms. In an isolated atom each orbital, which can hold no more than two electrons, has a well-defined energy. When two atoms are brought together close enough for their orbitals to overlap, each energy level is split in two; adding a third atom splits the atomic levels into three components. In a solid made up of perhaps 10^{23} atoms the levels are so numerous and so closely spaced that they effectively form a continuous band. The width of the band is determined by the extent to which the orbitals overlap. In a solid made up of n atoms each band can accommodate $2n$ electrons. If each of the atoms in a solid has only one electron at some energy level (rather than the maximum of two), the corresponding band will be only half full.



BAND STRUCTURE OF A SOLID determines most of its electrical properties. No conduction can take place in an empty band because it holds no electrons. Conduction is also impossible in a full band because the total population of electrons in such a band can have no net motion. A metal is a material with a partially filled band, where the electrons can be given a net velocity by shifting some of them to infinitesimally higher energies within the band. An insulator has only completely filled bands and completely empty ones, with a large energy gap between them. A semiconductor has a band structure much like that of an insulator but has some mobile charge carriers. The carriers can be introduced by impurities, by defects, by departures from stoichiometry or by the excitation of electrons from the highest filled band (the valence band) to the lowest empty one (the conduction band). What distinguishes a semiconductor from an insulator is that the gap between bands is comparatively small in a semiconductor. In a semimetal the valence band is filled and the conduction band is empty, but these bands overlap in energy. As a result electrons redistribute themselves to create two partially filled bands.



SCALE OF CONDUCTIVITIES covers an enormous range of values: copper and silver conduct some 10^{24} times better than insulators such as Teflon and polystyrene. Most highly conductive linear-chain materials have conductivities between 10 and 20,000 mhos, or reciprocal ohms, per centimeter. The values given for each material are typical, but there is considerable variation among specimens; the range for polyacetylene, an organic polymer, is exceptionally broad. All the conductivities except the one for superconducting lead are measured at room temperature, and for linear-chain materials the measurement is made along the favorable axis.



BASIC SUBUNITS of linear-chain conductors are assembled either in columnar stacks or in linear chains. The disklike TCP unit consists of a platinum ion surrounded by four cyanide groups (CN^-). The molecules abbreviated TCNQ, TTF and NMP are also planar structures, but they are organic molecules. They are the constituents of the salts (TTF)(TCNQ) and (NMP)(TCNQ), in which the two kinds of molecules form segregated stacks. Poly(sulfurnitride) and polyacetylene are polymers in which the subunits form extended, parallel chains.

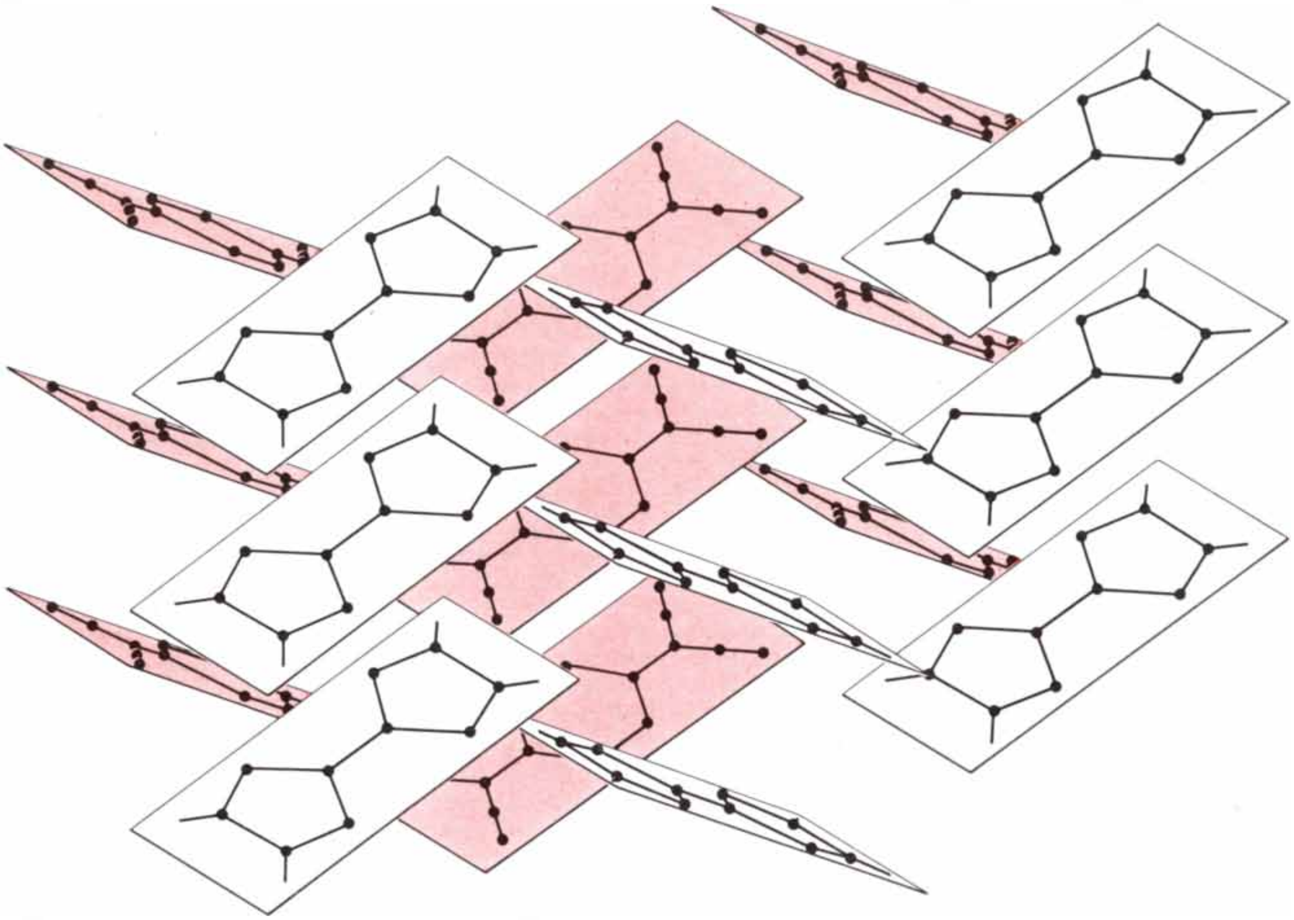
tering is the interaction of conduction electrons with vibrations of the atomic lattice. As the temperature is reduced the lattice vibrates less. For this reason the conductivity of a metal increases as the temperature decreases. The ultimate limit of the conductivity is determined by scattering from imperfections in the crystal structure, from impurity atoms and from other sources such as the surface of the crystal.

Semiconductors and insulators differ from metals in that every energy band is either completely full or completely empty. In a semiconductor conduction is possible when an electron is excited by heat or light to move from a filled band to one of the vacant levels of an empty band. Such a change in the energy of an electron effectively creates two charge carriers: the electron itself and the "hole" left behind in the previously filled band. The minimum energy needed to create these charge carriers is equal to the width of the gap between the bands. Charge carriers can also be supplied by defects or by impurities, such as the dopants that are diffused into silicon semiconducting devices.

Two factors determine the conductivity of a semiconductor: the concentration of charge carriers and their mobility, which is a measure of the average length of time they can move before being scattered. Because of the thermal excitation of electrons across the band gap the concentration of charge carriers increases with temperature, but so does scattering from lattice vibrations, thereby reducing the mobility of the charge carriers. In most semiconductors at practical temperatures the availability of charge carriers is the dominant influence. As a result the conductivity decreases as the temperature decreases, in contrast to what happens in a metal.

An insulator, like a semiconductor, is a material that has only completely filled and completely empty bands, but the gap between bands is generally much wider, so wide, in fact, that almost no charge carriers can be created by thermal excitation. The bulk conductivity that is observed in such materials results mainly from defects and impurities in the structure, and it is usually best described by models other than the band model.

There is one further category of band structure. Some materials are formally described as having only filled and empty bands, but the bands overlap in energy. In a solid of this kind the electrons, in order to minimize their energy, are redistributed to create two (or more) partially filled bands. Such materials are called semimetals, and their electrical properties are not substantially different from those of a metal; in particular their conductivity increases as the tempera-



HERRINGBONE PATTERN is formed by the planes of stacked molecules in the salt (TTF)(TCNQ). Within each stack the molecules are parallel to one another, but they are inclined to the axis of the stack. Orbitals that extend above and below the plane of each molecule overlap, giving rise to an electronic conduction band along the stacks. It has been found that on the average .59 of an electron per molecule is transferred from TTF to TCNQ, which creates partially filled bands in both molecules; as a result conduction takes place in both kinds of stack. In TCNQ the carriers of charge are electrons and in TTF they are holes, which represent the absence of electrons.

ture is reduced. $(SN)_x$ is an example of a semimetal.

The overlapping of selected orbitals explains the formation of anisotropic bands in linear-chain materials, but for high conductivity it is also necessary that one of the bands be only partially filled. This can be accomplished by incorporating in the solid charged groups that do not have a simple integer ratio; the resulting compound is said to be nonstoichiometric. The formula given above for the potassium-bromide salt of TCP indicates that each molecule includes .3 of a bromide ion. What this means is that for every 10 TCP molecules three bromide ions are incorporated in the crystal. Other TCP salts exhibit similar nonstoichiometries.

In (TTF)(TCNQ), which is a stoichiometric crystal, charge is transferred between the two conducting columns. TTF is an electron donor and TCNQ an electron acceptor, and on the average .59 of an electron is transferred from each TTF molecule to the TCNQ stacks. The result is to partially empty a filled band and partially fill an empty one.

The mechanism of conduction in $(SN)_x$ has already been described: $(SN)_x$ is a semimetal, where filled and empty bands overlap in energy. Polyacetylene, when it is untreated, is a poor conductor because of a scarcity of charge carriers. It is made highly conductive by doping it with impurities that act as donors or acceptors of electrons. Heavy doping introduces large numbers of electrons into the conduction band or holes into the valence band, giving the material many of the properties of a metal. When polyacetylene has a low concentration of dopants (less than 1 percent), it acts in most respects like a traditional semiconductor.

One of the most informative properties of linear-chain materials is the variation of their conductivity with changes in temperature. Three broad categories of temperature-dependent behavior are observed. Materials that we shall designate as Class I have a room-temperature conductivity between 10^{-6} mho per centimeter and 10 mhos per centimeter; when the temperature is reduced, the conductivity falls off steadily and in some cases quite rapidly. An example is the cesium salt of TCNQ, with the formula $Cs_2(TCNQ)_3$.

Class II materials have a room-temperature conductivity of roughly 100 mhos per centimeter. As the temperature is reduced the conductivity rises slightly to reach a broad maximum, then declines sharply as the temperature is reduced further. Another salt of TCNQ, formed with *N*-methylphenazinium and abbreviated (NMP)(TCNQ), is typical of this group.

Class III is represented by (TTF)(TCNQ). At room temperature its conductivity is roughly 650 mhos per centimeter. As the temperature decreases the conductivity increases steadily until it reaches a sharp peak of 10,000 mhos per centimeter at 58 degrees K. As the temperature decreases further the conductivity then decreases rapidly.

The temperature dependence of conductivity in Class I materials is comparatively easy to characterize. These substances seem to be semiconductors in which thermal energy is needed in order to excite electrons or holes and thereby activate the charge carriers. The conductivity is lower at lower temperatures because fewer charge carriers are available. Similar behavior is seen in undoped semiconductors such as silicon.

Among the explanations proposed for Class II systems, two models postulate a phase transition from a metallic state at high temperature to a semiconducting state at lower temperature. An early model was based on an idea introduced in 1949 by Sir Nevill Mott of the University of Cambridge and given a more quantitative formulation in 1963 by John Hubbard of the Atomic Energy Research Establishment at Harwell in England. The hypothesis concerns the occupation of the energy bands in solids. Although each orbital can hold two electrons of identical energy, Mott pointed out that an electrostatic repulsion arises between any two electrons occupying the same site. If the energy of this repulsion were greater than the width of the energy band in which the electrons move, then the band would be split into two parts. The lower half would be filled when every orbital had acquired a single electron; after that any more electrons added to the solid would have to join orbitals that already had one electron. Adding this second electron to each orbital would require a somewhat higher energy in order to overcome the repulsion. In other words, there would be a gap between the halves of the energy band.

When the band is split into two parts, the lower half is completely filled and the upper half is empty, so that the material acts as a semiconductor and shows a characteristic decline in conductivity at reduced temperature. Above a threshold temperature, however, the repulsion between electrons in the same orbital is overcome by their thermal motion and the divided halves of the conduction band merge again.

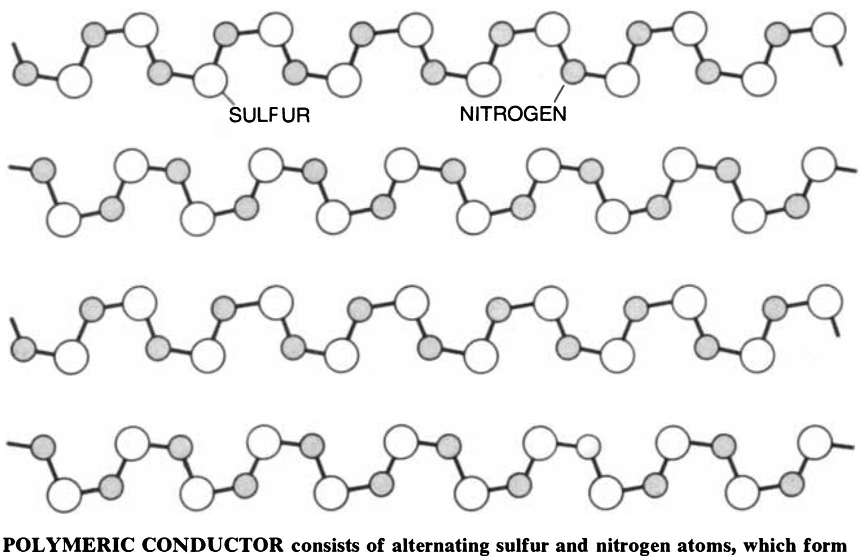
An application of this theory in the early 1970's to the (NMP)(TCNQ) system supposed that each TCNQ molecule carries one negative charge and each NMP molecule one positive charge. The observed maximum in the temperature dependence of the conductivity was attributed to a Mott-Hubbard transition. It is now known that less than one unit of charge is transferred between NMP and TCNQ and that the Mott-Hubbard transition is not an appropriate model for this solid. Nevertheless, the role of repulsion between electrons in determining the electrical, magnetic, optical and other properties of narrow-band, linear-chain conductors remains a topic of intensive study.

Another kind of electronic phase transition was proposed in 1953 by R. E. Peierls of the University of Birmingham. He pointed out that any one-dimensional conductor is susceptible to an instability that alters the periodic spacing of the crystal lattice. A tacit assumption of the simple band model of one-dimensional materials is that the spacing between units in a column or chain is uniform. Under some conditions, however, the energy of the electrons can be reduced if an energy gap is created at the highest occupied level of a partially filled energy band. The result is to convert a metal into a semiconductor, with only filled and empty bands. Such a gap can be introduced by a periodic distortion of the lattice. In the simplest case, where a band is initially half-filled, the units in the chain would clump together in pairs, creating alternating narrow and wide spaces between units. The resulting distortion would introduce a gap between the highest occupied level and the lowest vacant one, turning a metal into a semiconductor. The distortion would also strain the lattice, thereby increasing its energy. Hence the Peierls transition takes place only if the energy of the electrons is reduced enough to compensate for the increase in the strain energy of the lattice. In a quasi-one-dimensional solid that condition can be met at low temperature; at high temperature the uniform lattice is the preferred configuration. It has been suggested that a Peierls transition takes place in a methylated derivative of (TTF)(TCNQ).

Several Class II systems have significant crystallographic disorder, and some models exploit that fact to explain the variation of conductivity with temperature. Disorder has the effect of confining the electrons of a one-dimensional system, so that they cannot roam freely throughout the length of a strand but tend to remain in a comparatively small region. In such a disordered one-dimensional state conduction would be limited by the rate at which electrons could hop from one local region to another. The hopping would be facilitated by lattice vibrations, and so conductivity would increase at higher temperature. Models of this kind have been discussed by workers at Johns Hopkins University, at the University of Chicago and in the U.S.S.R. A similar idea was introduced by investigators at Brown-Boveri. They proposed to represent a TCP compound by a model in which conducting strands are interrupted at intervals by perfectly insulating defects. Thermal activation would be required for electrons to bridge the gaps. None of these models, however, is able to account systematically for the exact form of the conductivity-v.-temperature curve.

The effort to understand the mechanism of conduction in Class III materials, and particularly in (TTF)(TCNQ), has spawned a multitude of theoretical models. Many theorists believe some of the data cannot be explained by an ordinary band model, in which the charge carriers move independently of one another. Instead, mechanisms of collective charge motion that are reminiscent of superconductivity are invoked.

In three-dimensional metals superconductivity is accounted for by a theory set forth in 1956 by John Bardeen, Leon N. Cooper and J. Robert Schrieffer of the University of Illinois; today it



POLYMERIC CONDUCTOR consists of alternating sulfur and nitrogen atoms, which form long strands that line up parallel to one another. Interactions between the strands are significant, with the result that the conductivity along the chain axis is only about 10 times as great as the perpendicular conductivity. The material, which has the formula $(SN)_x$, is a semimetal.

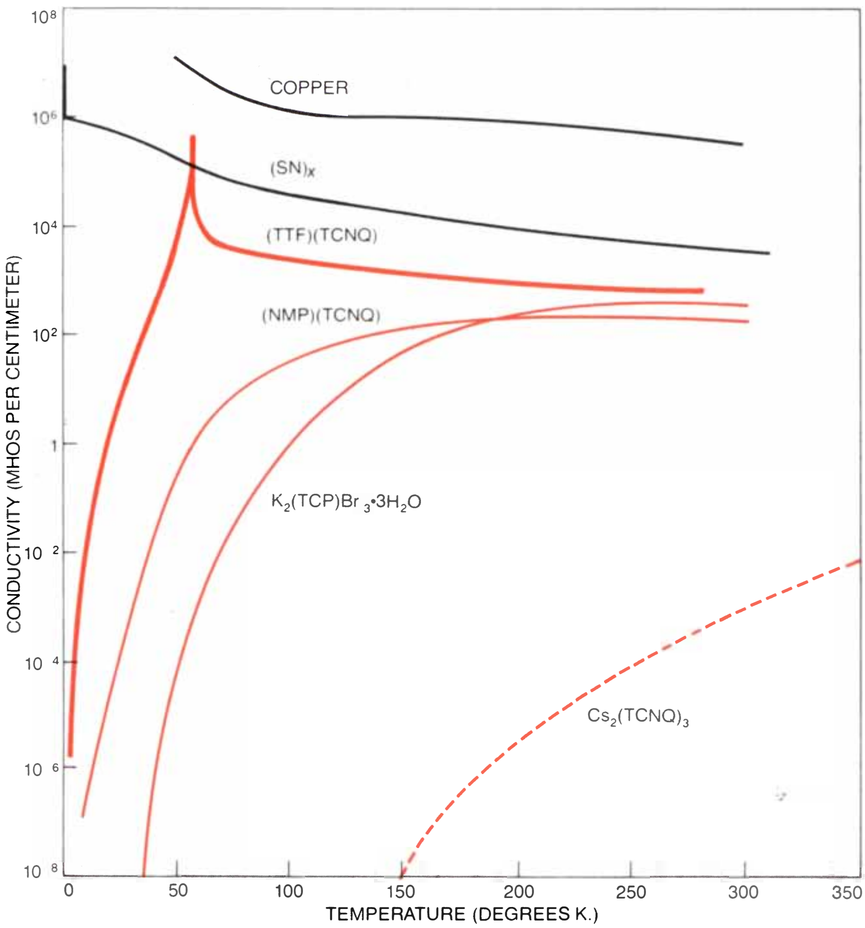
is called the BCS theory. The resistanceless current in this model is carried not by individual electrons but by pairs of electrons. The pairs are stabilized by an indirect attraction: the negative charge on one electron attracts the positive metal ions in the lattice surrounding it and thereby creates a region of enhanced positive charge. It is this charge that attracts the second electron. The phenomenon is observed only at extremely low temperature because the pairs are easily broken up by thermal agitation.

The form of superconductivity hypothesized by Little also depends on the pairing of electrons, but in a different context. The electron pairs move along a conducting spine and the indirect attractive interaction between them is mediated by side groups that are easily polarized.

A possible explanation of the conductivity of Class III systems also depends on an interaction between conduction electrons and the surrounding lattice, but it does not require the pairing of electrons. The model was first applied to (TTF)(TCNQ) by Bardeen, but the mechanism itself was first proposed by Herbert Fröhlich of the University of Liverpool.

As we have seen, the Peierls transition that converts a metal into a semiconductor comes about through a periodic distortion of the lattice: the chain is compressed in one region and stretched in another. The distortion tends to impede ordinary conduction of the band type, but it can enhance conductivity through another mechanism. The conduction electrons, at once creating the distortion of the one-dimensional lattice and responding to it, tend to become concentrated in the regions of greater positive charge, where the chain is contracted. The variation in electron concentration is called a charge-density wave. Once the electrons have gathered in this way on a single chain, and before the relative position, or phase, of the wave becomes locked in place, the electrons can move along the strand as a group in response to an applied voltage. The aggregation of electrons is carried along by the moving wave in much the same way that a surfboard is carried by an ocean wave. The molecules of the lattice execute only periodic motions around their average positions, but the electrons move forward along the strand.

The movement of a charge-density wave can enhance conductivity because it stabilizes the electrons in their state of motion. An electron moving independently is easily scattered by a defect or a lattice vibration. An electron trapped in a charge-density wave is subject to the same influences, but unless the disturbance is large the electron can be restored to its original trajectory by the



TEMPERATURE DEPENDENCE of conductivity is an important indicator of the electronic structure of a solid. In metals such as copper and in semimetals such as $(SN)_x$ conductivity increases as the temperature is reduced. Semiconductors generally have the opposite behavior: conductivity falls off as the temperature is reduced. Linear-chain molecular conductors can be sorted into three classes according to the temperature dependence of their conductivities. Class I materials, typified by a cesium salt of TCNQ, resemble ordinary semiconductors in that their conductivity declines steadily when they are cooled. Class II is made up of materials such as (NMP)(TCNQ) and the potassium salts of TCP that have a broad maximum in conductivity. (TTF)(TCNQ) is a Class III material: as it is cooled from room temperature its conductivity rises to a sharp peak at 58 degrees Kelvin, then it declines as the temperature drops further.

attractive potential that surrounds it.

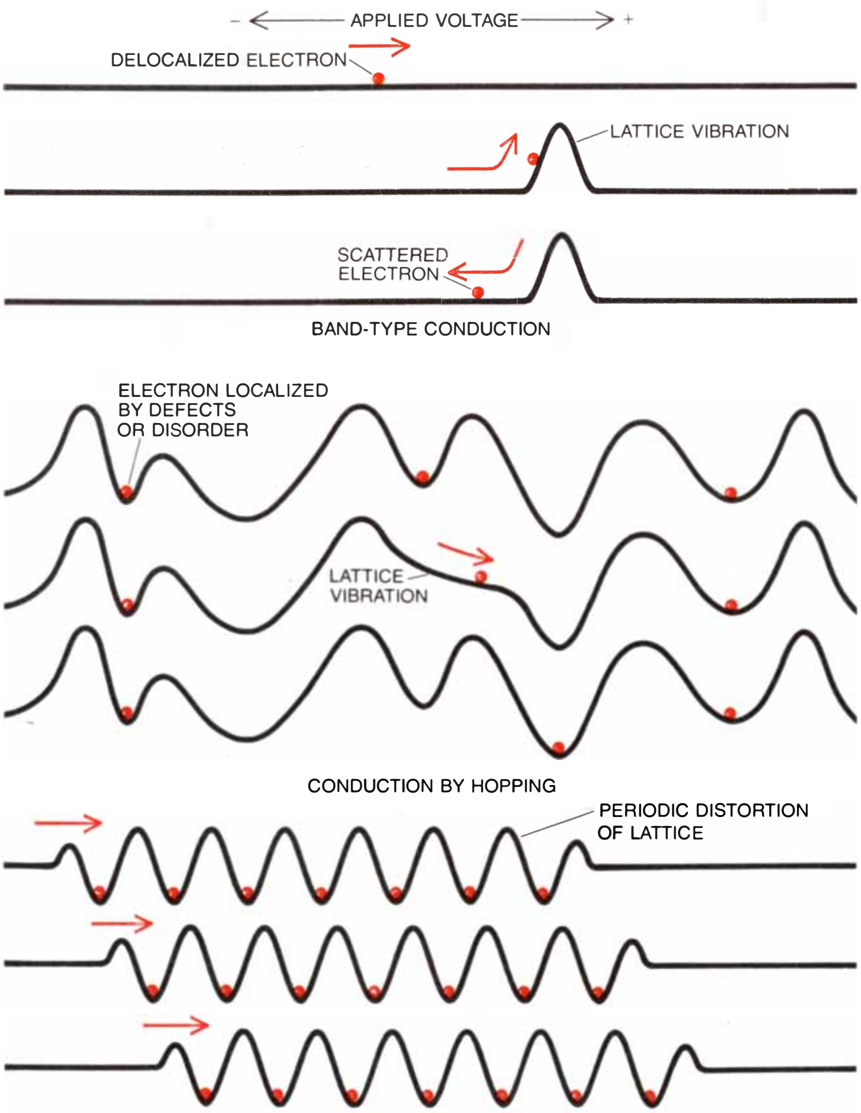
At low temperature the electrostatic interaction between neighboring charge-density waves increases, with the result that the waves become fixed in the crystal. A wave can also be pinned on a defect or on any other irregularity of the lattice, so that the conductivity is sensitive to the quality of the specimen. The pinning of charge-density waves might explain the sudden decline in the conductivity of (TTF)(TCNQ) below 58 degrees K. The Fröhlich model has also been applied to the inorganic quasi-one-dimensional substance niobium triselenide.

Several other theories have been proposed to account for the conductivity of Class III materials without invoking collective charge transport. For example, workers at the International Business Machines Corporation have suggested that the distinctive properties of these materials might be attributed to the scattering of electrons by other electrons. A group at the Hebrew University of Jerusalem has formulated an explanation based on the scattering of electrons by oscillatory librational motions of the TCNQ molecules, and a group at the University of Paris has cited the possibility of electron scattering by magnetic excitations.

A model developed at Xerox is applicable to linear-chain organic conductors that exhibit all three types of conducting behavior. In order to embrace all three kinds of material it is necessary to confine the range of temperatures being considered to those above about 70 degrees K. Below that temperature transport by independent, conduction-band electrons is sometimes inadequate to explain the data.

The model focuses on the basic factors that determine the conductivity of a semiconductor: the concentration of charge carriers, which is represented by the symbol n, and the mobility of those carriers, which is symbolized by μ. The conductivity is the product of these quantities and a constant, the charge of the electron, e; hence the conductivity is given by $ne\mu$. If this expression is to be useful in understanding conductivities, however, it must first be recognized that the values of n and μ vary both with the nature of the material and with the temperature. The model that has been developed at Xerox analyzes these variations in the linear-chain organic molecular conductors.

We found that in many cases the temperature-dependent variation in the value of n could be represented as an exponential expression in which the exponent is a ratio of two energies. One of the energies is half the gap energy, the interval between the highest filled level and the lowest vacant energy band; the other is the thermal energy, which is merely the temperature converted into energy units. The form of the expression implies that any decrease in the gap energy or increase in the temperature will cause the concentration of charge carriers to rise exponentially. That relation is in accord with intuition, since the charge carriers are created by thermally exciting electrons to cross the gap.



MECHANISMS OF CONDUCTION in quasi-one-dimensional materials include some that cannot be observed in three-dimensional solids. In metals (band-type conduction) charge is transported by mobile electrons and conductivity is limited by their scattering. The most important sources of scattering are thermal vibrations (shown as a peak that the moving electron cannot surmount). Vibrations become less common at low temperature, so that metallic conductivity is improved by cooling. The hopping of electrons across energy barriers has been proposed by some workers to explain peculiarities of anisotropic conduction. The hopping is facilitated by lattice vibrations, which may transfer enough energy to a localized electron for it to overcome a barrier and migrate to a neighboring site. A charge-density wave is a possible mechanism of collective electron motion. The wave forms when electrons distort the one-dimensional lattice in a periodic manner, creating positively charged regions where electrons gather. When a voltage is applied, the electrons move along with the charge-density wave.

The mobility, we found, is proportional to the absolute temperature raised to the power $-\alpha$, where the exponent α is a constant whose value depends on the nature of the material and on its quality. Any increase in the temperature diminishes the mobility, which also agrees with intuition, since the mobility is limited by the scattering of electrons from thermal vibrations of the lattice. The constant α determines how steeply the mobility declines as the temperature rises.

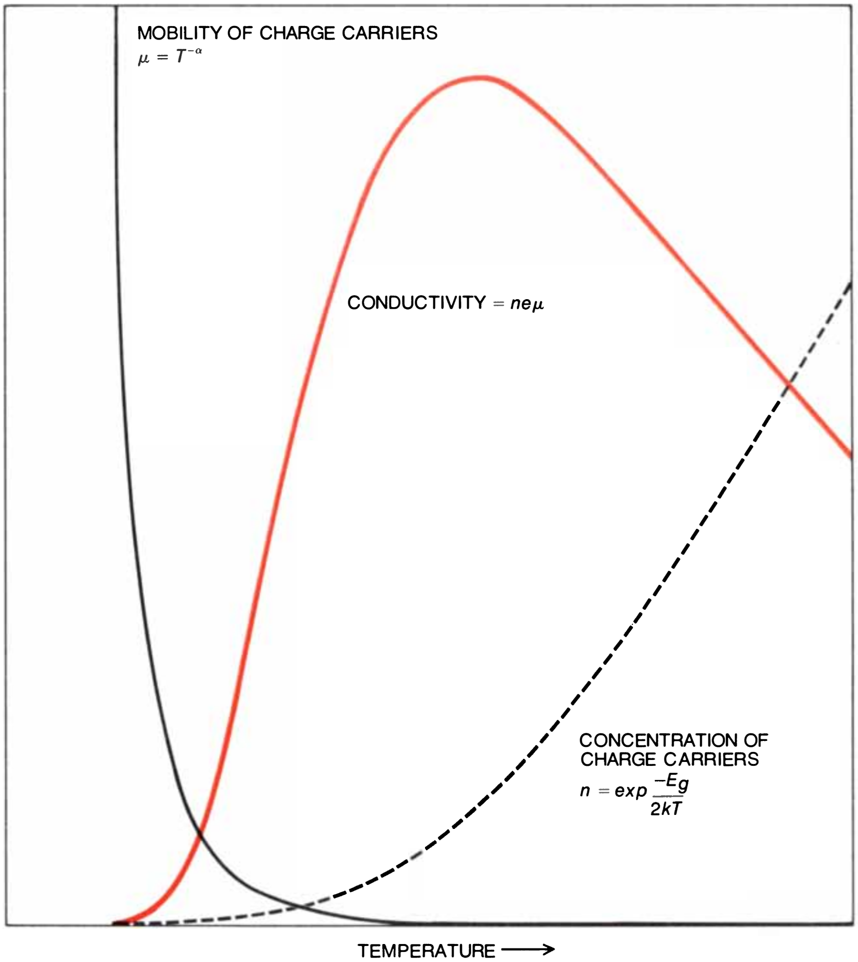
A graph of these functions shows that n rises steeply with increasing temperature and at the same time μ declines steeply. The conductivity as a function of temperature is given by the curve that traces the product of n and μ. The product curve is found to have a maximum at some finite temperature.

The form of this theoretical conductivity curve can be recognized as fitting the experimental data for Class II materials. They have a broad maximum in conductivity at temperatures somewhat below room temperature. The model can also be applied to Class I and Class III systems. Materials in Class I have comparatively low conductivity, which rises with increasing temperature. This pattern can be accounted for in the model if it is assumed that the energy gap in these materials is substantially larger than it is in Class II systems. Increasing the band gap depresses the concentration of charge carriers at any given temperature, and it defers the maximum in conductivity to higher temperatures, which are experimentally inaccessible.

Class III materials can be interpreted as having a band structure with a negligible energy gap. The availability of charge carriers is therefore essentially constant, regardless of the temperature, and the conductivity is governed mainly by the mobility. Mobility is greater at lower temperatures, and so the conductivity rises when the material is cooled, as in a metal. Below 70 degrees K. there may be an additional contribution to the conductivity of (TTF)(TCNQ) as a result of the motion of charge-density waves. Below 60 degrees K. a Peierls transition converts the material into a semiconductor.

This model simplifies considerably the classification of linear-chain systems. What had been perceived as three distinct kinds of behavior can now be understood to result from the continuous variation of a single parameter: the width of the band gap. For Class II systems the model has also been applied with success to other physical properties of the materials, such as their optical properties and their thermoelectric power. In addition the mobility of the charge carriers in Class II and Class III substances has been calculated directly at Xerox from the constants that define the coupling between electrons and lattice vibrations. The results are consistent with the model.

In order to test the applicability of this model and of other models to Class II materials a novel system was devised at Xerox in which both the filling of the energy bands and the disorder of the crystal can be varied continuously. The system is a variant of (NMP)(TCNQ) in which some proportion of the NMP molecules (up to 50 percent) are replaced by molecules of phenazine. Phenazine is the same size and shape as NMP but it remains electrically neutral, whereas NMP takes on a charge. Hence for each NMP molecule replaced the (NMP)(TCNQ) system is missing one electron. The random intermixture of the two kinds of molecules also increases the disorder of the crystal. An extensive series of measurements revealed consistent Class II behavior at all concentrations of NMP studied. The substance could not be made metallic by changing the stoichiometry but remained a semiconductor with an effective band gap proportional to the square of the NMP content. Recent X-ray studies in France suggest a nonuniformity in the lattice spacing, and so the gap may be of the Peierls type. The results are not consistent with the models that attribute the decline in conductivity at low temperature to the effects of disorder.



MOBILITY MODEL describes the behavior of materials in all three classes of linear-chain organic molecular conductors at temperatures above 70 degrees K. The conductivity is defined as the product of three quantities: the charge of the electron (*e*), the concentration of charge carriers (*n*) and the mobility of the carriers (μ). Because thermal vibrations both activate charge carriers and scatter them, the concentration of charge carriers increases with the temperature but their mobility decreases. The product of these curves has a maximum at the optimal combination of concentration and mobility. In Class II materials the maximum is observed directly. Class I materials are thought to have a larger gap between energy bands, with the result that the predicted temperature of maximum conductivity is beyond the range of observation. In Class III materials the gap is thought to be negligible, so that mobility governs conductivity, which therefore increases at lower temperature. In the equations that define the curves *T* is the temperature in degrees K., *kT* gives the temperature in energy units, E_g is the energy gap, α is a constant determined by the scattering mechanisms of the specimen and "*exp*" is a notation indicating that the base of natural logarithms is to be raised to the power given by the exponent.

As linear-chain conductors have come under intensive scrutiny in the past few years much progress has been made in understanding their physical properties. At the same time chemists have quickened the pace at which new materials are introduced. Recently, for example, a great many linear-chain crystals and polymers have been doped with halogens, such as iodine and bromine, in order to increase their conductivity. They include $(SN)_x$, polyacetylene, phthalocyanines and other organic compounds.

An intriguing anisotropic conductor that has been synthesized at McMaster University in Ontario and has also been studied at the University of Pennsylvania is a compound of mercury, arsenic and fluorine, Hg_3AsF_6. It has a linear-chain structure in two dimensions. Conducting chains of mercury atoms form layers with all the strands in each layer oriented parallel to one another. Successive layers, however, are turned by 90 degrees. As a result the conductivity of the compound is high along two perpendicular axes but is lower by a factor of 100 along the third axis. At temperatures below four degrees K., Hg_3AsF_6 has been reported to become a superconductor.

It seems unlikely that highly conductive linear-chain materials will soon replace copper wiring, but there is little doubt the materials will eventually have practical applications. The technological potential of polyacetylene and similar polymers is particularly noteworthy because doped films of this polymer have an extremely wide range of conductivities and also have mechanical properties typical of polymers, such as the ability to stretch.

Some applications would make direct use of the semiconducting properties that are intrinsic to many of the linear-chain materials. At the University of Pennsylvania films of polyacetylene doped with two different impurities have been formed into a simple diode. Workers at the Rockwell International Corporation have demonstrated a solar cell made of $(SN)_x$. Finally, at least one commercial application has already been reported. At the Eastman Kodak Company a linear-chain organic conductor has been embedded in an insulating film to make a photoconducting belt for a copying machine.

The Acoustic Microscope

Research on ultrasonic imaging technology has given rise to a new experimental tool capable of generating pictures with a resolution comparable to that of light micrographs

by Calvin F. Quate

Sound waves vibrating in the atmosphere at frequencies lower than about 20,000 hertz (cycles per second) are most familiar as a means of communication. Human speech, for example, is in this range. The atmosphere is not a good medium for the propagation of acoustic waves at higher, ultrasonic frequencies; these inaudible vibrations are quickly dissipated in gases such as air. Waves of ultrasound can be made to propagate for appreciable distances with moderate attenuation only in liquids and solids. Acoustic waves vibrating in such condensed mediums at ultrasonic frequencies are widely exploited not for communicating but for imaging. Ultrasonic "cameras" based on the generation and detection of acoustic waves with frequencies in the range near one megahertz (a million cycles per second) are currently available for the study of underwater objects, internal structural features in materials and organs inside the human body. The fetus in the womb of a pregnant woman, for example, can now be harmlessly examined in this way [see "Ultrasound in Medical Diagnosis," by Gilbert B. Devey and Peter N. T. Wells; SCIENTIFIC AMERICAN, May, 1978].

An extension of ultrasonic imaging technology is the acoustic microscope. The ultrasound waves that are the basis of this experimental device have frequencies in the vicinity of one gigahertz (a billion cycles per second), some 1,000 times higher than the frequencies typical of macroscopic ultrasonic imaging systems. In terms of wavelength the comparison is particularly striking. The wavelength of the sound generated by the human voice is measured in meters; the operating wavelength of a macroscopic ultrasonic camera is measured in millimeters; the operating wavelength of the most advanced acoustic microscope is measured in micrometers. In other words, the vibrational waves employed in this new imaging device are comparable in length to the electromagnetic waves of visible light.

It should come as no surprise, therefore, that the acoustic microscope is beginning to yield images that are comparable in resolution to the images obtained with conventional light microscopes. What might seem surprising is that workers in this field are striving not just to match the resolution of the optical microscope but to surpass it. Resolution is not the whole story, however. The source of contrast in the acoustic system is quite different from that in the optical one, and previously inaccessible properties of microscopic objects are coming into view in the acoustic images. In time it is expected that the acoustic microscope will find its place along with the optical microscope, the electron microscope and the forthcoming X-ray microscope as one of a complementary set of tools for exploring the world of the very small.

The idea of harnessing acoustic waves in the gigahertz range for a system of microscopy was first put forward more than 40 years ago by S. Y. Sokolov in the U.S.S.R. The technology necessary to generate such waves, however, was not available until the early 1960's. Much of the early work on very-high-frequency ultrasonic devices was done by Hans E. Bömmel and Klaus Dransfeld at Bell Laboratories. A little more than a decade ago James S. Imai and Isadore Rudnick of the University of California at Los Angeles managed to generate in liquid helium acoustic waves with a wavelength of only .2 micrometer.

In recent years several groups, following somewhat different approaches, have taken up the task of trying to develop a practical system for acoustic microscopy. Researchers at the Zenith Radio Corporation, led by Adrianus Korpel and Lawrence W. Kessler, were prominent in the early stages of the effort. (Kessler has since gone on to form Sonoscan, Inc., which manufactures a laser-scanning version of the acoustic microscope.) Our own group at Stanford University has been actively pursuing this goal for five years. We have succeeded in reducing the operating wavelength of our device from five micrometers to half a micrometer. That is not the end, however. We believe future instruments will operate with wavelengths shorter than those achieved by Imai and Rudnick.

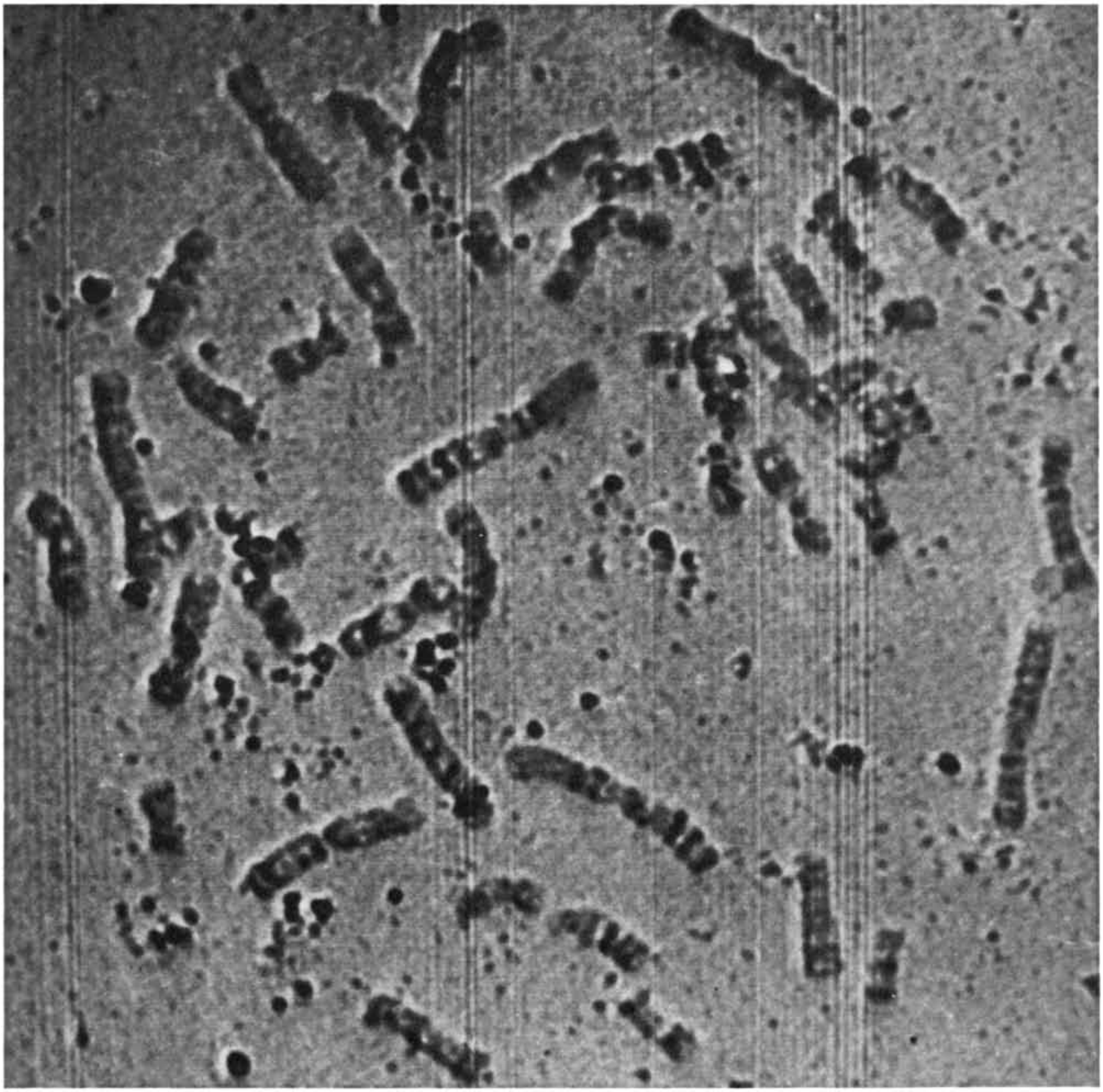
What have we learned so far? We have learned that an acoustic microscope can be built with a focusing element that is simpler than the lens of an ordinary optical microscope. We have found that the resolution of an acoustic microscope is limited only by its operating wavelength, and that aberration does not play an important role. We have demonstrated that the detail in acoustic micrographs can be as fine as the detail in optical micrographs. We have confirmed that it is an easy matter to measure the thickness of semiconductor and metal films with acoustic waves, and that it is also possible to study the adhesion of these layers to a substrate. In the case of biological specimens we have found that we can obtain high-contrast images of various tissues without staining them beforehand. We feel confident this last capability will soon make it possible to glean by means of acoustic microscopy new information from living cells.

How are the images formed in an acoustic microscope? As in the optical microscope, the key is the change in the velocity of the waves as they cross an interface between two different materials. For acoustic waves this change is much greater than it is for light waves. The index of refraction, the standard measure of the relative velocities of light waves in any two mediums, is never greater than 1.9. Indeed, its value is usually less than 1.5. In contrast, the velocity of sound waves can decrease by a factor of 10 in traversing a suitable solid-liquid interface.

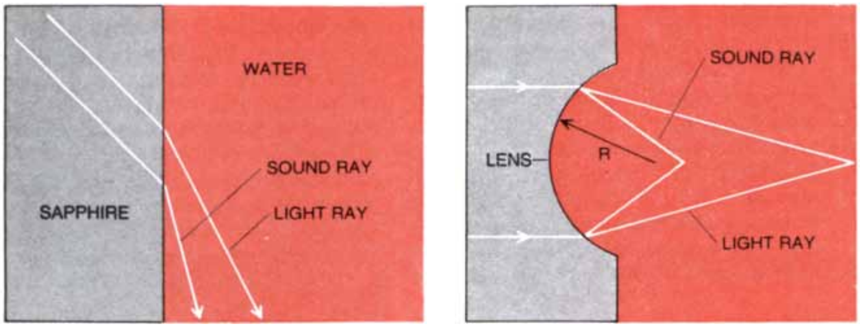
The primary focusing element in the acoustic microscope developed by our group at Stanford is a concave spherical interface, approximately 40 micrometers in radius, formed between sapphire and water. At that interface the sound rays are bent sharply inward as they enter the liquid. In our system the entire set of rays converges to a narrow "waist" that is near the center of curvature of the sphere. The change in velocity for light waves at such an interface would be small and the convergence angle of the focused beam would be shallow; the resulting aberration could be overcome only by resorting to a compound lens with multiple surfaces.

In the acoustic system a spherical lens with a single surface is almost ideal. The sharp convergence angle keeps aberration to a minimum. Moreover, antireflection coatings similar to those commonly employed with optical lenses can be applied to the sapphire to reduce the reflection of the acoustic waves from the interface. Residual reflections cannot be completely eliminated, however, and they continue to be a minor source of interference. Fortunately the different reflections can be distinguished with the aid of a test signal in the form of a momentary pulse of ultrasound. The part of the pulse that travels through the liquid will be reflected from the object and will arrive at the receiver later than the part that is reflected from the lens. This procedure enables us to identify the source of the reflected waves, since they are recorded at different times.

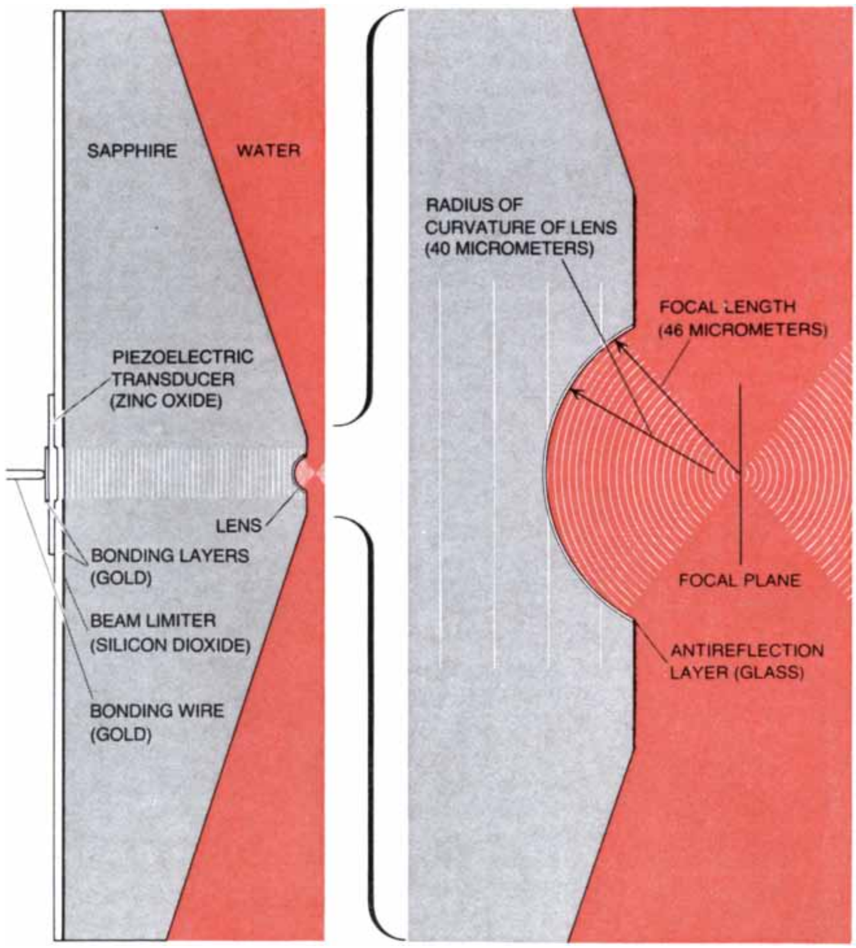
The main components of the acoustic



ACOUSTIC MICROGRAPH of a set of human chromosomes was made by the author and his colleagues in the department of applied physics at Stanford University as a test of the resolution of their new ultrasonic imaging device. For this purpose both the specimen slide and the lens assembly of the instrument were immersed in liquid argon and maintained at a temperature of about 85 degrees Kelvin (minus 188 degrees Celsius) by surrounding the liquid-argon container with a bath of liquid nitrogen. The ultrasound signals recorded through the lens assembly were processed electronically to form the image of the chromosomes on the screen of a television set; in this photograph of the television image the vertical lines are the result of defects in the scanning process. The chromosome spread was prepared in the laboratory of Howard M. Cann at the Stanford University Medical Center. The protein coating of the chromosomes was removed, and they were routinely stained to make their banded structure stand out in light micrographs. Enlargement is 3,500 diameters.



SOUND RAYS, LIKE LIGHT RAYS, are refracted, or bent, when they cross an interface between two different materials in which the velocities of the propagating waves are not the same. The degree of bending is much greater in the case of sound rays (*diagram at left*). A concave spherical surface ground into the surface of the material with the higher index of refraction (that is, the one in which the waves move faster) will serve to focus both kinds of rays (*diagram at right*). In this situation the focused sound rays converge at a much sharper angle than the focused light rays, and the resulting aberration in the focused image is correspondingly less. For the sound waves the focal point is 1.3 times the radius of curvature of the lens (*R*); for the light rays the focal point is about three times the radius of curvature of the lens.



ACOUSTIC LENS designed by the author's group at Stanford consists of a tiny concave spherical interface between sapphire and water, formed at the beveled tip of a short sapphire rod a quarter inch in diameter (*left*). The piezoelectric film that is deposited on the back surface of the sapphire lens element converts electrical signals into acoustic ones during the exposure stage of the image-making process; the same film serves to convert the acoustic signals back into the electrical ones during the subsequent readout stage. The enlargement at the right shows the approximate profile of the focused beam of ultrasound in the water; the parallel white lines in both views are the acoustic wave fronts. The flat face of the sapphire rod is rough-ground in order to scatter stray sound waves. In addition a glass antireflection coating is usually applied to the front surface of the lens in order to reduce internal acoustic reflections.

system are fairly simple compared with the prodigious array of electronic equipment needed to generate the signals and prepare them for final display. Our system is designed so that we can move the signals into and out of the acoustic cell as quickly as possible. The reason for this design is simply that at the present state of the art electrical signals are much easier to manipulate than acoustic ones.

The first step in the image-making process is to convert an electrical signal into an acoustic one by means of a piezoelectric film deposited on the back surface of the sapphire lens element. The film consists of many small crystals, and each crystal has a preferred direction: its long axis is usually oriented at right angles to the surface of the substrate on which it is deposited. An alternating electric field imposed along this axis will compress and expand the crystals as the polarity of the field fluctuates. Piezoelectric films of various types have been available for a number of years, and they are an integral part of many commercial acoustic devices. The particular piezoelectric film we favor, zinc oxide, enables us to convert electromagnetic energy into acoustic, or mechanical, energy with an efficiency that approaches 50 percent.

The beam of ultrasound thereby generated is focused by the acoustic lens on a very small spot in the plane of the object. The image is formed by moving the spot across the object plane point by point and line by line in a raster pattern, in much the same way that the image focused on the photosensitive surface of a television camera is recorded by scanning the surface with an electron beam. In the acoustic microscope, however, the scanning is done mechanically and the rate of scanning is much slower. It takes several seconds to scan a single frame, making it necessary to store the recorded signals for some time before they can be utilized to construct an image of the object.

The signals reflected from the object and stored in the electronic memory are finally amplified and made to modulate the intensity of the electron beam in an ordinary television receiver. The image is formed by scanning the electron beam across the screen in synchrony with the motion of the acoustic beam across the object (or, in our present model, the object across the focal point of the acoustic beam). A one-to-one correspondence is maintained between the position of the electron-beam spot on the television monitor and the position of the acoustic-beam spot on the object. If the electron beam is displaced across the television screen by one centimeter and the acoustic beam is displaced across the object by 10 micrometers, the resulting image will be magnified by 1,000 diameters. The ratio of these two displacements is easily adjusted, and the images can be made to appear with magnifications ranging from 100 to several thousand diameters.

In our present system the mechanical scanning is accomplished as follows. The sample is attached to a round disk fitted into a holder that is driven horizontally by a loudspeaker. The loudspeaker vibrates at 60 hertz, and it drives the sample through a horizontal displacement of between 100 and 200 micrometers. The entire assembly is displaced in the vertical direction by mounting the speaker on vertical slides. The vertical motion, which is driven by a direct-current motor, is much slower than the horizontal motion.

At this point visitors to our laboratory often ask two questions: (1) Does the relative motion of the sample and the lens not introduce troublesome vibrations in the solid parts of the apparatus or turbulence in the liquid? (2) Why is the mechanical scanning method not replaced with an electronic method that would eliminate the need for moving all these components and also be much faster?

The answer to the first question is that vibrations are sometimes present in the sample, but they appear to be quite small, probably because the mechanical resonance modes of the object are at much higher frequencies than the 60-hertz frequency of the scanning process. So far turbulence has not been encountered in the liquid. As we refine the instrument and observe finer details these factors may become important, but they do not concern us at this stage. The images we have made with the instrument are in focus with sharp edges across the entire field of view.

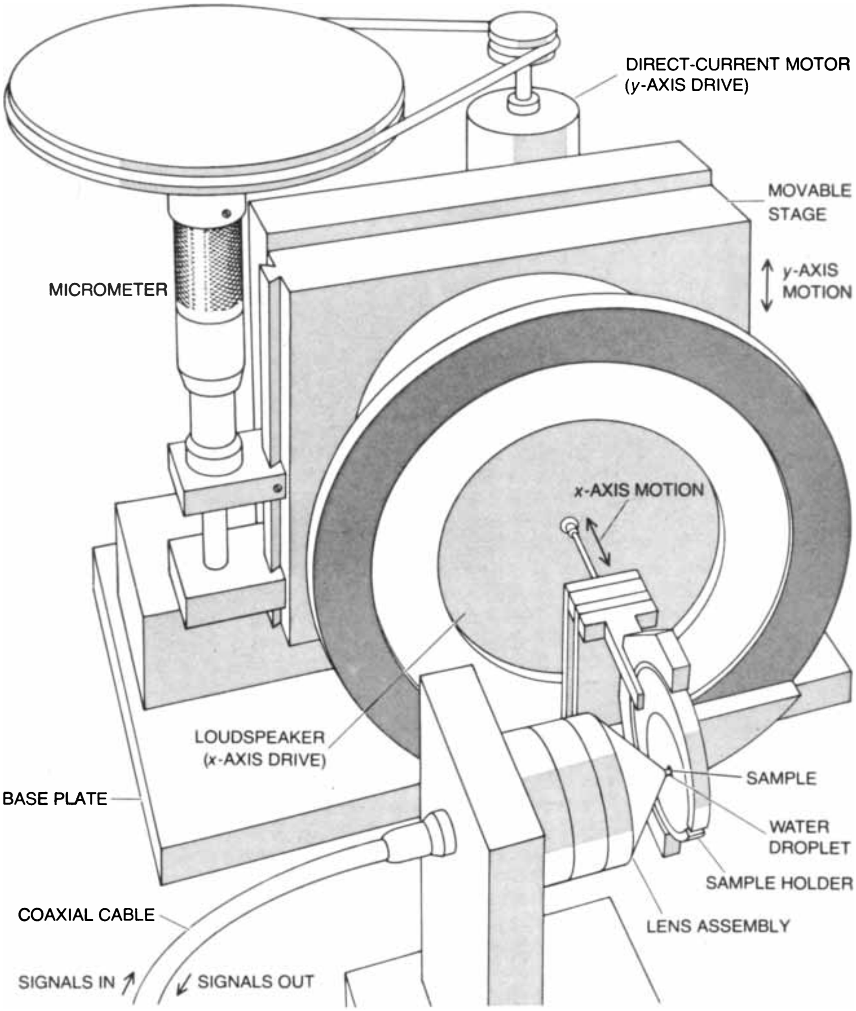
The answer to the second question has to do with the pace of invention. We certainly appreciate the potential advantages of electronic scanning, but we have not yet been able to devise a workable method for nonmechanical scanning with a tightly focused beam.

Solid surfaces are examined with our instrument by exposing the surface to the converging spherical waves from the acoustic lens and then monitoring the reflected diverging waves with the aid of the same lens operating in reverse. The images generated by the reflected waves have special properties. It is important to understand the physical basis for the difference between these images and the images produced by optical systems. Since the situation is easier to visualize in terms of plane waves, I shall first describe the essential difference in those terms and then discuss the more complex process involving converging and diverging spherical waves.

Any plane wave striking a flat, elastic surface at right angles will be reflected back in the opposite direction with a phase that is equal to that of the incident wave. When the direction of the incoming plane wave is tilted with respect to the perpendicular, however, a new wave can be excited that appears to cling to the surface. When this happens, the reflected wave returns at an angle equal to the angle of incidence but with a phase that is different from that of the incident wave [*see illustration on next page*]. This phenomenon, called total internal reflection, has been intensively studied in the case of light waves impinging on transparent surfaces by P. K. Tien of Bell Laboratories and others; it forms the basis of the new technology of guided-wave optics [see "Guided-Wave Optics," by Amnon Yariv; SCIENTIFIC AMERICAN, January].

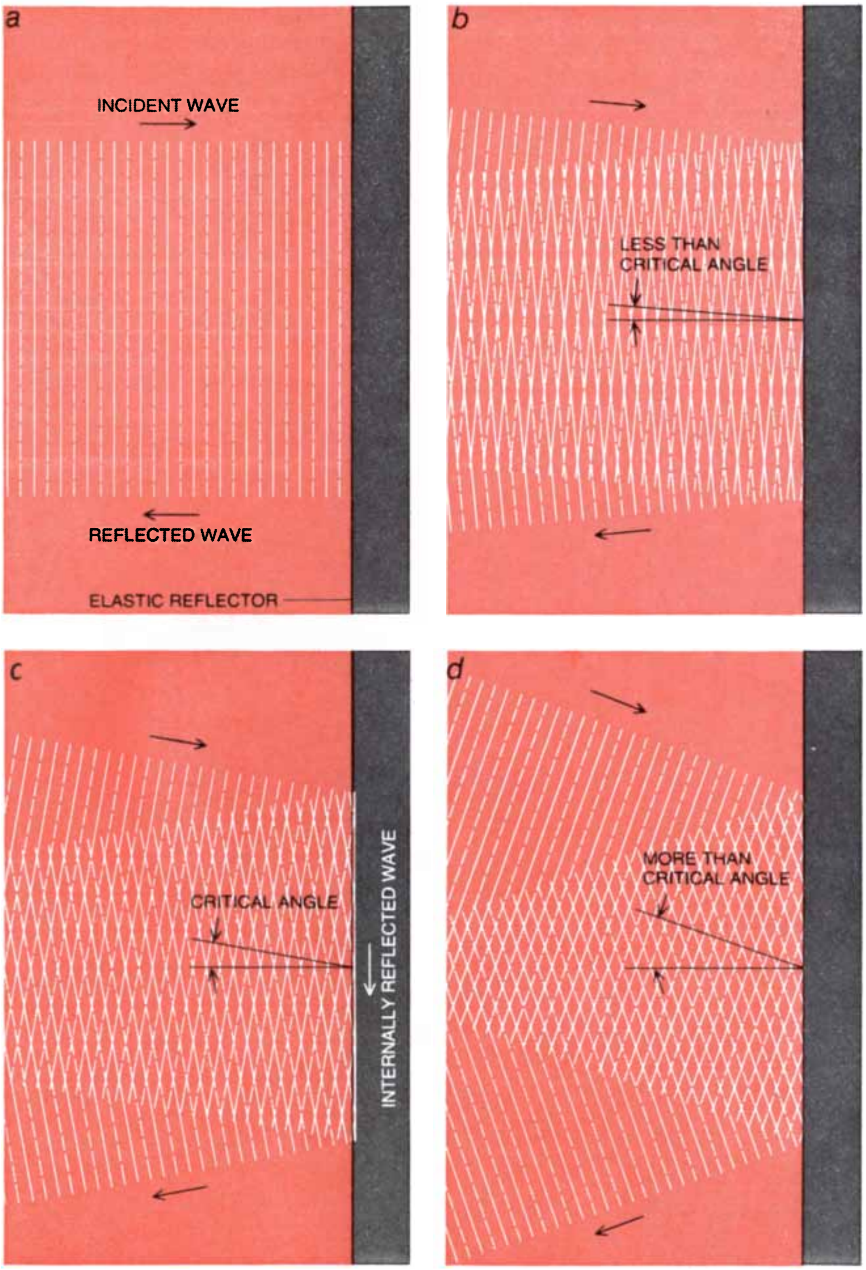
In acoustic systems these special waves travel along all surfaces and interfaces. They can be excited by adjusting the angle of the incoming wave to the critical angle. For angles of incidence less than or greater than the critical angle the wave is totally reflected in phase with the incoming wave and no vibrational energy is transmitted into the solid. For sound waves impinging on a liquid-solid interface the critical angle can be as small as 10 degrees.

The total internal reflection of high-frequency sound waves has been exploited by K. W. Andrews and R. L. Keightly of the British Steel Corporation to examine various solid materials. The elastic properties of materials are known to change with increasing stress, and one would expect the change to af-



MECHANICAL SCANNING of the sample in two dimensions is accomplished in the present version of the Stanford acoustic microscope by means of the apparatus depicted here. The round disk holding the sample is displaced horizontally between 100 and 200 micrometers by a loudspeaker vibrating at a frequency of 60 hertz (cycles per second). The vertical displacement of the sample and the loudspeaker assembly is driven by a direct-current electric motor. (The vertical motion is much slower than the horizontal motion.) The reflected acoustic signals from the sample are processed electronically and used to modulate the intensity of the electron beam in a television monitor. The image is formed by scanning the electron beam across the screen in synchrony with the motion of the sample across the focal point of the acoustic beam. It is the ratio of displacements of the two beams that determines the magnification of the final image.

fect the velocity of sound waves traveling along the surface of the material. Andrews and Keightly have shown that the critical angle for total internal reflection varies not only with the applied stress but also with the composition of the material. In their apparatus the critical angles are determined by exposing the surface to a beam of plane sound waves tilted with respect to the perpendicular by various degrees. The piezoelectric transducer for receiving the reflected signals is also tilted with respect to the perpendicular by the same angle, but it is on the opposite side of the perpendicular. With this experimental system one can examine inhomogeneities in the elastic properties of materials that result, for example, from compositional changes in alloys.



CRITICAL ANGLE for internal reflection, an important concept in the understanding of the physical basis of the special properties of the images formed in acoustic microscopy, can be visualized in simplified terms with the aid of this illustration, in which the incident and reflected waves are represented as having plane wave fronts. When such a wave strikes a flat surface at right angles (*a*), it will be reflected in the opposite direction with a phase that is equal to that of the incoming wave. (Plane wave fronts are represented here by alternating solid and broken white lines; the solid lines denote crests of waves, the broken lines troughs.) When the direction of the incoming plane wave is tilted with respect to the perpendicular by a certain angle (*c*), an internally reflected wave will be excited along the surface of the reflecting material and the phase of the reflected wave will be altered. (In this case the reflected wave is shown exactly a half cycle out of phase.) At angles less than or greater than the critical angle (*b, d*) the waves are reflected in phase. In acoustic microscope internally reflected waves travel along all surfaces.

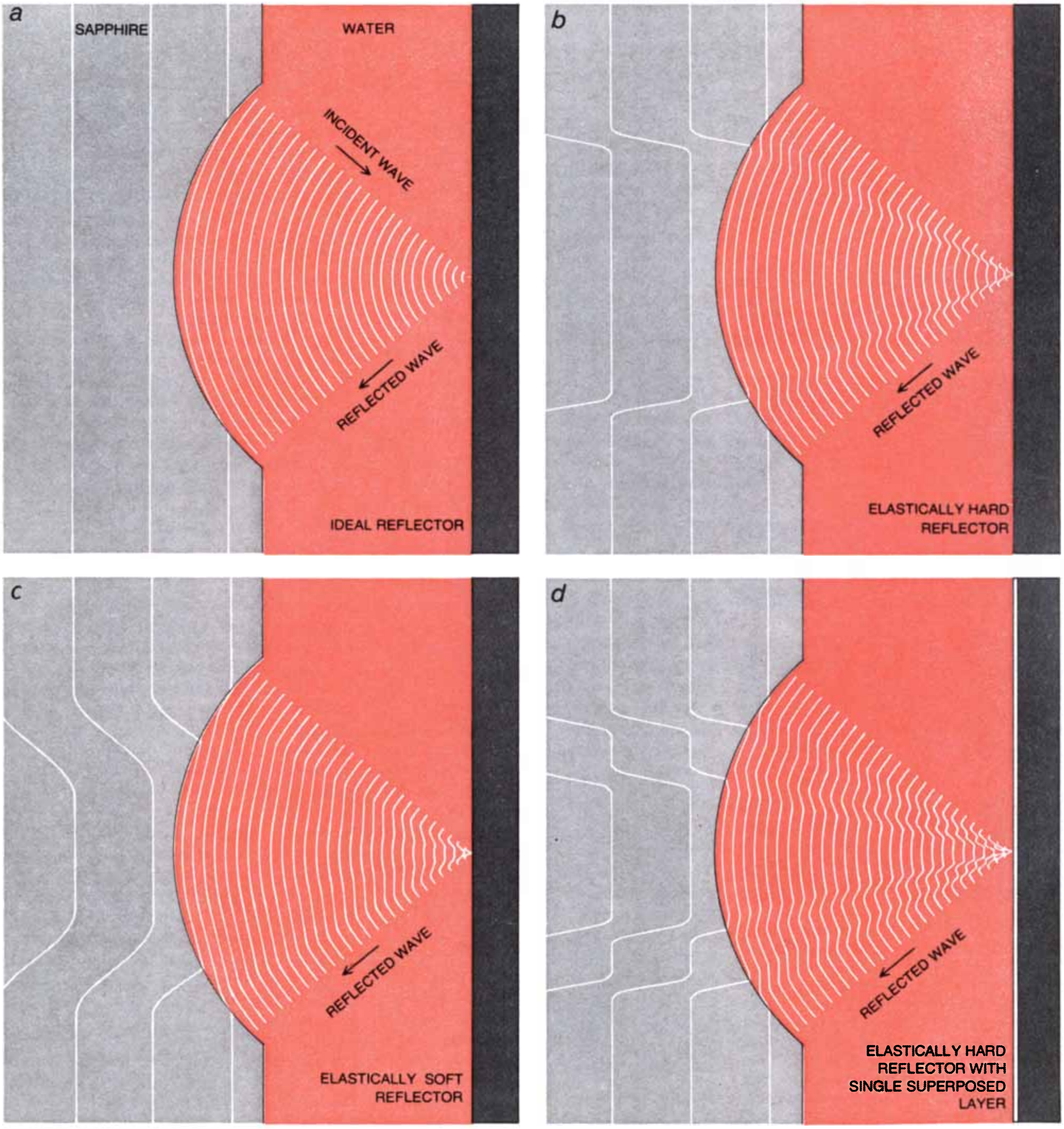
As I have indicated, in our acoustic microscope the incoming ultrasound waves are spherical and converging, and the contrast that appears in the images results from the complex details of the reflection process for spherical waves. In simplified terms, the incoming wave fronts are reflected in phase except for the sector of the wave near the critical angle; because of total internal reflection that sector is reflected with a different phase. The consequences of this subtle shift in phase are substantial.

Consider the wave fronts after they have been reflected [*see illustration on opposite page*]. Assume that at the transition in phase near the critical angle for elastic reflectors the wave fronts of the reflected wave are displaced through a full wavelength from one crest to the next. For a hard material such as sapphire the transition is abrupt and the reflected wave fronts remain nearly spherical. For a soft material the transition is gradual and the reflected wave fronts are distorted; they bear little relation to the original spherical waves. In the case of a substrate with a single superposed layer there are two reflecting interfaces and two critical angles, the first angle influenced by the material of the superposed layer and the second angle influenced by the material of the substrate. Since the reflected wave fronts must be transposed through two successive wavelengths, the distortion is even larger than it is with a single soft surface.

(The waveforms in the illustration have been simplified for the sake of clarity. Actually the amplitude of the reflected wave is never equal to the amplitude of the incident wave, and the actual behavior of the waves is accordingly more complex. Although we have not measured the actual waveforms, we are sure this model conveys the essence of the matter.)

How do we detect the distorted wave fronts? The piezoelectric transducer does it for us. A thin-film piezoelectric transducer with a broad dimension that is much greater than a wavelength acts as a phase-sensitive detector. Its response will be greatest for plane waves with wave fronts parallel to the plane of the transducer; in that situation the entire area of the transducer will be excited in phase. If the wave fronts are tilted, some regions of the transducer film will be exposed to components of the variable-phase wave that will cancel one another. The lens will convert the reflected spherical wave fronts into plane waves if their center of curvature coincides with the focal point of the lens. If the wave fronts are distorted, the plane waves approaching the transducer will be correspondingly distorted. Any variation in the shape of the wave fronts will alter the response of the transducer. It is this variation in the detected signal, originating with inhomogeneities in the elastic properties of the object, that serves as the source of contrast in the acoustic images.

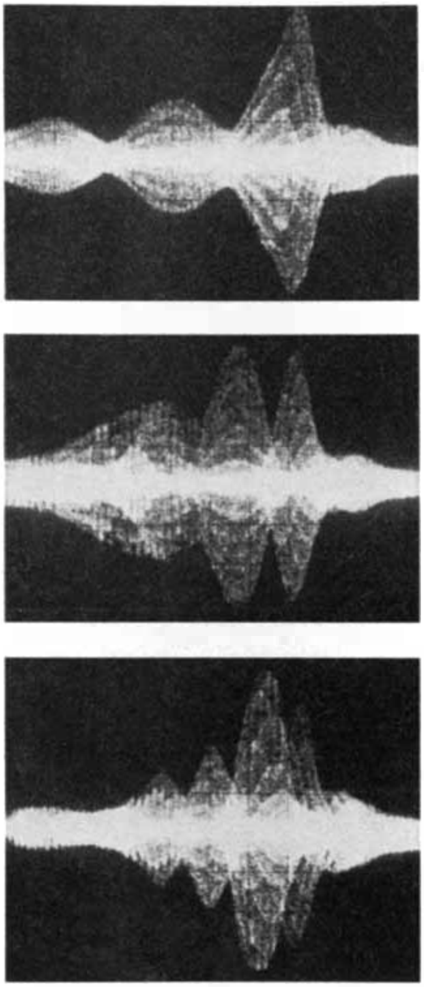
Between the lens and the focal point the incident acoustic beam is, as I have indicated, spherically converging, but beyond that point the wave fronts are spherically diverging. The curvature of the incident wave front at the surface of the reflecting object changes as the spacing between the lens and the object is varied. A converging wave will be reflected with the phase changes I have described. At the focal point the incident wave will be neither converging nor diverging. Beyond the focal point the diverging wave will have a different character in reflection. Oscilloscope traces showing the transducer response as a function of the spacing between the lens and the reflecting object are revealing on this point. Abdullah Atalar of our group has shown, for example, that materials with different elastic constants will pro-



SPHERICAL CONVERGING WAVES of ultrasound formed by the sapphire-water lens in the acoustic microscope are shown after being reflected from four different reflecting surfaces. In the case of an ideal reflector (*a*) the diverging reflected wave fronts are exactly coincident with the converging incident wave fronts. The other three surfaces reflect the incoming waves in phase except in the sector of the wave near the critical angle; because of total internal reflection the waves in that sector are reflected with the opposite phase. In the case of an elastically hard material (*b*) the full-cycle transition at the critical angle is abrupt and the reflected wave fronts remain nearly spherical. In the case of an elastically soft material (*c*) the transition is gradual and the reflected wave fronts are distorted. In a substrate with a single superposed layer (*d*) the distortion is even greater. (Actually behavior of reflected waves is more complex than is shown here.)

duce strikingly different oscilloscope traces. The difference in the critical angle for total internal reflection accounts for the difference in the oscillating portions of these traces.

The transducer response for a substrate with a single superposed layer produces a particularly interesting oscilloscope trace. For example, there is a dramatic change in the shape of the trace for a pure silicon surface when a one-micrometer layer of aluminum is deposited on the substrate and an even greater change when a two-micrometer layer is deposited [*see illustration on this page*]. In this example there is a lens-to-object spacing where there is a peak in the response for a two-micrometer layer and a valley in the response for a one-micrometer layer. With this method small changes in the thickness or the elastic properties of a thin film are obviously easy to monitor. Robert G. Wilson and Rolf D. Weglein of the Hughes Research Laboratories have found, for instance, that the acoustic process gives rise to striking variations in contrast in images of multilayered objects such as microelectronic circuits.



AMPLITUDE of the reflected signal from the acoustic microscope is displayed on the screen of an oscilloscope as a function of the spacing between the lens and the reflecting surface. At the peak of each of these traces the reflector is positioned at the focal point of the lens. To the left of the peak the reflector is exposed to a converging beam; to the right the reflector is exposed to a diverging beam. The differences in the shape of the oscillating segments of these traces arise from differences in the elastic properties of the reflecting surfaces. The top trace is for a pure silicon surface. The middle trace is for a silicon surface covered with a one-micrometer layer of aluminum. The bottom trace is for a silicon surface with a two-micrometer layer of aluminum. Exceedingly small changes in thickness can easily be detected in this way.

As I have mentioned, the operating wavelength of the best acoustic microscopes is about a micrometer, not too different from the wavelength range of visible light. As far as I can tell this is sheer coincidence. If the limiting wavelength for the sound system had been 10 micrometers, acoustic microscopy would have attracted little attention except for specialized applications. If the limiting wavelength had been .1 micrometer, acoustic microscopy would have had the entire field to itself. The role of a microscope is defined by, above all other factors, the limiting resolution. What determines this limit for the acoustic microscope? It is the absorption of sound along the acoustic path in the liquid. Liquids in some form must be used if the object is to be moved with respect to the lens. Although liquids make short wavelengths possible, they do absorb energy. In many liquids the molecules have vibrational modes that are excited by high-frequency sound waves. The absorption for these liquids is so high that their use is out of the question.

So far the preferred liquid has been water. Fortunately water is compatible with living biological specimens. The velocity of sound waves in water is about 1,500 meters per second, and the absorption is less than it is in most other liquids, although it does increase at the higher frequencies.

Argon is another attractive liquid for acoustic microscopy. Its acoustic properties are similar to those of water, except for the velocity of sound, which in argon is 60 percent of its value in water. Argon has one disadvantage: it exists in the liquid state only in a narrow temperature range near 85 degrees Kelvin (minus 188 degrees Celsius). It is therefore a cryogenic fluid, and special equipment must be designed for instruments that operate with such fluids. This has been done for the acoustic microscope, and Daniel Rugar of our group has obtained high-quality images with the system.

Liquid argon is also important as an intermediate stage in the transition from water to liquid helium. The absorption of sound in the liquid phase of the isotope helium 4 is so small that it becomes negligible. The velocity of sound in this medium is one-eighth that of water. Joseph E. Heiserman of our group is currently planning an acoustic microscope with liquid helium as the sound-transmitting medium. He is confident that the resolving power will be superior to that of the optical instrument.

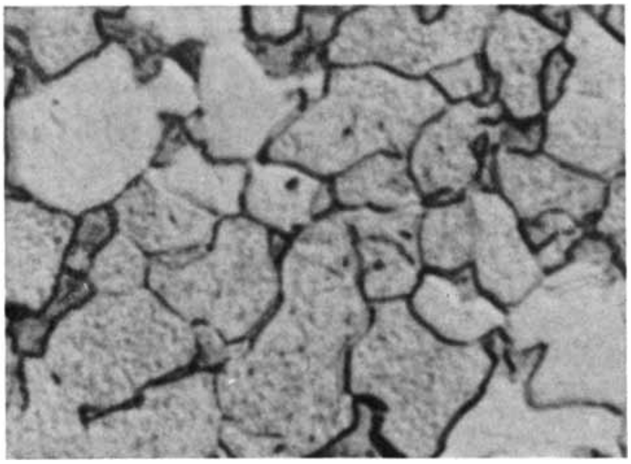
All this background information is of course prologue to the images themselves. In the past visitors to our laboratory were first shown the images. They were then asked if they wanted to see the instrument. This approach often saved us a lot of time. The acoustic micrographs were at that stage inferior to optical micrographs, and they were looked on as merely a curiosity. Visitors seldom asked to tour the laboratory and see the instrument in operation. That is no longer true. The improved quality of the images has generated an intense interest in the instrument, and we now have to take the time to explain the details of its operation.

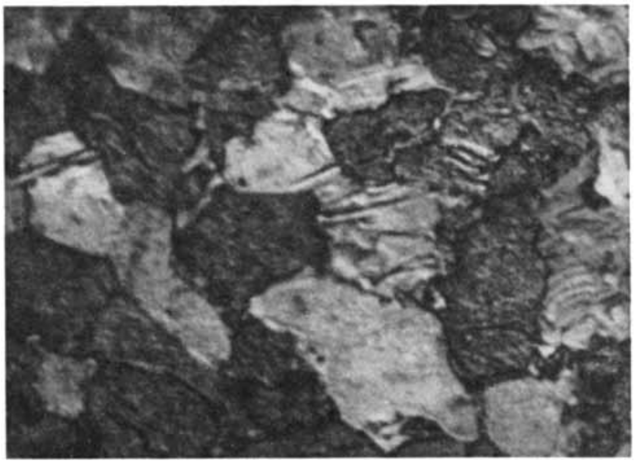
Nevertheless, the essence of any program for microscopy is still to be found in the images. The acoustic micrographs accompanying this article are, we believe, competitive with their optical counterparts. The images fall into three categories: materials, microelectronic circuits and biological cells. The common element is size: all the objects have details with dimensions measured in micrometers.

Metals such as steel look uniform to the unaided eye, but they are actually quite complex structures, consisting of many individual crystal grains cemented together along the grain boundaries. An acoustic micrograph of a steel surface looks quite different from its optical counterpart [*see upper illustration on opposite page*]. The greater contrast in the acoustic image gives a better idea of the relative orientation of the crystals at the surface. Here the surface has been polished and etched so that the grain boundaries appear clean in the optical micrograph, but this step is not necessary with the acoustic system.

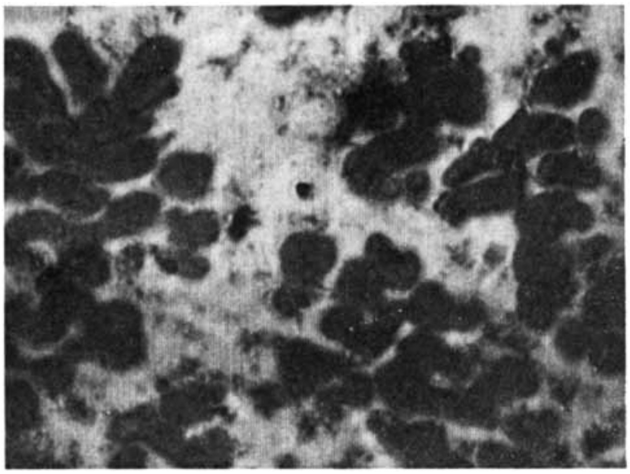
Alloys have a different kind of inhomogeneity. In a cobalt-titanium alloy, for example, the mixture can condense in several phases, each with a different ratio of cobalt to titanium. The phases differ in their elasticity, and the difference accounts for the source of contrast in the acoustic image [*see lower illustration on opposite page*]. The properties of such alloys depend critically on the size and distribution of their phases.

Microelectronic circuitry is another good source of objects for acoustic micrography. The architectural details of these devices often show up with better contrast than they do in the rather uniform-looking optical micrographs [*see top illustration on page 70*]. Moreover, with the acoustic instrument one can ob-





STEEL SURFACE seen in an acoustic micrograph (*right*) is quite different from its optical counterpart (*left*). The greater contrast between the individual crystal grains appearing in the acoustic image is attributable to slight variations in the elastic properties of the grains.



ALLOY of cobalt and titanium is seen in these matching micrographs, one made with an optical microscope (*left*) and the other with an acoustic microscope (*right*). The splotchy forms in the acoustic image are regions where the alloy has a different composition.

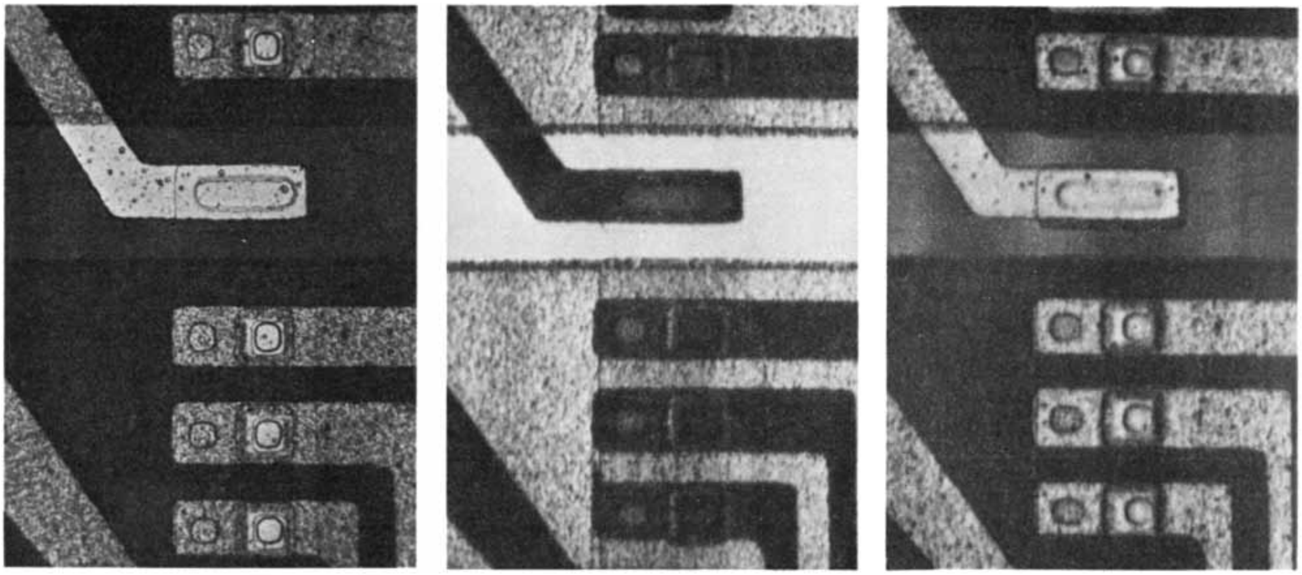
tain several different images of the same object, each corresponding to a different spacing between the object and the lens. Each image provides a new level of detail. This "in depth" information has attracted more outside interest in our research than any other feature.

In the biological world there are many cellular structures on the scale of micrometers. They are routinely viewed in the optical microscope, but not without difficulty. Often there is a lack of visible contrast. The problem is that the structures of the cell are quite transparent: they absorb only a tiny amount of light. Extraordinary efforts have been made to compensate for this deficiency and increase the contrast. The selective staining of cells is a common method. It is a highly developed art and it adds color to the images, but it is time-consuming and cannot be done with living cells. The optical techniques associated with phase-contrast microscopy and polarized-light microscopy make it possible to introduce subtle changes in a way that increases the contrast, but there is still room for improvement.

Our innocence in things biological at the beginning of this part of our investigation can be conveyed by an anecdote. When we first inquired about obtaining white blood cells from a biology laboratory at Stanford, we were asked how many cells we needed. We wanted to make only a few images, and so we asked for 10 cells. That response could have come only from nonbiologists. We were told that the minimum number of cells that could be separated conveniently was a million, and all of them could be placed on a single glass slide! In many instances it would be desirable to study the entire assemblage of these cells, but that is no easy matter. The cost of the photographic film required to record the images of all the cells would be staggering. One such exposure costs five cents, and it is common practice in biology laboratories to record 20 images of each cell. The cost of the film therefore runs about $1 per cell. Even if one were to study only 1 percent of the cells on the slide, the cost ($10,000) would be prohibitive. With the aid of special circuits connected to our acoustic instrument this cost can be reduced. We shall soon be able to record such images directly onto videotape. The playback time in a videocassette recorder (at a rate of 30 frames per second) would be one hour for 10,000 cells, and the cost of the tape would be less than $20.

So far we have made acoustic micrographs of three main classes of biological objects: lymphocytes (white blood cells), fibroblasts (connective-tissue cells) and chromosomes. The fact that we have obtained such good images of various cellular structures without staining suggests that we shall be able to see some of the same structures inside living cells.

The images we have chosen to illustrate this article are in many ways ideal for acoustic microscopy, and they demonstrate what can be done today.

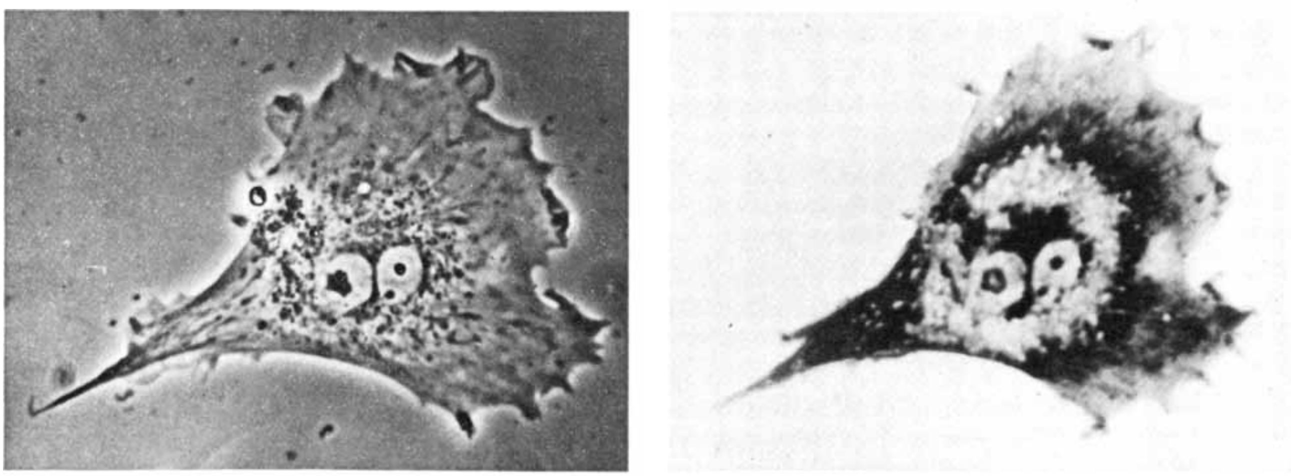


MICROELECTRONIC CIRCUIT ELEMENTS fabricated on the surface of a silicon "chip" are portrayed three times, once in the form of an optical micrograph (*left*) and twice in the form of acoustic micrographs made with the object at different focal positions (*center and right*). The device consists of several interconnected transistors. The fingerlike extensions contacting the transistors are strips of vapor-deposited aluminum film. The smooth band across the upper part of each image is the pure silicon substrate; the pebbly texture elsewhere is caused by superposed layers of polycrystalline silicon. The reversal of contrast in the acoustic images assists in the identification of detailed structures on the chip. The device was supplied by Roderick D. Davies of the Integrated Circuits Laboratory at Stanford University.

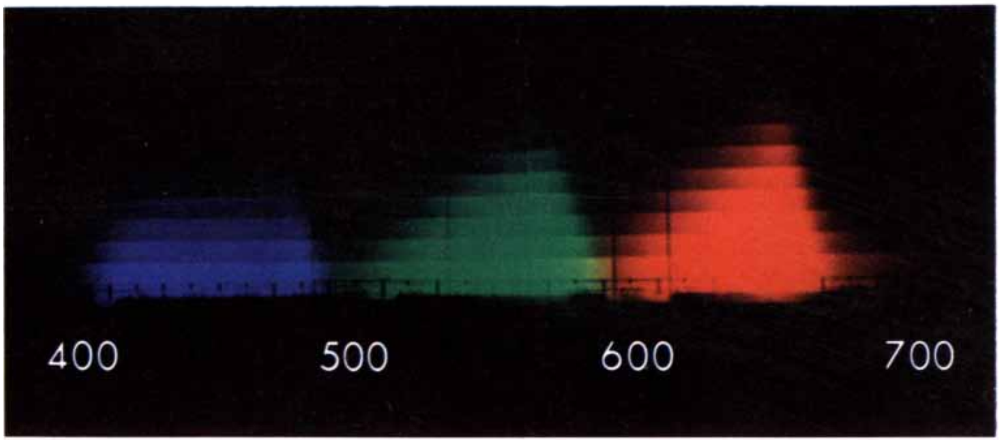
They are very different from the images that were obvious choices to us when we undertook this line of inquiry. As we look back we realize that those early examples (such as the bonding between opaque materials) were simple extensions of the work that was being done with acoustic imaging devices at lower ultrasonic frequencies. The examples shown here were not included in our original list. They evolved in a way that was both natural and unexpected.

We felt sure at one time that the reflection of sound waves from a pure sapphire surface could be determined only by the contour of the acoustic beam; it would be at a maximum with the reflector at the focal point, and it would decay at a rate that conformed to the profile of the beam as the reflector was moved away from this plane. Little did we suppose that the elastic properties of the reflecting surface were encoded in the diverging waves. At another time we worked hard to perfect an acoustic microscope of the transmission type. We never expected that images in reflection, particularly of soft biological specimens, could be so crisp.

Our future research program is laid out for us in the form of a series of questions. How far will we be able to go with cryogenic fluids? Will it be possible to find a noncryogenic, low-absorption liquid that will enable us to achieve shorter wavelengths and higher resolutions at room temperature? Will we be able to penetrate the liquid-solid barrier and focus the microscope beam inside solid materials? Will we be able to image microscopic patterns of stress in such materials? Will living cells thrive in the acoustic microscope and yield information important to biologists? It is as though we had just begun.



SINGLE CELL, a fibroblast cell obtained from the connective tissue of a chick embryo, is represented by an optical image (*left*) and a corresponding acoustic image (*right*). The cell was grown on a collagen-coated glass slide and was fixed with methanol by Randal N. Johnston of Stanford. Both micrographs reveal the network of internal fibers that holds the cell together. The contrast in the optical image results from a selective-staining technique; the contrast in the acoustic image results from variations in the elastic properties of structures in the cell. The fact that the acoustic technique works without staining suggests that it may be possible to visualize such structures in living cells.



Kodak instant color film really is different.

See why it catches small hue differences for technical work?

Note how little the spectral responses of the three color-sensing emulsions overlap in the wedge spectrogram.

Picking the spectral responses we wanted for our film was a basic strategic decision. There is no one correct choice. No technology can perfectly reproduce your personal color perceptions from every scene and subject. A film manufacturer can only aim to maximize happiness among the users of the product. One manufacturer's film almost always gives a different rendition from another make of film.

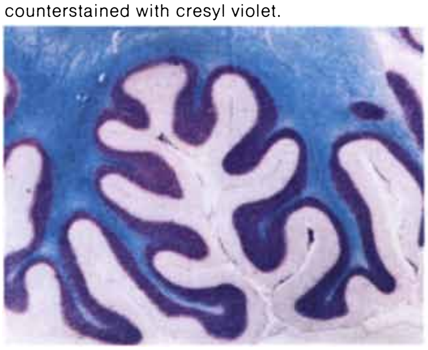
For technical photography personal taste generally means less than does ability to record small color distinctions.

If you would like good color differentiation, try KODAK Instant Color Film. We now use carriers that are more active than were in our initial product. (Those are the molecules that release the image dyes to the image-receiving layer in accordance with exposure to the three emulsions.) More active carriers permit thinner layers. This buys several improvements in performance, including sharpness, which now justifies the use of a hand magnifier in studying prints.

Do not be surprised to find that the outfit from whom your microscope or other photographic recording instrument was purchased has been expecting you to bring up the subject of adapting to KODAK Instant Color Film.

Kodak

Cat cerebellum. Stain: Luxol fast blue counterstained with cresyl violet.



Tongue, foliate papillae with taste buds. Stain: Hematoxylin and eosin.

Cat brain stem. Stain: Davenport's silver for nerve fibers.

Photos smaller than actual size.

Martin Scott, Scientific Photography, Kodak, Rochester, N.Y. 14650, who did these photomicrographs on Kodak instant color film PR10, would rather have you judge their quality from an original he wants to make for you at the microscope. Drop him a note if interested.

SCIENCE AND THE CITIZEN

Couples Want Fewer Children

A dramatic decline in fertility is under way in many parts of the underdeveloped world. The downturn has been suggested for several years by United Nations estimates based on vital statistics, but for many countries such data are considered unreliable. Now a series of 15 single-country reports on carefully designed household surveys based on national probability samples has documented the trend. The downturn is clearly evident in 12 of the countries and is indicated for two more. The primary cause of the decline is a massive change in individual aspirations: couples want fewer children. There is still a significant gap, however, between family-size preference and family-planning capabilities. Most women have heard about effective contraceptive methods, but on a worldwide basis only a minority of them are practicing contraception.

The new data are the first findings of the World Fertility Survey, an international research program being coordinated by the International Statistical Institute with UN support. The survey, perhaps the largest social-science research project ever undertaken, has currently enrolled 41 underdeveloped countries and 20 developed ones; more countries are expected to join the program. Individual governments conduct their own national surveys, working with standardized questionnaires and methods so that valid international comparisons of data can be made, with technical and financial assistance from the central survey organization for the underdeveloped countries. The first national survey was undertaken in Fiji in 1974. Now the findings of the first 15 national reports (for nine countries in Asia and the Pacific and six in Latin America) have been summarized by Maurice Kendall, the director of the World Fertility Survey, in *Population Reports,* a publication of the Population Information Program of Johns Hopkins University.

The salient finding of declining fertility is derived from a direct comparison of past fertility (the average number of children born to surveyed women aged 45 through 49) and the current total fertility rate: the average completed-family size expected if the present fertility rates among women in each age group are maintained. The largest change was recorded for Costa Rica, where women now 45 through 49 had an average of 7.2 live births and younger women are expected to average only 3.8 births by the end of their reproductive years: a decline of 47 percent in family size. Almost equally impressive declines were registered in Thailand, Sri Lanka and Colombia. To some extent the decreased fertility is the result of later marriage: in all but four of the countries the mean age at first marriage was found to have been higher for women 25 through 29 years old than for those 40 through 44.

The major reason for the declining fertility, however, is apparently simply that people want to have smaller families. In 10 of the 15 countries more than half of the married women of reproductive age reported that they want no more children; in six of those countries the figure was more than 60 percent. One analysis has shown that the desire to have no more children is substantial in all socioeconomic groups; for women with four children the difference between urban and rural women, between educated and uneducated ones and between those married to men with high-level occupations and those married to men with low-level ones is less than 10 percentage points in most of the countries. This suggests, according to Kendall, that the desire to stop having children "may be less influenced by socioeconomic development" than has been thought.

There is extensive awareness of family planning. In all but one of the countries at least 75 percent of the married women have heard of a method of contraception, in most cases a method that is considered effective; in all but three countries the best-known means of contraception is the pill. There is a contrast, however, between awareness and practice. Among women who are or have been married the fraction who have ever practiced contraception ranges from 82 percent in Costa Rica down to 4 percent in Nepal. In Bangladesh and Pakistan 82 and 75 percent respectively of the ever-married women have heard of a contraceptive method but only 14 and 10 percent respectively have ever used one. Even among "exposed" women (currently married, of reproductive age and not pregnant) who specifically state that they want no more children, there is clearly a large "unmet need." In Nepal, Bangladesh and Pakistan 90 percent or more of such women are not practicing any modern means of contraception; in most of the other countries the fraction is more than 50 percent. Vast numbers of women apparently either do not know of any sources of family-planning services or have no easy access to such sources. Kendall concludes that "increasing knowledge and access to family-planning services can contribute substantially to reducing the present gaps between knowledge and practice, between the number of women who want no more children and the number who use effective methods of family planning."

Wet Solar Cells

The National Photovoltaics Act of 1978 authorized the expenditure of $1.5 billion over the next 10 years for research, development and demonstration of photovoltaic (solar-cell) systems for converting sunlight into electric power. In fiscal year 1979 the Department of Energy will spend more than $100 million on solar-cell research and development under its Photovoltaic Systems Program. The objective is to lower the initial cost of solar-cell modules capable of generating one kilowatt of electricity at high-noon illumination from roughly $12,000 today to $2,800 by 1982, to $700 by 1986 and to between $500 and $150 by 1990. The goal for a complete photovoltaic system in 1990 is between $1,500 and $1,000 per kilowatt, a cost at which solar power could begin to compete with fossil-fuel and nuclear power.

The conventional solar cell is a solid-state device in which a junction is formed between single crystals of silicon separately doped with impurity atoms in order to create *n* (negative) regions and *p* (positive) regions, which are respectively receptive to electrons and to "holes" (the absence of electrons). Light falling on the cell creates electron-hole pairs. Electrons migrate preferentially from the *p* region to the *n* one; holes move in the reverse direction. The electrons are prevented from recombining with the holes until they have traveled through an external circuit. Single-crystal solar cells made of silicon have achieved an efficiency of about 20 percent. Single-crystal cells made of gallium arsenide, which are more expensive to manufacture, have reached an efficiency of 26 percent. Costs could be much reduced if solar cells could be made from thin films of polycrystalline materials rather than from single-crystal ones, but such a substitution cuts the efficiency of the cell to 10 percent or less.

A different type of photovoltaic cell in which the junction is formed between a semiconductor and a liquid electrolyte has recently stimulated much effort at laboratories in the U.S., West Germany, Israel and Japan. Such cells, known as electrochemical photovoltaic cells, promise to be easier to fabricate than all-solid-state cells. They are simpler than conventional cells because no doping of the semiconductor is required; a junction forms spontaneously when a suitable semiconductor, such as gallium arsenide, is brought in contact with a suitable electrolyte. Originally such


LOCHAN ORA
LIQUEUR
IMPORTED
LOCHAN ORA
LIQUEUR
CHIVAS BROTHERS LTD.
Hardly your everyday liqueur.
Some things in life are too precious for every day. Lochan Ora is one. An imported golden liqueur with a unique taste all its own, it's to be lingered over—sipped slowly, savored fully. You'll be proud to share it—but not with just anybody.
Lochan Ora. Imported Liqueur.
70 PROOF • IMPORTED BY GENERAL WINE & SPIRITS CO., N.Y., N.Y.

The Hollis Hedberg. Gulf's own floating computer center.

"This computer center happens to be on the *Hollis Hedberg*," says Gulf Research Geophysicist John McDonald. "It's the only research vessel in the industry that has complete data processing right on board. And that makes all the difference.

"The typical oil exploration ship is equipped only to record raw data from whatever equipment it has on board to detect oil deposits under the ocean floor. Usually they

"On-board computers help you learn a lot more, a lot faster."

have seismic sounding gear, occasionally magnetometers and gravity meters. The raw data gets sent to an onshore computer center for analysis, and by the time analysis is complete, the ship could be a thousand miles away.

"By contrast, the *Hedberg* has a full complement of recording equipment, including hydrocarbon 'sniffers,' and with computers right on board, we can make a preliminary analysis immediately. If it looks good, we go back for a second look at that location.

"It could take five or six years between the time you find an oil deposit and the time you actually start drilling for oil, so any time we save in exploration puts the country that much closer to a new supply of petroleum.

"It's a real challenge, trying to find that invisible spot under the seabed that's likely to produce oil. But the *Hedberg* is one of the best ways there is of finding it."

Gulf

Gulf people: meeting the challenge.

Gulf Oil Corporation

cells looked unpromising because they were subject to three major problems, all of which have recently been solved.

The first problem was that the electrolyte rapidly oxidized the surface of the semiconductor. It has been solved by selecting electrolytes that oxidize faster than the semiconductor does. The second problem was that ions were exchanged between the semiconductor and the electrolyte, forming a layer on the surface of the semiconductor that blocked the flow of charge carriers across the junction. It has been solved by selecting a semiconductor and an electrolyte with common ions, for example cadmium selenide in contact with a selenide or sulfoselenide solution. The third problem was the deposition of ions of impurities on the surface of the semiconductor. Such impurities introduced new surface states that short-circuited the junction, causing electrons and holes to recombine. Attempts were made to use superclean solutions, but the problem was finally solved by introducing special impurities into the electrolyte that are strongly adsorbed ("chemisorbed") on the surface of the semiconductor. The impurities cover the surface of the semiconductor with a stable ultrathin layer that does not block the movement of charge carriers. An effective substance for covering the surface of gallium arsenide is the ruthenium ion released from aqueous trivalent ruthenium solutions.

The three problems were solved primarily through an international exchange of ideas among workers in a small group of laboratories: Mark S. Wrighton of the Massachusetts Institute of Technology, Adam Heller of Bell Laboratories, Joost Manassen of the Weizmann Institute of Science in Israel, Heinz Gerischer of the Fritz-Haber Institute of the Max Planck Institutes in Berlin, Kenichi Honda of Okayama University in Japan and Allen J. Bard of the University of Texas at Austin. At Bell Laboratories, Heller, Barry Miller and their colleagues have recently demonstrated single-crystal gallium arsenide semiconductor-liquid photovoltaic cells with an efficiency of 12 percent and polycrystalline cells of the same type with an efficiency of 7 percent. Heller emphasizes, however, that because the present cells contain toxic substances (such as selenium) they are not currently candidates for large-scale production. Their significance lies in demonstrating that formidable problems associated with electrochemical photovoltaic cells, which are simple and potentially cheap, can be solved.

Such cells also lend themselves to configurations in which the cell could become an integral part of a storage system. In one mode of operation the cell might directly decompose water into hydrogen and oxygen, which could be stored and recombined in a fuel cell to generate electricity as it is needed. Other molecular species, such as hydrogen bromide, could be similarly decomposed and reconstituted. Another possibility is to add a third electrode to the cell to form a battery in which energy could be stored.

Aboriginal Aborigines

Recent fossil discoveries in Australia have pushed the date of man's arrival on the island continent substantially further into the past and have added a hitherto unknown element to the gene pool of the Australian aborigine. Twenty years ago the aborigines were thought to have reached Australia only a few thousand years ahead of the Europeans. This view was overturned in the 1960's by the discovery at Kow Swamp in the eastern Australian state of Victoria of a skull that was between 10,000 and 15,000 years old and aborigine in physical type. Since then fossil aborigine remains as much as 20,000 years old have been found at sites along the Murray River in Victoria. Human fossils have now been found buried in the sands of a dry-lake region in western New South Wales only 50 miles north of the Murray River. They appear to be more than twice as old as the Kow Swamp skull.

The new fossil evidence, collected in the Willandra Lakes area in general and at one dry bed, Lake Mungo, in particular, consists of the remains of 50 individuals. Carbon-14 tests of hearth charcoal, bone and shell indicate a maximum age of 30,000 years for the remains. Both living and fossil aborigines are characterized as robust: they have a heavily built skeleton and a long, low skull with a projecting face, a slanting forehead and large teeth. The Willandra lakes people were in contrast gracile: they had a lightly built skeleton and a small, delicately formed skull with a flattish face, a rounded forehead and small teeth. In these respects their remains are reminiscent of a skull discovered decades ago in Niah Cave in Borneo and believed to be some 30,000 years old, and also of gracile human fossils unearthed near Wadjak in Java.

The occupational debris associated with hearths at Willandra Lakes indicates that the inhabitants were hunters and fishermen. They caught the large Murray River cod, perch and crayfish, gathered mussels in quantities sufficient to build up middens of shells, hunted small kangaroos and collected the eggs of the emu. They buried their male dead in the sand, sometimes sprinkling red ocher over the body. The two female burials uncovered give evidence of a far more complicated procedure. The body was first partially cremated on a funeral pyre. Then after a time the bones were collected, smashed and burned again. The remains from the second burning were buried in small pits.

Where did these people come from? The principal investigator, Alan G. Thorne of the Australian National University, suggests that their original homeland was in mainland Asia, possibly as far away as southern China. What became of them? When the Willandra Lakes began to dry up some 15,000 years ago, the inhabitants of the area were the aborigines. Thorne suggests that by then the gracile people had commingled with their robust neighbors, who may have entered Australia several thousand years earlier than the gracile people did. Thus the basic aborigine genetic heritage would have come to include an admixture of the ancient Willandra Lakes strain, just as recent aborigines have interbred with migrants from Indonesia and New Guinea.

Sixty-Cycle Chorus

The radiation emitted by power-transmission lines leaks into the earth's magnetosphere and triggers very-low-frequency (VLF) electromagnetic disturbances, write J. P. Luette of Sandia Laboratories and C. G. Park and R. A. Helliwell of Stanford University in *Journal of Geophysical Research*. It has long been known that natural processes are responsible for some VLF-wave disturbances, which have been given such names as whistler, hiss and chorus because of the way they sound on radio at audible frequencies. A whistler lasts a second or two and sweeps (usually downward) in frequency from about 30,000 hertz (cycles per second) to a few hundred. Some whistlers have proved to be signals generated by a lightning stroke in one hemisphere of the earth and conducted to the opposite hemisphere through a natural waveguide formed by the lines of force of the earth's magnetic field. Hiss, which sounds like static, can originate with spontaneous plasma instabilities in the magnetosphere. Only recently, however, has VLF-wave activity been traced to a man-made source.

The radiation from power-transmission lines interacts with energetic electrons that have been trapped by the magnetic field of the earth. The dynamics of this wave-particle interaction can boost the power of the radiation by a factor of 1,000 and can propel the interacting electrons into the top of the atmosphere, where they lose energy in collisions with gas molecules. Satellite surveys of VLF-wave activity link the power-line radiation with chorus emissions, which are rising tones whose amplitude starts at the background noise level but increases rapidly at a rate of between 200 and 2,000 decibels per second. The surveys reveal that the starting frequencies of many chorus emissions over North America are harmonics of 60 hertz, which is the fundamental frequency of power transmission in North

America. Over Europe, where power is transmitted at 50 hertz, the starting frequencies of many chorus emissions are harmonics of 50 hertz. And over the Pacific between Alaska (which transmits power at 60 hertz) and New Zealand (which transmits it at 50 hertz) the starting frequencies are chiefly harmonics of both 60 and 50 hertz.

The connection between power-line radiation and chorus emissions is compelling, but does the radiation create the emissions or merely control them? To put it another way, does power-line radiation trigger chorus emissions under conditions that are not conducive to the spontaneous generation of the emissions or does the radiation simply establish the starting frequency of emissions that would have occurred even in the absence of the radiation? The evidence seems to point to the first alternative. Chorus emissions are more prevalent over industrialized areas, which of course consume the most power. Moreover, observations made in Antarctica show that chorus emissions reach a pronounced low on Sundays, when power transmission is at a minimum.

Clean Coal

A major problem associated with the shift to coal as a source of energy is controlling the noxious substances it emits when it is burned, chiefly sulfur dioxide (SO_2), oxides of nitrogen (NO_x) and various particulates. The standard approach has been to remove or diminish the substances in the exhaust gas of coal combustion. Much work is being done to improve this approach and to explore three others: cleaning the coal before it is burned, improving the efficiency of combustion so that fewer oxides and particulates are produced and transforming the coal into a gaseous or liquid fuel while simultaneously removing potential pollutants.

The idea of cleaning coal before it is burned is quite old, but it has not been widely applied because it makes the fuel considerably more expensive. Now the combined effect of diminishing coal quality (because of more intensive mining) and stricter environmental regulations is improving the economics of coal cleaning. The prospects are described in *EPRI Journal,* a publication of the Electric Power Research Institute.

The first step is to break up the coal to separate it from rock and other debris. Then the coal is cleaned, usually by the principle of gravity separation, which takes advantage of the fact that coal has a lower specific gravity than most of its impurities, so that the impurities settle out faster. With improved techniques, according to *EPRI Journal,* cleaning could remove up to 65 percent of the ash and from 30 to 35 percent of the sulfur in coal.

The principal avenue of exploration in the effort to improve the efficiency of combustion is the fluidized-bed technique, in which a gas (such as air) is forced up through a bed of fine coal, causing it to rise and become suspended in a virtually fluid condition. The technique is efficient in transferring heat and can be coupled with methods of controlling sulfur dioxide and nitrogen oxides. It therefore offers a way of burning low-quality coal without preprocessing and with little or no need to clean the exhaust gases.

The usual method of extracting sulfur in the fluidized-bed combustion of coal is to add limestone to the coal. The sulfur dioxide that is formed by combustion reacts with the calcium in the lime to form calcium sulfate. Workers at the Brookhaven National Laboratory have been experimenting with Portland cement as a substitute for the limestone. Cement is a calcium silicate, and both the calcium and the silica participate in removing sulfur. The method is advantageous because the cement can be reused a number of times.

Workers in the Energy Laboratory of the Massachusetts Institute of Technology are also looking into fluidized-bed combustion to see if it can be modified to achieve further reductions in the emission of nitrogen oxides. The opening is provided by the fact that the oxides are formed at the highest rate near the base of the bed, where the fluidizing air enters; as the combustion products rise, reactions occur that destroy part of the nitrogen oxides and lead to the formation of molecular nitrogen (N_2). The work is described in *e-lab,* a publication of the laboratory.

Another project under way at the M.I.T. laboratory has to do with pyrolysis, a method of obtaining clean gaseous and liquid fuels from coal. In pyrolysis coal is heated in the absence of air until it decomposes chemically, yielding a mixture of gaseous and liquid fuels and char, a solid residue. The laboratory workers have been exploring hydropyrolysis, which is pyrolysis in the presence of hydrogen. According to *e-lab,* "gaseous molecular hydrogen reacts chemically with the coal to give larger quantities of gaseous and liquid fuels" than are obtained with conventional pyrolysis; moreover, "the liquid hydrocarbons are lighter, the solid residue is reduced and the overall reaction occurs more quickly."

As for improving the techniques employed to remove noxious substances from the exhaust gas of coal combustion, consideration is being given to adding nitrogen-removing devices to the procedures that now "scrub" sulfur dioxide (usually by means of an aqueous solution of lime or limestone) from stack gases and trap particulate matter in the stack (usually by electrostatic precipitators). *EPRI Journal* describes "dry" and "wet" processes for reducing the stack-gas emissions of nitrogen oxides. Dry processes employ a gaseous reagent that changes nitrogen oxides in the flue gas into molecular nitrogen. Wet processes rely on an aqueous reagent that first absorbs nitrogen oxides and then reduces them for disposal in liquid form.

Mendel Was No Mendelian

The myth of the lone scientist stumbling on a law of nature while puttering in his laboratory is as persistent as it is false. Historians of science have amply demonstrated that science is a complex cultural phenomenon and does not operate in a vacuum: it builds of necessity on the knowledge of the past and the intellectual currents of the present. Nevertheless, the myth of the lone scientist continues to endure in the figure of Gregor Johann Mendel, the Moravian monk who is often credited with single-handedly originating the science of genetics in the 1860's by systematic breeding experiments with pea plants. The traditional view is that Mendel's work was "premature" and hence was forgotten for 35 years before being rediscovered around the turn of the century, when biology was sufficiently advanced for his ideas to be comprehensible. Robert Olby of the University of Leeds has now questioned this interpretation. Writing in the journal *History of Science,* he places Mendel's work squarely in the mainstream of 19th-century biology.

Olby shows that Mendel was working directly in the tradition of well-known plant hybridizers such as Carl von Gärtner. These naturalists studied hybrids because they believed they were a source of new species. By the same token Mendel did not set out to formulate the laws of heredity by studying hybrids; he was investigating the role of hybridization in evolution. His systematic experiments were designed to resolve the then current controversy over whether or not hybrids bred true, that is, gave rise to offspring identical with themselves. If they could be shown to breed true, they might mark the first stage in the genesis of new species. Concurring with the historian L. A. Callender, Olby believes Mendel supported the hypothesis of true-breeding hybrids. Mendel's discovery of segregating traits in pea hybrids must therefore have been a disappointment to him, mitigated only by the finding of true-breeding forms among the hybrid progeny that showed new combinations of characters.

As is well known, Mendel distinguished between those characters that "pass into hybrid association entirely or almost entirely unchanged, thus themselves representing the characters of the hybrid," which he called dominant, and those that become latent in hybrid association, which he called recessive. It was not the genetic elements that Mendel de-

Hitchin has witnessed centuries of history. Now it's making some.

Last year, not far from the heart of historic London, everyday phone calls began to follow a remarkable new route.

Instead of telephone wires, the calls were sent on *beams of light*—over hair-thin threads of glass called optical fibers.

This was no laboratory experiment. The 6-mile-long optical fiber system was created by the people of ITT to handle regular, everyday phone calls. (If you were to make a call between Hitchin and Stevenage, your call would very likely go this way.)

Why try to send phone calls any way but by wire?

Because telephone systems are carrying more and more information nowadays. Not only phone calls, but television and computer data of all kinds.

And for good technical and economic reasons, optical fibers will let telephone companies handle this ever-growing traffic, without continually adding more and more telephone lines.

Lots of people have been saying that for reasons like this, optical fibers are the wave of the future.

For the people of Hitchin, they're already here.

The best ideas are the ideas that help people. **ITT**

fined as dominant and recessive, however, but the physical characters themselves. He did not distinguish explicitly in his writings between genetic elements and characters, and it appears that he did not have the distinction clear in his own mind.

Indeed, Olby presents evidence that Mendel did not understand that pairs of mutually exclusive genetic elements (later called alleles) determine the contrasting characters. Mendel never made any statement about the number of elements per character, and in his writings he used two letters (*Aa*) for the pairs of unlike elements (heterozygotes) but only one letter (*A* or *a*) for the pairs of like elements (homozygotes). Mendel's statement that "only the differing elements are mutually exclusive" is actually in conflict with classical Mendelian genetics. If it were true, the number of like elements determining a character would increase every time the germ cells fused in fertilization. Moreover, Mendel erroneously postulated that three contrasting characters for the same trait could coexist in the same hybrid individual but be mutually exclusive in the germ cells. Ironically the rediscoverers of Mendel's work invested his theory with a significance born of hindsight, extending the explanatory level of his principal paper on the basis of their own recent knowledge of cells and chromosomes. For example, the later investigators represented the homozygote in Mendel's scheme by double letters, contending that Mendel had surely been aware of the existence of paired genetic elements but that he had employed a single letter as an abridged form of expression.

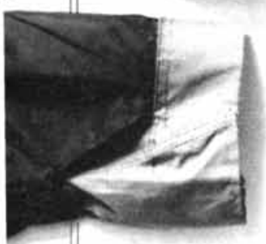
Olby's revisionist interpretation explains in part why Mendel's work had so little impact in his lifetime. Viewed in the context of 19th-century biology, Mendel's discussion of hybrids, although rigorous and systematic, did not break entirely new ground. The laws of inheritance were of concern to Mendel only insofar as they bore on his analysis of the evolutionary role of hybrids: they provided a conceptual framework wherein he could set out the conditions under which some hybrids would breed true and others would not. Mendel was more interested in the mathematical laws governing the characters of pea plants than in the pursuit of the genetic elements in the cell that were subject to those laws.

Olby concludes: "If we arbitrarily define a Mendelian as one who subscribes explicitly to the existence of a finite number of hereditary elements, which in the simplest case is two per hereditary trait, only one of which may enter one germ cell, then Mendel was clearly no Mendelian. This... is not just an idiosyncratic piece of iconoclasm directed at a now defenseless Moravian monk but a long-overdue reinterpretation."

How to tell the from the imitations.

Back in 1968, we made our first 60/40 parka. And while we knew it was an exceptionally well-built, good-looking and functional jacket, we had no idea it would become one of the most imitated pieces of outdoor clothing ever designed.

Inside the sleeve of a Sierra Designs 60/40 parka. Note the quality of stitching.



Inside the sleeve of a competitive 60/40 parka. Note the quality of stitching.

Yet that's exactly what happened. Today, there are no fewer than 15 manufacturers making copies of the Sierra Designs original.

Like the original, most have seven pockets that hold as much cargo as a small daypack. Most have velcro-adjustable cuffs, a built-in hood, large-tooth zippers. And most are made with rugged, lightweight 60/40 cloth—the name we coined when we introduced the material.

But while many have nearly duplicated the outward appearance of the Sierra Designs version, a close comparison reveals that none has duplicated our tailoring, attention to detail, and meticulous stitching.

Have a look at the cuffs, for example. On the outside, our competitors' look very much like ours. But when you turn them inside out, they don't. Where others are content to partially secure the lining to the outer material, our lining is completely sewn, so there's nothing to snag or rip.

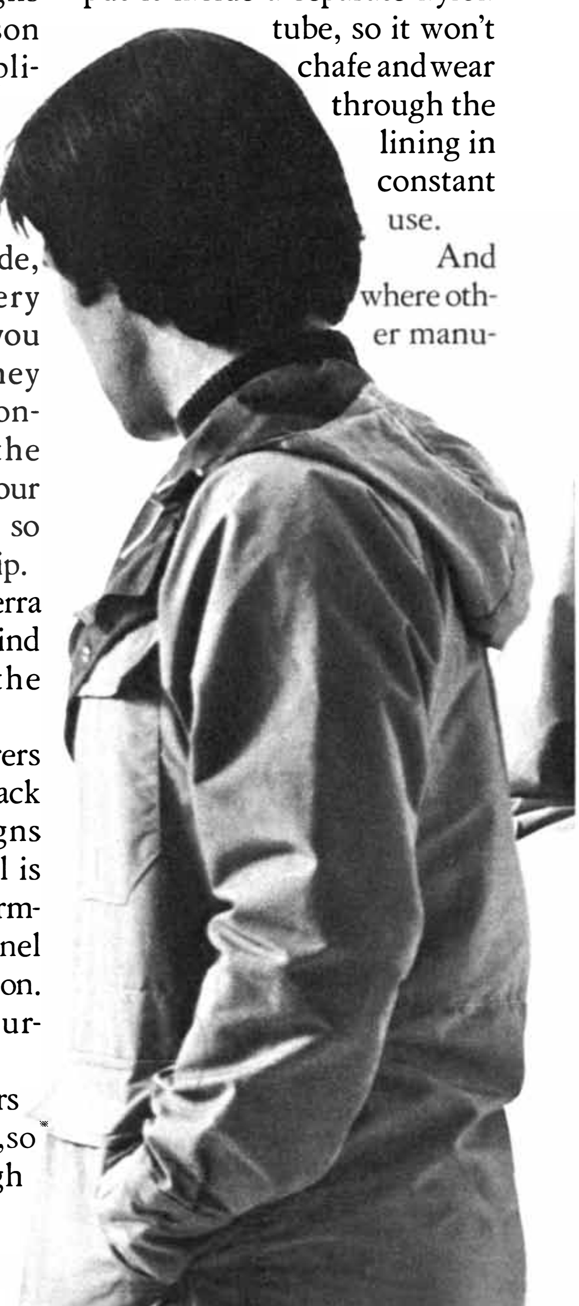
That's one example of Sierra Designs' detailing. You'll find many more throughout the parka.

Where other manufacturers use one panel to line the back of the jacket, Sierra Designs uses two. The upper panel is 60/40 cloth, to keep you warmer and dryer; the lower panel is light, friction-resistant nylon.

Where other manufacturers finish the hood with a single seam, we reinforce ours with four rows of stitching, so it will hold its shape through years of rough weather.

Where other manufacturers run the waist drawcord through the space behind the lining, we put it inside a separate nylon tube, so it won't chafe and wear through the lining in constant use.

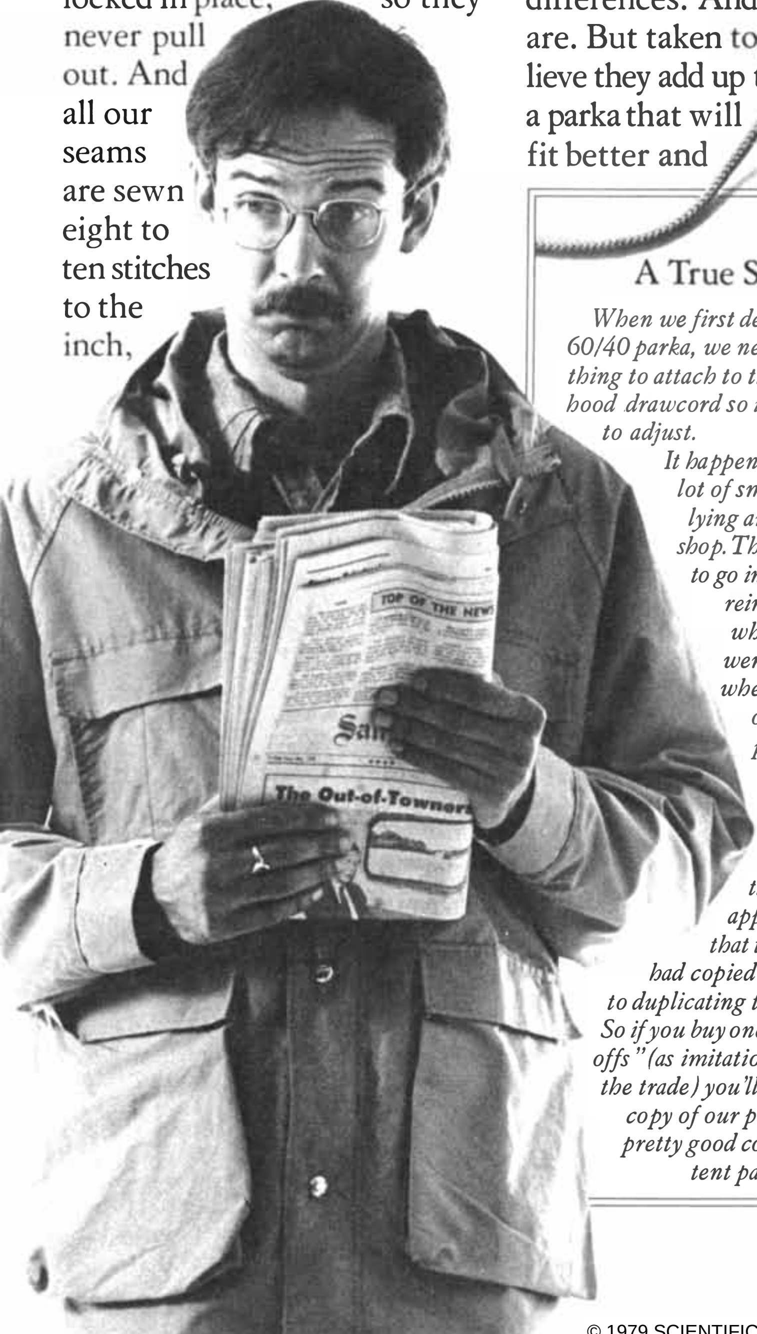
And where other manu-



original 60/40 parka

facturers are somewhat casual about the quality of their sewing, Sierra Designs is almost fanatical. All our stitches are locked in place, so they never pull out. And all our seams are sewn eight to ten stitches to the inch, because we've found that fewer stitches weaken the seam, while more stitches weaken the fabric.

These may seem like small differences. And perhaps they are. But taken together, we believe they add up to a parka that will fit better and last longer.



A True Story.

When we first designed the 60/40 parka, we needed something to attach to the ends of the hood drawcord so it would be easy to adjust.

It happened that we had a lot of small leather discs lying around our workshop. They were designed to go inside our tents, to reinforce the places where the guylines were anchored. But when we tried them out, they worked perfectly as little toggles for the drawcord.

When one of the first imitations of our parka appeared, we found that the manufacturer had copied ours right down to duplicating the leather discs. So if you buy one of these "knock-offs" (as imitations are known in the trade) you'll get more than a copy of our parka. You'll get a pretty good copy of one of our tent parts to go with it.

We're so sure of that, in fact, that we attach a Full Lifetime Warranty to every 60/40 parka we make. It reads: "We unconditionally warrant this product against defects in materials and workmanship for your lifetime. If any stitching ever wears out, tears loose, or breaks, return the product to us. We will repair or replace it, free of charge."

When it comes right down to it, this lifetime warranty may be the easiest way to distinguish a Sierra Designs parka from an imitation.

Because only the original has it.

For a list of dealers near you, call 800-227-2082, toll-free. Or write us at 247 Fourth Street, Oakland, California, 94607.

Sierra Designs

Legionellosis

The mysterious 1976 Legionnaires' disease epidemic has been traced to a previously unknown bacterium. Legionellosis, the infection the bacterium gives rise to, has turned out to be not very rare after all

by David W. Fraser and Joseph E. McDade

The 58th annual convention of the American Legion's Pennsylvania Department was held at the Bellevue-Stratford Hotel in Philadelphia from July 21 through 24, 1976. In most respects it was a typical Legion convention, with some 4,400 delegates, members of their families and other conventioneers on hand for a parade, meetings and the variety of social activities, formal and informal, that are customary on such occasions. What no one realized was that this convention was destined to be recorded in the annals of medicine. Between July 22 and August 3, 149 of the conventioneers developed what appeared to be the same puzzling illness, characterized by fever, coughing and pneumonia. This, however, was an unusual, explosive outbreak of pneumonia, with no apparent cause. Because most of the conventioneers had returned to their homes before becoming ill it was not until August 2 that reports to the Pennsylvania Department of Health made it clear that an epidemic had developed among those who had attended the convention.

That day marked the start of one of the largest and most complex investigations of an epidemic ever undertaken. Legionnaires' disease, as the illness was quickly named by the press, was to prove a formidable challenge to epidemiologists and laboratory investigators alike. Only after many months was the causative agent discovered, and more months were to pass before the agent was characterized as a previously unknown bacterium. Meanwhile Legionnaires' disease had been shown to be not a rare illness but a relatively common one, responsible for other outbreaks and for sporadic individual cases in many parts of the world both before the Philadelphia epidemic and after it.

Epidemiology

The initial step in the investigation of any epidemic is to determine the characteristics of the illness, who has become ill and just where and when. The next step is to find out what was unique about the people who became ill: where they were and what they did that was different from other people in the same general group who stayed well. Knowing such things may indicate how the disease agent was spread and thereby suggest the identity of the agent and where it came from. In order to gather such information as quickly as possible from conventioneers dispersed throughout the state 23 epidemic intelligence officers from the Center for Disease Control (CDC) of the U.S. Public Health Service collaborated with scores of public-health workers from the Pennsylvania and Philadelphia health departments.

We quickly learned that the illness was not confined to Legionnaires. An additional 72 cases were discovered among people who had not been directly associated with the convention. They had one thing in common with the sick conventioneers: for one reason or another they had been in or near the Bellevue-Stratford Hotel. All together 221 people had become ill; 34 of them died of pneumonia or its complications.

Older conventioneers had been affected at a higher rate than younger ones, men at three times the rate for women. Legionnaires who had stayed overnight at the Bellevue-Stratford showed a higher illness rate than those who stayed elsewhere, and among the latter those who became ill had spent more time at that hotel than those who remained well. The only family contacts who became ill had themselves been in Philadelphia for the convention, indicating that the disease had not been spread from person to person. The obvious possibility that the disease might have been spread by food or drink was ruled out. Conventioneers who became ill were shown to be no more likely than those who remained well to have eaten at particular restaurants, to have attended particular functions where food and drink were served or to have drunk water or used ice in the hotels.

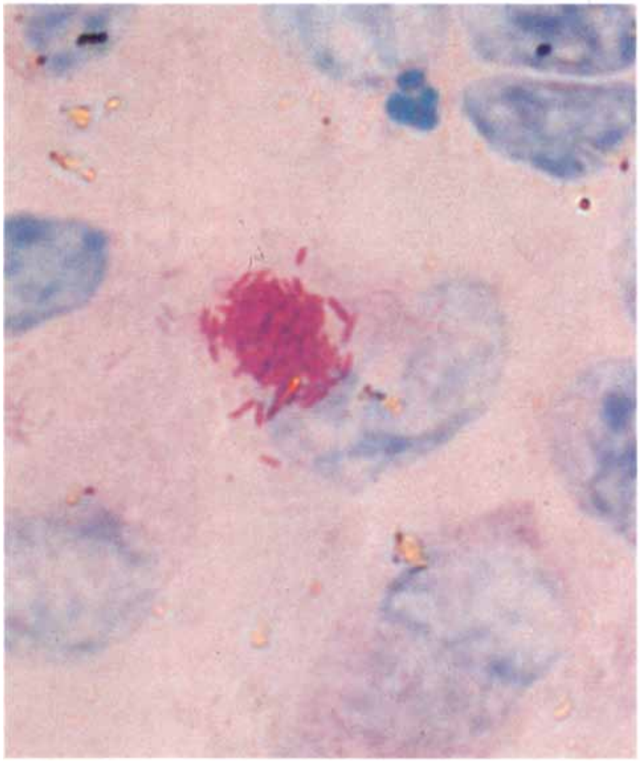
Certain observations suggested that the disease might have been spread through the air. Legionnaires who became ill had spent on the average about 60 percent more time in the lobby of the Bellevue-Stratford than those who remained well; the sick Legionnaires had also spent more time on the sidewalk in front of the hotel than their unaffected fellow conventioneers. If the disease agent had contaminated the air in the lobby and passed through the front door to the sidewalk, it could have affected Legionnaires and other hotel visitors in just that pattern. It appeared, therefore, that the most likely mode of transmission was airborne. What agent had caused the disease, however, remained undetermined for some time.

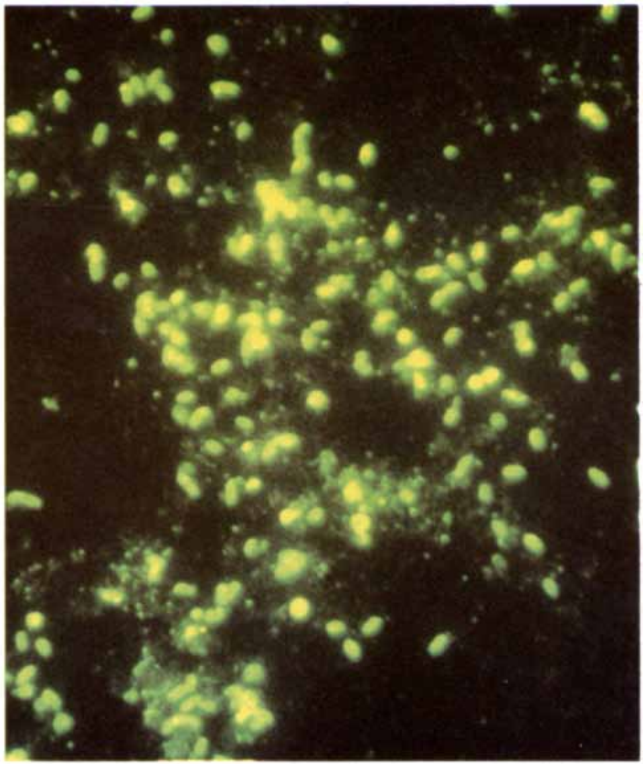
Finding the Agent

The clinical symptoms could have been caused by a wide variety of agents, including heavy metals, toxic organic substances and infectious organisms. During the first weeks of the investigation specimens obtained from patients and tissues taken at autopsy were subjected to intensive testing. Comparison of specimens from cases and controls failed to reveal the consistent presence in patients of unusual levels of metallic or toxic organic substances that might be related to the epidemic; no known microbial agents were found that could explain the outbreak. The testing continued, however, with attention directed to the possibility that some unknown organism had been responsible but had somehow escaped detection. That supposition turned out to be correct.

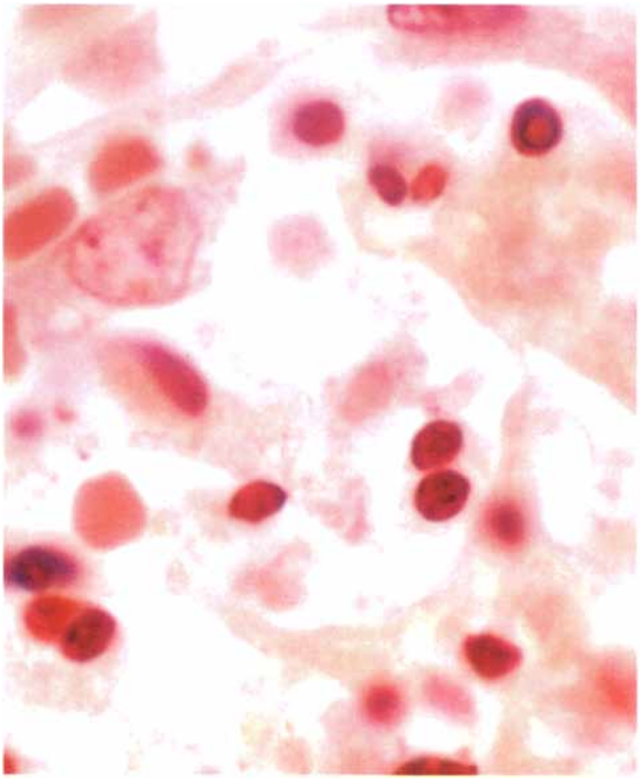
In January, 1977, in collaboration with Charles C. Shepard and others at the CDC, we isolated from the tissues of Legionnaires who had died of the disease a previously unrecognized pathogenic bacterium and then identified it as the causative agent. How the bacterium had escaped earlier detection became clear only after it had been isolated and characterized.

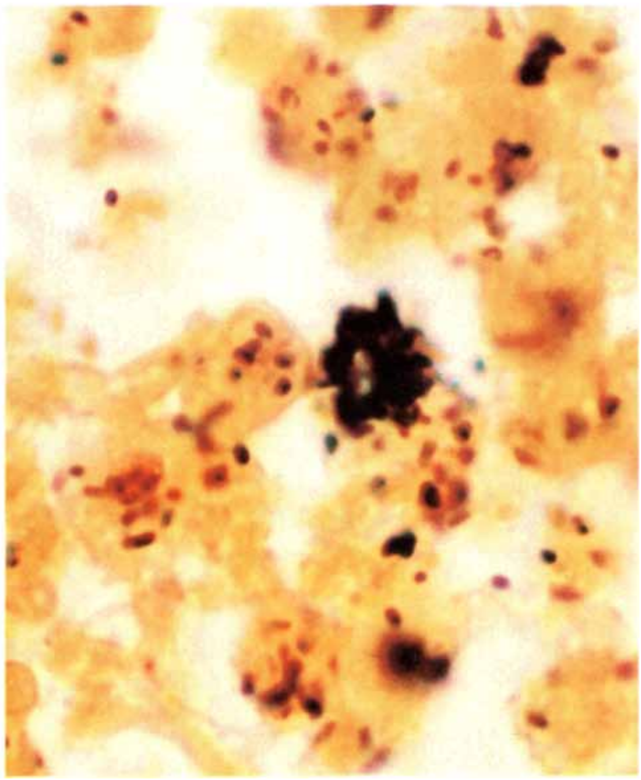
The agent was discovered not during an attempt to isolate a bacterium but rather by a procedure designed to isolate a rickettsia, a very small bacterium-like organism. Rickettsias are special in



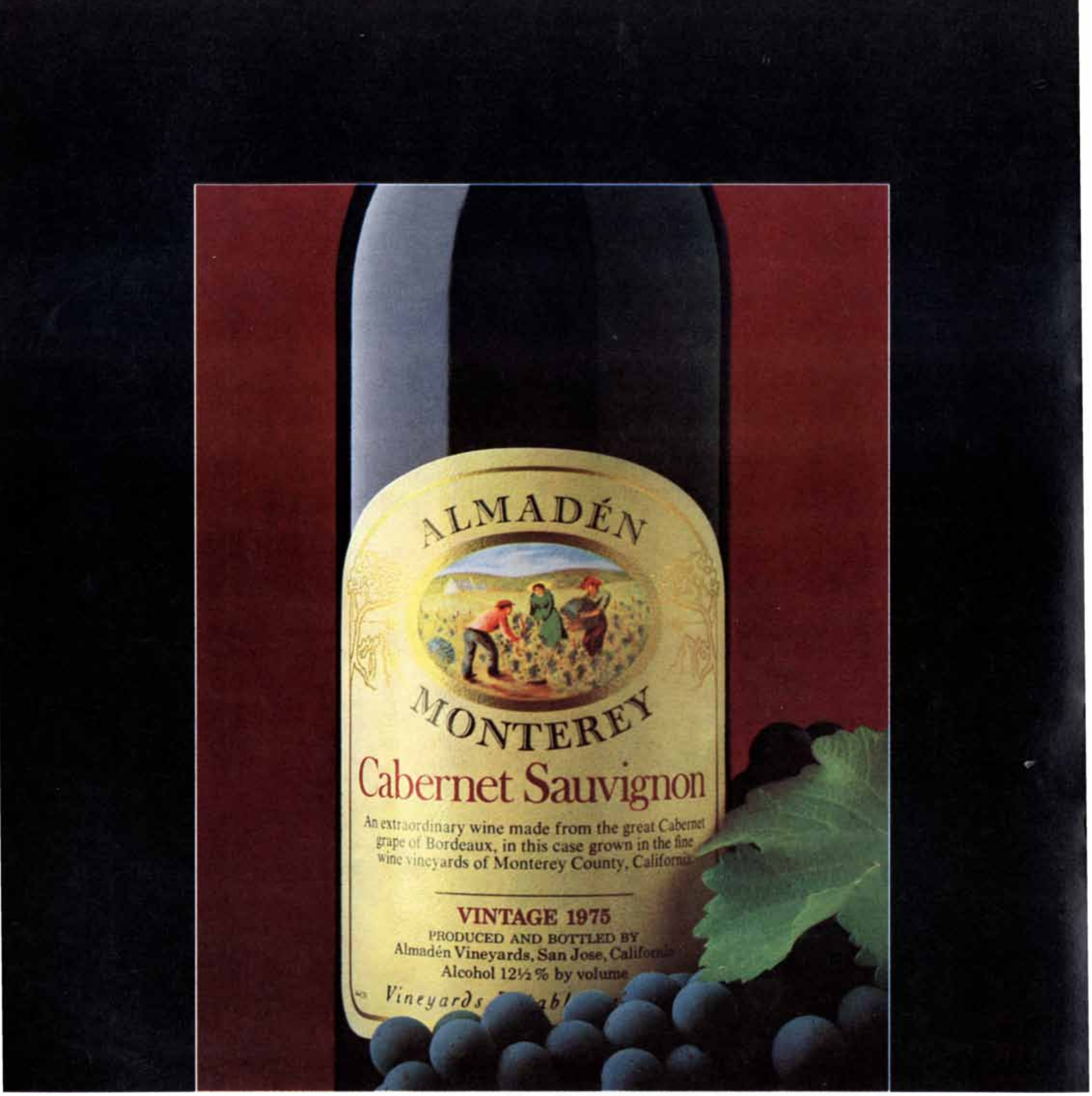


AGENT OF LEGIONNAIRES' DISEASE was first observed in a smear of the spleen of a guinea pig inoculated with lung tissue from a patient who had died of the disease. The photomicrograph (*left*) reveals a clump of rod-shaped organisms (bacilli) stained red by the Giménez method. The organisms were subsequently shown to be bacteria when they were successfully grown on a modified bacteriologic medium. They were shown to have caused Legionnaires' disease by indirect fluorescent-antibody tests (*see illustration on page 90*). In a similar (but direct) test, suspensions of diseased lung tissue are stained with a fluorescein-linked rabbit antibody to the Legionnaires' disease agent. The antibody becomes fixed to the bacterium and the bacteria glow yellow green when they are viewed in the ultraviolet microscope (*right*). Both preparations have been enlarged 2,500 diameters in these photomicrographs, which were made by J. Donald Howard.





CONVENTIONAL STAINS for visualizing bacteria in a pathologic specimen (tissue from a patient) do not usually work for the Legionnaires' disease bacterium. A lung section stained by the commonly used Brown and Brenn method reveals no organisms (*left*). Francis W. Chandler of the Center for Disease Control found that a modified version of the Dieterle silver-impregnation method would stain the bacterium, as is shown in a micrograph (*right*) of a section from same lung. Both sections have been enlarged 2,500 diameters.



We are proud of the birthdates of our children, the grapes of Almadén.

Yes, we are very proud. And we show our pride by putting a vintage date on our classic varietal wines.

Because each vintage—each year's harvest—is like a new child. A child we lovingly care for as it develops and matures. And we give each its just recognition by giving it a vintage date.

For you, the vintage date adds yet another dimension to the enjoyment of our classic varietal wines. Because each vintage has a different personality. Subtle differences you may taste and enjoy.

ALMADÉN

Also, while all our wines are ready to drink when purchased, you may sometimes wish to set a few bottles aside. Here you will find the vintage date helpful in giving the wine the additional aging you desire.

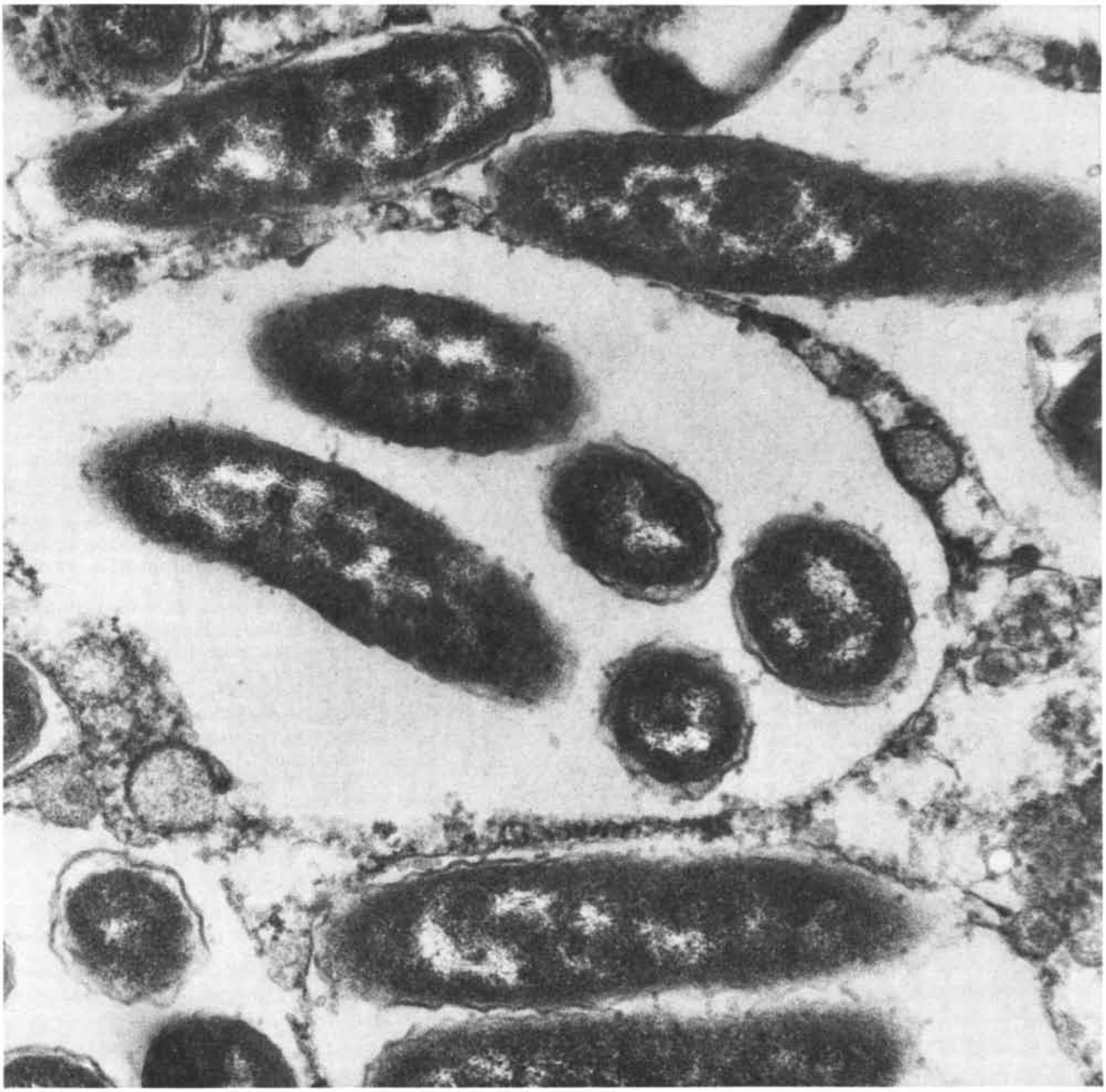
We invite you to enjoy the incomparable quality of our vintage-dated wines. An incomparable quality you will find in our entire family of Almadén wines.

that they do not grow on synthetic culture mediums as most bacteria do but grow only in a living host such as the arthropods many of them parasitize, experimental animals such as guinea pigs, and embryonated eggs. One rickettsia species, *Coxiella burnetii*, causes a type of pneumonia called Q fever. Attempting to isolate *C. burnetii*, we inoculated guinea pigs with tissue specimens obtained from patients at autopsy. The guinea pigs developed fever; after several days they were sacrificed, and microscopic examination of stained tissue samples yielded a first result. A spleen specimen stained by the Giménez technique contained some small rod-shaped organisms.

That observation in itself did not mean very much. The presence of a few organisms in the guinea pig spleen was far from establishing that they had caused the patient's death, but it did provide a promising lead. In order to enhance the growth of the organisms and any others present in the animals' tissues, suspensions of guinea pig spleen were inoculated into the yolk sac of embryonated hen eggs. The embryos died several days later. The yolk-sac membranes were harvested and smears of the membranes were stained by the Giménez method. An apparently pure culture of rod-shaped organisms (bacilli) was visible in the yolk-sac smears. We could not tell whether they were large rickettsias or small bacteria, but we immediately proceeded to find out if they were the cause of Legionnaires' disease.

Establishing Cause

We did that by testing the serums of the Legionnaires' disease patients for antibodies to the newly isolated organism. When a patient's serum is found to contain antibodies specific for the antigen molecules on a given organism, it



LEGIONELLA PNEUMOPHILA is stained with uranyl acetate and lead citrate and enlarged some 50,000 diameters in an electron micrograph of the yolk sac from an infected embryonated egg. The organism, seen in longitudinal and transverse sections, is structurally similar to some other bacteria but appears to be unrelated to any previously described species or genus; hence the name *L. pneumophila*.

is said to be "seropositive"; one can assume then that the patient has been exposed to that particular organism (or possibly one with similar antigens) at some time. If the patient's antibody level is found to have risen in the course of convalescence from a given illness, "seroconversion" has been demonstrated and it is very likely that the illness was caused by the particular organism.

When we tested serums from the Legionnaires' disease patients for antibodies to the newly isolated organism by indirect fluorescent-antibody tests [*see illustration on page 90*], more than 90 percent of the serum specimens taken during convalescence turned out to be seropositive. In more than 50 percent of the cases—most of those for which suitably timed serum specimens were available—we were able to demonstrate seroconversion, indicating that the patients had recently been infected by this particular bacillus.

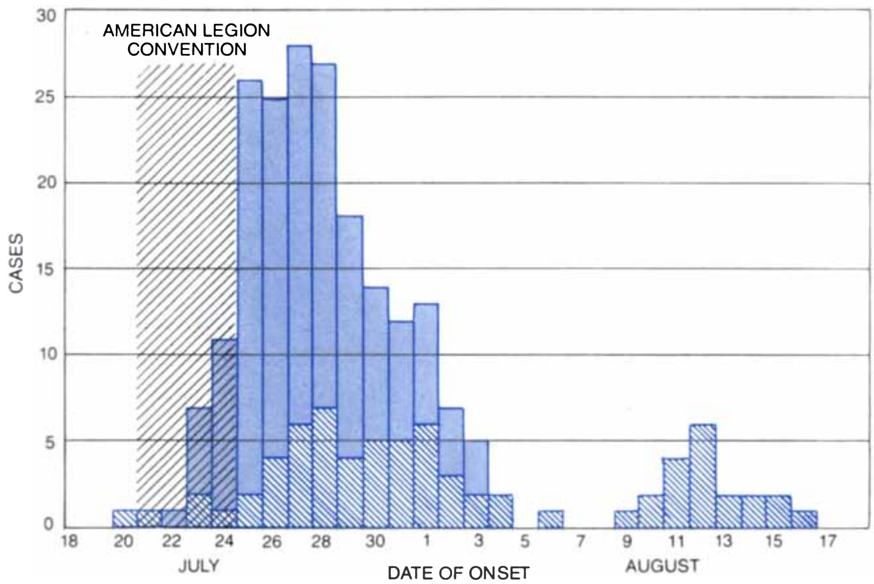
Although the results of the indirect fluorescent-antibody tests showed that we had isolated the causative agent of Legionnaires' disease, for several weeks we did not know whether the organism was a rickettsia or a bacterium. Morphologically it appeared to be a bacterium, but it had been isolated by techniques more appropriate to rickettsias, and it failed to grow on a variety of culture mediums on which bacteria normally thrive. Then, finally, Robert E. Weaver of the CDC found a bacteriologic medium on which the bacillus would grow. It was similar to the one on which gonococci are isolated, and it was prepared by adding 1 percent hemoglobin and 1 percent of a commercial supplement (IsoVitaleX) to a standard medium (Mueller-Hinton agar). When Weaver inoculated this medium with a heavy suspension of infected yolk-sac material, distinct bacterial colonies were observed growing on the culture plates after several days of incubation. The newly isolated bacillus was thereby shown to be not a rickettsia but rather a bacterium.

It was an exceedingly fastidious bacterium, with very specific growth requirements. James C. Feeley and his associates at the CDC defined the requirements. Among other things, the bacterium will not grow unless the medium contains a high enough concentration of the amino acid cysteine (of which IsoVitaleX has a large amount) and of iron (which was supplied by the hemoglobin). Such specific nutritional constraints are unusual. Many bacteria require supplementary cysteine or iron for growth, but the double requirement is rare.

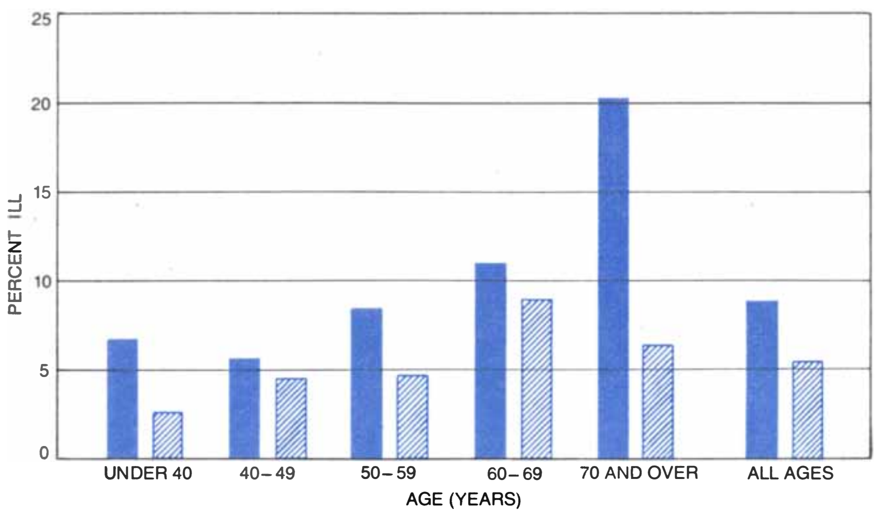
Another unusual characteristic of the Legionnaires' disease bacterium is that in pathologic specimens (lung tissue from a patient, for example) it cannot consistently be stained by techniques that normally are effective for bacteria. Even after it was established (by our successful isolation of the organism from lung tissue) that the bacteria were present in diseased lungs, the organisms could still not be seen in lung sections stained by the usual methods. Francis W. Chandler and his colleagues at the CDC thereupon tried a number of less conventional methods. Eventually they found that the bacterium could be stained dependably by a modification of the Dieterle silver-impregnation technique, a procedure developed more than 50 years ago for staining the spirochete of syphilis. The many unusual properties of the Legionnaires' disease bacterium suggested it was an unknown species, but more detailed characterization of the organism was necessary before that could be established.

Earlier Outbreaks

Meanwhile another line of research being pursued in the course of the intensive CDC investigation was producing evidence that the bacterium was new only in the sense that it was unfamiliar to laboratory workers. An important question presented by the original discovery of the agent was whether or not the unfamiliar bacillus had been responsible for earlier outbreaks of respiratory disease for which no cause had ever



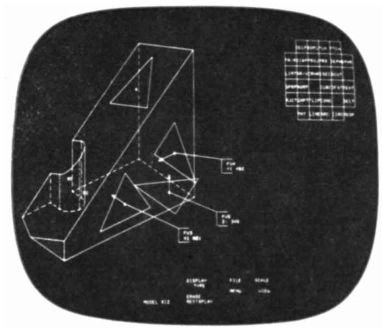
ATTACK RATE in the 1976 Legionnaires' disease epidemic varied, among delegates to the Philadelphia convention, with age and hotel of residence. The incidence was higher for older Legionnaires than it was for younger ones and also higher for delegates who stayed at Bellevue-Stratford Hotel (*solid color bars*) than it was for those who stayed at other hotels (*hatched bars*).



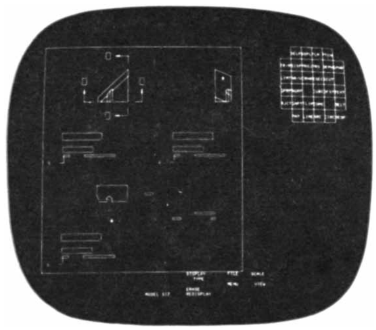
DISTRIBUTION OF CASES in the Philadelphia epidemic was typical of an outbreak arising from a common source: the number of cases among conventioneers (*color*) and among other people (*color hatching*) rose rapidly from July 22 through July 25 and later declined more slowly. The relation between the dates of the convention and the onset of illness among Legionnaires indicated that the incubation period of Legionnaires' disease ranged from two to 10 days.

Only from MCAUTO...

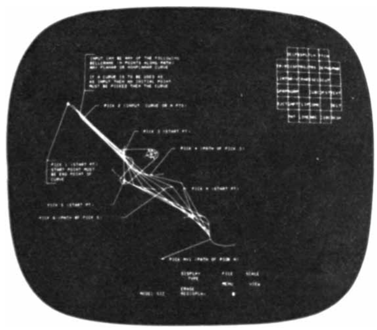
Computer-aided design and drafting with independent view and construct.



True 3-D Design Capability. Interactively create a fully-surfaced three-dimensional geometric model in a real-time environment.



Part Drawing System. Create a two-dimensional picture of either a fully-surfaced or wire-framed part—generating, arranging, and labeling section cuts as you require.



Kinematics Analysis. Analyze and display intermediate and terminal positions of bell cranks, rods and slides; critical positions and intermediate positions are displayed at increments you define.

Introducing CADD (we pronounce it "caddy"), from MCAUTO. A remarkable combination of applied mathematics, computer science, and proven terminal console procedures incorporated into a true three-dimensional, real-time graphics software system. A system which allows you to actually see your product develop—and change it on the spot if you wish—as it is conceived.

With CADD, you are creating, and have available for recall, an easily accessed base of geometrically accurate data which will provide the initial building block in your product development. By computerizing analytical geometry, by leaving all mathematical calculations to the computer, CADD preserves design-engineer time for *designing.*

You can expect better product designs, reduced lead times, improved productivity and job satisfaction, plus reduced design costs. And since CADD part design can be interfaced with analysis and NC programming, you save in the manufacturing process as well.

Send today for all the facts about CADD. Or phone us at (314) 232-8021.

MCAUTO
MCDONNELL DOUGLAS
Setting the standard in data processing

W. R. Vickroy, MCAUTO, Dept. KI-500
Box 516, St. Louis, MO 63166
☐ Send me information on CADD.
☐ Have your representative phone me for an appointment.

Name ____ Title ____
Company Name ____ Phone ____
Address ____
City ____ State ____ Zip ____

been discovered. Perhaps the most notorious of those outbreaks was the outbreak of "Pontiac fever" in 1968. Beginning on July 2 of that year 95 of the 100 people who worked in a single building of the Oakland County Health Department in Pontiac, Mich., developed high fever, headache and muscle aches; other common symptoms were diarrhea, vomiting and chest pain, but there was no pneumonia. The illness lasted for only three or four days and all patients recovered. A team of CDC epidemiologists headed by Thomas H. Glick and Michael B. Gregg went to investigate the outbreak, and several of them became ill a day or two after entering the building. Guinea pigs exposed to the air in the building by Arnold F. Kaufmann developed pneumonia, whereas control animals placed in adjacent buildings remained healthy.

It was noted that of all the people who entered the building, the only ones spared were those who were there only when the air-conditioning system was turned off. Investigation of the air-conditioning system disclosed a defect. Mist generated by the evaporative condenser was not being exhausted properly; holes in an exhaust vent and a fresh-air vent had allowed the mist to condense and accumulate in a puddle in the fresh-air duct. Guinea pigs exposed to an aerosol of water from the puddle developed pneumonia. The conclusion reached at the time was that the agent of Pontiac fever was in the evaporative-condenser water and had been spread by the air-conditioning system. No pathogenic organism could be isolated, however, and the cause of the illness remained undetermined.

Serum specimens collected from Pontiac fever patients had been preserved at the CDC. In 1977, nine years later, they were tested for antibodies to the newly isolated Legionnaires' disease bacillus. Serum specimens from 31 of 37 cases showed seroconversion to the organism during convalescence, indicating that Pontiac fever had been caused by the same agent that caused Legionnaires' disease.

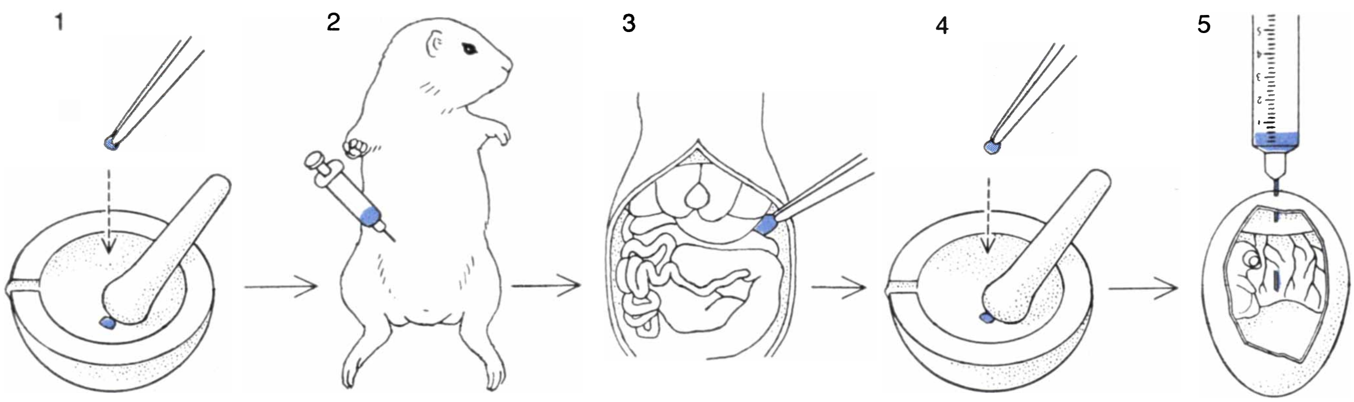
Another earlier outbreak traced to the same organism dated back to July and August of 1965, when 81 patients at St. Elizabeth's Hospital, a chronic-care facility in Washington, D.C., developed pneumonia and 14 died. The cause could not be determined. In 1977 paired acute-phase and convalescent-phase serum specimens from 23 patients were tested against the Legionnaires' disease agent. For 17 patients there was a substantial rise in antibodies to the bacterium. Similar results associated the bacterium with several other earlier outbreaks, including one that had affected Scottish vacationers at a Spanish resort in 1973.

Publication of a report on the characteristics of the Legionnaires' disease bacterium also led to the production of direct evidence of earlier infection by the same agent. F. Marilyn Bozeman of the U.S. Food and Drug Administration noticed that the bacterium was similar in some respects to four hitherto unclassified rickettsialike agents she had described years before. Like the newly isolated bacterium, each of them had been isolated in embryonated eggs from guinea pigs inoculated with clinical or autopsy specimens and each could be stained by the Giménez method. From the time of Bozeman's investigations the unidentified agents had been stored in a deep freeze. In 1978 various tests revealed that three of the four rickettsialike agents were not related to the Legionnaires' disease agent but that the fourth was virtually identical with it. The identical bacterium had been isolated in 1947 from a guinea pig inoculated with blood from a patient who had a febrile respiratory illness.

Characterizing the Bacterium

Meanwhile the effort to learn more about the Legionnaires' disease bacterium continued. Before a bacterium is designated as a new species one must show convincingly that it is significantly different from all previously described species. A team of CDC microbiologists examined the newly isolated bacterium's morphology, physiology and staining characteristics, its susceptibility to various antibiotics, its antigens and its fatty-acid composition. These characteristics were then compared with those of known bacteria. The Legionnaires' disease agent turned out to have quite a few properties in common with other bacteria, but the overall pattern of its properties was quite different from that of any known species.

Convincing evidence that the organism was a new species came when Donald J. Brenner, Arnold G. Steigerwalt and their CDC colleagues compared the genetic material of the Philadelphia bacterium with that of other bacteria by means of the elegant DNA-hybridization technique. The chromosome of the Legionnaires' disease bacterium, like that of other bacteria, is a double-strand DNA molecule composed of two single strands linked by base pairing between complementary nucleotides. Under the



ISOLATION of the bacterium was first accomplished by means of a method generally reserved for isolating nonbacterial microorganisms such as rickettsias. A bit of lung tissue from a patient was ground up (*1*), suspended in a buffer solution and inoculated into the abdominal cavity of guinea pigs (*2*). The animals became ill. At autopsy (*3*) a bit of spleen tissue, which presumably contained the still-unknown disease agent, was made into a suspension (*4*) that was inoculated into the yolk sacs of embryonated eggs (*5*) to enhance the growth of any organism present. The embryos died in four or five days. The yolk-sac membrane was removed from dead embryos (*6*), smeared on a microscope slide (*7*) and stained by the Giménez method (*8*). Rod-shaped organisms were visible under the microscope (*9*). Then, in an attempt to culture the organism, a yolk-sac membrane was made into a suspension (*10*), which was streaked on various bacteriologic mediums (*11*). Colonies eventually grew on an enriched Mueller-Hinton agar (*12*), showing that the organism was a bacterium and also making it available in quantity for further experiments. Supplemented mediums that meet the specific requirements of this fastidious organism have since been devised, so that the bacterium can now be isolated directly rather than by this long process in guinea pigs and eggs.

appropriate experimental conditions single strands of DNA from related species can combine to form hybrid double-strand molecules; the degree of hybridization depends on the closeness of the relation [*see illustration on page 92*].

Attempts to hybridize single strands of DNA from the Legionnaires' disease bacterium with single strands from many other bacteria failed to identify any species that were related to the newly isolated organism. Although additional comparisons are still being made, the available evidence seems to justify the designation of the bacterium as a new genus and species. It has provisionally been named *Legionella pneumophila*.

In retrospect it is easy to understand how *L. pneumophila* had remained undetected even though it had been causing illness for decades. First of all, the disease it causes is clinically similar to several types of nonbacterial ("atypical") pneumonia, and in searching for the agents of such diseases investigators commonly add antibiotics to the specimens to reduce contamination by bacteria. This makes it easier to isolate a virus or rickettsia, if either is the culprit, but it usually precludes the recovery of bacteria such as *L. pneumophila*. The bacterium was finally isolated in the course of a rickettsial isolation procedure in which no antibiotics were included. Second, the peculiar growth requirements of the bacterium had militated against its earlier identification. The mediums on which the usual pneumonia-causing bacteria or fungi are ordinarily isolated do not contain the optimal concentrations of cysteine and iron. Finally, the inability to consistently visualize the bacterium in tissue stained by the usual methods had surely been a major obstacle for workers investigating earlier cases of atypical pneumonia that were caused by *L. pneumophila*.

Clinical Course and Therapy

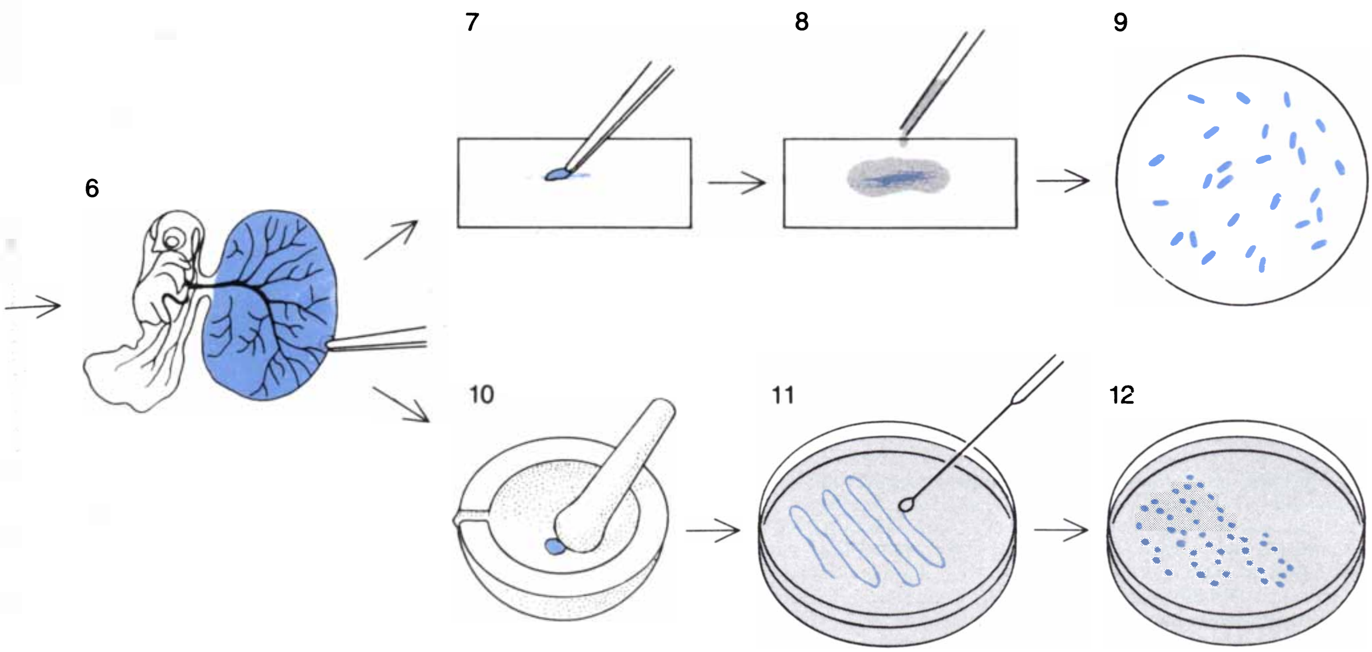
Legionellosis, as illness caused by *L. pneumophila* is now designated, has so far been seen in two basic forms: Legionnaires' disease and Pontiac fever. The clear distinguishing features are the incubation period, the attack rate and the presence of pneumonia.

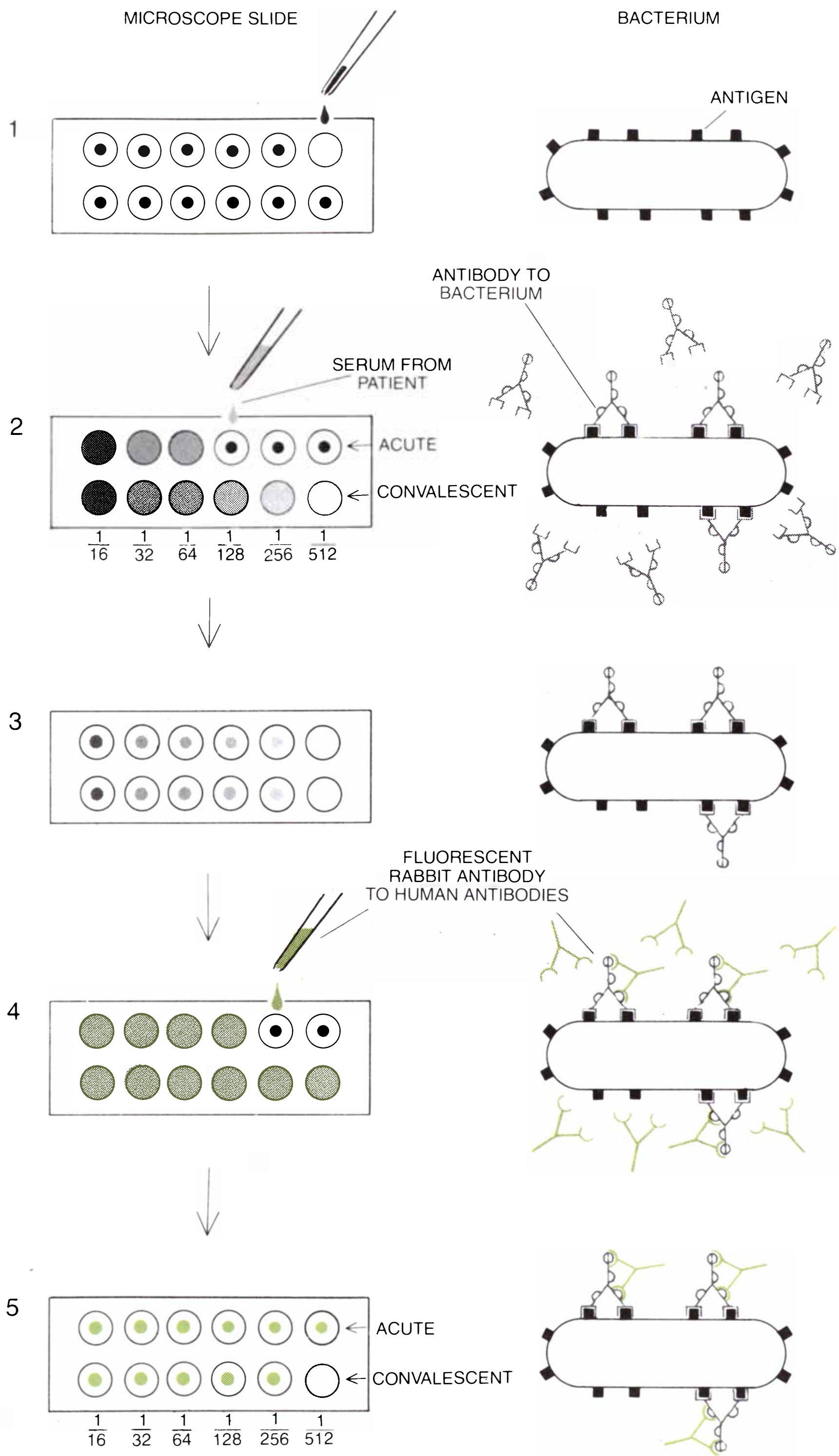
Legionnaires' disease typically begins from two to 10 days after exposure. A general feeling of malaise is accompanied by muscle aches and a slight headache. The fever rises rapidly and is usually accompanied by coughing, chest and abdominal pain, diarrhea and shortness of breath. When patients are first seen by a physician, most of them have a temperature of from 102 to 105 degrees Fahrenheit and pulmonary rales (abnormal breathing sounds heard with a stethoscope, suggesting pulmonary infection). Some patients appear confused or stuporous, suggesting that the central nervous system may be affected. Tests usually show a moderately elevated white-blood-cell count, an increased number of immature cells, a high red-cell sedimentation rate, protein and red cells in the urine and evidence of some abnormalities of liver and kidney function. In 90 percent of the patients chest radiographs show pneumonia.

Patients with Legionnaires' disease usually require hospitalization, and for several days after admission their condition deteriorates. In the absence of specific treatment about 20 percent of these patients die of progressive pneumonia (which interferes with the oxygenation of the blood) or from shock (the mechanism of which is not known, but which may result from the presence of bacteria in the blood). The other 80 percent recover gradually over a week or more, during which time some of them may require mechanical help in breathing or dialysis to deal with kidney failure. Recovery is usually complete, although patients may feel weak for several months and there may be a certain amount of permanent lung damage.

Pontiac fever begins much as Legionnaires' disease does, with headache, muscle aches and fever, but typically the symptoms appear only a day or two after exposure. Symptoms include coughing, chest pain, diarrhea, vomiting, sore throat and confusion, but ordinarily these symptoms are not prominent. Neither pneumonia nor shock has been seen in Pontiac fever, nor have the kidneys or the liver been involved. Patients have fever and feel very sick for from two to five days, but they all recover, apparently completely.

It is still not clear which antibiotics are most effective in *L. pneumophila* infections; no randomized trials with controls have been done with human beings. Claire V. Broome and her colleagues at the CDC did observe that during a 1977 epidemic in Vermont only one person of 22 who were treated with erythromycin died, whereas 10 of 41 who did not receive erythromycin died. Clyde Thornsberry found that growth of the bacterium on agar is inhibited by rather low concentrations of several antibiotics, but the organism manufactures an enzyme that inactivates most antibiotics of one group, the cephalosporins. Infected chick embryos were best protected by erythromycin, rifampin, gentamicin,





INDIRECT FLUORESCENT-ANTIBODY TEST, which established that the isolated bacterium had caused Legionnaires' disease, is described here at the experimental level (*left*) and at the level of the organism (*right*). One drop of a suspension of infected yolk-sac membrane was placed in each of 12 wells on a slide (*1*) and fixed in acetone. Six successive dilutions of serum taken from a patient during the acute and the convalescent phase were placed in each set of six wells (*2*); any antibody (*gray* Y-*shaped structures*) to the bacterium present in the serum became fixed to antigens on the outer bacterial membrane during a 30-minute incubation period. The slides were washed, removing unfixed antibody and other serum components (*3*). A fluorescein-labeled rabbit antibody (*color*) to human antibodies was added (*4*); the rabbit antibody attached to whatever patient antibodies had remained fixed to the bacterium. Excess rabbit antibody was washed away and each well was examined under an ultraviolet microscope. Bacteria with antibodies attached emitted a yellow green glow under the ultraviolet microscope (*5*). Antibodies were present in higher dilutions of convalescent-phase serum than of acute-phase serum, showing that the patient had developed antibodies to the bacterium in the course of his illness and thereby incriminating the bacterium as the agent of patient's disease.

streptomycin, sulfadiazine or chloramphenicol. With Theodore F. Tsai and others we found that infected guinea pigs recovered after treatment with erythromycin or rifampin. At this stage physicians are advised to treat patients who may have Legionnaires' disease with erythromycin, with rifampin therapy added for those who do not respond. Antibiotic therapy is apparently not required for Pontiac fever.

Diagnosis

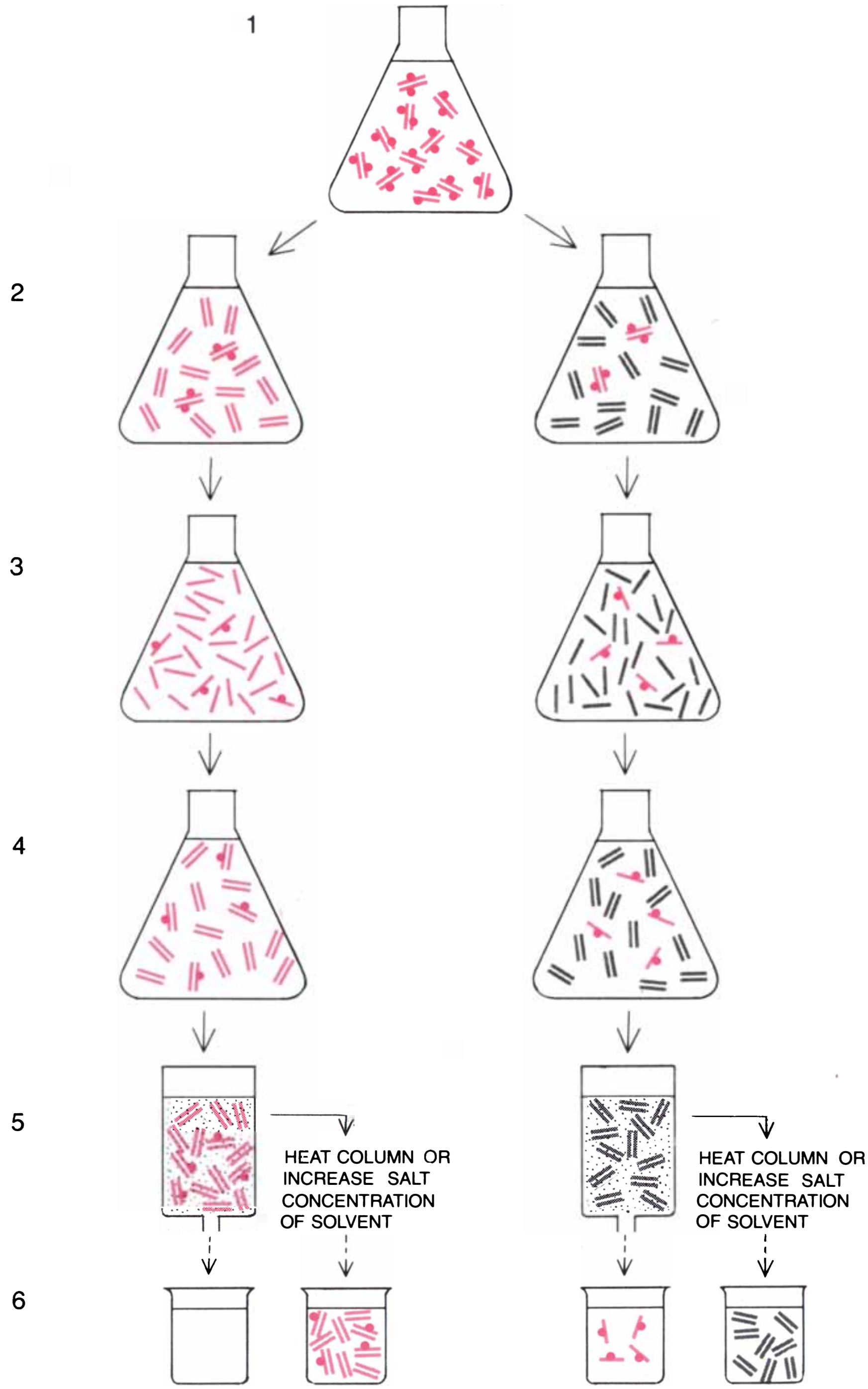
The clinical features that help to distinguish Legionnaires' disease from other pneumonias include diarrhea, confusion, lack of runny nose or sore throat, red blood cells in the urine, abnormalities of liver function, absence of the usual pneumonia-causing bacteria in the sputum and, of course, the failure to respond to courses of treatment that have been shown to be ineffective against *L. pneumophila*.

A specific diagnosis of legionellosis can be made only by isolating the organism from the patient, by showing a rise in the level of antibodies to the bacterium as the patient recovers or by demonstrating the presence of *L. pneumophila* itself in the patient's tissues or body fluids. The organism has been isolated from blood, sputum, pleural fluid and lung tissue by the inoculation of either guinea pigs or the appropriate bacteriologic mediums. There are practical limitations, however, on attempts to isolate the bacterium directly on a culture medium. It can be hard to obtain suitable specimens, the organism is not readily cultured and it takes several days for the cultured bacterium to grow.

By means of the indirect fluorescent-antibody technique high concentrations of antibody can be demonstrated by the end of the second week of illness in about half of the patients tested, but some patients do not develop antibodies until the sixth week and about 15 percent of them never achieve significant levels. There are other techniques for demonstrating antibodies, but their sensitivity and specificity are not well established. Another problem is that there are several different serologic groups of *L. pneumophila*, each one carrying a different set of antigens. An infection may be caused by a member of any group, so that the patient's serum must be tested against representatives of each group. Moreover, a diagnosis based on a rise in the level of antibodies to the organism must necessarily be retrospective. It is usually only during convalescence that the antibody level is high enough to be firmly diagnostic, and that is obviously too late to help the physician choose the right therapy.

Direct fluorescent-antibody testing, on the other hand, can make diagnosis possible within a few hours, early in the


SEAGRAM DISTILLERS CO., N.Y., N.Y.
Seagram's
Crown Royal
Fine De Luxe
Blended Canadian Whisky
A PRODUCT OF
Joseph E. Seagram & Sons Limited
WATERLOO · ONTARIO · CANADA
80 PROOF 750 ML (25.4 FL. OZ.)
Have you ever seen a grown man cry?



DNA-HYBRIDIZATION TEST capitalizes on the fact that the two strands of double-strand DNA are linked by bonds between specific complementary nucleotides, so that single strands from the same species reassociate readily and strands from different species reassociate in proportion to the species' relatedness. Double-strand DNA of *L. pneumophila* (*light color*) is labeled with a radioactive isotope (*color dot*) and broken into short lengths (*1*). A very small amount of labeled DNA is mixed with short pieces of unlabeled DNA (*2*) from the same bacterium (*left*) and from another bacterium (*gray*) being tested for relatedness (*right*). The mixtures are heated, breaking the bonds between complementary nucleotides and forming single strands of DNA (*3*). When the mixtures are cooled and then incubated, complementary strands collide and reassociate (*4*). In time essentially all the labeled DNA pieces mixed with homologous DNA "find" a complementary unlabeled strand (*left*), but the labeled DNA pieces mixed with DNA from a different bacterium do not (*right*). Labeled DNA is present in such a low concentration that random reassociation between two labeled strands is unlikely; all labeled double-strand DNA must be a hybrid of labeled and unlabeled strands. The single- and double-strand DNA's are separated by passing them through a column containing hydroxyapatite gel (*5*) and are collected (*6*). Measuring the relative amounts of radioactivity in the single-strand and in the double-strand DNA provides a measure of the extent of hybridization. Here there is no hybridization between the *L. pneumophila* DNA and the DNA of the bacterium against which it is being tested, demonstrating an absence of relatedness between the two organisms.

course of infection. The patient's sputum or other respiratory secretion is exposed to antibodies to *L. pneumophila* (prepared by immunizing rabbits with inactivated bacteria) that have been conjugated to a fluorescent material. The antibodies become fixed to the *L. pneumophila* antigens, revealing the presence of the bacterium. As many as 70 percent of the patients may show positive fluorescent staining, and false-positive results seem to be infrequent. Preliminary work suggests that antigens of the organism can be detected in patients' serum and urine by a sophisticated procedure called enzyme-linked immunosorbence. If this test proves to be sufficiently sensitive and specific, it should be of great value in making early diagnoses.

Occurrence of the Disease

In the few years since legionellosis was recognized as a new disease entity cases have been discovered almost wherever they were sought: in more than 40 states of the U.S. and also in Australia, Canada, Denmark, Greece, Israel, Italy, the Netherlands, Norway, Spain, Sweden, Switzerland, West Germany, the United Kingdom and Yugoslavia. No significant geographical pattern has yet emerged, but information from many regions of the world is still lacking or is sketchy. Sometimes there are clusters of cases that appear to have a common source and sometimes there are isolated single cases; almost all the clusters and most of the sporadic cases have been reported in the summer.

Legionnaires' disease is most often seen in middle-aged and elderly people, although children as young as 16 months have been affected. More men are affected than women. The disease is commoner in cigarette smokers (but not in cigar or pipe smokers) and in heavy drinkers. Underlying medical problems, such as cancer, leukemia, lymphoma and kidney failure, and drugs that impair the immune system appear to predispose a person to Legionnaires' disease (and to other infections).

Gregory A. Storch and William B. Baine of the CDC have found that sporadic cases are more frequent among travelers, construction workers and people living near sites of excavation or construction. It may be that such people are more likely to be exposed to airborne dust, which may carry the bacteria. Outbreaks of legionellosis have usually been concentrated in or near a building, often a hotel or a hospital. In some instances the building appears to have been simply a place where large numbers of susceptible people congregated, but in other instances the central air-conditioning system has been implicated as the source of the bacteria (for reasons that will be made clear below).

The overall incidence of legionellosis

Harris technology on the job

In information systems, Harris produces data processing terminals, general-purpose computers, word processing systems and supervisory control systems.

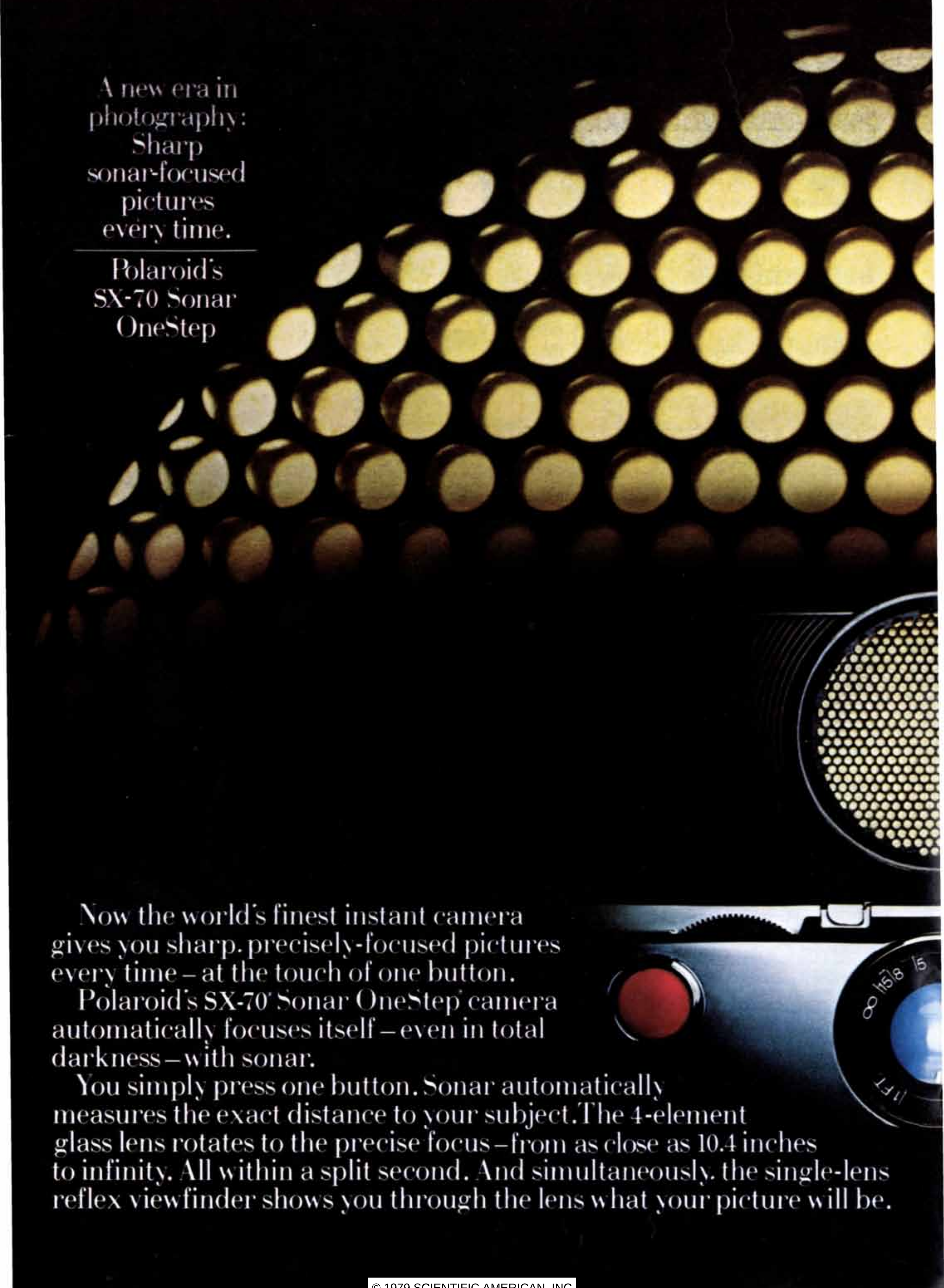
Harris 1680 distributed data processing system at General Foods

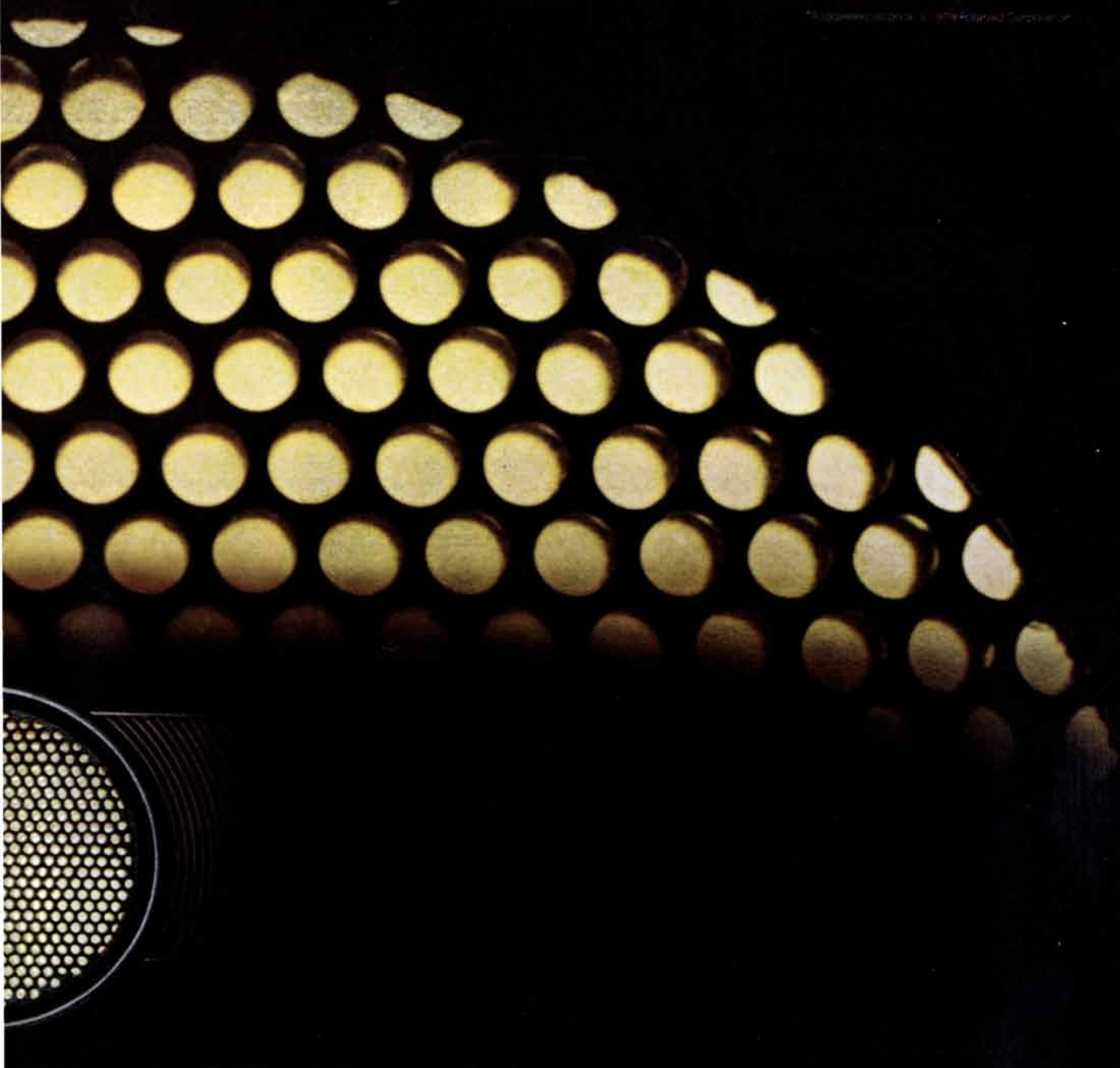
Multi-function distributed data processing systems from Harris provide remote access to central computers *plus* on-site computing power, helping business and government handle growing data processing workloads.

Harris technology works worldwide — in communication equipment, information systems, government systems, semiconductors and printing equipment. For information, write: Harris Corporation, Melbourne, Florida 32919.

HARRIS

COMMUNICATION AND INFORMATION PROCESSING


A new era in photography: Sharp sonar-focused pictures every time.
Polaroid's SX-70 Sonar OneStep
Now the world's finest instant camera gives you sharp, precisely-focused pictures every time – at the touch of one button.
Polaroid's SX-70 Sonar OneStep camera automatically focuses itself – even in total darkness – with sonar.
You simply press one button. Sonar automatically measures the exact distance to your subject. The 4-element glass lens rotates to the precise focus – from as close as 10.4 inches to infinity. All within a split second. And simultaneously, the single-lens reflex viewfinder shows you through the lens what your picture will be.

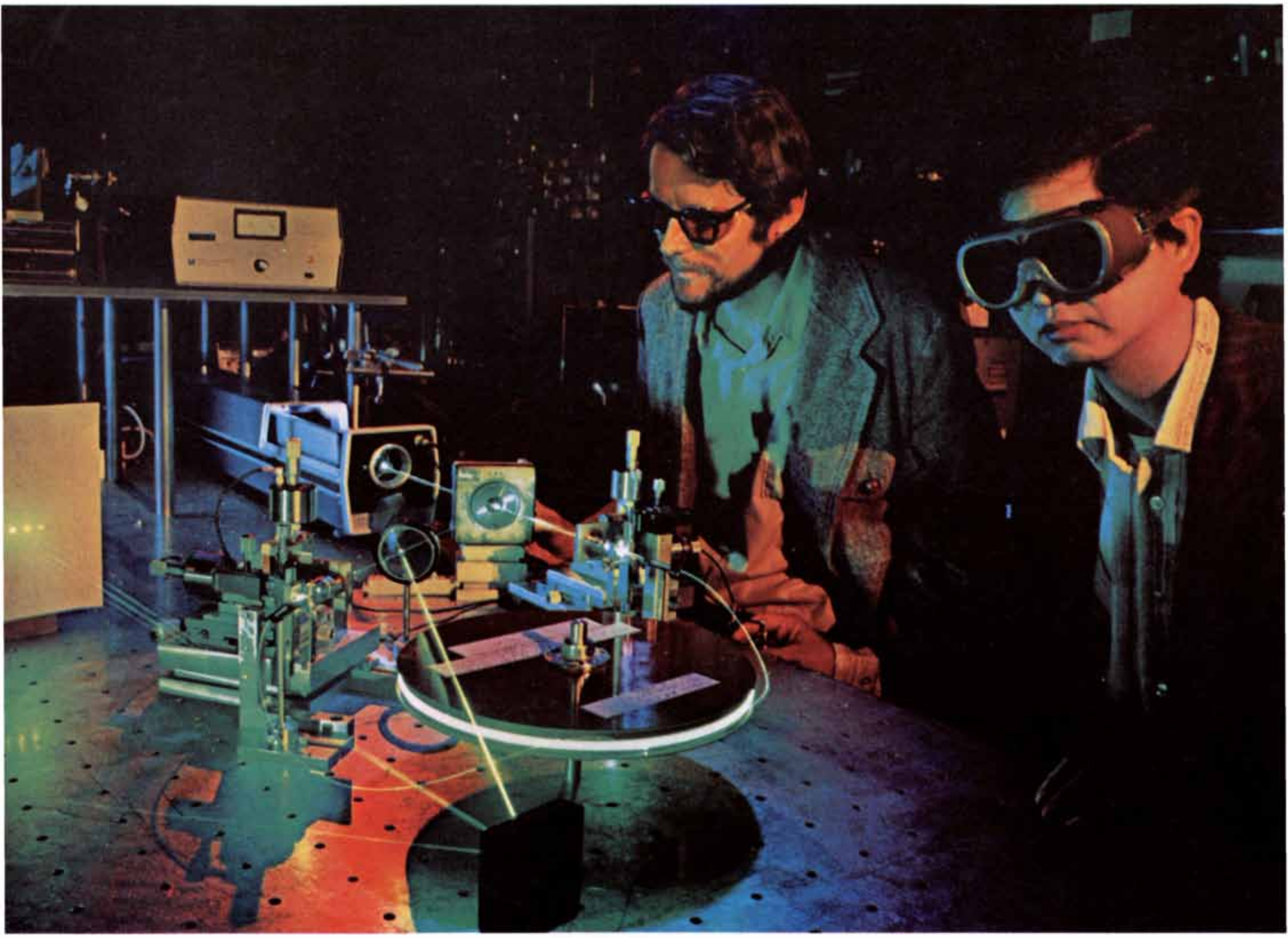


In fact, sonar can change focus so quickly, you can shoot as fast as every 1.5 seconds.

Polaroid's SX-70 Sonar OneStep camera sets all exposures automatically, too. Including automatic time exposures up to 14 seconds. And flash pictures up to 20 feet. Its built-in motor even hands you the developing picture.

There are two Polaroid® Sonar OneStep Land cameras: the folding SX-70 Sonar OneStep (described here), in genuine leather and brushed chrome finish. And the Pronto® Sonar OneStep, at $99.95.* No other cameras make precision photography so automatic.

The longest in our long line of laser firsts...



Bell Laboratories
Murray Hill, New Jersey 07974

Bell Labs scientists Roger Stolen and Chinlon Lin work with a fiber Raman laser, one of a new class of light sources that use optical fibers—up to a kilometer long—to produce tunable laser light. At left, the laser's output—which contains multiple Raman-shifted wavelengths—is taken off a beam splitter and dispersed by an external grating to show the broad range of wavelengths that can be tuned.

	LEGIONNAIRES' DISEASE	PONTIAC FEVER
ATTACK RATE	1–5 PERCENT	95 PERCENT
INCUBATION PERIOD	TWO TO 10 DAYS	ONE TO TWO DAYS
SYMPTOMS	FEVER, COUGH, MUSCLE ACHES, CHILLS, HEADACHE, CHEST PAIN, SPUTUM, DIARRHEA, CONFUSION	FEVER, COUGH, MUSCLE ACHES, CHILLS, HEADACHE, CHEST PAIN, CONFUSION
EFFECTS ON LUNG	PNEUMONIA, PLEURAL EFFUSION	PLEURITIS, NO PNEUMONIA
OTHER AFFECTED ORGAN SYSTEMS	KIDNEY, LIVER, GASTROINTESTINAL TRACT, NERVOUS SYSTEM	NONE
CASE-FATALITY RATIO	15–20 PERCENT	NO FATALITIES

LEGIONELLOSIS (disease caused by *L. pneumophila*) has been seen in two basic forms: Legionnaires' disease and Pontiac fever. The major distinguishing features are the attack rate (much higher in Pontiac fever), the incubation period (shorter in Pontiac fever) and the presence of pneumonia in Legionnaires' disease but not in Pontiac fever, a much milder illness.

is not known with precision. When Hjordis M. Foy of the University of Washington School of Public Health and her co-workers did a prospective study of pneumonia among people enrolled in a health program, their results indicated that *L. pneumophila* was the cause of 12 cases per 100,000 of population per year. If this rate applies nationwide, there are some 26,000 cases of legionellosis in the U.S. per year. Between .3 and 1.5 percent of the serum specimens from pneumonia patients submitted to the CDC have shown evidence of legionellosis. If those specimens are broadly representative of all pneumonia patients, from 7,000 to 36,000 of the nation's 2.4 million annual cases of pneumonia are in fact cases of Legionnaires' disease.

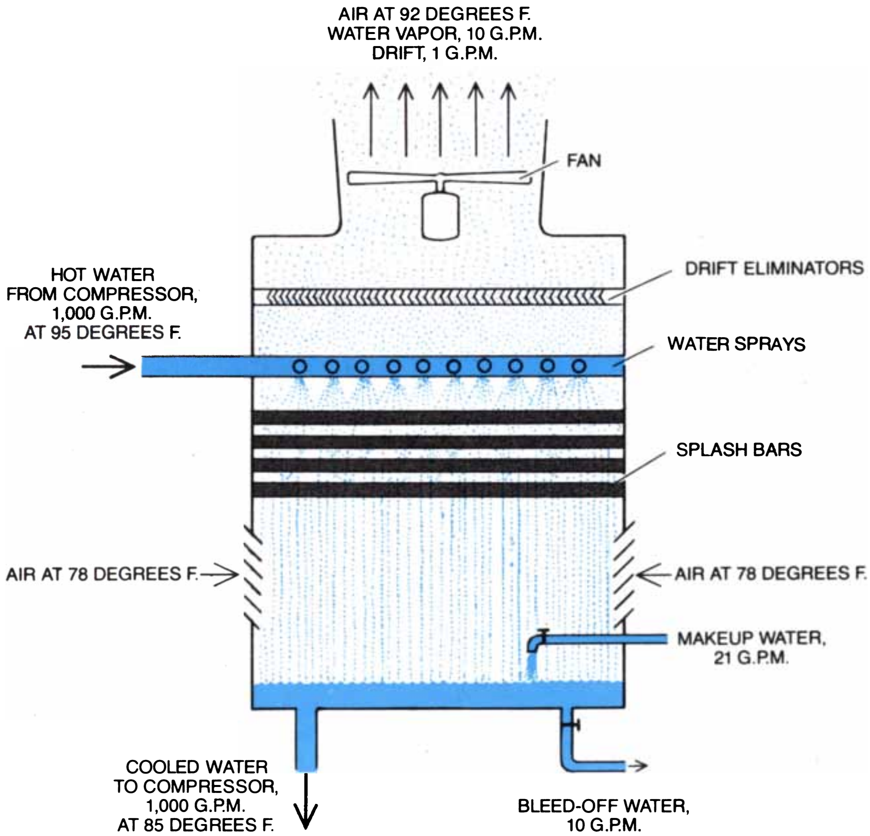
Ecology of the Disease

Most infectious agents that cause epidemics of pneumonia but do not spread from person to person are found to spread through the air from some characteristic ecological niche in the nonhuman environment. A common example is the fungus *Histoplasma capsulatum,* which lives in the soil and causes disease when contaminated dust is stirred up and inhaled. It seemed likely that *L. pneumophila* might live in such an inanimate environment; this hypothesis was strengthened by laboratory studies showing that the bacterium could survive (although apparently not proliferate) for more than a year in tap water.

The 1965 outbreak at St. Elizabeth's Hospital suggested that *L. pneumophila* might live in the soil. During the summer of that epidemic several sites on the hospital grounds had been excavated for the installation of a new lawn-sprinkler system. Stephen B. Thacker and John V. Bennett of the CDC found that it was mainly patients whose beds were near windows in buildings close to excavation sites who became ill. The cases were clustered in time, with each cluster coming five or six days after an excavation site had been filled in, suggesting that contaminated dust raised in the process had spread through the air to infect the patients. No attempt was made in 1965 to recover a disease agent from the soil (and, as we have shown, it would have been difficult to isolate *L. pneumophila* from such specimens in any case), but the organism has since been isolated from mud at other locations.

In three recorded outbreaks the water in the cooling towers or evaporative condensers of air-conditioning systems has been implicated as the source of infection. In a cooling tower the water that has been warmed by heat from a compressor is sprayed over splash bars of wood or other material as air is drawn past the falling droplets by a fan. Some of the water evaporates, cooling the rest, which is recirculated. A small amount of water is entrained as drift, or minute droplets in the exhausted air; the drift carries with it whatever is in the cooling-tower water, including any bacteria that may be present, and it can be carried a considerable distance. In an evaporative condenser the water is sprayed over metal coils containing the refrigerant, which is cooled directly by the evaporation of the water.

In two outbreaks associated with evaporative condensers and one associated with a cooling tower, patients were



COOLING TOWERS for air-conditioning systems have been implicated as spreaders of *L. pneumophila.* A typical cooling tower of a size appropriate for a hotel, hospital or large office building may handle about 1,000 gallons of water per minute. Coolant water from the compressor is sprayed over splash bars; a fan draws fresh air through the spray to maximize evaporation, which cools the water. Some water is exhausted as vapor and a small amount of drift (water droplets); the rest is recirculated. Airborne bacteria could be drawn in by fan, entrained in spray water and exhausted in drift. In one case drift from tower contaminated conditioned air.

shown to have been exposed to drift. In all three instances *L. pneumophila* was isolated from water found in the cooling device. George K. Morris of the CDC has also recovered the bacterium from towers that are not known to have been associated with any cases of legionellosis, and it therefore seems likely that both the towers and the condensers are quite commonly contaminated with airborne soil bacteria gathered from the air by the spraying water. If this is the case, the conditions remain to be defined under which certain towers and condensers that are contaminated but for a time do no harm come to disseminate the bacteria. Meanwhile studies are under way to find antibacterial agents that are effective against *L. pneumophila* and could serve to decontaminate air-conditioning equipment.

Unanswered Questions

It seems paradoxical that an organism so fastidious in its nutritional requirements and so hard to grow in the laboratory apparently persists and succeeds so well in an inanimate environment. *L. pneumophila*'s precise ecological niche and nutritional supply have not been defined. Learning its natural history would be a big step toward understanding how people are exposed to the bacterium in individual cases of legionellosis and under what conditions outbreaks develop.

It is still not known why *L. pneumophila* causes two such different illnesses as Legionnaires' disease and Pontiac fever, or why the bacterium affects only 1 percent of the people apparently exposed to it in the case of the former illness and as many as 95 percent in the case of the latter one. It may be that strains of the bacterium differ in some critical way that determines which illness develops. Laboratory testing, including a search for toxins secreted by the bacterium, has yet, however, to show any significant difference between the Pontiac strain and the Philadelphia one.

It is quite possible that Legionnaires' disease and Pontiac fever are not the only disease syndromes caused by *L. pneumophila.* Clinicians have only recently begun to investigate the possibility that the bacterium may be responsible, for example, for certain cases of lung abscess and endocarditis (inflammation of the lining of the heart). Moreover, only about two-thirds of the 2.4 million annual U.S. pneumonia cases can be linked to a known viral or bacterial agent; what causes the 800,000 or so other cases is not known, and identifying the agents is currently a major challenge. It is conceivable that some of the unknown agents are bacteria related to *L. pneumophila* that have not yet been detected. Continuing investigation of legionellosis should provide an opportunity to find any such related organisms.

Painting by Robert Giusti

The Photosynthetic Membrane

The conversion of light energy into chemical energy by green plants is accomplished in the thylakoid membrane of the plant cell. Electron microscopy reveals the asymmetry that makes the conversion possible

by Kenneth R. Miller

The earth is a planet bathed in light. It is therefore not surprising that many of the living organisms that have evolved on the earth have developed the capacity to trap light energy. Of all the ways in which life interacts with light the most fundamental is photosynthesis, the biological conversion of light energy into chemical energy. In an energetic sense all living things are ultimately dependent on photosynthesis, which is the source of all forms of food and even of the oxygen in the earth's atmosphere. The vast majority of living cells, from simple algae to large and complex terrestrial plants, are photosynthetic. Much has recently been learned about the design and function of the biological structure in which light energy is initially trapped: the photosynthetic membrane.

The overall chemistry of photosynthesis can be expressed in a deceptively simple equation. Six molecules of carbon dioxide are taken up from the environment together with six molecules of water. In the presence of light these molecules are converted into a single molecule of the six-carbon sugar glucose and six molecules of oxygen are released. The products of the reaction hold more chemical energy than the reactants do, and so in a sense the energy of sunlight is captured in the glucose and oxygen that are produced in the process.

For purposes of analysis photosynthesis is most conveniently divided into two steps: a light reaction and a dark reaction. The light reaction captures energy from the sun in two comparatively unstable molecules: adenosine triphosphate (ATP) and nicotinamide adenine dinucleotide phosphate (NADPH). In the dark reaction ATP and NADPH supply the energy needed to form glucose from carbon dioxide.

Both of these reactions, the light and the dark, take place in regions of the plant cell known as chloroplasts. The chloroplast is a complex organelle that has its own genetic material and is able to synthesize at least a few of the proteins needed for its own functioning. It is filled with membranous sacs called thylakoids, which in many chloroplasts are piled on top of one another in stacks called grana. Thylakoid membranes carry out the light reaction of photosynthesis; the dark reaction takes place in the soluble, or nonmembranous, component of the chloroplast.

Because the enzymes that carry out the conversion of carbon dioxide into glucose are soluble and can be easily studied by the biochemist, the dark reaction was the part of photosynthesis that yielded to experimentation first. Melvin Calvin and his associates at the University of California at Berkeley unraveled the enzymatic steps of the dark reaction some 20 years ago. The dark reaction actually occurs most often in the light (when leaves are illuminated); it is called the dark reaction because it does not require the direct participation of light. The reaction can proceed in total darkness if the appropriate compounds are supplied.

The mechanism by which the light reaction achieves the initial conversion of light into chemical energy is the inner sanctum of photosynthesis. Only very recently have workers in several fields of biology and biochemistry begun to arrive at a comprehensive picture of how the light reaction is organized in the photosynthetic membrane.

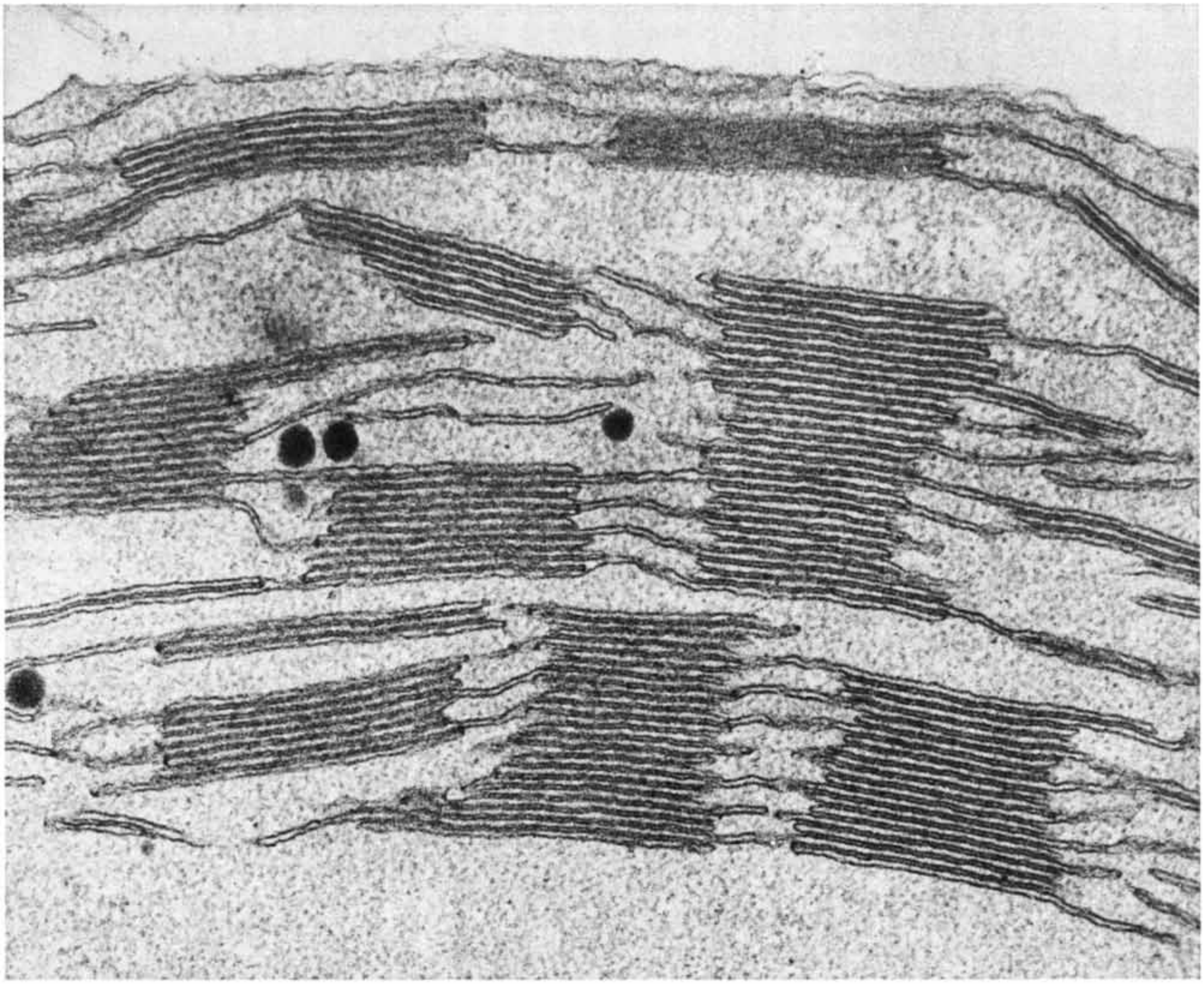
The green pigment chlorophyll is the central molecule of photosynthesis. Chlorophyll and accessory pigments such as beta-carotene and xanthophyll absorb light over much of the visible spectrum. The absorption of a single photon of light by a chlorophyll molecule gives rise to a separation of positive and negative charges in a special region of the membrane known as a reaction center. After the absorption event the reaction center, which has lost an electron, has a positive charge, and the free electron, now the recipient of much of the energy of the absorbed photon, is capable of reducing (that is, adding an electron to) an appropriate molecule in the membrane.

The ability to maintain this charge separation is one of the critical features of the system. Although positive and negative charges are plentiful in the membrane, a mechanism exists to prevent the energized electron from recombining with positive charges and giving up its energy as heat or fluorescence. As we shall see, the thylakoid membrane is specifically designed to harness the energy available in the excited electrons.

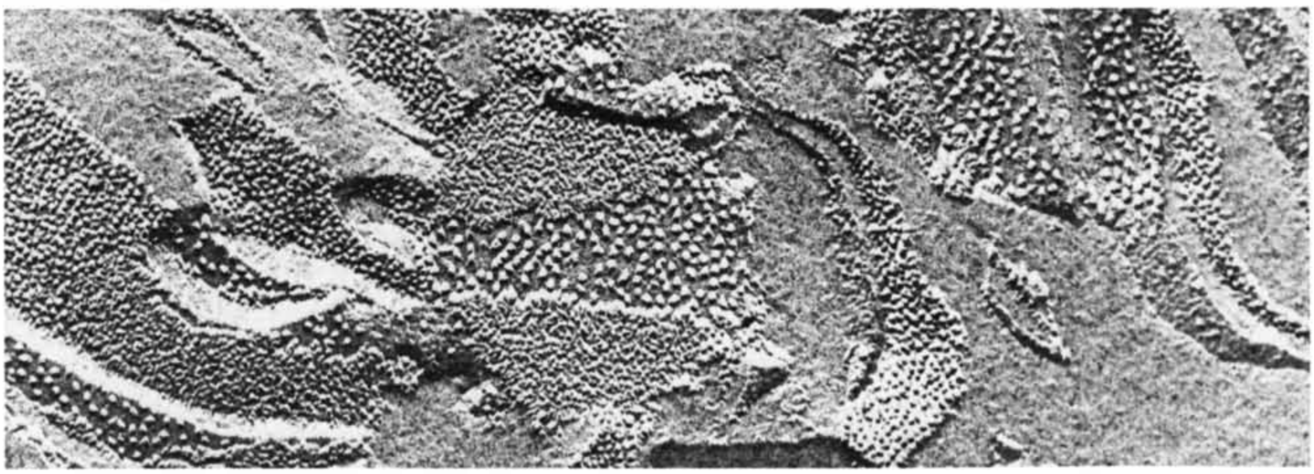
An analogous process of excitation and charge separation takes place in a man-made solar cell based on crystalline silicon. In such a photovoltaic cell *p* (positive) and *n* (negative) regions are created in the silicon crystal by the controlled addition of small amounts of impurities. In effect the *n* region is made more hospitable to negative charges and the *p* region more hospitable to positive ones. If the photovoltaic cell is wired into a circuit, the accumulation of excited electrons can be tapped to yield a small electric current.

The chloroplast lacks the luxury of a wiring system, but it does have a special array of molecules near the reaction center that prevent the energy stored in excited electrons from being wasted. These molecules, some protein and some lipid, form an electron-transport chain that carries electrons from the reaction center and ultimately utilizes them to convert $NADP^+$ into NADPH. $NADP^+$ is the oxidized form of NADPH. On the addition of a pair of energized electrons and a proton $NADP^+$ is reduced to NADPH, one of the molecules required for the dark reaction. Since electrons cannot flow continuously to $NADP^+$ from the reaction center without being replenished, the membrane incorporates a system to reduce the reaction center once it has been oxidized (that is, depleted of an electron). The system removes electrons from water and in the process releases from the membrane protons (hydrogen ions) and oxygen molecules. Electron flow in the thylakoid therefore follows a linear pathway from water to NADPH.

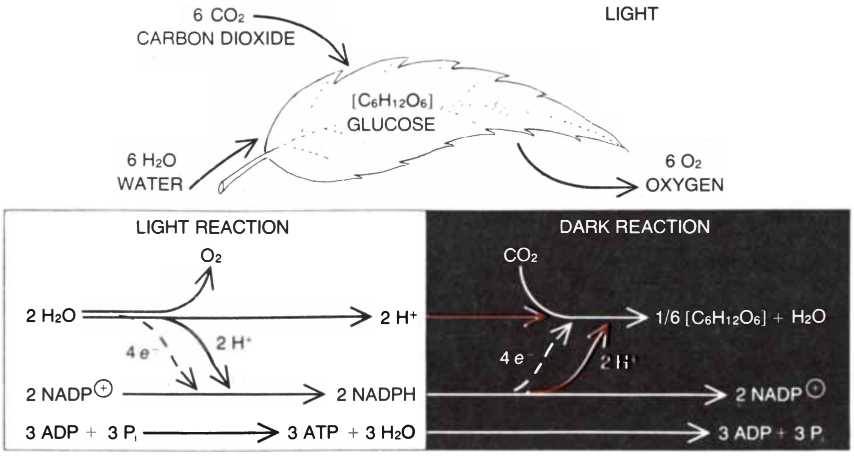
One of the essential features of this linear flow is that it is vectorial, or unidirectional. The movement of electrons is from inside the thylakoid sac to the



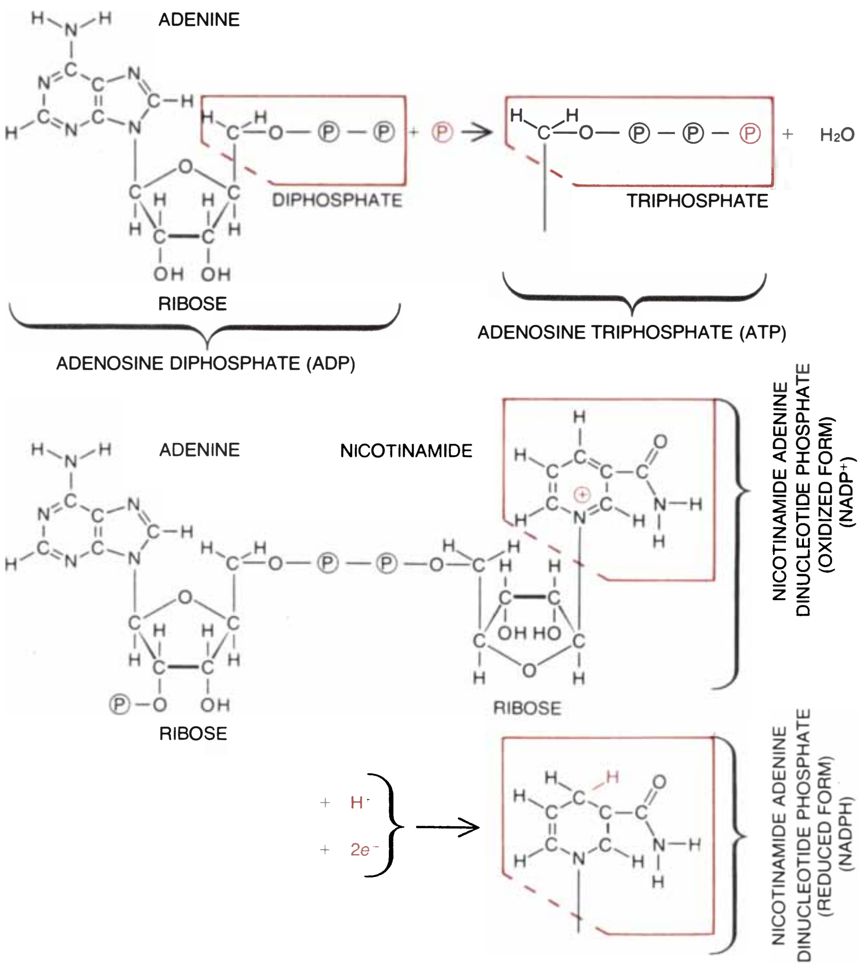
SITE OF PHOTOSYNTHESIS in the green-plant cell is the organelle known as the chloroplast. This electron micrograph shows part of a chloroplast from a leaf cell of a corn plant magnified 70,000 diameters. The closely spaced membranes, which in many cells are piled into stacks called grana, conduct the initial step in photosynthesis: the "light reaction" that traps the photons of light. The capture is carried out by pigment-protein complexes embedded in the membranes. They can be revealed by a technique in which the bilayer membrane structure is split into two halves by freezing and fracturing. Splitting one of the membranes of a closely spaced pair exposes large particles of the kind identified in the micrograph below. Unless it is otherwise indicated all the micrographs in this article were made by the author.



VARIETY OF EMBEDDED PARTICLES are exposed when grana are frozen and fractured. The fractured surface is "shadowed" with platinum and coated with carbon, producing a replica that can be separated from the frozen sample and examined in the electron microscope. The complex contours arise because the fracture tends to jump from one level to another as it passes through the membranes. Large particles in middle are exposed when one membrane of a close pair in a granum is fractured. The magnification is 70,000 diameters.



OVERALL CHEMISTRY OF PHOTOSYNTHESIS is simply expressed (*top*) as the synthesis of glucose (a six-carbon sugar) from water and carbon dioxide with the release of oxygen. The products of the reaction store more energy than the reactants do, so that the energy of light is captured chemically. The process has two distinct steps (*bottom*). In the first step (the light reaction) photons captured by chlorophyll and other pigments supply the energy to split two molecules of water into a molecule of oxygen, four protons (or hydrogen nuclei, denoted H^+) and four electrons (e^-). Simultaneously the four electrons and two of the protons convert two molecules of nicotinamide adenine dinucleotide phosphate ($NADP^+$) into its reduced form (NADPH). Another portion of the light energy helps to add inorganic phosphate (P_i) to adenosine diphosphate (ADP) to form a molecule that stores more energy: adenosine triphosphate (ATP). In the dark reaction (so named because it can proceed in the absence of light) the energy stored in NADPH and ATP is extracted to convert carbon dioxide and hydrogen into glucose and water. Half of the hydrogen and all the electrons that are needed are supplied by NADPH.



STRUCTURAL SIMILARITIES are evident in the two molecular systems that provide temporary chemical storage for the light energy trapped by the photopigments. In the ADP-ATP system energy is stored in bond of the added phosphate group. In $NADP^+$-NADPH system energy storage is achieved by the addition of a proton and two electrons to nicotinamide group.

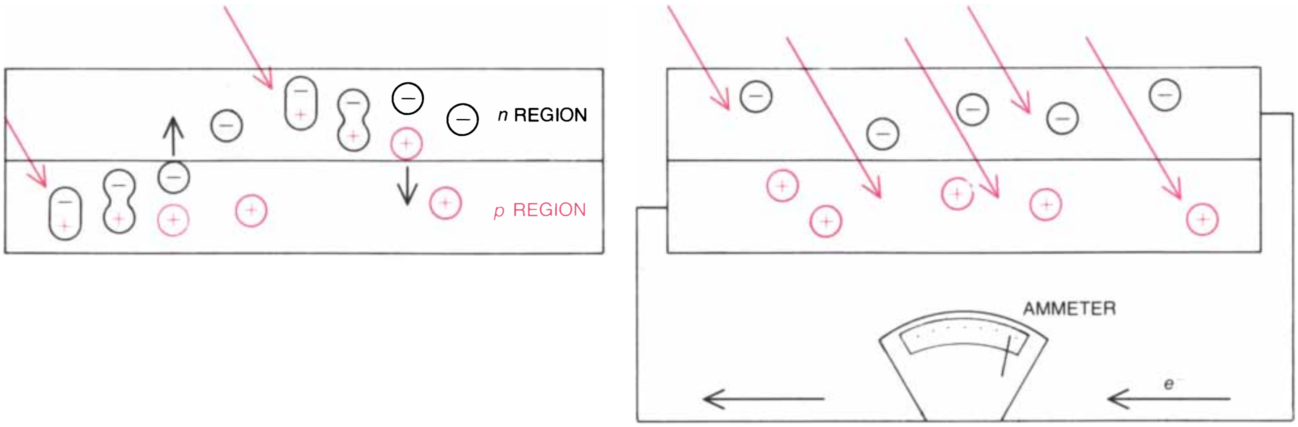
outside. Therefore the inside of the sac quickly becomes positively charged and the outside becomes negatively charged. This has two important consequences. First, the displacement of electrons creates an electric potential across the membrane, similar in many respects to the potential developed across the *p-n* junction of a silicon photovoltaic cell. Second, the concentration of protons is now dramatically different on the two sides of the membrane, with the interior being much the richer. Because the random forces of diffusion tend to drive molecules in such a way as to equalize their concentrations on both sides of a barrier, the difference in concentration, or proton gradient, is a source of potential energy. The larger the proton gradient, the greater the amount of potential energy.

One can now begin to see why membranes are associated with the photosynthetic process. It is likely that the light reaction is associated with a membrane precisely in order to hold the reaction centers and electron-transport systems in an asymmetrical orientation capable of establishing a proton gradient. The membrane also provides a tough barrier to maintain the difference in charge that has built up, and to store that energy for a very brief period.

Although there is (just barely) enough energy in a single absorbed quantum of light to drive electron flow uphill from water to $NADP^+$, the thylakoid membrane evidently requires the absorption of two quanta at separate reaction centers, linked in series, to accomplish the task. The two reaction centers, generally called photosystem I and photosystem II, were discovered because the chlorophyll molecules at the two centers absorb light of slightly different wavelengths. Why does the plant expend an additional photon on the transport of electrons? The answer may lie in the way the other molecule required by the dark reaction—ATP—is synthesized by the thylakoid.

The transport of electrons is accompanied by an acidification of the internal space of the thylakoid sac. The hydrogen ions responsible for the acidification seem to come from two sources: the protons released during the removal of electrons from water and the protons actually pumped across the thylakoid membrane in the course of electron transport. Much of the energy trapped in electrons excited by the absorption of light seems to be consumed in this shuttling of protons across the membrane, with much of the remaining energy accounted for by the formation of NADPH.

The British biochemist Peter Mitchell was the first to realize that the buildup of protons in the thylakoid sac might serve as a source of energy to drive phosphorylation: the adding of a phos-

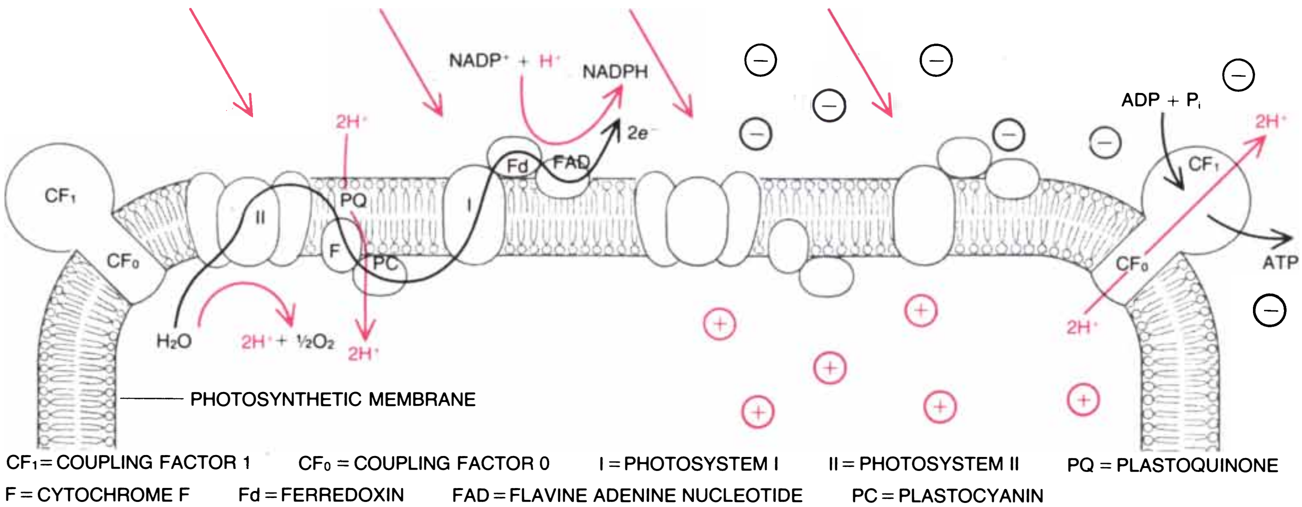


MECHANISM OF CHARGE SEPARATION is exploited in man-made solar cells. Such photovoltaic cells consist of crystalline silicon grown in the form of a junction in which two adjacent layers are separately doped with trace amounts of other elements. These "impurities" make one layer hospitable to positive charges (*p*) and the other layer hospitable to negative charges (*n*). When a photon enters the *n* region of the cell, an energized positive charge (actually a "hole," or the absence of an electron) passes into the *p* region. Similarly, the absorption of a photon in the *p* region releases a negative charge into the *n* region (*left*). Under steady illumination (*right*) the buildup of positive and negative charges on opposite sides of the *p*-*n* junction can eventually reach a level as high as .6 volt. The effectiveness of the photovoltaic cell rests on its inherent asymmetry and the ability of the junction to keep the positive and negative charges separate.

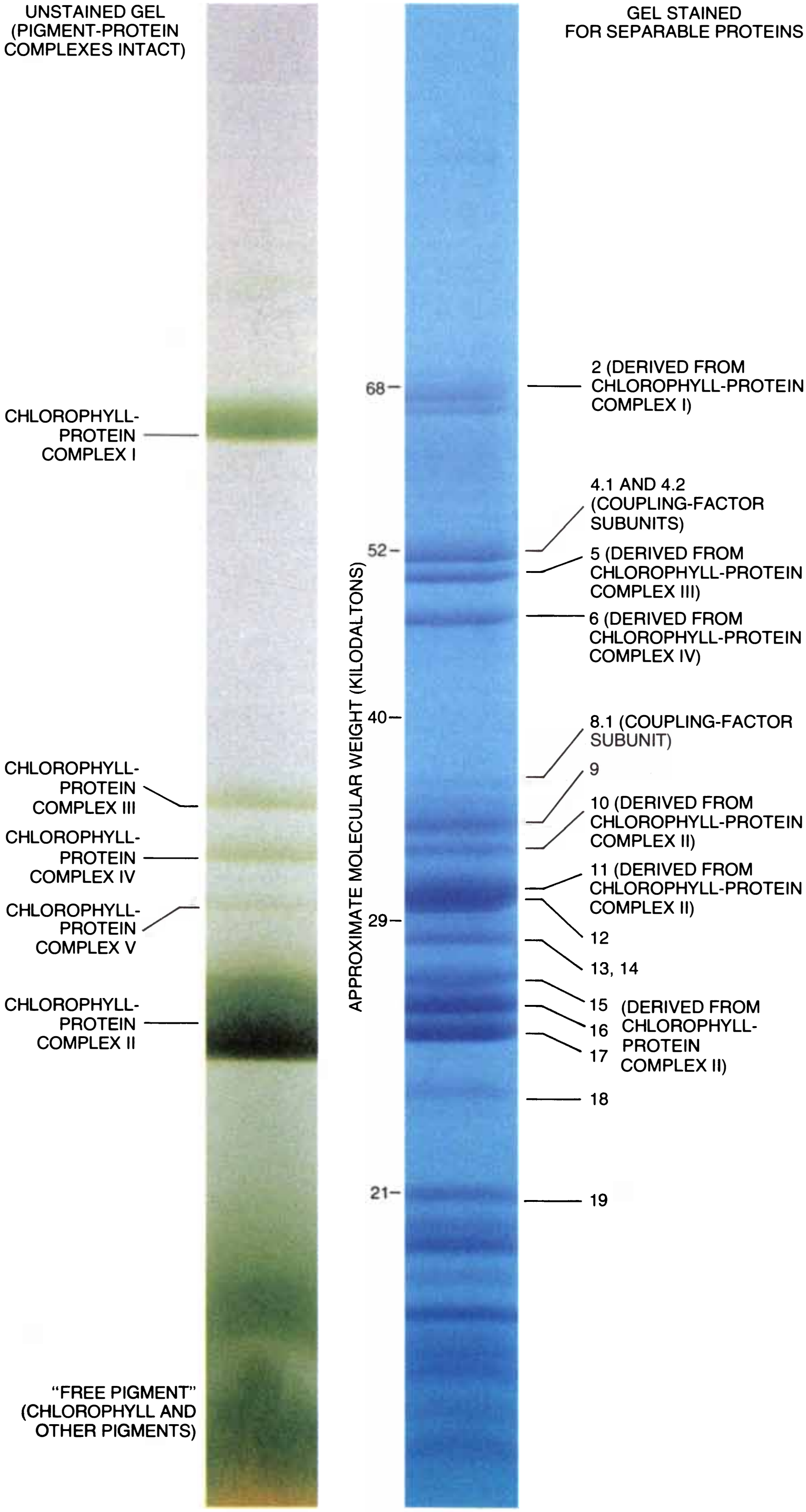
phate group to adenosine diphosphate (ADP) to form ATP. The phosphorylation of ADP to ATP calls for a source of energy and therefore does not occur spontaneously. Mitchell recognized that an electrochemical gradient created across a membrane by the transport of protons might serve as just such a source. He suggested that specialized enzymes in the thylakoid membrane might be able to extract the energy inherent in the gradient to drive the synthesis of ATP. At the time he proposed his ideas few biochemists were willing to consider them. Now the essential elements of Mitchell's proposals have been accepted by most investigators.

The key predictions of Mitchell's model, which is often called the chemiosmotic theory, were confirmed by experiments in many laboratories. Among them was an experiment by André T. Jagendorf of Cornell University and Ernest G. Uribe of Washington State University showing that ATP could be synthesized in total darkness if a proton gradient was artificially created across the thylakoid membrane. Work in Efraim Racker's laboratory at Cornell demonstrated that a large protein molecule at the outer surface of the membrane was critical to photophosphorylation. Although electron transport was not disturbed when the molecule was removed from the membrane, ADP could no longer be phosphorylated. When the molecule was added back to the membrane, photophosphorylation could again proceed. Because this molecule couples electron transport to photophosphorylation it is generally called the coupling factor.

It should be noted that one of Mitchell's critical predictions is that the thylakoid membrane should be asymmetri-



CHAIN OF REACTIONS powered by photons is exploited by the photosynthetic membrane inside the chloroplast to accomplish a separation of charges. In the process $NADP^+$ is reduced to NADPH. The voltage differential that results from the separation of charges also helps to drive the conversion of ADP into ATP. The two processes, which constitute the light reaction, proceed simultaneously in the membranous sacs called thylakoids. For the purposes of clarity the processes are depicted here as taking place separately within a thylakoid that is part of a large granum. At the left in the illustration two reaction centers, designated photosystem I and photosystem II, absorb light of slightly different wavelengths. In photosystem II water is dissociated into two protons (H^+), oxygen and two electrons (e^-). The electrons are shuttled through an electron-transport chain involving plastoquinone, plastocyanin, photosystem I, ferredoxin and flavine adenine dinucleotide, which finally enables the electrons, with the aid of a proton, to reduce $NADP^+$ to NADPH at the outside surface of the thylakoid membrane. At the right positive charges have accumulated on the inside of the membrane from two sources: the protons that are left behind after dissociation of water and positive charges that seem to be shuttled across the membrane from the exterior in electron transport. The outflow of protons through the coupling factors (CF_0 and CF_1) drives the synthesis of ATP from ADP.



COMPONENTS OF THE PHOTOSYNTHETIC MEMBRANE can be separated by dissolving the membrane in a detergent (lithium dodecyl sulfate) and allowing the protein components to migrate through a polyacrylamide gel under the influence of an electric field, the technique known as electrophoresis. The rate of migration of a molecule is determined largely by its size. The green color of several of the fractions in the chromatogram at the left shows that chlorophyll forms complexes with several different proteins in the membrane. The chromatogram at the right has been stained with a blue dye that brings out all the polypeptide components (proteins and protein subunits) of the membrane. Bands are more numerous and appear at positions different from those in the gel at the left because the pigment-protein complexes were dissociated by heating. The chromatograms were made by Nam-Hai Chua of Rockefeller University.

cal, both structurally and biochemically. A perfectly symmetrical membrane could not build up the differences of proton concentration that are required for photophosphorylation, nor could it effectively utilize such gradients. As we shall see, this prediction has been dramatically realized in studies of the structural organization of the thylakoid membrane.

The purple membrane of the salt-loving bacterium *Halobacterium halobium,* the simplest energy-transducing membrane, has only a single protein for shuttling protons in response to light. The thylakoid membrane of the chloroplast is far more complex. In electrophoretic gels, where proteins and shorter chains of amino acids (polypeptides) can be separated into bands according to their molecular weight, the proteins and polypeptides of the thylakoid membrane can be resolved into at least 40 molecular species, of which only a few have been identified biochemically. The identified bands include some of the components of the coupling factor and also a few of the electron-transport components.

One of the main questions engaging the interest of investigators working on the thylakoid membrane is how the green pigment chlorophyll is dispersed within it. The membrane basically consists of a bilayer of lipid molecules, with protein molecules either embedded in it or associated with its surface. Chlorophyll can freely associate with a lipid bilayer, so that there is at least a possibility that some chlorophyll molecules are freely mobile in the membrane. Workers in several laboratories have now found, however, that most of the pigment seems to be associated with particular proteins. When the membrane is carefully dissolved in certain detergents, several proteins bind chlorophyll and appear as green bands in an electrophoretic gel. Intensive work is now under way in several laboratories in the effort to learn more about such chlorophyll-protein complexes. Several of the complexes have been shown to be associated directly with either photosystem I or photosystem II.

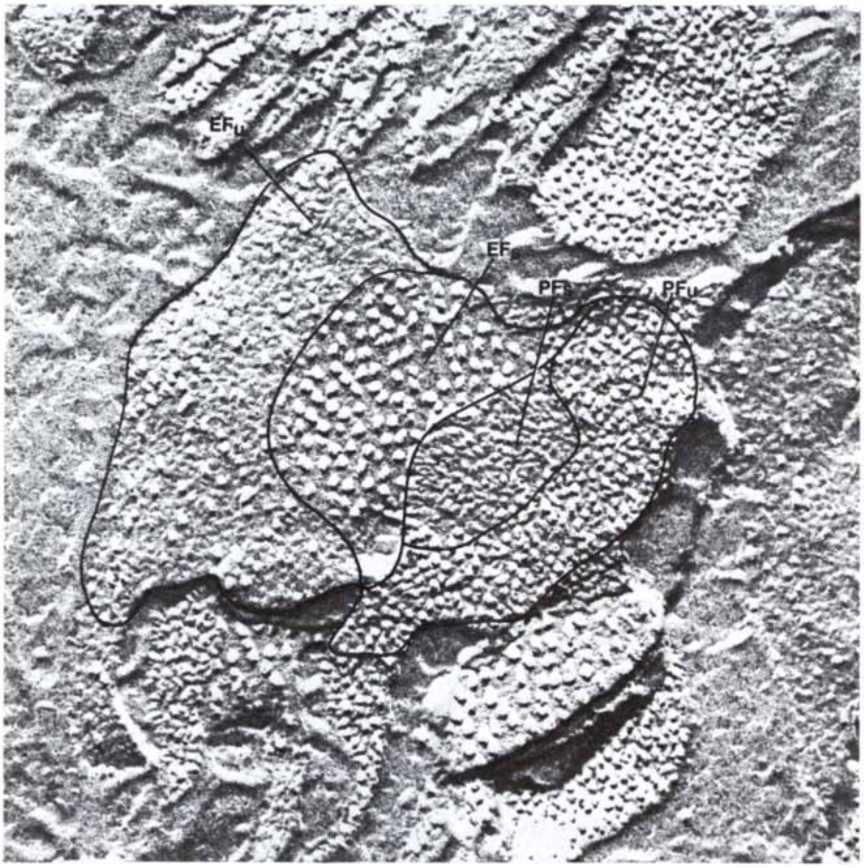
As I have mentioned, the two photosynthetic reaction centers are driven by slightly different wavelengths of light. Does this imply that the two centers exist in separate structural centers within the membrane? Apparently it does. When thylakoid membranes are dissolved in very mild detergent, large complexes can be recovered from the membrane that contain chlorophyll, together with several polypeptides, and that display activity characteristic of either photosystem I or photosystem II. Experiments such as this one, which have been done with a wide variety of detergent agents, are evidence for two physically distinct reaction centers.

In my laboratory at Harvard University we have been concerned with developing a structural description of the thylakoid membrane that will deepen understanding of the light reaction of photosynthesis. Although conventional techniques of preparing specimens for electron microscopy are useful, they have not enabled us to investigate the nature of structures within the photosynthetic membrane itself. We have therefore resorted to the technique of freeze-fracturing, in which a sample of photosynthetic membranes is rapidly frozen and then fractured in a vacuum chamber. The surface exposed by the fracture turns out to be of considerable interest for studying the organization of a biological membrane. A replica is made of the surface by covering it with a thin layer of platinum and carbon, which conforms to the shape of the fractured surface and can be removed from it and placed in the electron microscope. Such replicas reproduce very fine detail from the original specimen. Moreover, by placing the platinum source to one side of the sample, shadows are created that indicate the relative heights of projecting structures as if they were seen in three dimensions.

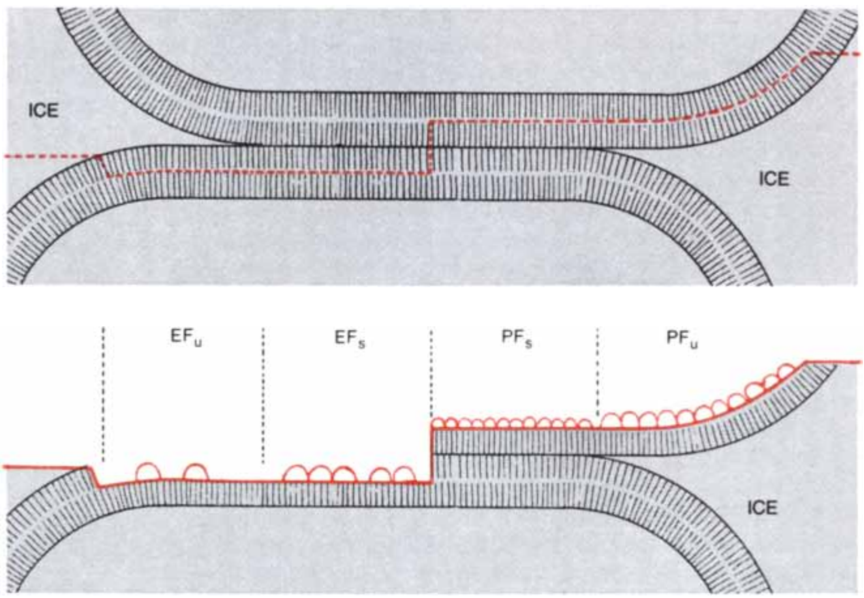
Freeze-fracture replicas can be prepared in two slightly different ways. When a membrane specimen is simply fractured, the lipid bilayers of membranes in the path of the fracture split along a central plane, thereby separating the two layers. For certain purposes the freshly fractured surface can be "etched" by allowing ice to sublime away from the frozen membrane before the replica is made. Etched preparations are useful for revealing the outer surfaces of a membrane that would otherwise be covered with ice.

Whereas natural or artificial membranes with little or no protein associated with them exhibit large, smooth surfaces when they are frozen and fractured, the photosynthetic thylakoid membrane reveals a bewildering variety of particles of different sizes and shapes. Indeed, the complexity is so great that several years passed before those of us involved in freeze-fracture studies were able to agree on what we were seeing. The complexity of the structure is indicative of the complexity of the membrane's function. It has turned out, for example, that the internal organization of the membrane is different in regions where two membranes are stacked on top of each other. The stacking, which is characteristic of the thylakoid membrane, gives rise to the piles of grana visible in the intact chloroplast.

In regions where membranes are stacked one can see a dense concentration of large particles on one of the fracture faces. In adjacent regions of the same membrane that are not stacked the density of the same type of particle is



PARTICLES OF VARIOUS TYPES are visible in photosynthetic membranes of barley that have been exposed by the freeze-fracture technique and magnified 100,000 diameters. The fracture path jumped from the middle of one membrane to the middle of an adjacent closely stacked membrane, so that four different "fracture faces" were exposed, designated EF_u, EF_s, PF_s and PF_u. EF and PF stand for exoplasmic face and periplasmic face; s and u, for stacked and unstacked. An explanation of how the different faces are formed is diagrammed below.

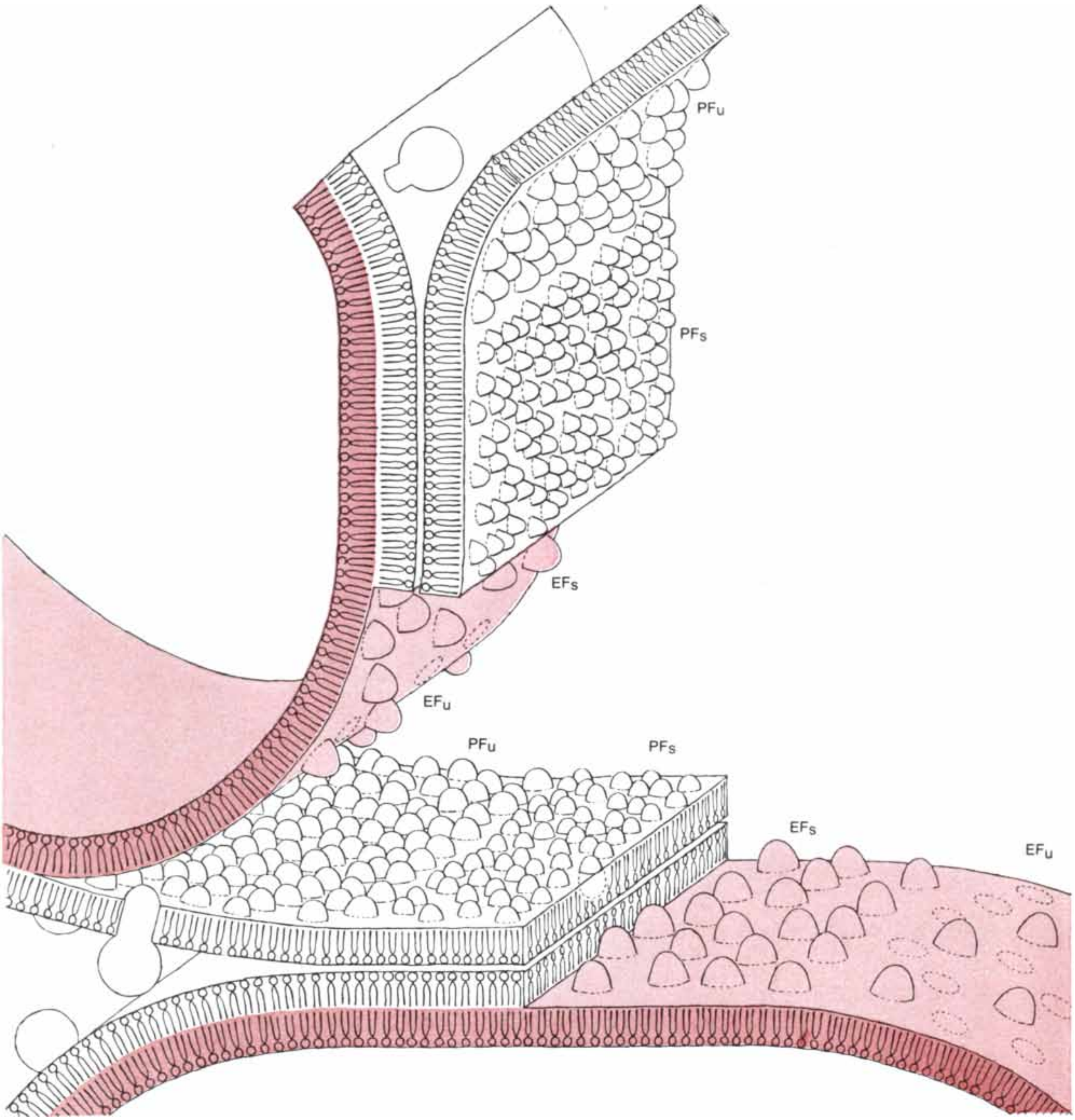


FRACTURE PATH that exposed the four types of fracture face in the micrograph at the top of the page is depicted schematically. The largest particles (EF_s) are on the inner half of a thylakoid membrane that closely abuts another membrane. Where the two membranes are not in contact, at the left, the large particles are reduced in size and number; the exposed face is one designated EF_u. Where the fracture has split the upper membrane of the adjacent pair, at the right, the exposed faces are different again. Here the exposed surface is the inside of the outer half of the thylakoid membrane rather than the inner half. Exposed faces are those designated PF_s in the region where the membranes touch each other and PF_u where membranes do not.

much lower. In 1966 Seikichi Izawa and Norman E. Good of Michigan State University noted that when thylakoid membranes are suspended in a solution of low ionic strength, the stacked regions disappear and the membranes appear freely unfolded. When such membranes are studied by freeze-fracturing, internal surfaces that have dense populations of large particles are no longer seen.

Thylakoid membranes that have been studied by the etching technique show that Mitchell's prediction of a highly asymmetrical photosynthetic membrane is borne out. The inner and outer surfaces of the membrane are remarkably different. The outer surface is covered by a mixture of large particles (about 12 nanometers across) and smaller ones (about eight nanometers across). The inner surface is densely covered by boxcar-like particles with an overall diameter of about 18 nanometers, seemingly composed of four or more subunits.

Several years ago, while I was working in L. Andrew Staehelin's laboratory at the University of Colorado, I removed the coupling factor, which synthesizes ATP from ADP, from the photosynthetic membrane in order to see how its absence might change the appearance of the freeze-etched specimens. Our results and those from several other laboratories showed that the large particles on the outer surface of the membrane are indeed molecules of coupling factor. In studying the effect of



ASYMMETRY OF THYLAKOID MEMBRANE is demonstrated by analysis of the faces exposed by freeze-fracturing. The particles on the exposed inner half of the fractured membrane (*color*) are different in appearance from the particles on the exposed outer half (*black*). In other words, the membrane has a distinct inside-outside polarity. Large particles on the EF_s and EF_u faces seem to be multiunit structures associated with photosystem II. Some PF_s and PF_u particles may be associated with photosystem I. Light-harvesting components, portions of electron-transport chain and membrane-bound elements of coupling factor may also form parts of the small PF particles.

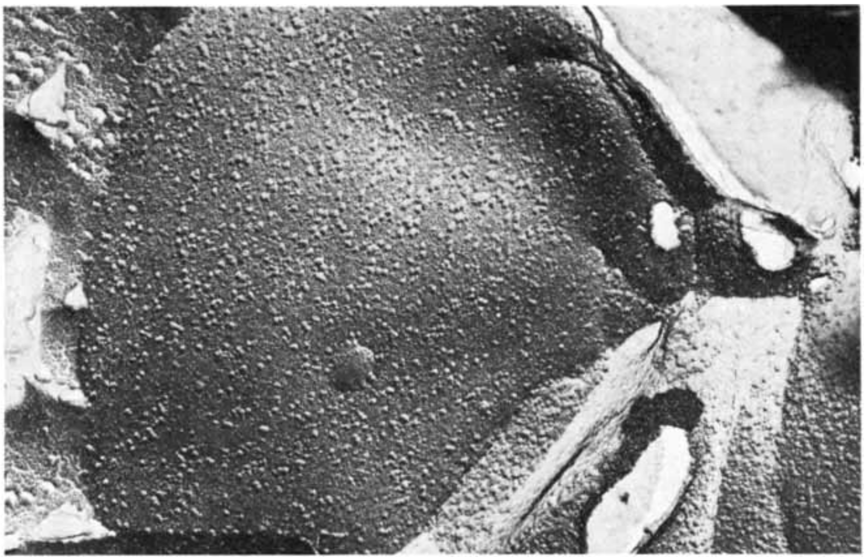
membrane stacking and unstacking on the density of these particles we found, to our great surprise, that they seemed to be excluded from the outer surface of the thylakoid, where the membranes came together to form part of a granum; they were found only on surfaces in unstacked regions. This result is further support for Mitchell's chemiosmotic theory because it shows that the ATP-synthesizing system need not be directly connected to an electron-transport chain in order to take advantage of the light-induced proton gradient.

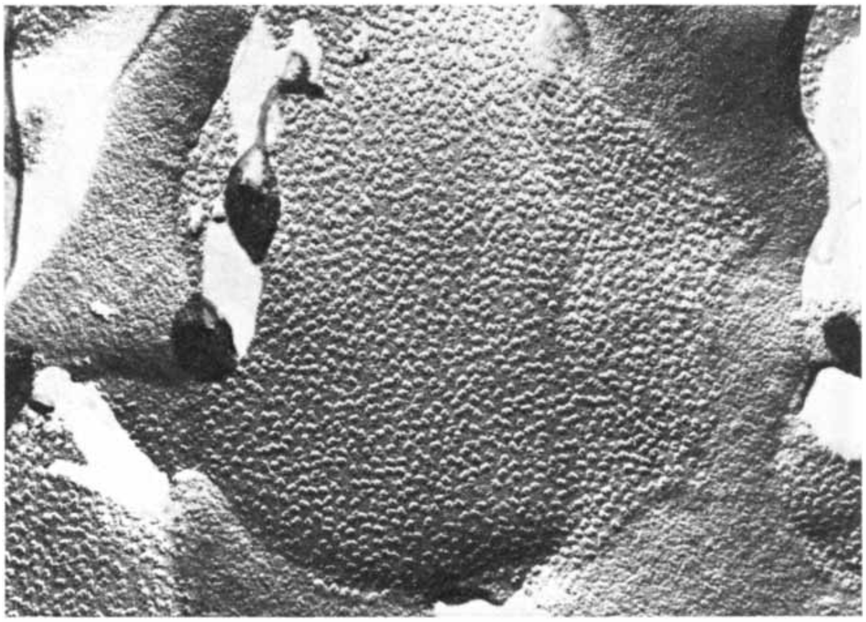
The inner surface of the membrane does not show coupling-factor molecules but rather is populated by the curious particle with four or more subunits. Under certain conditions, which we have only begun to understand, these particles seem to "crystallize" into a regular two-dimensional lattice. Such lattices are found not only at the inner surface of the membrane but also at the outer one and in fractured preparations. The particle is visible at both surfaces of the thylakoid and in the interior membrane as well, so that one must conclude that the particle spans the membrane. What is the nature of this unusual structure, and what is its association with photosynthesis?

Independent studies conducted in my laboratory and in the laboratory of Charles J. Arntzen at the University of Illinois at Urbana gave the first evidence that a particular chlorophyll-protein complex (CP II) is associated with the membrane-spanning structure. Each study demonstrated that the absence of CP II resulted in a reduction in the diameter of the large membrane-spanning particle; each group therefore concluded that this particle must be the reaction complex of photosystem II. Evidently the photosystem II reaction center is at the core of the particle and CP II molecules are arrayed around it.

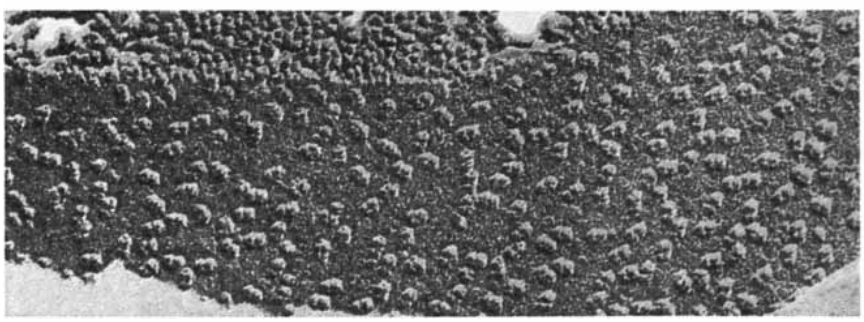
More direct evidence has recently been obtained in a study of a mutant tobacco plant that is deficient in photosystem II activity. Robert Cushman and I found that the thylakoid membranes of the mutant deficient in photosystem II are almost completely devoid of the large particle and that all the other structures of the membrane are apparently unaltered by the mutation. The fact that the large membrane-spanning particles seem to be concentrated in stacked regions of the thylakoid-membrane system is also consistent with the idea that they represent photosystem II reaction complexes. Studies in which the membrane system is broken up into grana and stroma (nongrana) fractions show that photosystem II activity is much higher in the grana membranes, which contain large numbers of the particles.

Recent studies suggest that membrane stacking may also be mediated by CP II

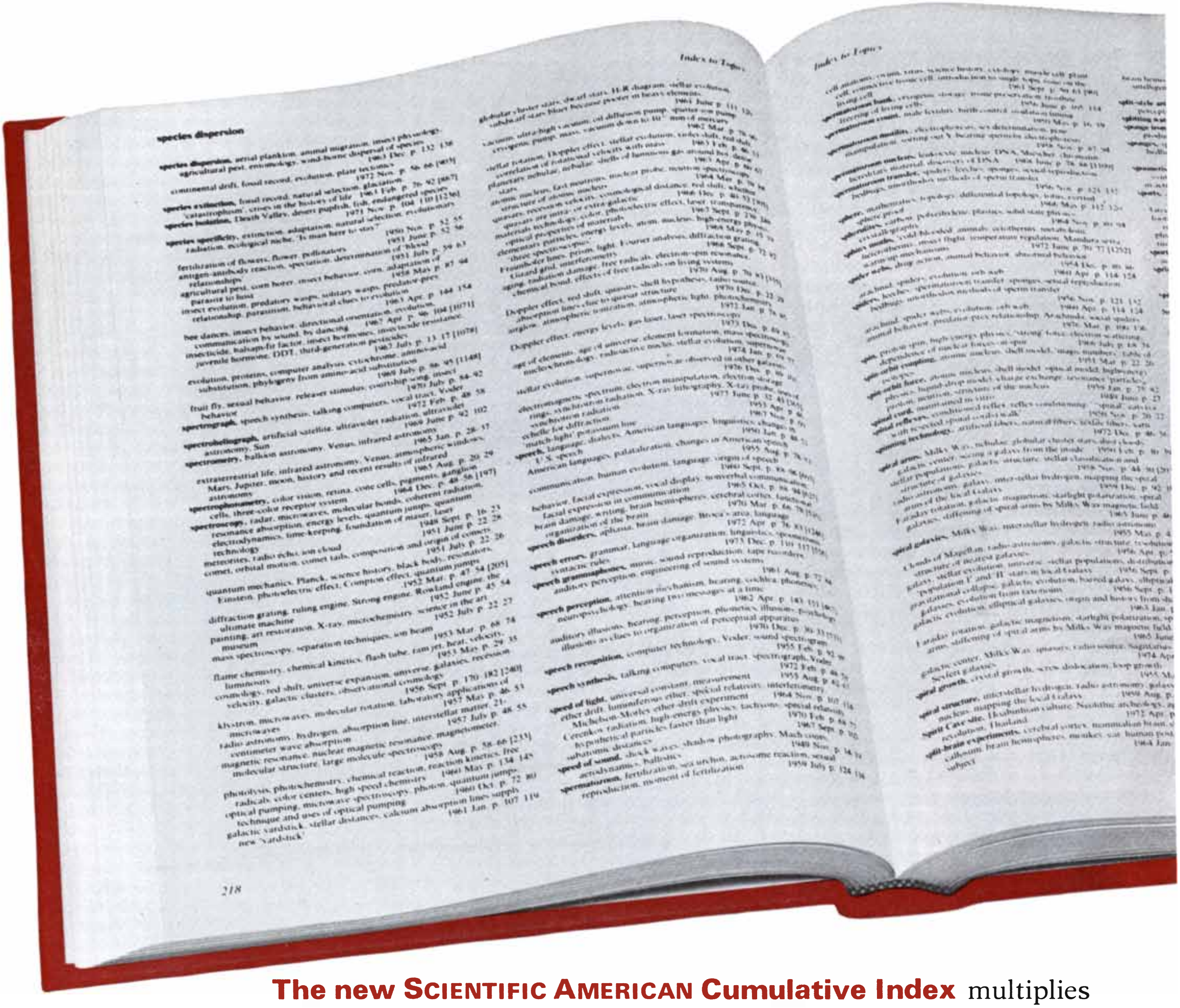




EXTERIOR SURFACES OF THE THYLAKOID MEMBRANE are made visible by freeze-etching. In this method surfaces of membranes that have not been fractured but that were originally covered by ice are exposed by allowing the ice to sublime away. The exposed surfaces are then shadowed with platinum and replicated as usual. The outer surface of the photosynthetic membrane (*top*) is covered by a mixture of large and small particles, 12 nanometers and eight nanometers in diameter. The larger particles have been identified as coupling-factor molecules. Inside surface of membrane (*bottom*) is studded with a boxcar-like particle that seems to have four or more subunits. Magnification in both micrographs is 55,000 diameters.



MAGNIFIED VIEW OF INSIDE SURFACE of the thylakoid membrane shows the boxcar-like particles in greater detail. The magnification here is 135,000 diameters. The particles penetrate the membrane and can also be recognized on the outside surface. These membrane-spanning structures correspond to the photosystem II particle EF_s in illustration on opposite page.



The new SCIENTIFIC AMERICAN Cumulative Index multiplies many times the value and usefulness of the collected issues of the magazine. The Index is actually eight different indexes in one:

Index to Topics—a rotated key-word index of the subject matter of all articles and Science and the Citizen features (see above).

Index to Proper Names—the names of all persons mentioned in the articles and Science and the Citizen and of places and institutions featured in a primary role.

Listing of Tables of Contents—a chronological listing of the complete contents of all 362 issues that provides ready identification, by titles and authors, of the articles cited in the two foregoing indexes.

Index to Authors—an alphabetical listing of authors of all articles.

Index to Titles of Articles—an alphabetical listing of all articles by the first key word in each title, and by the other key words in the title.

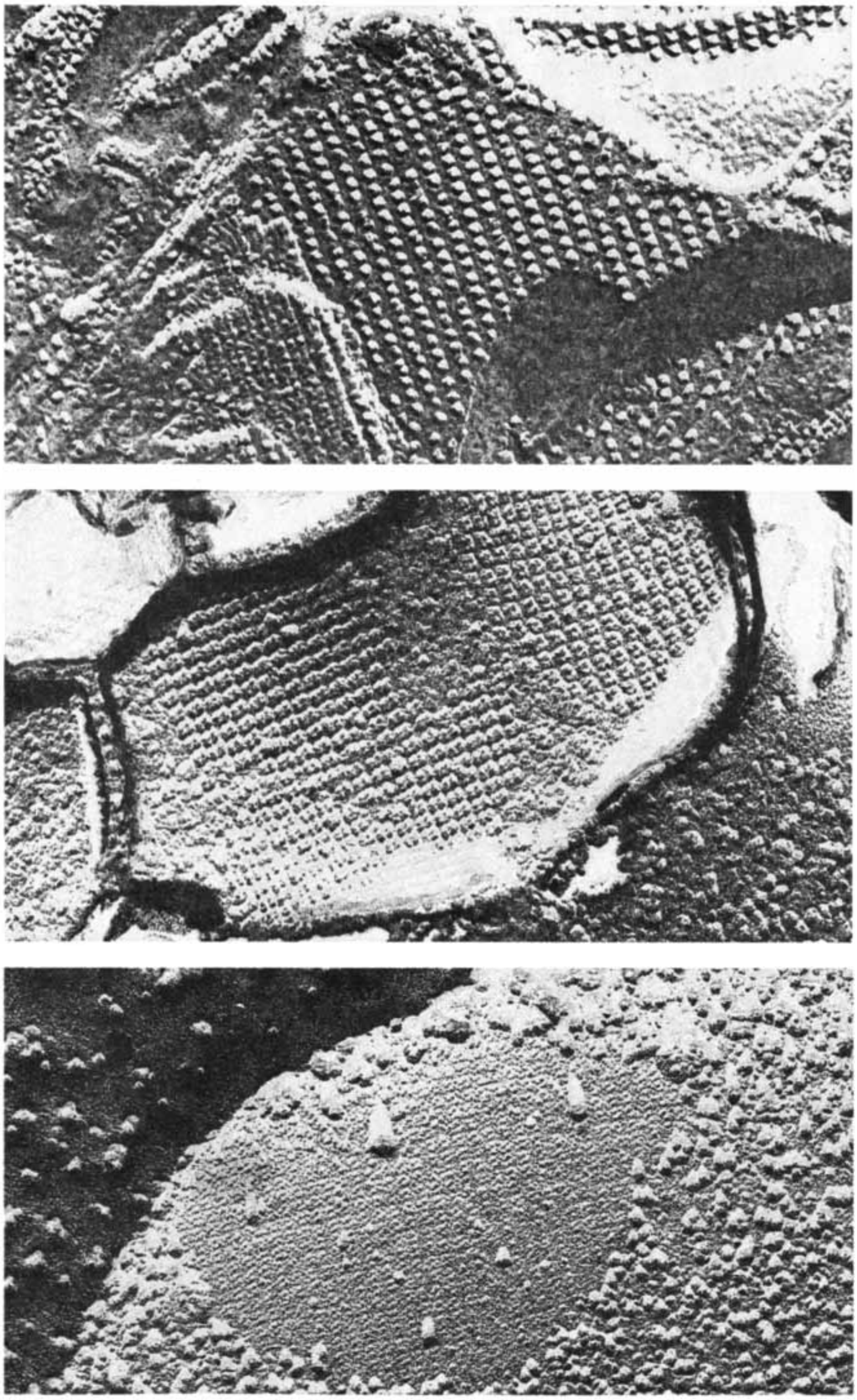
Index to Book Reviews—a listing of the longer book reviews in three parts: authors of books, titles and reviewers.

Index to Mathematical Games—a listing of the puzzles, games and diversions presented in this department since its inception in January 1957, under the editorship of Martin Gardner.

Index to The Amateur Scientist—a listing of all the projects, experiments and demonstrations presented in this department from April 1952 through February 1976 under the editorship of the late C. L. Stong and from July 1977 through June 1978 under the editorship of Jearl Walker.

molecules associated with this particle. Evidently some interaction between inorganic salts and CP II is responsible for the stacking process; it may also help to arrange the particles into regular lattices. For example, when purified molecules of CP II are incorporated into artificial membrane systems, the molecules arrange themselves into regular lattices.

What is the location of photosystem I in the thylakoid membrane? Does it also appear as a distinct particle? And if it does, how does it interact with photosystem II? Preliminary studies in our laboratory of mutants deficient in photosystem I indicate that it is associated with a subset of the smaller particles found in the photosynthetic membrane after freeze-fracturing. Other studies support this idea. For example, in hybrid systems consisting of a synthetic lipid bilayer to which the purified photosystem I reaction center has been added, particles with a diameter matching that of a class of particles found in the natural membrane are observed.



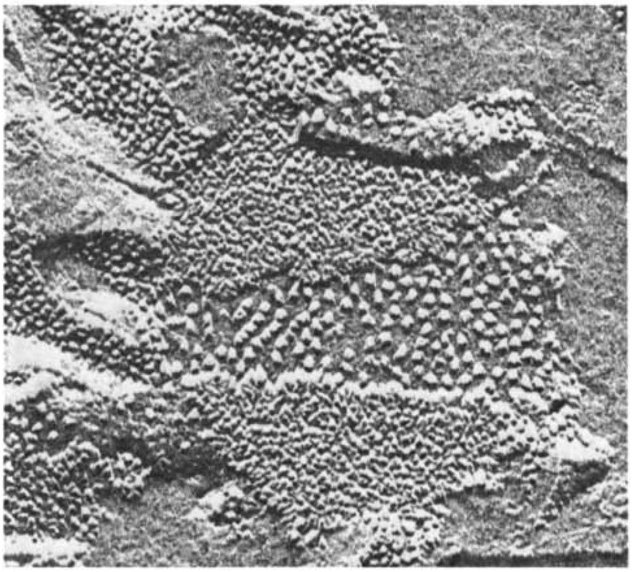
PARTICLES FORM REGULAR LATTICES under certain conditions. These three micrographs, all at a magnification of 135,000 diameters, depict such lattices inside a fractured membrane (*top*), on the inner surface of the intact membrane (*middle*) and on the outer surface of the intact membrane (*bottom*). The lattice has the same dimensions in all three pictures. It is least evident in bottom micrograph but can be seen if image is viewed at a shallow angle. Lattices are apparently formed by membrane-spanning particle associated with photosystem II.

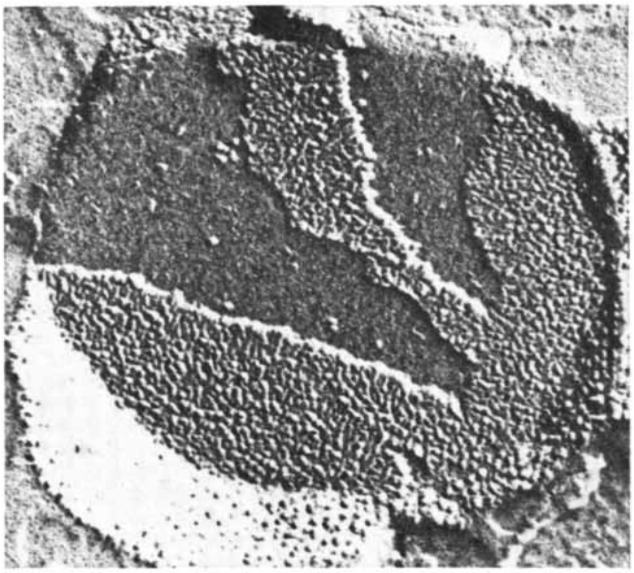
It is clear that a number of distinct substructures within the thylakoid membrane have important roles to play in the light reaction of photosynthesis. One of the disappointing aspects of this story is that little can yet be said about the details of molecular organization within the complex particles revealed by freeze-fracturing. One would like to know, for example, the detailed arrangement of the components of the electron-transport chain, the arrangement of chlorophyll molecules around the two reaction centers and the arrangement of the special proteins that seem to be important to various activities of the membrane.

At present detailed molecular structures are known only for two systems that are related to the photosynthetic membrane. One of these systems is the proton-pumping protein in the purple membrane of the bacterium *Halobacterium halobium*. The bacteriorhodopsin molecules in this membrane are arranged in a regularly repeating hexagonal lattice. That feature enabled Richard Henderson and P. N. T. Unwin of the Medical Research Council Laboratory of Molecular Biology at Cambridge in England to combine electron microscopy, electron diffraction and X-ray diffraction to map the structure of the protein to a resolution of .7 nanometer, which is sufficient to reveal such molecular details as the foldings of amino acid chains.

The second molecular structure now known in detail is that of the water-soluble bacteriochlorophyll-protein complex found in the photosynthetic bacterium *Chlorobium limicola*. Roger E. Fenna and Brian W. Matthews of the University of Oregon have studied crystals of this pigment-protein complex by X-ray diffraction. Their detailed picture of the bacteriochlorophyll-protein complex provides an exciting glimpse of the intricate arrangement of chlorophyll molecules and amino acid chains needed to capture solar energy at the submicroscopic level.

There is reason to believe, however, that neither of these studies will be able to tell us much about the photosynthetic membrane, because in each case the work was done on a system quite distinct from a true photosynthetic one. The purple membrane of *Halobacterium* participates in the light-activated pumping of protons, but it does so without electron transport and without the release of molecular oxygen. Further, although the structure of the chlorophyll protein from *Chlorobium* is most interesting, the protein is a water-soluble one





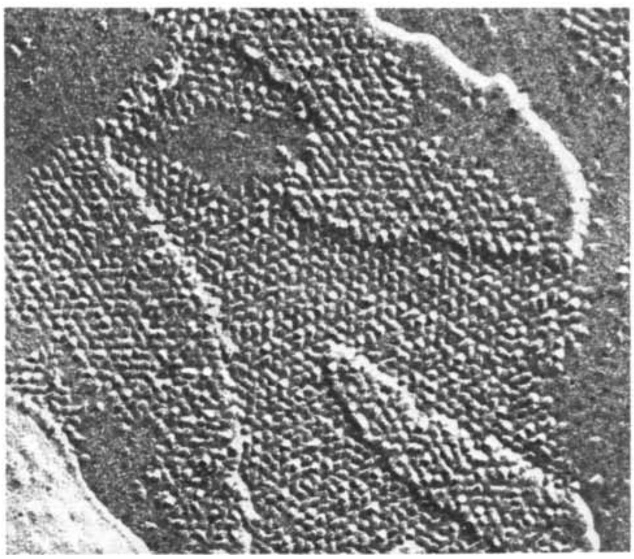
MEMBRANE-SPANNING PARTICLES are the large particles at the left in the freeze-fracture replica of a photosynthetic membrane from a normal tobacco plant. In a mutant deficient in photosystem II a comparable replica of photosynthetic membrane shows no large particles (*right*). All the other structures of the mutant membrane seem to be unaltered. The replicas are magnified 120,000 diameters.

and not a true membrane protein such as the proteins associated with the thylakoid membrane proper. In all probability the structures of the chlorophyll-binding proteins of the photosynthetic membrane will be quite different.

My colleagues and I hope to learn more about the photosynthetic membrane by applying recently developed techniques in which diffraction images of ordered molecular systems are subjected to Fourier analysis. The techniques seem capable of yielding high-resolution images of how the membrane systems are put together. In this article I have described two examples of such ordered systems that should be amenable to the new analytical approach: the ordered lattices that occasionally appear in the thylakoids of higher plants and the regular lattices formed by purified CP II molecules in synthetic lipid membranes.

We have also begun to investigate regular structures in the photosynthetic membrane of the bacterium *Rhodopseudomonas viridis.* With these analytical methods we have obtained images of this membrane with a resolution of about three nanometers. It is the hope of an impatient investigator that not too many years will pass before these approaches and others will reveal the structure of the photosynthetic membrane at the molecular level. For now one is left to marvel at the intricate workmanship that has shaped this membrane and equipped it to operate with such exquisite efficiency. As remarkable and as varied as other living systems have become, it is always worth remembering that the photosynthetic membrane is the engine that has made such extraordinary diversity possible.

ARTIFICIAL MEMBRANE was prepared in the author's laboratory by adding photosystem I complexes, which were purified by Michael Newman, to bilayers formed from a commercially available lipid. When the membrane was frozen and fractured, particles identical in size with certain particles that are present in natural membranes were visible on the fracture face. Magnification is 68,000 diameters.



CHLOROPHYLL-PROTEIN COMPLEXES (CP II) that are associated with photosystem II form regular lattices in this micrograph of a freeze-fracture replica of an artificial membrane prepared by Annette McDonnel in the laboratory of L. Andrew Staehelin at the University of Colorado. The replica is magnified 125,000 diameters. The chlorophyll-protein complex also seems to promote grana formation.

The Lava Lakes of Kilauea

The eruptions of the Hawaiian volcano leave pools of molten basalt that can take as long as 25 years to solidify. They provide a natural laboratory for studying the nature of magma from the earth's mantle

by Dallas L. Peck, Thomas L. Wright and Robert W. Decker

Magma—molten rock—from the interior of the earth is responsible for a host of phenomena at the earth's surface. The flow of magma out of the mid-ocean rifts adds to and pushes apart the rigid plates that make up the earth's surface and carry the continents on their backs. All igneous rocks are by definition formed by the congealing of magma. If the magma is erupted at the surface as lava, it forms extrusive igneous rocks such as basalt; if it slowly crystallizes below the surface, it forms intrusive igneous rocks such as granite. In spite of the importance of magma, however, there is much that is not known about it. Most studies of the cooling, crystallization and other properties of magma have centered on the laboratory analysis of small samples and on theoretical extrapolations from already solidified lava. A group of us at the U.S. Geological Survey, Sandia Laboratories and several universities have taken a different approach. We have studied molten and solidifying lava in situ by examining three "lakes" of lava left in the wake of eruptions of the volcano Kilauea on the island of Hawaii.

The three lakes are filled with basaltic lava. Basalt is the commonest rock formed by the solidification of magma extruded to the surface of the earth, the moon and perhaps other bodies in the solar system. Rich in calcium, magnesium and iron, basaltic lava crystallizes to form dark rocks consisting mostly of the silicon oxides plagioclase feldspar, pyroxene and olivine (whose gem form is peridot), the iron oxide magnetite and the iron–titanium oxide ilmenite. Basalt is found on all the continents and covers huge expanses of land such as the Columbia River and Snake River plateaus in the northwestern U.S. and the Deccan plateau in western India. Basaltic lavas, erupting from the mid-ocean rifts to create the floor that underlies the sediments of the ocean basins, have poured forth throughout geologic time from the early Precambian to the present. Although basalts vary significantly in chemical and mineral composition, they have all formed at high temperatures. In studying the lava lakes we undertook, among other things, to examine the feasibility of obtaining geothermal energy directly from deep bodies of basaltic magma. In principle the high temperature of the molten rock makes it attractive as a source of energy, although in practice numerous obstacles stand in the way of tapping its heat.

Most of the more than 500 active volcanoes on the earth are entirely or predominantly basaltic, including the active volcanoes that make up the southern two-thirds of the island of Hawaii: Mauna Loa, Hualalai and Kilauea. Mauna Loa is the largest of the three, rising four kilometers above sea level and almost nine kilometers above the adjacent ocean floor. Kilauea is the smallest, rising only 1.3 kilometers above sea level. Shaped like an inverted saucer, Kilauea is indented at the summit by a caldera, or large crater, from which radiate two rift zones. The eruptions of Kilauea are usually limited to the caldera and the rift zones, particularly the eastern rift zone and Halemaumau, the "fire pit" in the caldera. Since 1952 Kilauea has erupted on the average at least once a year.

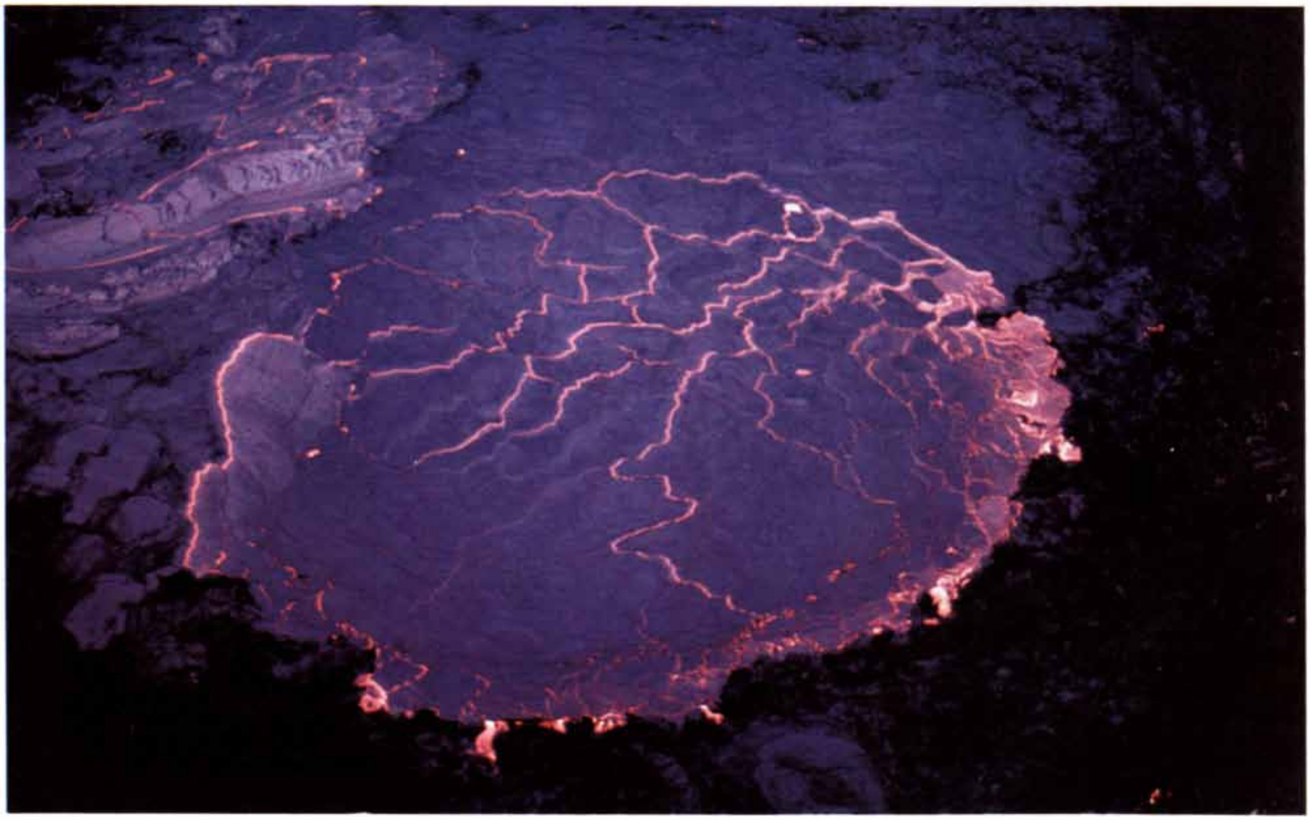
The basaltic magma that feeds the eruptions comes from the earth's mantle at depths of at least 50 kilometers below the surface. Geological and geophysical data suggest that magma rising from these depths is stored in an irregularly shaped reservoir about four kilometers below the summit of Kilauea. In the formation of a lava lake lava from the reservoir erupts to the surface and flows into a depression. Eventually the natural dikes that channel the lava into the lake collapse, and so the lake is cut off from a source of lava and starts to solidify.

The lava lakes on Kilauea that we investigated were one formed in 1959 in Kilauea Iki Crater, a second formed in 1963 in Alae Crater and a third formed in 1965 in Makaopuhi Crater. The depth of the lakes is substantial, so that it takes a long time for them to cool and solidify. The times range from 10½ months for the shallowest lake (in Alae Crater) to about 25 years for the deepest (in Kilauea Iki Crater). The lakes differ not only in depth but also in how they were filled, which affects their temperature, homogeneity and structure.

Kilauea Iki lava lake, which reaches depths of 120 meters, was formed in a spectacular event characterized by 17 separate eruptions over 36 days. Between eruptions volcanic activity stopped and some lava flowed out of the lake through a rift halfway up the side of the depression. Fragments of crust that had formed earlier then sank, making the upper levels of the lake markedly inhomogeneous. The eruption temperatures of between 1,200 and 1,215 degrees Celsius were high, and the abundant magnesium oxide favored the formation of olivine.

Alae lava lake, which is 15 meters deep, is poor in olivine. Its lava formed in the course of a three-day eruption at temperatures of between 1,140 and 1,160 degrees C. Makaopuhi lava lake is intermediate between the other two lakes in depth, temperature of formation and olivine content. Reaching a maximum depth of 83 meters, the lake formed in a 10-day eruption at temperatures of between 1,150 and 1,200 degrees. Our investigation of Alae and Makaopuhi lakes was abruptly halted in 1969 by new eruptions that inundated the lakes with fresh lava. Kilauea Iki lake has remained as it was, and it is still being investigated.

As the lakes cooled and solidified after they had been cut off from their source of lava we examined the process by drilling more than 50 holes in them between 1960 and 1977 and taking core samples of both solid and molten material. At first we worked with a small portable rotary drill powered by a chainsaw motor. Later we switched to a large trailer-mounted drill rig. The drilling re-

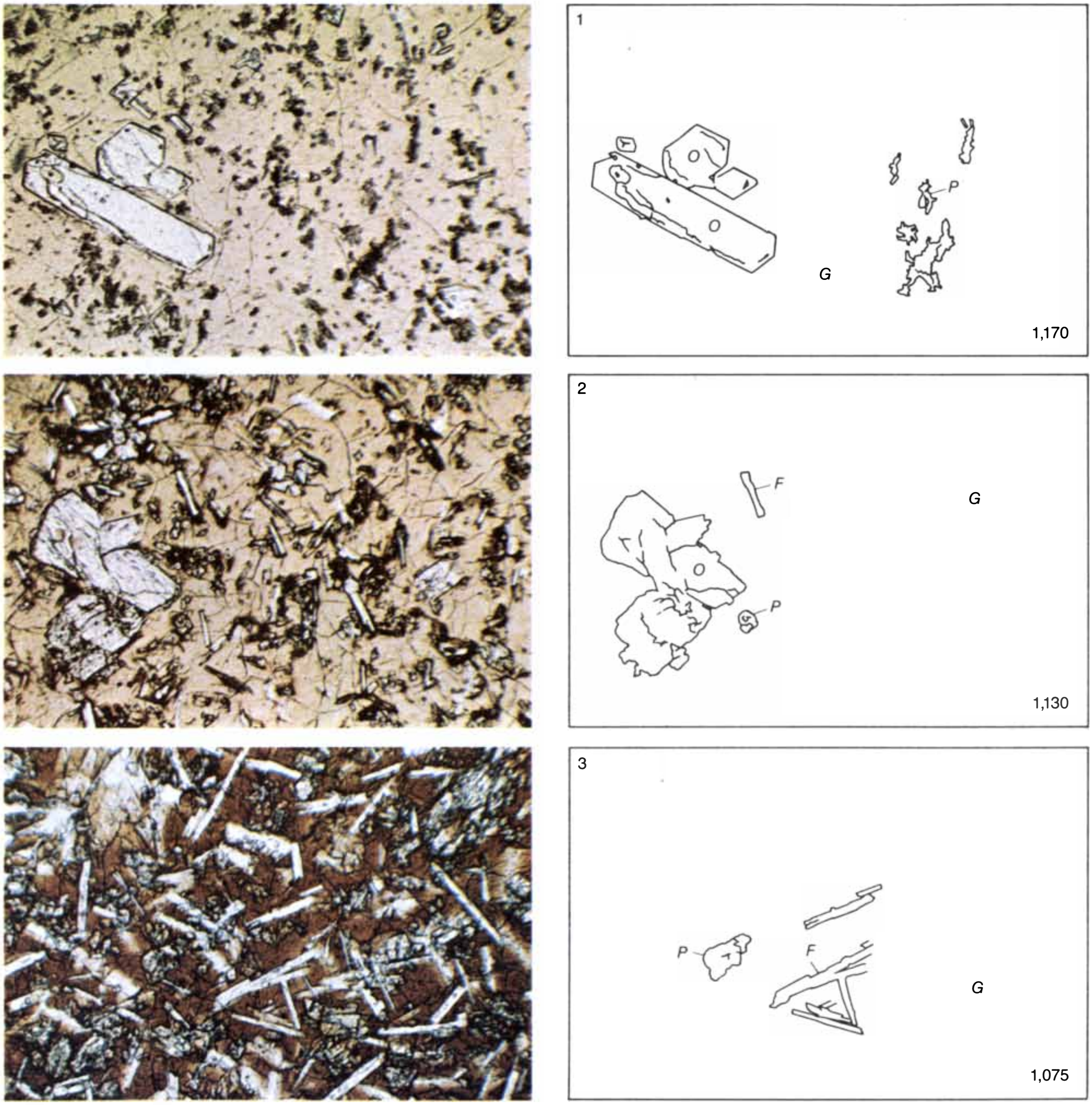


FOUNDERING OF SOLIDIFIED LAVA at the surface of Makaopuhi lava lake on Kilauea is shown in these photographs made in 1965. Gases exsolved from the underlying molten lava rose toward the surface and were trapped under the solidified but hot crust of the lake. In the top photograph the trapped gas gave rise to a gravitationally unstable situation that resulted in the foundering of a small piece of crust near the center of the lake's surface. The sunken lava was replaced by glowing molten lava (*center*). In the bottom photograph, made 10 minutes later, the foundering has extended to all the margins of the active part of the lake except the area at the upper right. The lava above 1,200 degrees Celsius is yellow. Between 900 and 1,100 degrees the lava is red, and below 900 degrees it is dark red or black.

vealed an interesting structure in depth. Like a freezing lake of water, a lava lake solidifies from the top down. The surface of a lava lake is cooled by air and particularly by rain, which falls copiously on Hawaii. (Over a four-year period 10 meters of rain fell on Alae lava lake.) Unlike a water lake, however, a lava lake also solidifies from the bottom up. That happens because the rock under a lava lake is cooler than the molten lava. As a result the molten lava is sandwiched between two layers of solidified crust and takes the shape of a lens.

The crust at the bottom of the lake is always thinner than the crust at the top because the rock under it has a low thermal conductivity and is not rapidly cooled by rainwater. The molten lens decreases in thickness as the lake solidifies from the top down and from the bottom up. A lava lake, like a lake of water, is shallower at the edges than it is at the center, so that the top and bottom crusts fuse at the edges of the lake, separating the molten lens from the rock enclosing the lake basin. As the top and bottom crusts continue to thicken the lens decreases in diameter and thickness until it disappears and the lake becomes a single body of solidified lava.

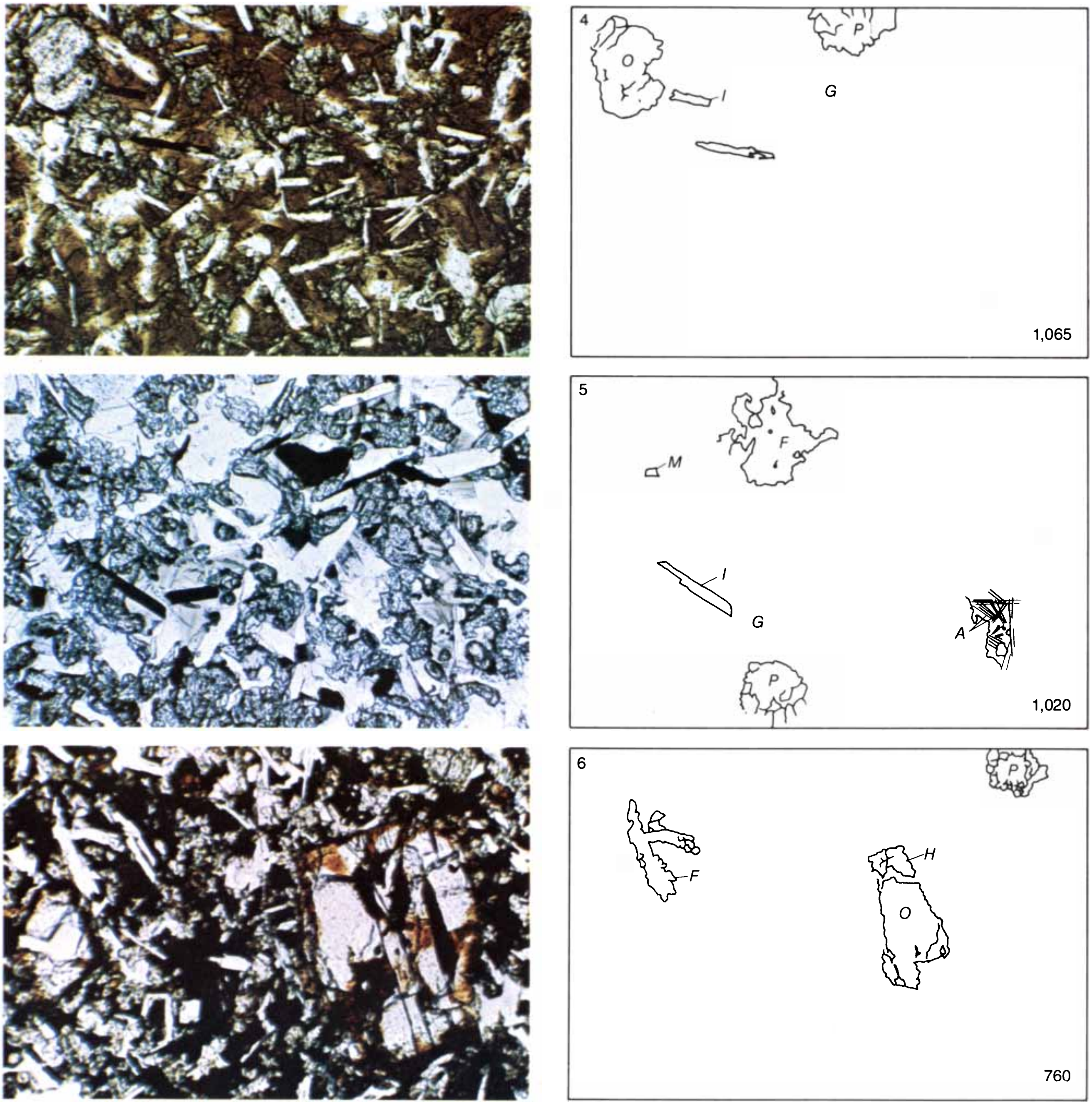
The ice on a lake of water is quite distinct from the water below it. That is



PHOTOMICROGRAPHS OF BASALTIC LAVA quenched from different temperatures in Makaopuhi lava lake show how the lava crystallizes. At the right of each photograph is a map identifying the minerals in the lava and a label indicating the temperature of the lava in degrees C. The minerals include the silicon oxides plagioclase feldspar (*F*), pyroxene (*P*) and olivine (*O*), the iron oxide magnetite (*M*), the iron–titanium oxide ilmenite (*I*) and the calcium fluoride–phosphate apatite (*A*). The maps also mark the location of lava in the glassy state (*G*): a supercooled liquid in which the silicon oxide molecules have not been organized into crystals. In photomicrographs *1* and *2*, which represent fluid lava at an intermediate depth in the lake, the crystals are sparse, dominated by coarse crystals of olivine (*O*). As the temperature of the lake decreases with decreasing depth, fine crystals of pyroxene (*P*) and pale, lathlike crystals of plagioclase feld-

not the case with a lake of basaltic lava, which consists not of one chemical compound but of different minerals that crystallize at different temperatures and rates. The solidified top crust grades slowly downward into the fluid lava through a region of partially molten crust several tens of centimeters thick. Within this region the temperature and the ratio of crystal grains to melt increase smoothly with depth. There is an interface, however, across which the physical properties of the partially molten lava change sharply. Above the interface the lava is solid enough to be drilled. Below the interface the lava is a fluid that yields like taffy when a drill probe is pushed into it. At the interface, whose temperature is 1,070 degrees C., crystal grains and melt are equally abundant. At 980 degrees the lava is entirely crystallized except for a small fraction in the glassy state: a supercooled liquid in which the silicon oxide molecules have not been organized into crystals.

The rate of thickening of both the top and the bottom crust decreases with time because the solidified material is a poor conductor of heat and so acts as an insulator. The top crust of all three lakes reached a thickness of .3 meter after a day and two meters after a month. After



spar (*F*) form in increasing abundance. In photomicrographs *3* and *4*, made from the solidified lava near the base of the crust of lava at the top of the lake, the crystals of pyroxene (*P*) and plagioclase feldspar (*F*) are larger and commoner, constituting about 50 percent of the lava. In photomicrograph *4* the first crystals of ilmenite (*I*) can be seen as opaque blades. In photomicrograph *5* the lava is mostly crystal, consisting of an interlocking network of crystals of pyroxene (*P*), plagioclase feldspar (*F*) and ilmenite (*I*) surrounding small quantities of a pale brown glass (*G*) containing needles of apatite (*A*) and a few tiny cubes of magnetite (*M*). In photomicrograph *6* the lava is entirely solid but not completely crystalline; it still consists of a few small pools of silicon-rich glass. The lavas of Alae and Kilauea Iki lakes are similar in composition to the lava of Makaopuhi lake, although there are nonetheless significant chemical differences among them.

PHOTOVOLTAIC RESEARCH MAY DRASTICALLY REDUCE THE HIGH COST OF SOLAR ENERGY CELLS.

IN 1960, PHOTOVOLTAIC ENERGY COST $2,000 PER PEAK WATT. TODAY THE SAME WATT COSTS $10. HONEYWELL RESEARCHERS THINK THEY CAN WHITTLE THAT COST DOWN TO JUST 50¢.

Converting sunlight directly into electrical energy is easy. All you need is a solar cell and some sunshine. The sunshine is free. The cost of the solar cell is the problem.

The Department of Energy wants a solution to that problem. They have established the goal of reducing the cost of solar cells to the point where they will be a practical alternative to fossil fuels in 1986. That means, 50¢ per watt.

Honeywell is working with the Department of Energy, under contract with Jet Propulsion Lab, on that problem. We think we're on the right track. We've developed a dip-coating process with the potential of providing solar-cell quality sheet silicon that costs less than the standard wafering technique.

In our new process, we first coat a ceramic substrate with a film of carbon. This enables the molten silicon to adhere. The resulting silicon layer exhibits grain structure much larger than the thickness of the layer, an important consideration for good solar-cell performance. In fact, we think this coating technology is so important we've patented the process.

But we've gone even further. We have a second patent on an idea for slotting the substrate in a way that makes it possible to contact 50% of the back surface of the silicon with a metallic electrode. That means we can approach the low internal resistance of conventional cells.

Dip-coating was a breakthrough, but it was only a first step. So we're developing a continuous coating process in which a long strip of substrate is coated by passing it over a molten silicon meniscus. Our objective is to increase the speed of layer production while keeping the silicon waste to an absolute minimum.

Recently, we reached a milestone with successful tests which proved the technical feasibility of the process.

The possibilities and opportunities this research is opening up are exciting. It could be the beginning of a whole new practical energy source.

If you are interested in learning more about Honeywell's research work in photovoltaics, you are invited to correspond with Paul W. Chapman, Honeywell Corporate Technology Center, 10701 Lyndale Avenue So., Minneapolis, MN 55420. If you have an advanced degree and are interested in a career in solid state electronics, sensors, or material sciences, please write to Dr. T. F. Hueter, Vice President Corporate Technology, Honeywell, Honeywell Plaza, Minneapolis, MN 55408.

Honeywell

Although Honeywell engineering is world-wide, the bulk of research and applied research is done in Minneapolis. The most recent *Quality of Life* Study conducted by Midwest Research Institute shows Minnesota to be one of the best places in the country to live and work considering cultural, social, economic, educational and political factors.

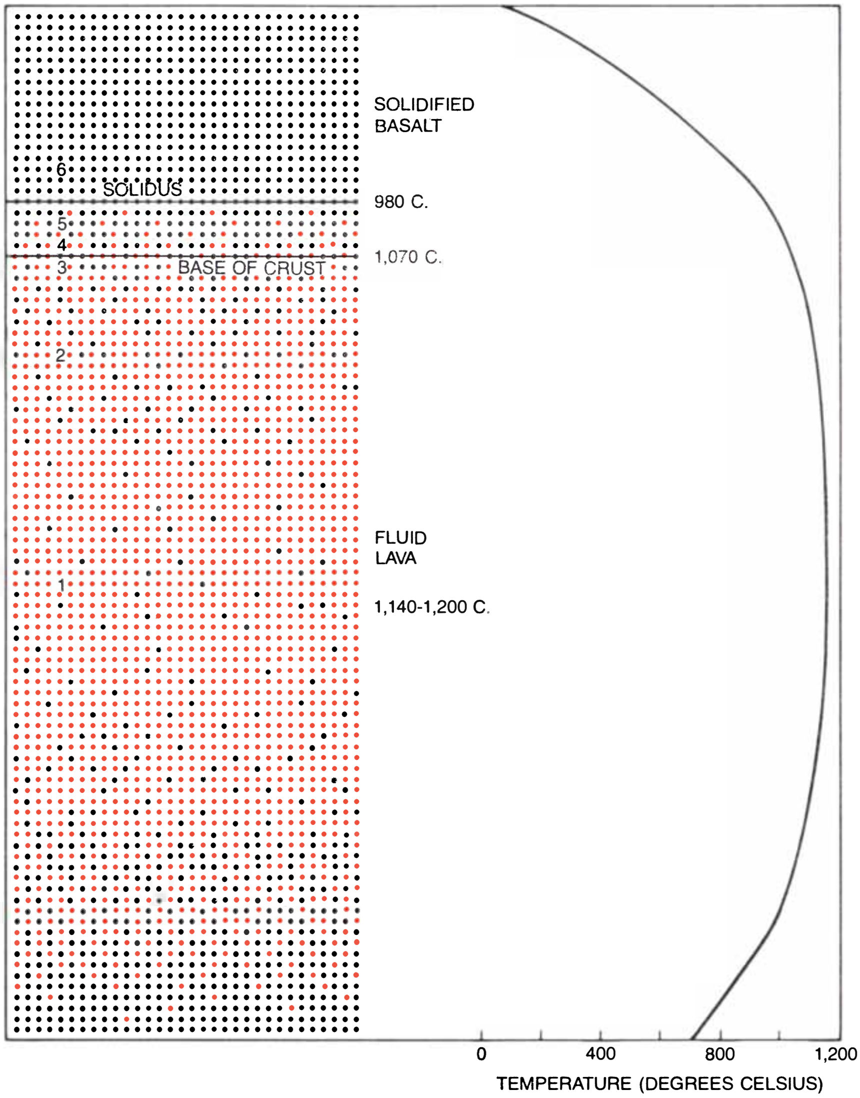
The Twin Cities area is the home of many of America's advanced technology companies employing a great many scientists and engineers. This ideal research environment is further enhanced by Honeywell's affiliations with universities across the country. We have an ongoing program of seminars with Berkeley, MIT, Stanford, Carnegie Mellon, the University of Illinois, Cornell, Waterloo of Ontario and the University of Minnesota.

a year the top crust of the two deepest lakes was eight meters thick. In 1976, 17 years after Kilauea Iki lava lake was formed, the top crust was 45 meters thick. From our core-sample data we found the rate of increase in thickness to be proportional to the square root of time until the maximum temperature in the lake falls appreciably below the initial temperature at the center of the lake. The thickness in meters is approximately equal to .4 multiplied by the square root of the time in days.

Since with the passage of time the top crust gets thicker at a lower rate, its temperature must also decrease at a lower rate. Measurements made with steel-sheathed thermocouples inserted into the drill holes indeed show that the rate of decrease in the temperature at any given depth in the top crust is lowered with time. With the aid of a computer we developed a numerical model for predicting the temperature at any depth in a lava lake. The model is based on Fourier's law of heat conduction: The flow of heat between two regions is proportional to the difference in temperature between those regions. The model incorporates the external processes that affect the flow of heat in a lava lake (cooling by rainwater), the latent heat of basalt (the heat released by crystallization) and variations in the density, porosity and thermal conductivity of the lava.

We applied this numerical model to the top crust of Alae lava lake, where we had regularly measured the temperature over a four-year period during which the lake had solidified and cooled to less than 100 degrees C. The average difference between the calculated temperature and the measured temperature was less than two degrees. Such close agreement tends to confirm the validity of the model, the main import of which is that the vaporization of rainwater from the hot crust of a lake has a large effect on the lake's decrease in temperature.

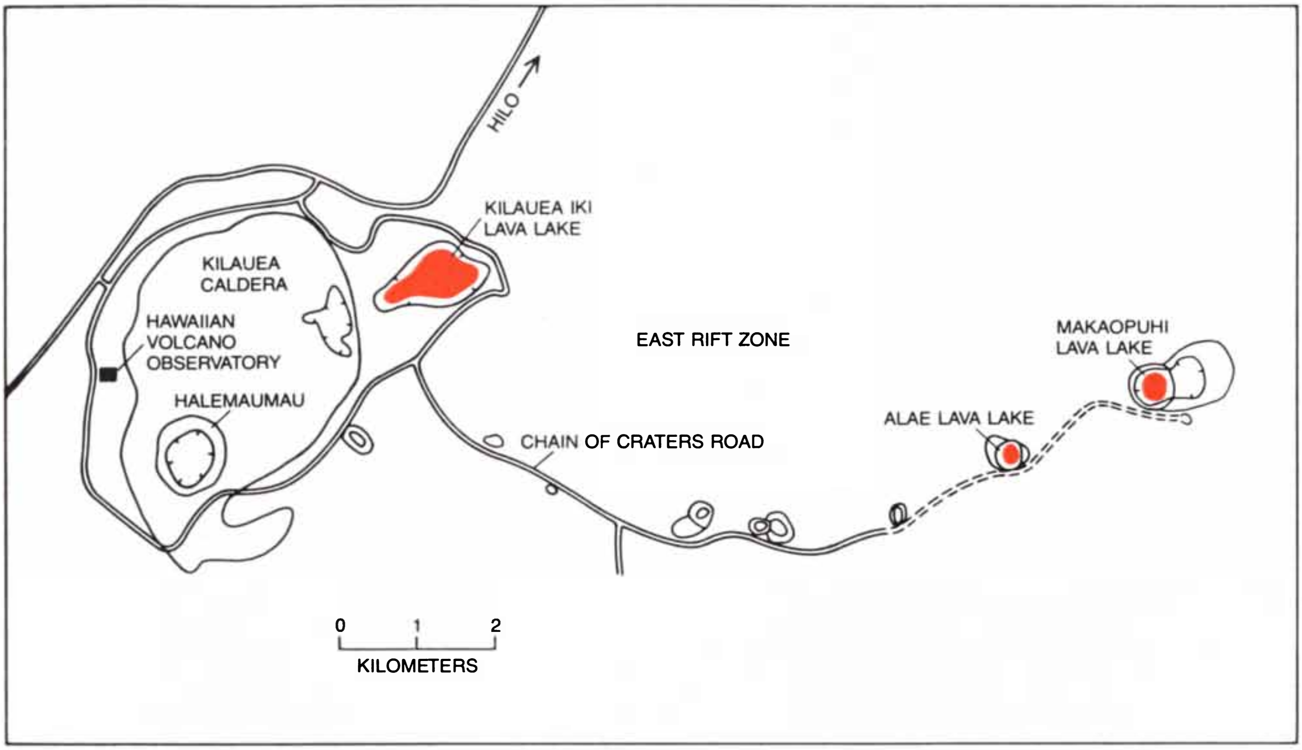
As molten lava cools, gas in the melt is driven out of solution. The gas either escapes into the atmosphere or remains in the lava as vesicles, or bubbles. The vesicles that were formed at the time of the eruption of the three Kilauea lava lakes and were frozen into the crust at shallow depths are chiefly spheres as much as a centimeter in diameter. They were apparently created at high temperatures when the lava became supersaturated with gas on the reduction of the confining pressure above it, just as bubbles appear in a bottle of soda water when the cap is removed. With increasing depth in the lava the confining pressure increases, and so the vesicles become smaller and scarcer. Below six meters in Alae lava lake most of the vesicles are minute angular pores less than a millimeter in diameter. They were apparently created when gas was driven out of solution by the crystallization of the cooling lava. The composition of the gases expelled from the lava also changed as the lake cooled: water vapor increased in abundance at the expense of the more rapidly exsolved gases of carbon and sulfur compounds.



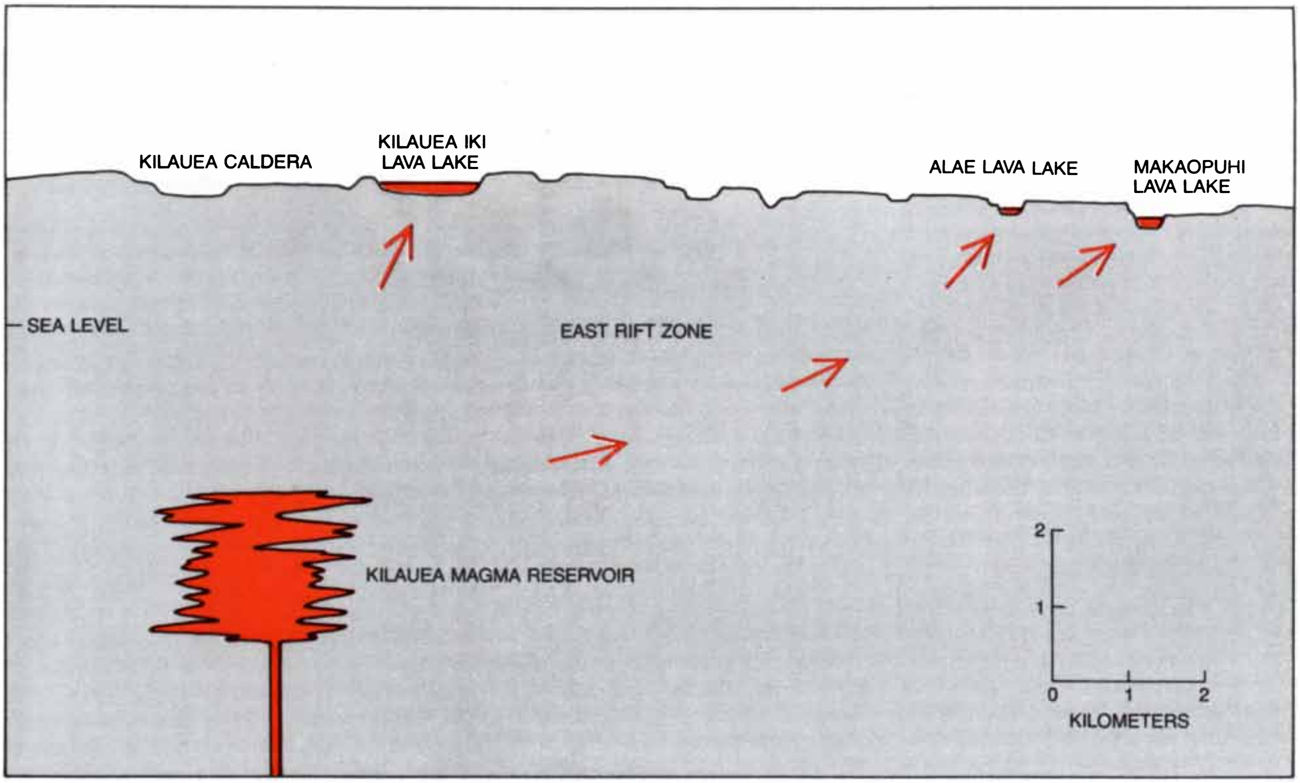
RATIO OF CRYSTAL TO MELT in a lava lake is shown in this highly schematic cross section (*left*). A lava lake solidifies from the top down and the bottom up. At temperatures below 980 degrees C. the lake is entirely solid. Between 980 and 1,070 degrees the lava is partially molten, but it is rigid because the melt is trapped in the interstices of a crystalline network of solidified lava. At 1,070 degrees the melt (*color dots*) and the crystal grains (*black dots*) are equally abundant. Above 1,070 degrees the lava is a fluid that yields like taffy when a drill probe is pushed into it. The numbers correspond to the positions of the lava shown in the photomicrographs on pages 116 and 117. At the right is a graph of the temperature as a function of depth.

Although a solidifying lava might be expected to sink because its crystals are denser than the melt, the lava in Alae lake became more buoyant as it solidified because it was filled with gas vesicles. Therefore as the lens of molten lava below the surface at the center of the lake solidified it pushed up the surface above it. At the edges of the lake, where there was no solidifying lens because the top and bottom crusts had fused, the surface subsided as the cooling lava thermally contracted and became denser; since this lava had already solidified, no bubbles were being formed in it that would make it more buoyant. The entire surface of Kilauea Iki and Makaopuhi lakes, on the other hand, has generally subsided as the lakes have cooled. Since these lakes are deeper than Alae lake, the higher pressures within them have hindered the formation of vesicular lava that would have pushed up the surface.

We investigated the crystallization of basalt by taking core samples of the semimolten lava at the base of the top crust, of the molten lava below the top crust and of the pumice and molten



MAP OF THE EAST RIFT ZONE of Kilauea shows the location of the three lava lakes the authors investigated: one formed in 1959 in Kilauea Iki Crater, a second formed in 1963 in Alae Crater and a third formed in 1965 in Makaopuhi Crater. Halemaumau is the fire pit in the caldera at the summit of the volcano, which is shaped like an inverted saucer. Over the past 10 years fresh lava flows have covered Alae and Makaopuhi lava lakes and the part of the Chain of Craters road indicated by the broken lines. Kilauea Iki lake was not covered.



MAGMA RESERVOIR lying about four kilometers below the summit of Kilauea is the source of the lava that erupted to the surface and flowed into depressions to form the lava lakes. Kilauea Iki lava lake, which reaches depths of 120 meters, was formed in a spectacular event characterized by 17 separate eruptions over 36 days. Makaopuhi lava lake, which has a peak depth of 83 meters, was formed in an eruption that lasted for 18 days. Alae lava lake, which is 15 meters deep, was created in an eruption that lasted for three days.

lava erupted to the surface when the lakes first formed. The samples enabled us to reconstruct the history of the crystallization of a lava lake. At first the crystals in the fluid lava are sparse, dominated by coarse grains of olivine. As the temperature of the lake decreases, fine crystals of pyroxene and pale, lathlike crystals of plagioclase form in increasing abundance. Near the base of the crust, at 1,070 degrees C., the crystals of pyroxene and plagioclase are larger and commoner, constituting about 50 percent of the lava. At 1,075 degrees iron and iron–titanium oxides are still in the glassy state, but at 1,065 degrees a few opaque blades of the iron–titanium oxide ilmenite can be seen. The abundance of such oxides gives the glass a rich brown color. Lava that has cooled to 1,020 degrees is mostly crystalline: an interlocking network of crystals of pyroxene, plagioclase and ilmenite surrounding small quantities of a pale brown glass containing needles of apatite and a few tiny cubes of magnetite. At 760 degrees the lava is entirely solid but not totally crystalline; it retains a few small pools of a silicon-rich glass.

The lavas of the three lakes are similar in composition, but there are nonetheless significant chemical differences among them and significant chemical variations within them. Three processes are responsible for these differences and variations. The first is the sinking of magnesium-olivine crystals, which are denser than the basaltic melt. The crystals migrate downward like grains of sand dropped into a pond. The effect of the migration is particularly apparent in an old solidified lake exposed in the walls of Makaopuhi Crater, where olivine is sparse at the top of the lake but abundant between six and 24 meters above the bottom. Most of the variation in the chemical composition of lava with more than 6 percent magnesium oxide is the result of variations in the abundance of olivine.

The second process, a separation of crystals of plagioclase and pyroxene in flowing lava resulting from differences in their density and shape, gives rise to subtler shifts in chemical composition. We first became aware of this separation in Makaopuhi lava lake when molten lava that had flowed into the drill casing was found to contain fewer crystals, particularly crystals of pyroxene, than the untapped fluid lava. The third process, which comes into play at lower temperatures, is the separation of interstitial melt from the network of crystals in the rigid but incompletely crystallized lava near the base of the top crust. Such melt fills fractures in the crust and oozes into drill holes.

The composition of the solidified but still hot basalt of the top crust of the lakes is also altered by the action of gas-

Having Vicky's tonsils out in Phoenix got her home to the ice cream two days faster. And two days cheaper.

She had one-day surgery in a Surgicenter.

If you think the cost of feeding and housing a family is out of sight, consider the cost of health care. In recent years it's also climbed faster than the rate of inflation itself.

That's why an ambulatory surgical center is such a boon to Phoenix's families. It's impressively reducing surgical costs by eliminating hospital stays for routine procedures whenever possible. Vicky's tonsillectomy, for instance.

Look at these figures. Based on a cost study of eight operations, using the Surgicenter's facilities costs from 42% to 61% less than the hospital's. Even surgeon's fees are lower, from 18% to 38%.

Now that's proof-positive health care costs can be contained. All it takes is the interest and commitment of aware communities like Phoenix.

We're really hopeful their success will inspire others to follow suit. Because as a major part of the nation's private health insurance system, which provides general health care coverage for 179,000,000 Americans and protection against major medical expenses for nearly 140,000,000, one of our deep commitments is to keeping health care affordable to the greatest number of people. And if costs can be kept down, premiums can, too.

And we're doing our part to help. For instance, many of our companies are cutting costs by paying for pre-hospitalization testing and second opinions for surgery.

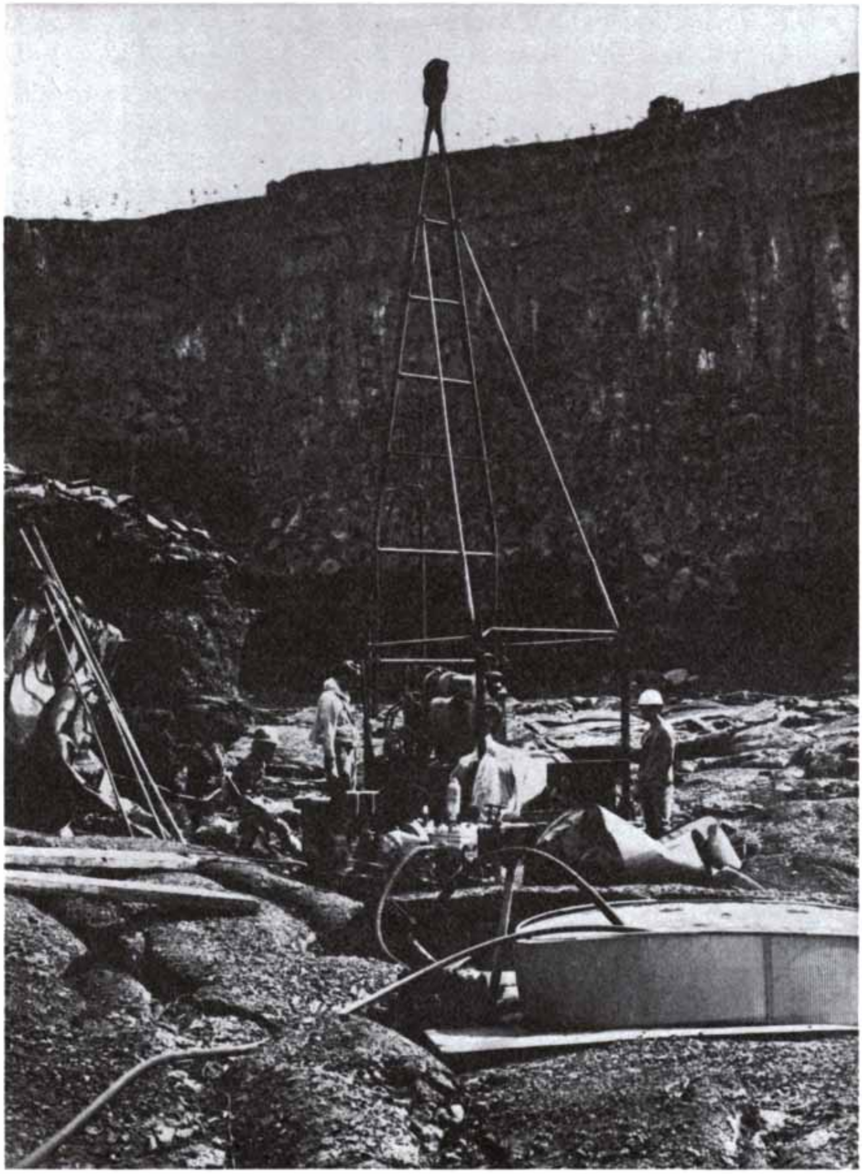
If you're interested in cutting health care costs in your corner of the world, we're interested in helping you out. We've put together a booklet on *Cutting Health Care Costs with Ambulatory Surgical Centers*. And it's free. Write us at Health Insurance Institute, Department 16, 1850 K Street, N.W., Washington, D.C. 20006.

THE HEALTH INSURANCE COMPANIES IN AMERICA

Let's Keep Health Care Healthy.

es formed in the underlying crystallizing lava that rise through fractures in the crust. The most striking alteration is the formation of red films of the iron oxide hematite in olivine and the replacement of silicate minerals by oxides at temperatures of between 500 and 800 degrees C. This alteration is due to variations in the concentration of free oxygen at different depths in the lakes. Where the oxygen is abundant the most oxidized form of iron, namely hematite, is highly stable. Where the oxygen is less abundant the less oxidized form of iron, namely magnetite, is stable. (The lack of free oxygen on the moon accounts for the absence there of hematite and the other principal oxide of iron, magnetite. On the moon the least oxidized form of iron, namely metallic iron, is stable.) At temperatures of between 550 and 750 degrees the concentration of oxygen was high and the accumulation of hematite was quite conspicuous. The high concentration of oxygen presumably resulted in part from the decomposition of water vapor in the gases rising through the fractures.

Another phenomenon that alters the cooling crust of the lava lakes is that large cracks develop in the crust as the cooling basalt contracts. Such cracks open a minute or so after incandescent lava appears at the surface in the course of an eruption. As the crust continues to



DRILL RIG powered by a chain-saw motor sits on the surface of Makaopuhi lava lake. The rig, which was lowered onto the surface by a helicopter, gathered samples of solid and molten lava from various depths in the lake. In front of the drill rig is a water tank, which supplied water to quench the samples of hot lava. The crater wall behind the drill rig exposes an older columnar-jointed lava lake that was subsequently covered by several thin horizontal lava flows.

cool and thicken, the cracks propagate downward by further fracturing and new cracks open that divide the surface of the crust into polygons. Within a few hours the polygons become deformed as their centers are elevated by the vesicular expansion of the solidifying lava under them. Then still more cracks open, which are later characterized by sulfur and gypsum deposited in them by rising gases. Repetitive mapping of the cracks on the surface of Alae lava lake indicates that the rate of cracking is greatest at times when the crust is being chilled by a heavy rain.

The Kilauean lava lakes provide a good opportunity to study the physical properties of basaltic magma. Herbert R. Shaw of the U.S. Geological Survey and two of us (Peck and Wright) measured the viscosity of fluid lava both in the laboratory and in Makaopuhi lava lake. Viscosity is defined as the degree to which a fluid resists flowing under an applied force. We placed a stainless-steel viscometer, a rotor that measures viscosity by resistance to its turning, in the drill hole at depths of between 2.6 and 3.2 meters within the melt. The lava there was at a temperature of between 1,130 and 1,135 degrees C. and contained 25 percent crystal and between 2 and 5 percent gas bubbles. Our results confirmed those obtained in the laboratory, where the viscosity rose from 100 poises at 1,300 degrees to about 320 poises at 1,200 degrees. (A poise is a dyne-second per square centimeter.) This change is comparable to the difference in viscosity between heavy oil and stiff honey at room temperature. At temperatures below 1,200 degrees the viscosity increases sharply because of the accumulation of crystals in the melt.

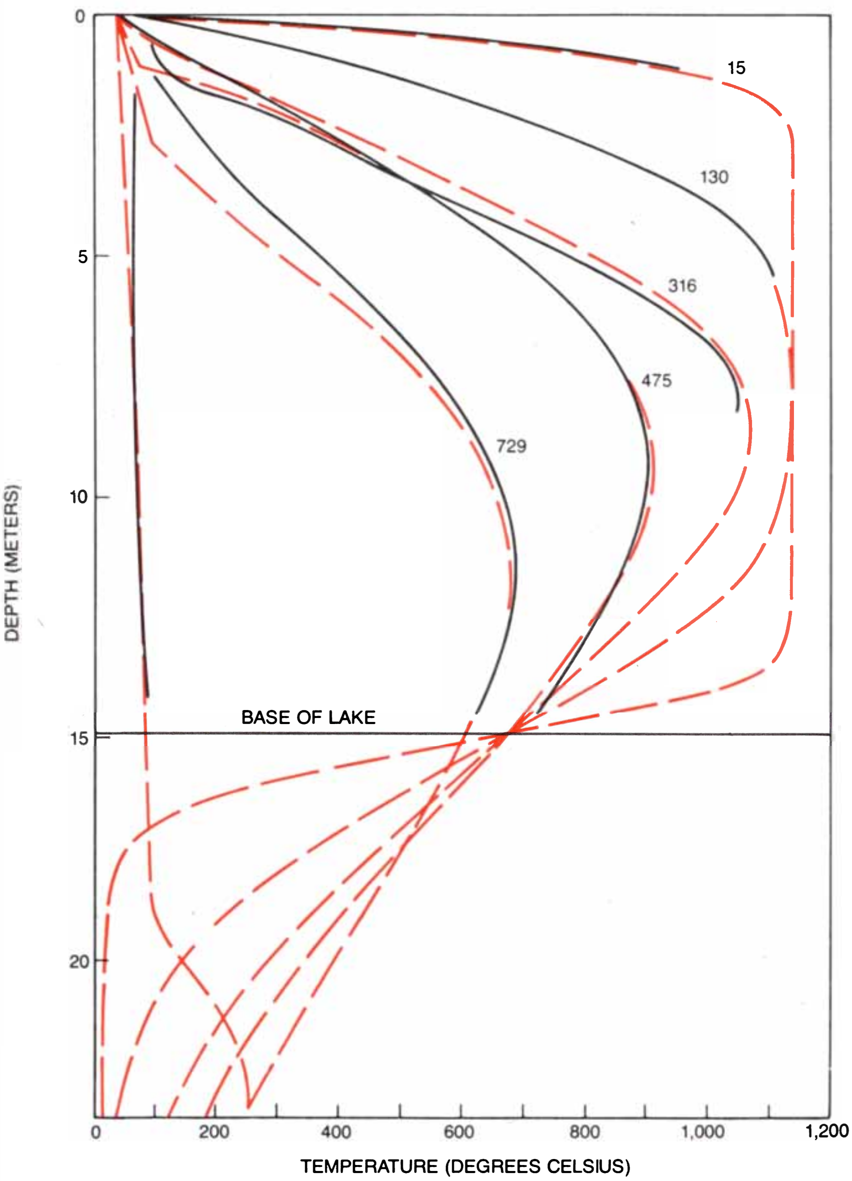
The presence of gas bubbles at these temperatures made the lava plastic rather than fluid. Fluid behavior is characterized by deformation under the smallest applied stress, plastic behavior by deformation only after the application of a certain minimum stress. At low rates of rotation of the viscometer the lava at 1,130 degrees had a viscosity of 7,500 poises. The gas bubbles "stiffen" the lava just as air stiffens egg whites when they are beaten. At higher rates of rotation the bubbles elongated to form planes of weakness that lowered the viscosity to 6,500 poises.

The investigations of the lava lakes in the 1960's were undertaken to discover more about the nature of magma. Over the past four years work on the lava lakes has intensified as a part of the search for alternative sources of energy. The magma underlying the U.S. could conceivably provide 13×10^{21} calories of heat per year, an amount equivalent to what would be released by the burning of 9,000 billion barrels of oil.

The obstacles to tapping this energy

are large. Methods have to be developed for finding a magma body and determining its size, shape, depth and heat content. Techniques are needed for drilling into a deeply buried magma body, putting a heat-transfer device into it and extracting the energy through the hole. The recovery of magma energy also presents a challenge to the materials sciences: materials must be found that can operate successfully for long periods in the hot, pressurized and corrosive environment of a magma.

Kilauea Iki lava lake provides a natural laboratory for work aimed at overcoming these obstacles. The lava in the lake was partially outgassed in the course of the eruption and is under relatively low pressure, so that the lake is not entirely representative of deeply buried magma. The magma at these shallow depths is much more accessible, however, and hence we have been able to apply what we had already learned about the physical and chemical properties of the melt in the lake to test poten-



TEMPERATURE IN ALAE LAVA LAKE is shown as a function of depth. The solid black curves are the result of measurements made with steel-sheathed thermocouples inserted into drill holes over a four-year period. Each curve is labeled according to the number of days that had elapsed since the lake formed. The broken colored curves come from a computer-generated numerical model of the thermal activity in a lava lake. The model incorporates the external processes that affect the flow of heat in a lava lake (cooling by the surrounding air and rock that are initially at a temperature of about zero degrees C. and cooling by rainwater), the latent heat of basalt (the heat released by crystallization) and variations in the density, porosity and thermal conductivity of the lava. The average difference between the calculated temperature and the measured temperature is less than two degrees. Such close agreement tends to confirm the validity of the model, the main import of which is that the thermal conductivity varies with the porosity and temperature of the lava and that the vaporization of rainwater from the hot solidified crust of a lava lake has a large effect on decreasing the temperature of the lake.

This is the computer...

that taught this boy...

TRS-80 Computer Consultant
Ralph Lipe of Fort Worth, Texas

to write this program!

TRS-80(TM)... from $499*(Level I, 4K)

Fourteen-year-old Ralph Lipe taught himself to program a TRS-80 Microcomputer System with Radio Shack's "plain language" instruction manual. Then he developed a program for completing the IRS 1040-A tax form that took top honors at his school and regional science fairs. *That was last year.*

This year Ralph is in high school and working part-time as a computer consultant. Everyone from grade schoolers to graduate students can use the world's best-selling microcomputer, the TRS-80. It doesn't do your thinking for you — it helps you learn to think. Think about it!

Meet TRS-80's Big Brother!

New TRS-80 Model II

Systems from $3450

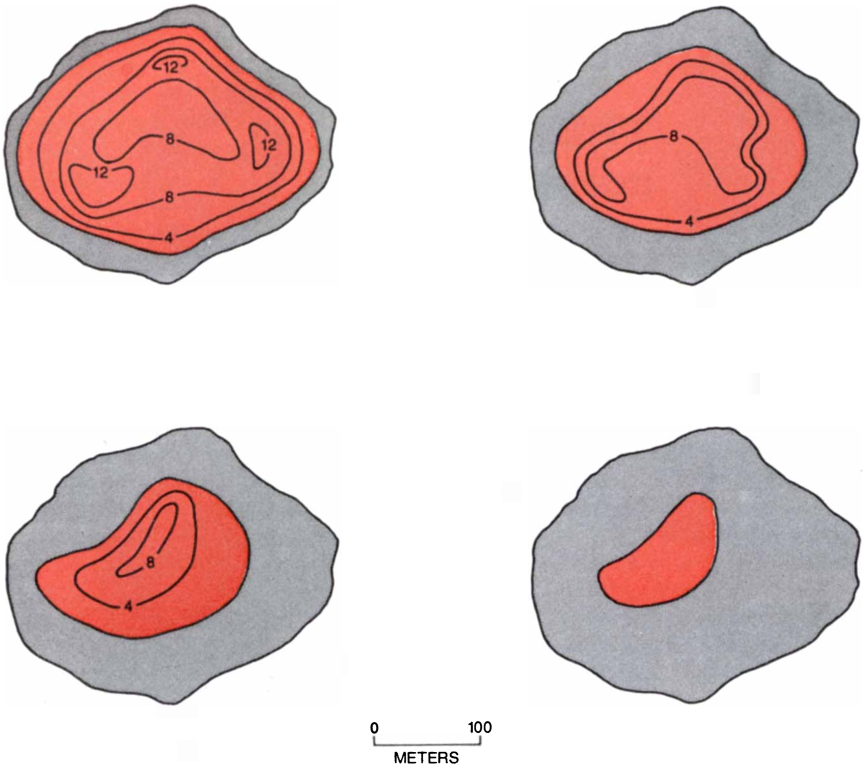
We've added a bigger, more powerful "brother" to the TRS-80 family. It's TRS-80 Model II, a completely new microcomputer for business applications. Its capabilities begin where Model I approaches upper limits. Storage capacity from ½ megabyte to 2 megabytes! For earliest delivery, order now from any of our 7000 locations.

Radio Shack®

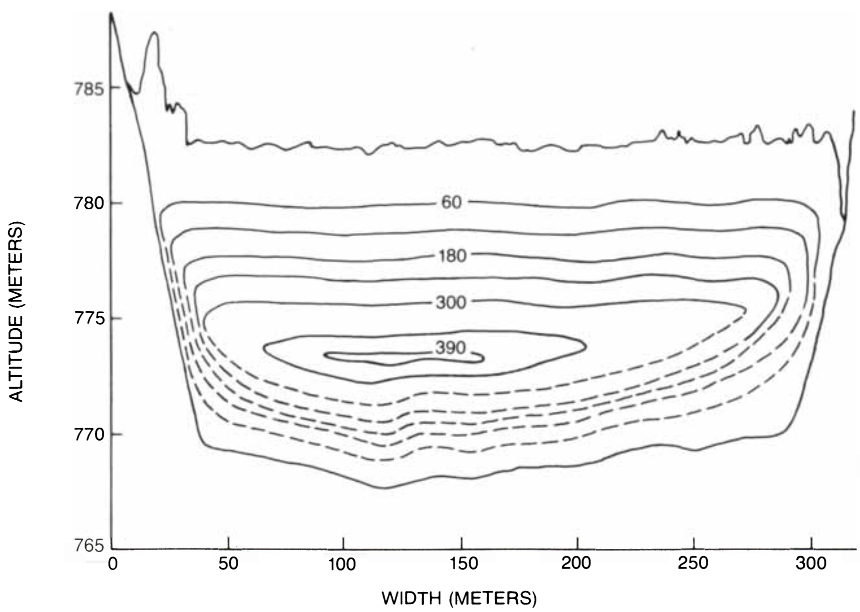
The biggest name in little computers(TM)

A Division of Tandy Corporation • Over 7000 Locations in 40 Countries

*Retail prices may vary at individual stores and dealers.



SURFACE OF ALAE LAVA LAKE is pushed up by gas bubbles that form in the underlying solidifying lava. A lava lake, like a lake of water, is shallower at the edges than it is at the center, so that the top and bottom crusts fuse at the edges (*black*), separating the molten lens from the rock enclosing the lake basin. The surface of the lake above the lens is colored. As the molten lens cools, gas in the melt is driven out of solution and pushes up the colored areas of the lake's surface. The contours indicate the rate of uplift of the surface in units of 10^{-2} centimeter per day. At the edges of the lake, where there was no solidifying lens because the top and bottom crusts were joined, the surface (*black*) subsided as the cooling lava became denser owing to thermal contraction; since this lava had already solidified, no bubbles were being generated that would make it buoyant. The contour diagram at the top left represents the lava lake from 34 to 104 days after it formed, the diagram at the top right from 236 to 293 days, the diagram at the bottom left from 293 to 345 days and the diagram at the bottom right from 345 to 400 days.



LONGITUDINAL CROSS SECTION OF ALAE LAVA LAKE shows how the isotherm, or surface of equal temperature, at 1,000 degrees C. changed as the lens of molten lava shrank. Each position of the isotherm is labeled by the number of days since the time the lake formed.

tial techniques for finding the position, size and shape of deeply buried magma bodies. Our investigation of the lake is entirely experimental, directed at filling specific gaps in knowledge and not at harnessing a particular Kilauean lava lake as a source of power.

In 1976 two holes were drilled through the crust of the lake to the top of the lens of molten lava. The properties of the lens were determined directly from core samples and from instruments lowered into the drill holes. We also tried to indirectly determine the same properties by remote techniques, such as seismic and electromagnetic measurements made at the surface of the lake; such techniques could in principle be applied to deeply buried magma inaccessible to direct drilling. An analysis of the results suggests that the seismic and electromagnetic techniques can determine the top and lateral margins of the lens of molten lava but cannot accurately estimate its thickness. It seems that these techniques perceived as melt not the entire lens of molten lava but only the hotter central part.

Working with seismometers, Bernard Chouet of the Massachusetts Institute of Technology and John L. Colp of Sandia Laboratories have mapped the approximate borders of the lens of the melt at Kilauea Iki lava lake. The seismic energy is provided by tiny earthquakes caused by the opening of thermal-contraction cracks in the regions with steep thermal gradients at the top and lateral margins of the lens; there are as many as 200 such shocks per hour. Electromagnetic studies of the lake done by Lennart A. Anderson and Charles J. Zablocki of the U.S. Geological Survey reveal approximately the same borders: a lens of molten lava 500 meters wide and 750 meters long.

The electromagnetic studies show variations in the electrical conductivity at different depths in the lake. The top layer of the top crust that is at temperatures below 100 degrees C. is quite conductive because the large fractures in it are filled with rainwater. Below this layer is a region of solidified basalt that at temperatures above 100 degrees is highly resistive because of the absence of liquid water and because the potentially conductive ions in the basalt are frozen in minerals. At depths below 40 meters lies the lens of molten lava, which is an excellent conductor because ions are free to move in it.

This year six new holes were drilled in Kilauea Iki lava lake. The lens of partially molten lava is now about 30 meters thick and consists of many crystals. It will solidify completely within five years at the most. Until Kilauea or some other volcano gives rise to a new lava lake, this unique opportunity to study magma in situ will soon be a thing of the past.


MARTELL
CORDON BLEU
Cognac

The Evolution of the Crocodilians

Some 200 million years after these reptiles arose and diversified three families remain: alligators, crocodiles and gavials. They are the last of the archosaurians, the ruling reptiles of the Mesozoic

by Eric Buffetaut

Over the past 200 million years one order of reptiles, the crocodilians, has repeatedly demonstrated a major feature of evolution: If a certain kind of ecological niche exists, a certain kind of organism is likely to fill it. For example, there are three families of living crocodilians: the alligators, the crocodiles and the gavials. Alligators have a blunt head, crocodiles a more elongated head and gavials an extremely narrow one. The same evolutionary adaptations of blunt and narrow heads arose repeatedly among crocodilians that are long extinct and are related only distantly, if at all, to those living today. Evidently the ecological niche favoring the evolution of long-snouted reptiles that feed on fish has been filled repeatedly by the adaptable crocodilians. As another example, the living crocodilians are in general adapted to a semiaquatic life in a warm environment. In the past, however, some members of the group adapted to niches on dry land that today are occupied by carnivorous mammals. At the same time others became entirely aquatic, with sharklike tails and limbs remodeled into swimming paddles.

Here I shall relate what is known about the evolution of this remarkable reptilian group since its appearance in the fossil record at the close of the Triassic, a period when reptiles were the leading forms of animal life on the earth, occupying the land, the sea and the air. When scientific paleontology came of age early in the 19th century, fossil crocodiles were among the first vertebrate animals to be unearthed, and even today new fossil crocodilians are discovered quite frequently. As a result, although gaps in the history of the crocodilians remain, it is clear that the Order Crocodylia diversified considerably over the past hundreds of millions of years, spreading to all parts of the world and adapting to a wide variety of ecological niches.

As early as 1875 Darwin's friend and defender T. H. Huxley recognized that some fossil crocodilians were distinctly different from others and also from living crocodilians with respect to two anatomical features. The first feature is the location of the internal nares: the inner openings of the nasal passage that admit air by way of the pharynx to the lungs. In living crocodilians a long bony partition, the secondary palate, divides the oral cavity from the nasal passage. Hence a substantial distance separates the external nares near the tip of the snout from the internal nares near the back of the skull, behind the secondary palate. In early fossil crocodilians the distance was much less.

The second anatomical feature is the shape of the vertebrae. Among living crocodilians the articulation of the spinal column is a ball-and-socket arrangement: the rounded back end of each vertebra fits into a hollow in the front end of the vertebra behind it. Like the rearward location of the internal nares, this was not the case among early crocodilians. Huxley proposed that there had been an evolution from a primitive anatomy to a more progressive one with respect to both features, and fossil discoveries since his day have confirmed the general validity of his observation.

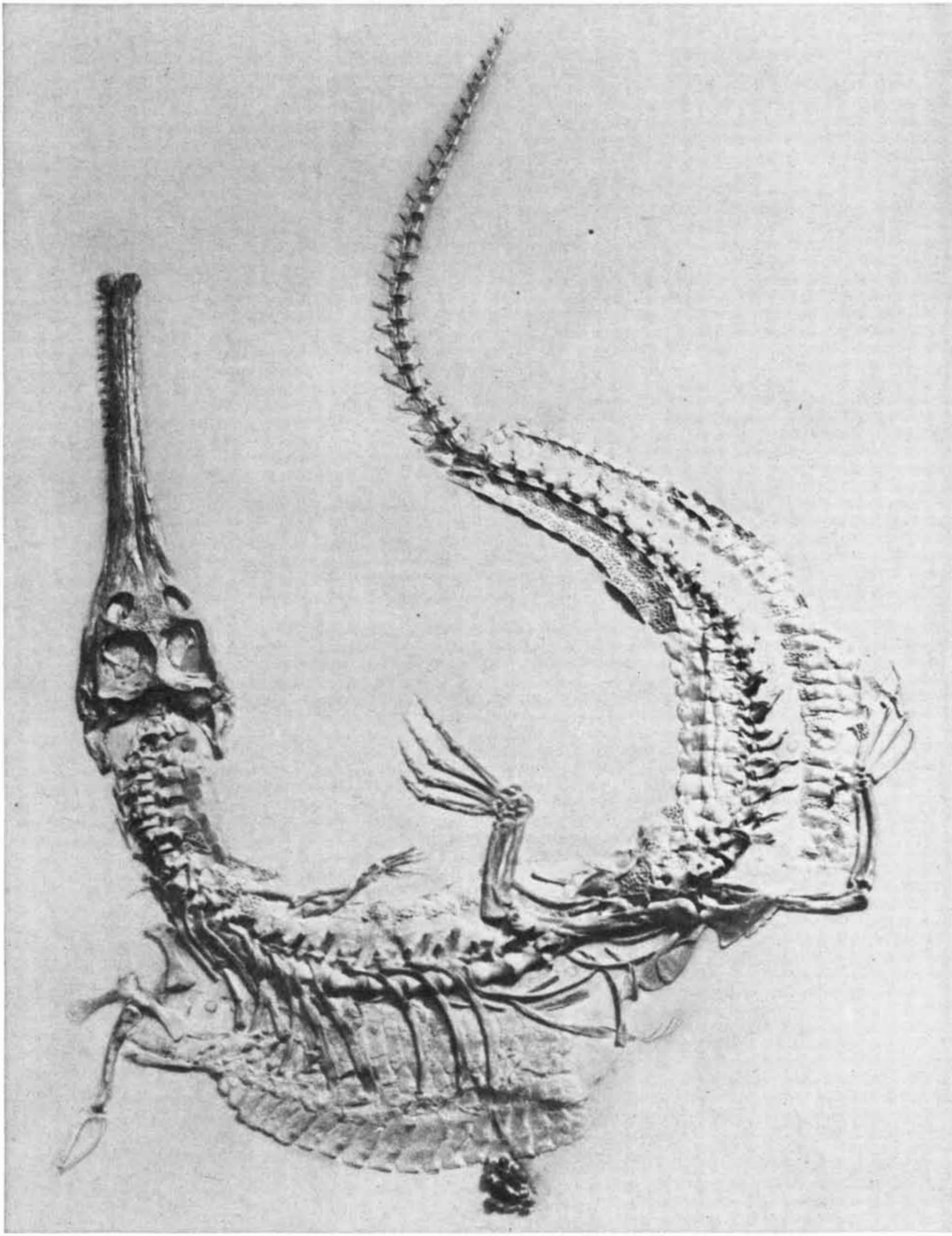
The crocodilians belong to the same reptilian subclass, the Archosauria ("ruling reptiles"), that includes the dinosaurs and the flying pterosaurs. The earliest crocodilians known, members of the suborder Protosuchia, are found as fossils in continental rock formations of late Triassic age in North and South America (Arizona and Argentina) and in Africa (South Africa and Lesotho). Fossil reptile specimens from Britain and China may also belong to this late Triassic suborder. The protosuchians were small animals, hardly more than a meter long; their snout was fairly short and their legs were fairly long. A strong armor of bony plates covered their back and belly. Their presence in continental rocks suggests that some of them were land carnivores, but among the several genera of protosuchians are others that already show adaptations to an aquatic life. The seemingly wide geographical distribution of the protosuchians, particularly when Britain and China are included, is partly illusory. In Triassic times the present continents were still united in a single land mass and the barriers to migration were few.

All protosuchians exhibit anatomical features reminiscent of primitive reptiles. One of them is the shape of the vertebrae that Huxley noted: each vertebra is slightly concave at both ends. Another is the presence of paired openings in the sides of the snout, forward of the eye sockets. The function of these openings is not clear. Finally, the secondary palate is made up of fewer bones than in later crocodilians. Nevertheless, the number of protosuchian genera and their differentiation suggest that by the late Triassic they had been evolving from an ancestral stock for some time.

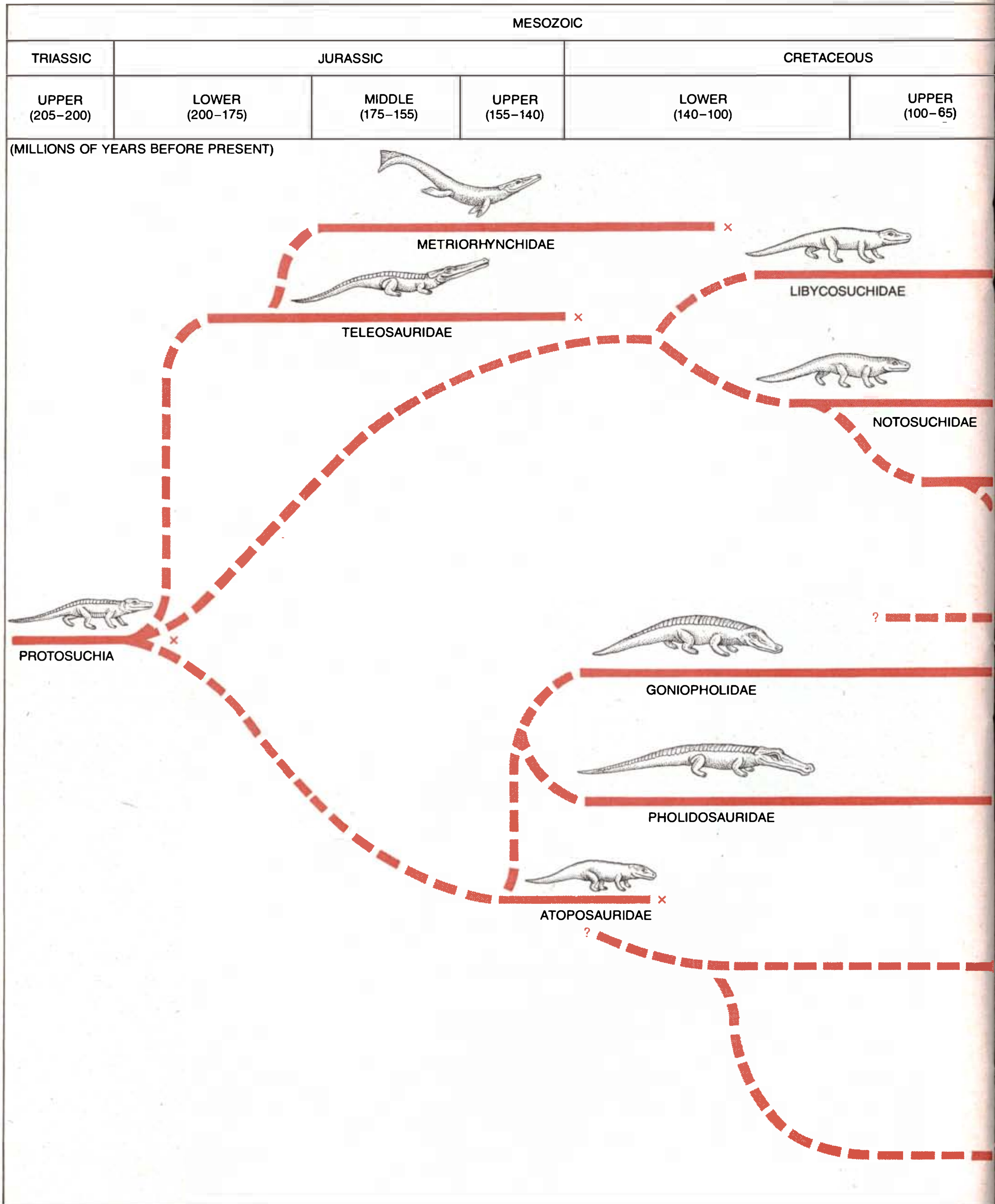
After the end of the Triassic there is a gap of some millions of years in the crocodilian fossil record. Not until the later Lower Jurassic do complete specimens appear, and they are very different from the protosuchians of the Triassic. Practically all the crocodilians from Lower and Middle Jurassic fossil beds are found not in continental rocks but in marine ones, laid down in the shallow seas that had invaded large continental areas at the time.

The earliest of the Jurassic crocodilians, members of the suborder Mesosuchia, were among the first vertebrate fossils known to science. Representatives of the family of teleosaurids, they were found in some abundance in rock formations in Britain, France and Germany as early as the 18th century. The teleosaurids, some of them small but others measuring up to four meters or more overall, are superficially reminiscent of living gavials; this is to say that they had a long snout and narrow jaws with numerous slender teeth. They must have been efficient predators of fish.

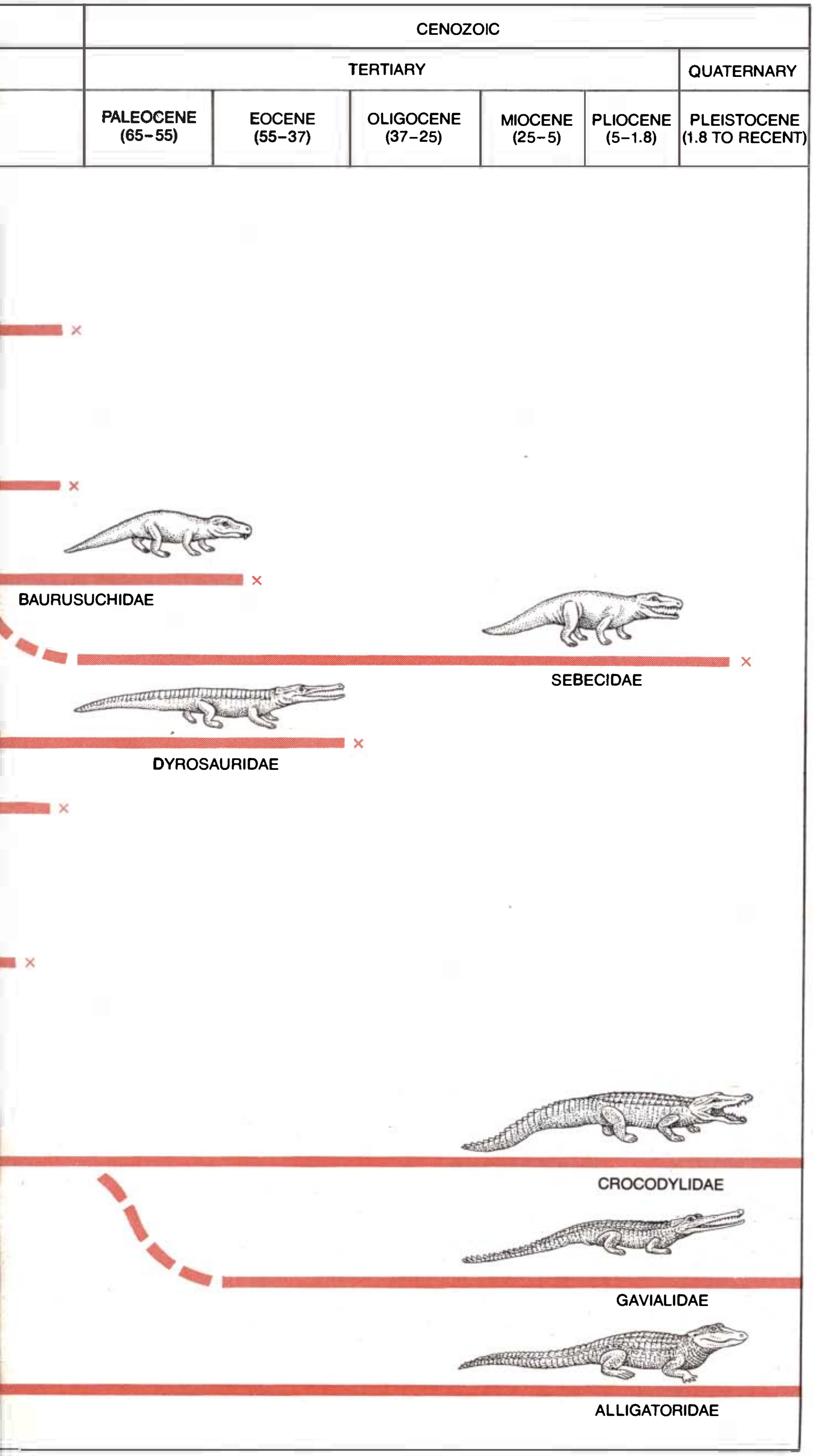
With respect to the two primitive anatomical features that characterized the protosuchians the Jurassic mesosuchians occupy an intermediate position: their vertebrae were still concave at



FOSSIL CROCODILIAN of the Lower Jurassic was found in a shale formation at Ohmden in southwestern Germany. The specimen is in the collection of the Paleontological Institute of the University of Tübingen. The three-meter animal belongs to the genus *Steneosaurus,* one of the teleosaurids, a family adapted to life in the shallow seas that invaded the margins of the continents in Jurassic times, starting some 200 million years ago. Superficially the teleosaurids resemble the narrow-snouted gavials of Asia, one of the three families of living crocodilians. They are, however, not related; the teleosaurids became extinct early in the Cretaceous, some 130 million years ago.



CROCODILIAN EVOLUTION over a span of more than 200 million years is traced along its main branches (*colored lines*). These are connected conjecturally (*broken lines*) or uncertainly (*question marks*). At the starting point is the protosuchian suborder, found in fossil formations of the Upper Triassic. Some of these small land animals exhibit an adaptation to aquatic life. A gap separates these first crocodilians from the suborder of mesosuchians: groups that appear in Jurassic and Cretaceous fossil formations. Only a few mesosuchian



families persisted into the Cenozoic era; the last members of the suborder may have survived in Australia until about a million years ago. The suborder of eusuchians, which comprises the three families of living crocodilians, first appeared during the Cretaceous, but the narrow-jawed gavials probably did not branch off from the crocodiles proper until early Tertiary times.

both ends but their internal nares were closer to the back of the skull and additional bones were included in the palate. Their bodies were still protected by an armor of bony plates, but their limbs, particularly the front ones, were short. In effect these crocodilians had adapted to occupy the niche of a fish-eating marine vertebrate. The teleosaurids apparently ranged worldwide; their remains are found not only in Europe but also in Asia, North Africa, Madagascar and in North and South America. They flourished in Europe until the end of the Jurassic but became extinct soon afterward.

A second group of marine crocodilians evolved during the Lower Jurassic, probably from teleosaurid ancestors. This was the family of metriorhynchids. They are the most specialized of all the members of the crocodilian order, being profoundly modified for an aquatic life. Their kinship to the teleosaurids is evident from their double-concave vertebrae, the structure of their palate and other features of their cranial anatomy. Beyond this, however, the skull is streamlined, the limbs are transformed into paddles and the end of the spine is bent sharply down. Impressions of the skin of metriorhynchids are preserved in certain Upper Jurassic fossil beds in Germany; they show that the animal's bent tail bore a sharklike caudal fin. The bony plate armor characteristic of the teleosaurids and their Triassic predecessors has completely disappeared.

The metriorhynchids must have been efficient swimmers; studies of the rocks that contain their remains indicate that unlike their shallow-water cousins they ventured a long way from shore. How they reproduced is a matter of speculation. They may have given birth to live offspring, as certain sea snakes do today. Or they may have emulated today's sea turtles and crawled ashore to lay their eggs. Their fossil remains have been found mainly in Europe, but Z. B. de Gasparini of the Museum of Natural History in La Plata in Argentina and her co-workers have recently described well-preserved metriorhynchid skulls uncovered in Jurassic formations both in Argentina and in Chile. The deep-sea metriorhynchids died out early in the Cretaceous; they were not replaced by any similarly adapted crocodilians.

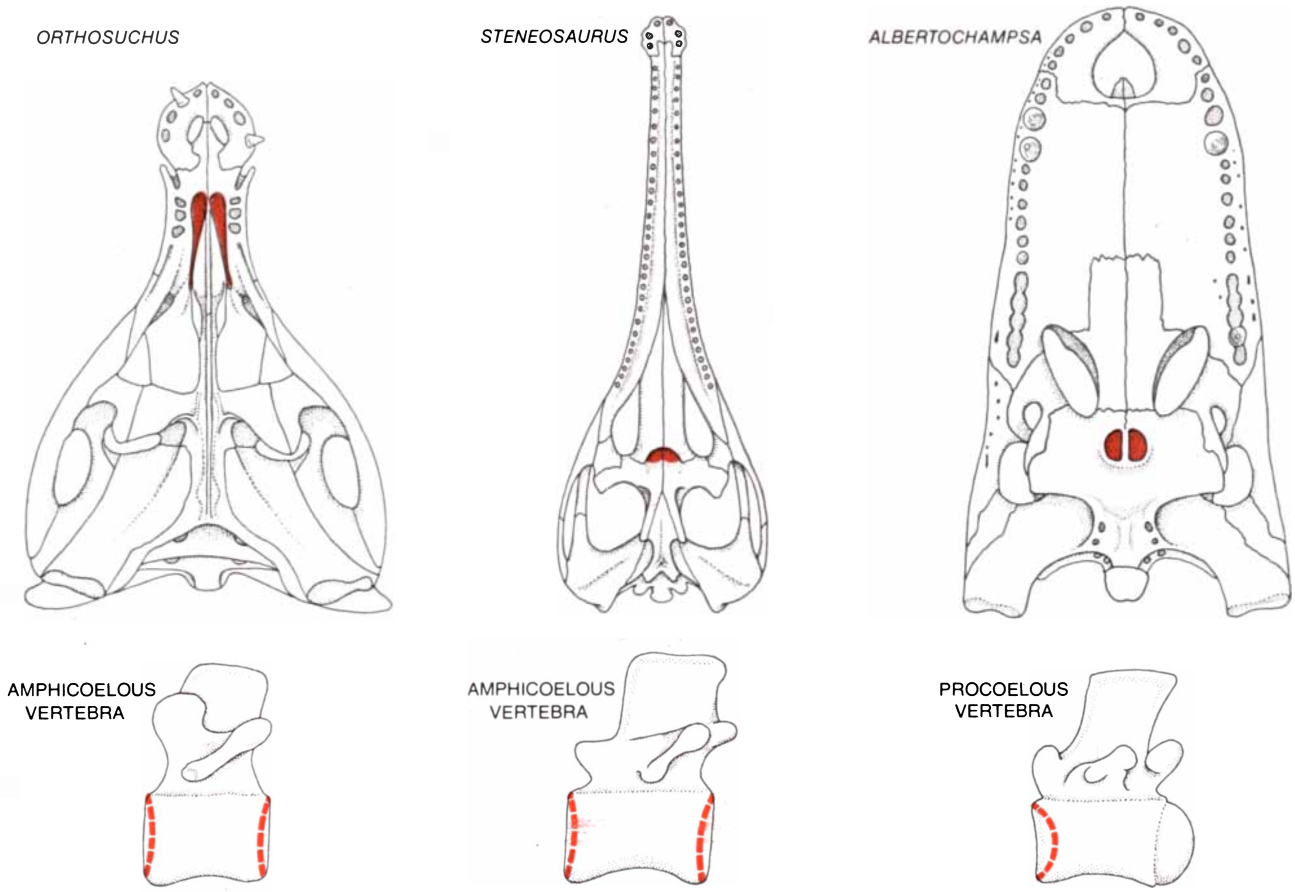
The end of the Jurassic was marked by a general retreat of the shallow seas that had invaded the continents of the world. One consequence of the retreat is that rock formations containing fossils of land and freshwater animals of Upper Jurassic and Lower Cretaceous age are comparatively abundant. Among these vertebrates of the latter part of the Mesozoic are a variety of crocodilians that were adapted to land, freshwater and seashore environments.

Their sudden appearance in the fossil record at this time bespeaks a long period of previous evolution and diversification, but in most instances their relations with earlier crocodilians remain obscure. Perhaps the new crocodilians had arisen deep in the interior of the continents; the absence of continental rock deposits formed during Lower and Middle Jurassic times makes the fossil past of these animals difficult to reconstruct. One of the new crocodilians, the family of atoposaurids, is by good luck well preserved in the lithographic limestones of Europe. A detailed study of their very complete skeletons made by P. Wellnhofer of the Bavarian State Collection for Paleontology and Historical Geology in Munich indicates that these small crocodilians, on the average no more than 40 centimeters long, were probably agile land carnivores.

A second family of newcomers, the goniopholids, were much like today's alligators and crocodiles in their general appearance and mode of life. They were fairly large animals with a broad and moderately long snout and strong teeth. Their vertebrae were of the primitive double-concave kind, and the roofing of their palate, although it was somewhat more progressive than that of earlier crocodilians, is still mesosuchian in pattern. The ecological niche occupied by the goniopholids must have been similar to the one exploited by living crocodiles. Their fossils are found in Upper Jurassic and Lower Cretaceous formations both in western Europe and in North America; the distribution is not surprising, because at that time the Atlantic had not yet completely opened and animals were free to move between North America and Eurasia.

Other crocodilians, mostly small ones, are preserved in these same fossil formations. Some of them had already evolved characteristically modern ball-and-socket vertebrae and rearward internal nares, but these features often remain combined with primitive characteristics so that it is difficult to classify the animals in a satisfactory way. One can only say in general that at the beginning of the Cretaceous several lineages of small mesosuchians showed a trend toward the acquisition of progressive features. Descended from one of these lineages, in all probability, were the eusuchians: the suborder that includes the living crocodilians. It is difficult, however, to decide which lineage it is. Possible candidates are *Theriosuchus*, a genus from the fossil-rich Weald beds in Britain, and a related form from North America now being studied by Wann Langston, Jr., of the University of Texas at Austin.

In the southern bloc of continents, known as Gondwanaland, the situation in early Cretaceous times was somewhat different. The various components of Gondwanaland shared a few groups of crocodilians with the northern bloc of continents, known as Laurasia. Among them were long-snouted pholidosaurids, but on the whole the crocodilians of the south appear to be more primitive. The pholidosaurid record for size is held by a huge specimen from Africa. P. Taquet of the National Museum of Natural



SKULLS AND VERTEBRAE trace the three main stages of crocodilian evolution. In the typical protosuchian, exemplified at the left by one species of the genus *Orthosuchus* from an Upper Triassic formation in Africa, the internal nares, or nostrils (*color*), are near the front of the skull and the vertebra is double-concave (*colored broken lines*). In the mesosuchian, at the center, a species of the gavial-like teleosaurid genus *Steneosaurus* from a Jurassic formation in Europe, the vertebra is still double-concave but the nares are much closer to the rear of the skull. In the eusuchian, at the right, a species of the alligatorid genus *Albertochampsa* from an Upper Cretaceous formation in Canada, the modern condition has evolved: the nares lie even more to the rear, opening within the pterygoid bones of the palate, and the vertebra is concave-convex. The specimens are not shown at the same scale: *Orthosuchus* was quite small and the others were larger.

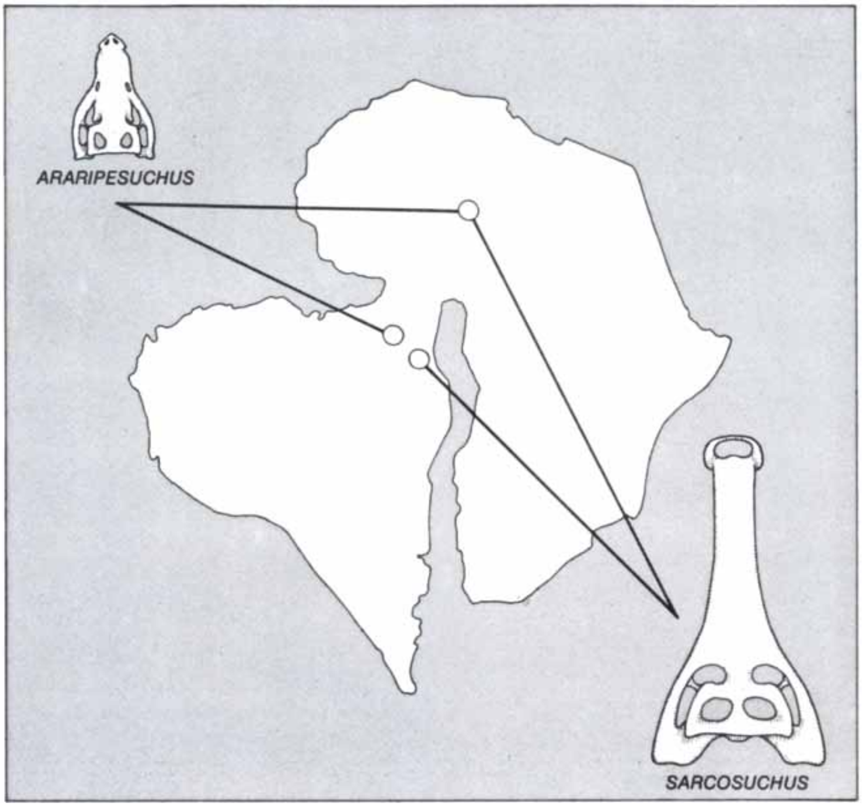
History in Paris has turned up complete skulls and partial skeletons of the species *Sarcosuchus imperator* in a remarkable Lower Cretaceous fossil locality in the Saharan part of the Republic of Niger. This gavial-like monster was 11 meters long overall, and its narrow skull measures 1.8 meters. It apparently lived in the rivers of this now-arid region.

The same freshwater fishes are found in Lower Cretaceous formations in both Africa and South America, one of several indications that at the time the two continents were not completely separated by the South Atlantic seaway. The crocodilians of Niger provide further support for this conclusion. As one example, Taquet's Niger fossil collection includes a very small mesosuchian that belongs to the genus *Araripesuchus;* the first member of the genus to be described was unearthed some two decades ago from a Lower Cretaceous formation in northeastern Brazil. A second example is provided by the remains of a very large, long-snouted crocodile from the Lower Cretaceous of the Bahia basin that was originally classified as a goniopholid. I recently reexamined this Brazilian specimen, and when Taquet and I compared findings, it was evident that it belonged to the Niger genus *Sarcosuchus.*

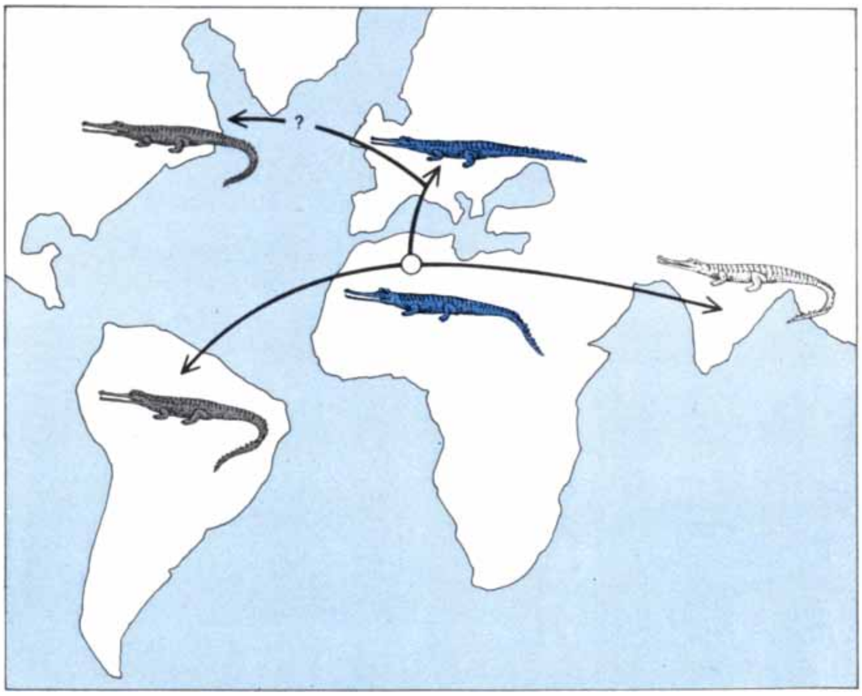
By the middle of the Cretaceous, Africa had become completely separated from South America, and the crocodilians of the two continents, although they retained some resemblances, increasingly diverged. At the same time evidence from both continents shows that ecological pressures continued to influence crocodilian evolution for the balance of the period. Small, short-snouted mesosuchians now arose in both Africa and South America. Their remains are found in Upper Cretaceous fossil beds, the libycosuchids in the Old World and the notosuchids in the New World. With their short jaws these animals are in a way reminiscent less of normal crocodiles than of carnivorous mammals or mammal-like reptiles. They almost certainly occupied a land predator's niche.

The origin of these remarkable crocodilians is still not clear. Their resemblance to the earlier protosuchians suggests the possibility that they may be direct descendants of those first crocodilians, by way of a common ancestor that resembled the early mesosuchian *Araripesuchus.*

The end of the Cretaceous, some 65 million years ago, was marked by mass extinctions of animals on a worldwide scale. The dinosaurs, the flying reptiles and the great marine reptiles (the plesiosaurs and the mosasaurs) all died out. Of all the archosaurian reptiles only the crocodilians survived, apparently little affected by whatever proved fatal

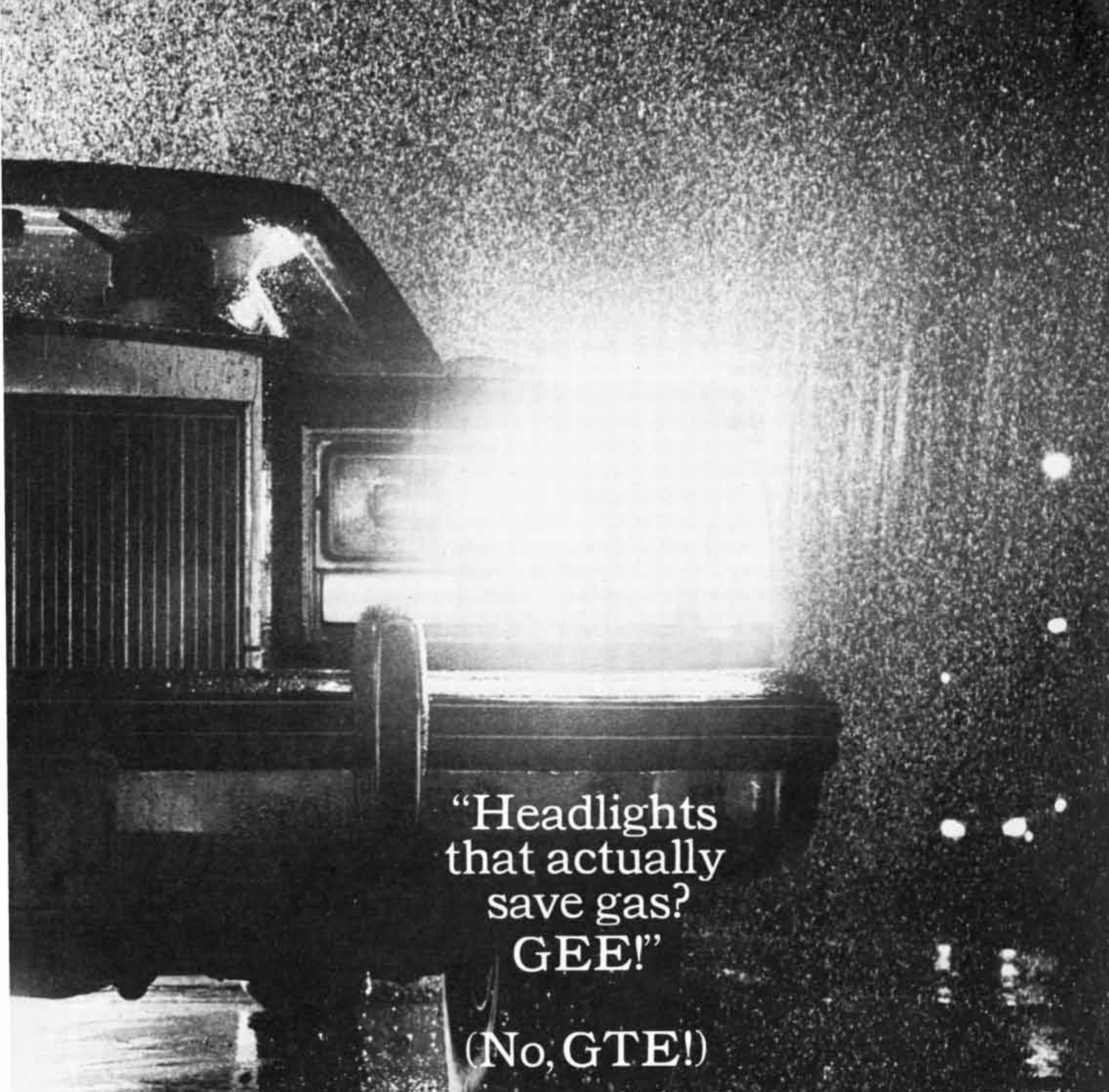


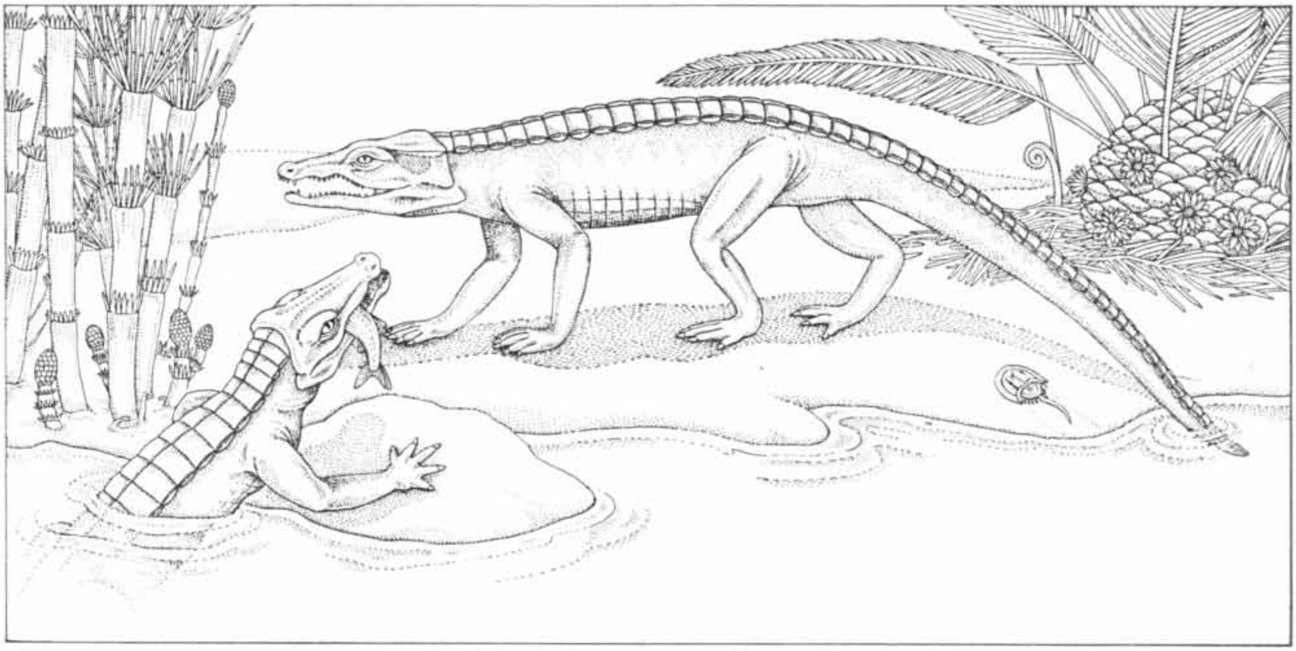
IDENTICAL CROCODILIANS of the Lower Cretaceous period are found in South America and in Africa, which were then still connected. One is the giant freshwater pholidosaurid, *Sarcosuchus,* first recognized in a Niger fossil formation. The other is *Araripesuchus,* a short-snouted land-dwelling mesosuchian first recognized in Brazil and also found in the Niger beds.



FOSSIL GAVIALS of Cenozoic age are found in both North and South America (*dark profiles*), but living gavials (*light profile*) are confined to the great rivers of Pakistan, India and Burma. This was difficult to explain until certain early Cenozoic long-snouted crocodilians in North Africa and Europe (*colored profiles*) were identified as primitive gavials. Very probably New World migrants arrived from Africa and Europe late in Eocene times, when the Atlantic Ocean was narrower than it is now. Gavials lived on in Europe until 20 million years ago.


One Stamford Forum, Stamford, Conn. 06904

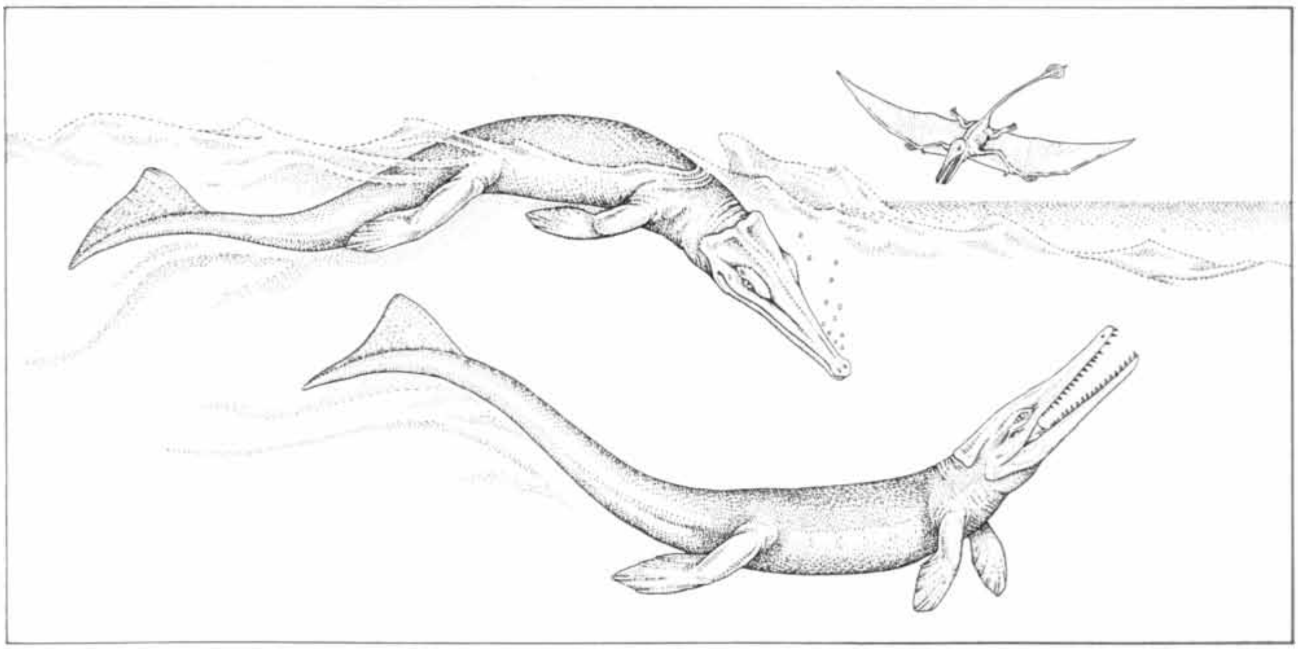

"Headlights
that actually
save gas?
GEE!"
(No, GTE!)



LONG-LEGGED PROTOSUCHIANS appear in a reconstructed Upper Triassic setting. They are the armored species *Orthosuchus stormbergi*, from fossil beds of that age in Lesotho in southern Africa, recently described by Diane S. Nash of the University of Bristol. Some 60 centimeters long, the species must have been an efficient terrestrial carnivore but was even then adapting to an aquatic life.

to many other animals. A few crocodilians, such as the goniopholids, did become extinct at that time, but they were probably victims of competition from the more progressive eusuchians rather than of the mysterious "great dying." (Two of the three families of living eusuchians, the alligators and the crocodiles, had by then appeared in Laurasia, notably in 70-million-year-old freshwater deposits in Canada and the western U.S.)

The cause of the mass extinctions has been a matter of much speculation, and no single explanation can be said to have gained general acceptance. The fossil record of the crocodilians does not reveal what actually happened, but it does contain some interesting negative information. For example, it has been suggested that the mass extinction was



AQUATIC MESOSUCHIANS are shown reconstructed in an open-sea setting of Jurassic times. They are of the metriorhynchid family, the most specialized of all the crocodilians. They have lost the plate armor of their ancestors and have converted their limbs into swimming paddles. The end of the spine is bent sharply down. Impressions of their skin, preserved in Upper Jurassic fossil beds in Germany, reveal that their tail supported a sharklike caudal fin. The metriorhynchids' adaptation to a marine predator's niche began in the Middle Jurassic; before they became extinct in the Lower Cretaceous some species had evolved to be from four to five meters long. No other crocodilians became so fully adapted to the niche they left vacant. Their range included both European and South American waters.

How to project a beautiful image.

To make the most of your slides, nothing does it like a Kodak Carousel projector. For bright, crisp, colorful images on the screen. With eleven models to choose from, you can get just what you want in the way of convenience. And in the way of price.

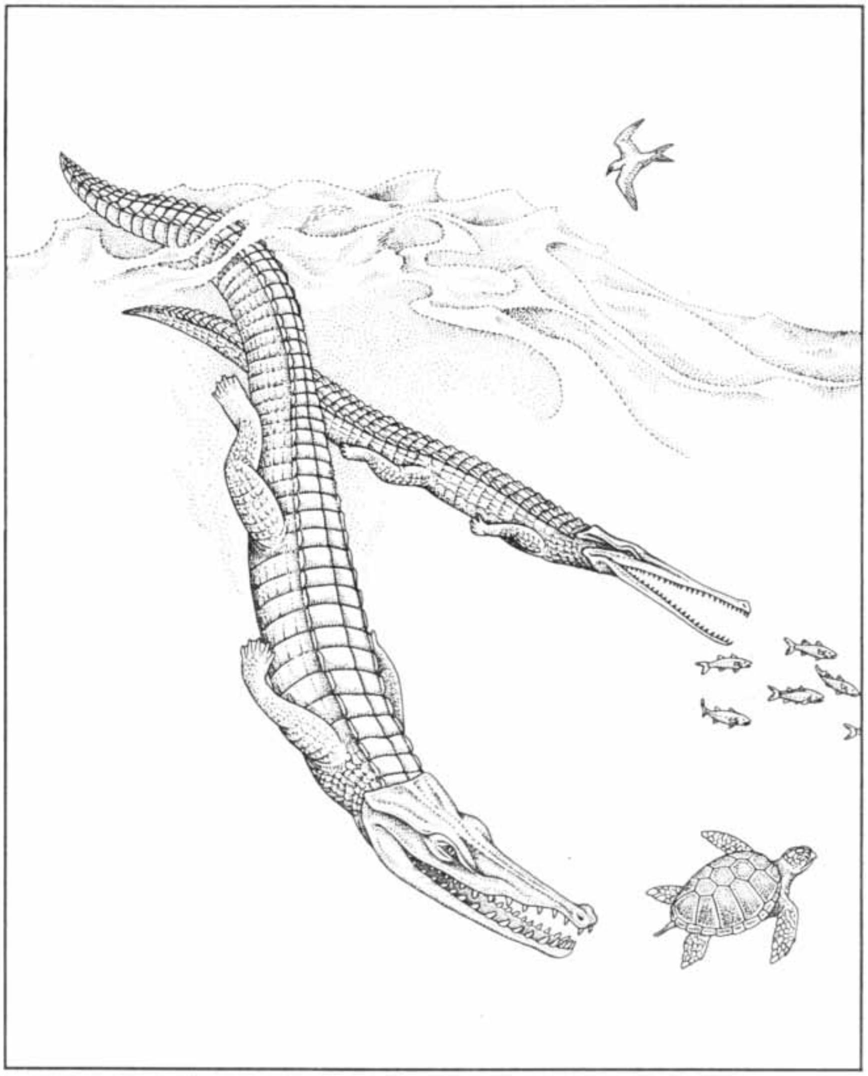
All Kodak Carousel projectors use dependable gravity to drop each slide in place. All feature a superb, $f/2.8$, curved-field lens for bright, sharp viewing. All are whisper-quiet.

Attention, international travelers. Kodak now offers an automatic multi-voltage Carousel projector for use in most parts of the world.

Kodak Carousel projectors

The dependable ones.

Kodak



CONTINUED DIVERSIFICATION in Cenozoic times affected one family of mesosuchians that appeared in Africa at the end of the Cretaceous. These were the dyrosaurids, which inhabited shallow coastal waters early in the Tertiary. Their fossils, abundant in the phosphate deposits of southern Tunisia, show that one species, *Dyrosaurus phosphaticus,* had a very long, narrow snout and pointed teeth suitable for capturing fish (*right*). Its cousin, *Phosphatosaurus gavialoides,* had a shorter, more robust snout (*center*). Its jaws carried large, blunt rear teeth adapted to the crushing of hard-shelled prey; the species may have fed on marine turtles. Both species were large: the fish eater reached six meters in length and the turtle eater reached nine.

caused by a cooling of the climate. Indeed, there is some evidence for a climatic deterioration at the end of the Cretaceous. The continued survival of the crocodilians, however, shows that world temperatures cannot have fallen very low, particularly since some of the survivors then inhabited high latitudes. Today's crocodilians are confined to warm climates. Even alligators, which have evolved mechanisms for hibernation, die in a prolonged cold spell. A large and extended drop in world temperatures at the end of the Cretaceous would certainly have added the crocodilians to the reptile casualty list.

The history of the crocodilians has now come up to the Cenozoic. This era is conventionally divided into one relatively long span of time, the Tertiary, and a brief late span, the Quaternary. (The terminology is left over from the early days of historical geology, when the rest of earth history was put into a Primary and a Secondary interval.) The crocodilians of the early Tertiary continued to some extent to reflect the north-south dichotomy that had arisen during the Mesozoic. The eusuchians were well established in North America, Europe and Asia; to the existing populations of crocodile forebears and alligator forebears were now added long-snouted gavial-like animals that inhabited coastal waters.

In the generally isolated southern continents a late group of mesosuchians now enjoyed a remarkable evolutionary success. These were the dyrosaurids:

large animals, measuring as much as nine meters overall, with long, narrow jaws and, in most members of the family, long, pointed teeth that mark them as adapted to the same fish feeders' niche that had been occupied by a succession of gavial-like crocodilians throughout the evolutionary history of the order. One dyrosaurid, however, was a variation on this common theme. The genus *Phosphatosaurus* is found in African fossil beds of the early Tertiary. Its jaws were more robust than those of its fish-feeding cousins and its teeth were very strong, particularly the back teeth, which had blunt, rounded tips. It would appear that *Phosphatosaurus* preyed on hard-shelled animals that had to be crushed, possibly marine turtles.

Dyrosaurid remains have been found in Brazil and New Jersey as well as in Africa; they are also known in Asia, where they occur in freshwater Tertiary deposits in Pakistan and Burma. Indeed, the last dyrosaurid fossil known is a single vertebra that is roughly 40 million years old, found in a fossil bed of the late Eocene in Burma. By then the family had already disappeared from Africa, possibly because of competition from the long-snouted northern eusuchians, adapted to the same gavial-like way of life, that became increasingly numerous in Africa during the Eocene.

Among the living crocodilians the gavials are now confined to the great rivers of Pakistan, India and Burma. All are of the same species, *Gavialis gangeticus*. The presence of ancestral eusuchian gavials in Africa during the Eocene helps to explain the past distribution of gavials from India to South America (by way of Africa and the Atlantic, which was a narrower ocean in the early Tertiary). The African population, having apparently replaced the gavial-like dyrosaurids in Eocene times, seems itself to have come under pressure from a peculiar crocodilian with long, narrow jaws that evolved in Africa later in the Tertiary and survived into the Quaternary. This was *Euthecodon*. Its competition for the same ecological niche may account for the disappearance of the gavial from Africa. For reasons still not known the American and European gavials also died out in Cenozoic times, so that the only surviving gavials are those that took shelter in the river systems of the Indian subcontinent.

None of the living crocodilians is nonaquatic, but still one more group of land crocodiles became prominent in South America during Tertiary times. They evidently arose from a group of peculiar mesosuchians, the baurusuchids, that first appeared in the late Cretaceous. By the beginning of the Tertiary their descendants, the sebecids, had appeared. They were large animals with a high, narrow snout. Their most remarkable feature was their dentition:

their teeth were long and laterally flattened and had sharp, serrated cutting edges, much like those of carnivorous dinosaurs.

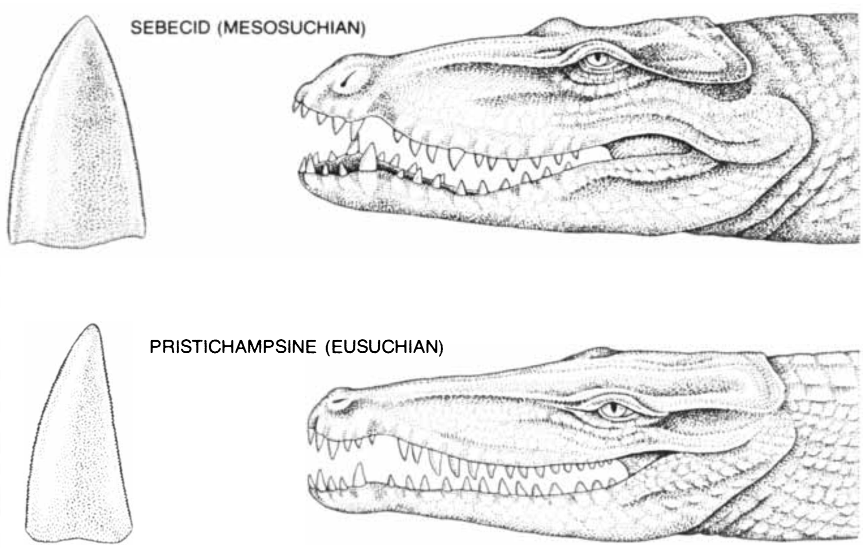
At the time the sebecids flourished South America was cut off from the other continents. When the carnivorous mammals that had become widespread elsewhere were able to enter South America toward the end of Tertiary times, their arrival seems to have been rapidly fatal to the predatory sebecids. It was even more rapidly so for another dinosaur-toothed group of terrestrial crocodilians, the pristichampsines, which evolved from more normal eusuchians somewhere in the Laurasian region at the beginning of the Tertiary. Although the pristichampsines were initially successful, with populations distributed from China to North America, they did not survive mammalian competition beyond Eocene times, at least in those regions. There is some evidence that both they and the sebecids survived into the Quaternary period in isolated Australia.

On the northern continents the late Tertiary history of the crocodilians is largely one of extinction. Alligators and gavials were still numerous in Europe at the beginning of the Miocene, but at the end of that epoch, some five million years ago, they had become very scarce. The last crocodiles of Europe seem to have disappeared during the following epoch, the Pliocene. As D. E. Berg of the University of Mainz has pointed out, the cause of the decline was doubtless a change in the European climate: conditions became increasingly cool and dry during both the Miocene and Pliocene, and environments suitable to crocodilians eventually vanished.

The same climatic change took place in North America. In Miocene times alligators (among them the probable ancestors of the living American form) inhabited what are now the north-central states of the U.S. Today the American alligator is confined to the southern states.

In the warmer regions of the world, however, crocodilians continued to flourish. During the late Tertiary a number of giant eusuchians evolved in South America. Although these oversize forms have since become extinct, the region's crocodilians remain varied to this day, as is indicated by the number of different species of crocodiles and their cousins the caimans found there. In Africa too the abundant crocodilian remains that are present in the fossil-rich formations of the Great Rift Valley demonstrate that the population of such animals was quite varied until well into Quaternary times. In addition to forms that are now extinct both the Nile crocodile (*Crocodylus niloticus*) and the West African long-snouted crocodile (*C. cataphractus*) were already present at the beginning of the Quaternary.

There is little doubt that predation by the Nile crocodile was a not negligible hazard for early man in Africa. The situation is now reversed: man is playing the leading part in the decline and fall of the crocodilians. In practically all parts of their range hide hunting has been decimating the crocodilian population for years. The time has come when man must choose between a few more handbags and shoes and the survival of the last remaining archosaurian reptiles.



CONVERGENT EVOLUTION produced flattened and serrated dinosaurlike teeth in two separate crocodilian lines. In South America the peculiar dentition arose among the sebecids (*top*), a lineage of mesosuchian land-dwelling predators that flourished until carnivorous mammals reached the continent at the end of the Tertiary. In the northern continents an almost identically shaped carnivorous dentition arose among a group of eusuchian predators, the pristichampsines (*bottom*). These died out in the northern part of their range during the Eocene.

On Photographing a Small Step

It is unearthly. The landscape underfoot. It is soft and powdery.

A silver beast stands there. His head is a shiny globe. His huge frankfurter fingers hold a small box that talks.

Click. Whir. Click. Whir.

A man is taking the first pictures on the moon. He takes them with a camera that gives him startling detail. It is a Hasselblad.

It is the first of twelve taken to the moon on six landings. And it is still there today.

Photograph by Hasselblad courtesy of NASA

Hasselblads have taken tens of thousands of photographs during millions of miles of travel in space.

In all, 60 Hasselblads on 27 flights with 67 astronauts have completed 3566 orbits of earth and 363 of the moon.

Every Hasselblad proved reliable on every flight under every condition.

Today, 10 years after, Hasselblad celebrates mankind's giant leap with a commemorative issue of a limited edition (earth version) of the 500EL/M motor drive camera that took those historic photographs.

The only medium format camera on earth or in space that can equal it, is another Hasselblad.

For even in an age of space, much of a year goes into building each Hasselblad; much of it crafted by hand.

100 Hasselblads from one

No Hasselblad can be any better than any other Hasselblad; only more suited to your purposes.

Hasselblad is a prodigious system of 4 cameras, 22 lenses, 8 viewfinders, 9 film magazines, and over 300 other accessories. And there is no obsolescence.

Simply by choosing the right pieces, you can *custom build* over 100 distinct Hasselblad cameras for almost any photographic application.

From a man's first steps on the moon to a child's first steps on earth.

An incredible motor drive

The 500 EL/M is the only hand held medium format camera with an integrated motor unit.

So unlike cameras with optional motor drive attachments, it is built to withstand the punishing demands of motorized operation.

And the possibilities are infinite.

There is a dial set in the body that will order the camera to fire in any of five different ways.

From single shot to speeded up continuous firing at the rate of one shot every 8/10th of a second.

That means you can walk away from the camera, devoting all your concentration to your subject, while firing with a remote cable.

Plug in the intervalometer, for instance, and you can trigger the camera automatically. At any moment you want, from 2 seconds to 14 minutes.

Or trigger up to four cameras simultaneously with a command unit.

Or, fitted for remote radio control, fire from a half mile away.

And the standard film frame on a Hasselblad is 2¼ inches square, *almost 4 times the area* of a 35mm frame. So the ability to hold sharpness and contrast is phenomenal. (For actual size of film frame, see box at lower right.)

The eyes of a Hasselblad

A Hasselblad can see with any one of 22 lenses; all multi-coated (except for two special purpose lenses) in focal lengths from 30mm to 500mm with angular fields from 180 degrees to 9 degrees.

There are lenses for sports, aerial, architectural, and fashion photography. Portrait, landscape, medical, underwater and news photography. Wildlife, laboratory, industrial and child photography.

A lavish brochure on the 500 EL/M and other Hasselblad cameras is available free if you write: Braun North America, Dept. SA10H, 55 Cambridge Parkway, Cambridge, MA 02142–a division of the Gillette Company and exclusive marketer of Hasselblad in the U.S.

There is even the 105mm Zeiss UV-Sonnar f4.3 which can photograph the ultraviolet portion of the electromagnetic spectrum with costly quartz elements that can detect radiations unseeable by the human eye.

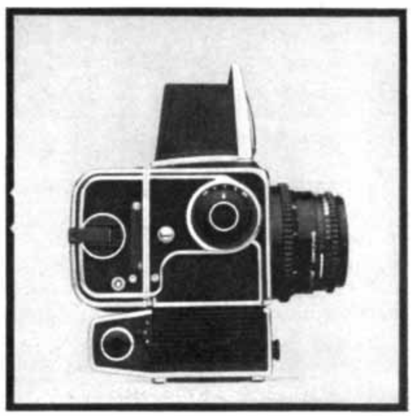
There are also lenses that adjust the aperture fully automatically as the light changes.

Even if you are not yet into these more exotic applications but simply recording travel and everyday happenings, it's nice to know that your camera produces the very best images in the medium format.

So, in effect, there is nothing you can't see with a Hasselblad. And photograph. From what is invisible to the naked eye to all that is visible in the universe.

Is it any wonder then that NASA came to Hasselblad for the camera which was to photograph space? And that top professional and serious amateur photographers still rely on Hasselblad.

HASSELBLAD®



The Motor-Drive Hasselblad 500 EL/M, made in Sweden. The standard lens is a multi-coated Zeiss Planar 80mm f/2.8T.*

Constructive Mathematics

This approach is based on the belief that mathematics can have real meaning only if its concepts can be constructed by the human mind, an issue that has divided mathematicians for more than a century

by Allan Calder

It is commonly held that if human beings ever encounter another intelligent form of life in the universe, the two civilizations will share a basic mathematics that might well serve as a means of communication. In fact, since the time of Plato it has been generally believed that mathematics exists independently of man's knowledge of it and thus possesses a kind of absolute truth. The work of the mathematician, then, is to discover that truth. Not all mathematicians, however, have shared this belief in a "God-given" mathematics. For example, the 19th-century German mathematician Leopold Kronecker maintained that only counting was predetermined. "God made the integers," he wrote (to translate from the German). "All else is the work of man." From this point of view the work of the mathematician is not to discover mathematics but to invent it.

The question of whether mathematics is discovered or invented is not an idle one. Depending on their answer to it mathematicians can have radically divergent views on how the work of mathematics should be conducted. In particular, the belief that mathematics is invented has given rise to a controversial theory known as constructive mathematics, which maintains that to prove that a mathematical object exists it is necessary to show how the object can be constructed. Most recently constructive mathematics has been put forward in a new and challenging form by Errett Bishop of the University of California at San Diego, but its origins can be traced to Immanuel Kant, who in the 18th century maintained that the ultimate truth of mathematics lay in the fact that its concepts could be constructed by the human mind. The consequences of this point of view for such fundamental issues as the validity of mathematical proofs and the meaning of mathematical existence can perhaps best be understood by beginning with the ideas and methods of ancient Greek mathematics, which have guided the development of mathematics over the past 2,500 years.

To the Greeks mathematics was essentially geometry, and the quintessence of the subject was Euclid's *Elements,* where the deductive methods of the Eleatic school of philosophy, the logical procedures of Aristotle and Platonic idealism came together in the notion of an axiomatic system. Euclid began with certain basically undefined objects such as points, lines and planes and a number of axioms, or presumably self-evident statements, about those objects and applied the Aristotelian rules of inference to obtain new statements called theorems. The underlying assumption that because the axioms were "true" and the rules of inference "logical" the theorems too would be true was made plausible by the intuitive nature of Euclidean geometry. Although objects such as a point (which has no size) and a line (which has length but not width) cannot be found in the natural world, most people feel they know what these objects are. Moreover, Euclid's theorems could be checked empirically, that is, an experiment could be done with the aid of pencil and paper to verify that, say, two triangles with sides of the same length have angles of the same size.

In the modern era mathematics has continued to depend on the axiomatic method, but in many cases this method has been applied to objects that lack the intuitive, or constructible, qualities of Euclidean geometry. To the constructivists such applications have robbed mathematics of its meaning. As a result a number of proposals for reinstating intuitive meaning have been put forward over the past 100 years, the latest being Bishop's new constructive mathematics. The earlier efforts played an important part in determining the development of mathematics, however, and so before discussing Bishop's program I shall review the history of this controversy. It is long and varied, but its fundamental issues can be understood by following the evolution of a single idea in mathematics, namely the concept of infinity.

The famous paradoxes of Zeno concerning issues such as infinite divisibility and continuity led the ancient Greek mathematicians to be extremely cautious about infinities. The fundamental object in Greek mathematics was length, and Euclid in particular referred not to unending lines but only to line segments that could be extended to any desired length. This notion of potential infinity—that for any line there is always a longer line—is a reasonable extrapolation from natural experience. The concept of completed, or actual, infinity—that there are actually infinitely long lines—is metaphysically quite different, and it requires a daunting leap of the imagination when it is first encountered. Hence it is not surprising that mathematicians avoided actual infinities as long as they could, limiting themselves to the consideration of potential infinities until well into the modern era.

The modern rebirth of mathematics came in the 17th century, when René Descartes and Pierre de Fermat, by the introduction of Cartesian coordinates, transformed geometric problems into arithmetic ones. This arithmetization of geometry made numbers rather than lengths the fundamental objects of mathematics and paved the way for the independent development by Isaac Newton and Gottfried Wilhelm Leibniz of the infinitesimal calculus: a set of rules for the manipulation of numbers that were infinitely small without being zero. The concept of infinitely small entities (the notion of limits was not made explicit until much later) was troubling in an age when mathematical objects were considered to have the same reality as, say, the electron is thought to have today. In fact, it was some time before the undeniable utility of the calculus forced mathematicians into an uneasy acceptance of it and of the actual infinities it implied. For the next 200 years mathematicians wielded the techniques of the calculus successfully while continuing to question their philosophical basis. Then at the end of the 19th century this irritating grain of sand within the mathematical oyster brought forth a pearl: the theory of sets.

The early development of set theory was due primarily to Georg Cantor, who defined a set as a collection of real or abstract objects, along with a rule by which it could be decided whether or not any given object belonged to the set. This deceptively simple idea had profound consequences for the concept of infinity. Sets such as the set of all real numbers are by their very nature actual infinities. Considering such objects led Cantor to the astonishing conclusion that just as the cardinality, or number of elements, of finite sets can vary, so can the cardinality of infinite sets. In other words, some sets are more infinite than others.

More precisely, two sets have the same size, or cardinality, if the elements of one of the sets can be paired one for one with the elements of the other. And a set is said to be countably infinite if it has the same cardinality as the set of natural numbers 1, 2, 3,.... For example, the even numbers 2, 4, 6,... can be matched one for one with the natural numbers and thus are a countably infinite set. (As this example indicates, an infinite set can have the same cardinality as a part of itself.) Cantor showed that for any given infinite set there is always another set with greater cardinality. For instance, he observed that no matter how the real numbers are paired with the natural numbers, there will always be some real numbers left unpaired, and so the cardinality of the real numbers is greater than that of the natural numbers. In fact, Cantor proved that beginning with the countably infinite sets an infinite hierarchy of infinities can be generated.

IN "THE SCHOOL OF ATHENS," painted in the early 16th century, Raphael depicts the great philosophers and mathematicians of ancient Greece: Plato (with one hand raised) and Aristotle stand at the center of the painting, Pythagoras contemplates his system of proportions at the lower left and Euclid draws a circle on a slate at the lower right. To the Greeks mathematics was essentially geometry, and it was in *Elements,* Euclid's treatise on the subject, that the notion of an axiomatic system was first laid out. In this system certain self-evident statements called axioms are assumed to be true and new statements called theorems are derived from them using the Aristotelian rules of inference. The intuitive nature of Euclidean geometry lends support to the assumption that because the axioms are self-evident and the rules of inference are logical the theorems too are true. The development of mathematics is still guided by the axiomatic method, but over the past several centuries the objects of mathematical inquiry have become increasingly more abstract and less intuitive. As a result some mathematicians have questioned their validity, and most recently a revival of interest in these ideas has been sparked by the work of Errett Bishop of the University of California at San Diego. Bishop proposes to rebuild mathematics without the use of nonintuitive concepts, an approach he calls constructive mathematics. In Bishop's system it is not sufficient to show that a mathematical object must exist logically. One must specify a procedure for actually constructing the object. Such a procedure does not have to be implemented, but it must terminate after a finite number of steps and at the end of each step it must be evident how to proceed to the next.

Today's Bunker Ramo operates more aircraft in a day than the average pilot will fly in a lifetime.

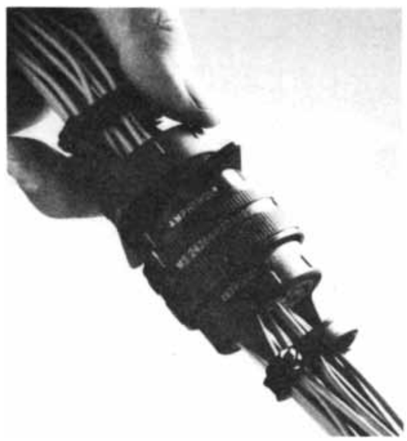
Among other things, a commercial airplane is a marvel of electrical, electronic and hydraulic components. All of which have to be connected to make that airplane operable.

That's why connectors become so critical. And when you consider that every commercial aircraft requires about 1500, you can easily understand why connectors are big business for Bunker Ramo.

Probably no connectors are better known than those marketed by Bunker Ramo under the Amphenol name. Ranging from low-cost simple designs to the highest-technology sophisticated environmental products for space and defense, Amphenol® connectors are also serving the needs of international customers in business equipment, instrumentation, federal technology, radio frequency communications and telephony.

To sustain and strengthen Amphenol leadership, Bunker Ramo invests heavily in new product development and in technologically advanced production equipment.

Today, through our well-known Amphenol connectors, Bunker Ramo has a leading position in the aerospace and aircraft industries, a strong foothold in telecommunications, and a growing role in fiber optics. In addition, we are a major producer of the data terminals and minicomputers that, together, comprise the information systems widely used by commercial banks, thrift institutions, insurance companies and brokerage houses.



Most important, we are a company with a reservoir of high technology, a solid base of innovative products, and a strong management team. All of which is why today's Bunker Ramo is a company that's on the move. For our latest financial reports, write Bunker Ramo Corporation, Dept. FR, 900 Commerce Drive, Oak Brook, Illinois 60521.

BUNKER
RAMO

Our Fish Story starts with Signals from 600 miles up.

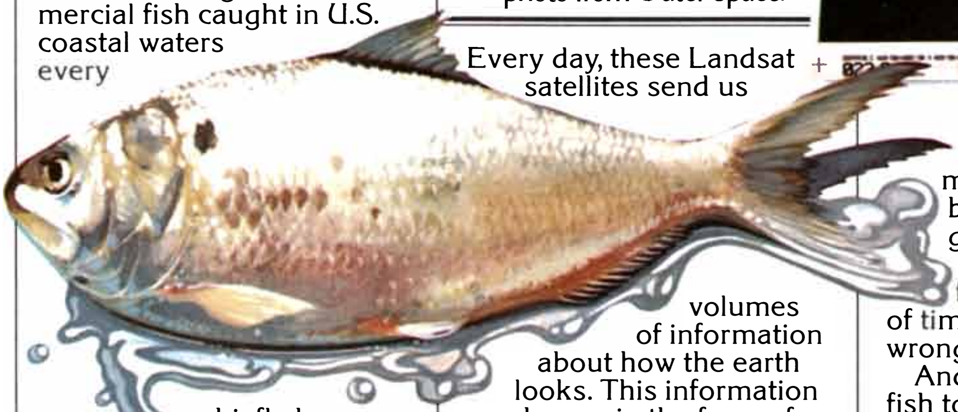
There's nothing new about a fisherman hoping for a little help from on high.

But when that help comes from a satellite, that's news.

Today, commercial fishermen are counting on—and getting—just that kind of help.

Signals from satellites 600 miles up, processed by a TRW electronic system, are helping fishermen find a homely little fish called the menhaden.

Meet the Menhaden.

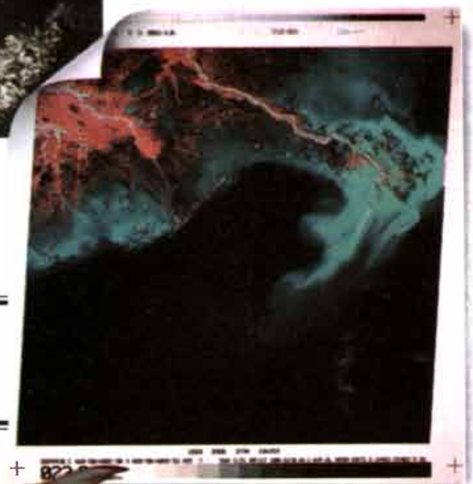


If you're like most people, you wouldn't know a menhaden if it jumped in a boat with you.

Yet it represents nearly half the tonnage of all commercial fish caught in U.S. coastal waters every year—chiefly because of the high-quality fish meal and oil it yields.

That fish meal is used as a prime nutrient in a variety of animal feeds. And menhaden fish oil finds its way into products ranging from margarine to paints to lubricants.

Landsat and TRW: Going fishing with a photo from Outer Space.



Every day, these Landsat satellites send us volumes of information about how the earth looks. This information reaches us in the form of billions of electronic signals. Then a TRW system translates and enhances them—and turns them into super-clear color photos, packed with data about the earth and its waters.

For example, these space pictures can help reveal ocean features like sediment layers, depth, salinity and chlorophyll content. And where there's chlorophyll there's very likely to be phytoplankton—the basic food for the menhaden.

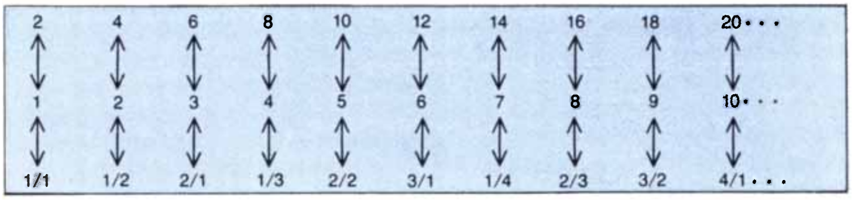
To a commercial fisherman, those photos can be worth their weight in gold—or fish.

They can save a whole fleet from wasting a lot of time looking in the wrong places for their catch.

And that means more fish to show for those long days at sea.

A COMPANY CALLED TRW

Eventually set theory would become the basis of modern mathematics, with sets—the new fundamental objects—providing mathematicians with a powerful tool for abstraction. At the outset, however, many of the leading mathematicians of the day found the very notion of an infinite set such as the set of all real numbers absurd. Even Cantor admitted that considering infinite sets as single entities was a concept to which he had been "logically forced, almost against my will." This erasing of the distinction between potential and actual infinities was "in opposition to traditions that had become valued." Yet the sets themselves were less disturbing than the uses to which Cantor put them, and mathematicians were particularly shocked by the novel method Cantor devised to prove there are infinitely many of the numbers that are called transcendental, or nonalgebraic.



THEORY OF INFINITE SETS, which was developed by Georg Cantor in the 1870's, installed sets as the fundamental objects in mathematics, a position that had been held first by lengths and then by numbers. Cantor pointed out that certain infinite sets, such as the even numbers (*top row*) and the rational numbers (*bottom row*), are countable in the sense that they can be put into a one-to-one correspondence with the natural numbers 1, 2, 3 ... (*middle row*), but other infinite sets, such as the set of real numbers, are not countable because they are essentially larger than the set of natural numbers. (The rational numbers, or the set of all fractions, have no natural order; they are shown here in order according to the sum of the numerator and the denominator.) Cantor devised a series of "transfinite" numbers to express the cardinality, or number of elements, in these infinite sets and then developed a general arithmetic for infinite as well as finite numbers. Before Cantor most mathematicians had limited themselves to the concept of potential infinity: the notion that for every number there is a larger number. Infinite sets forced them to confront the quite different concept of an actual, or completed, infinity: a collection of infinitely many objects that can be considered simultaneously. Constructive mathematicians believe that numbers, not sets, are the basis of mathematics and that since infinite sets do not seem to exist in nature, mathematicians can have no real, intuitive understanding on which to base judgments concerning them. Hence many of the most familiar and useful techniques of modern mathematics must be altered significantly to be constructively valid.

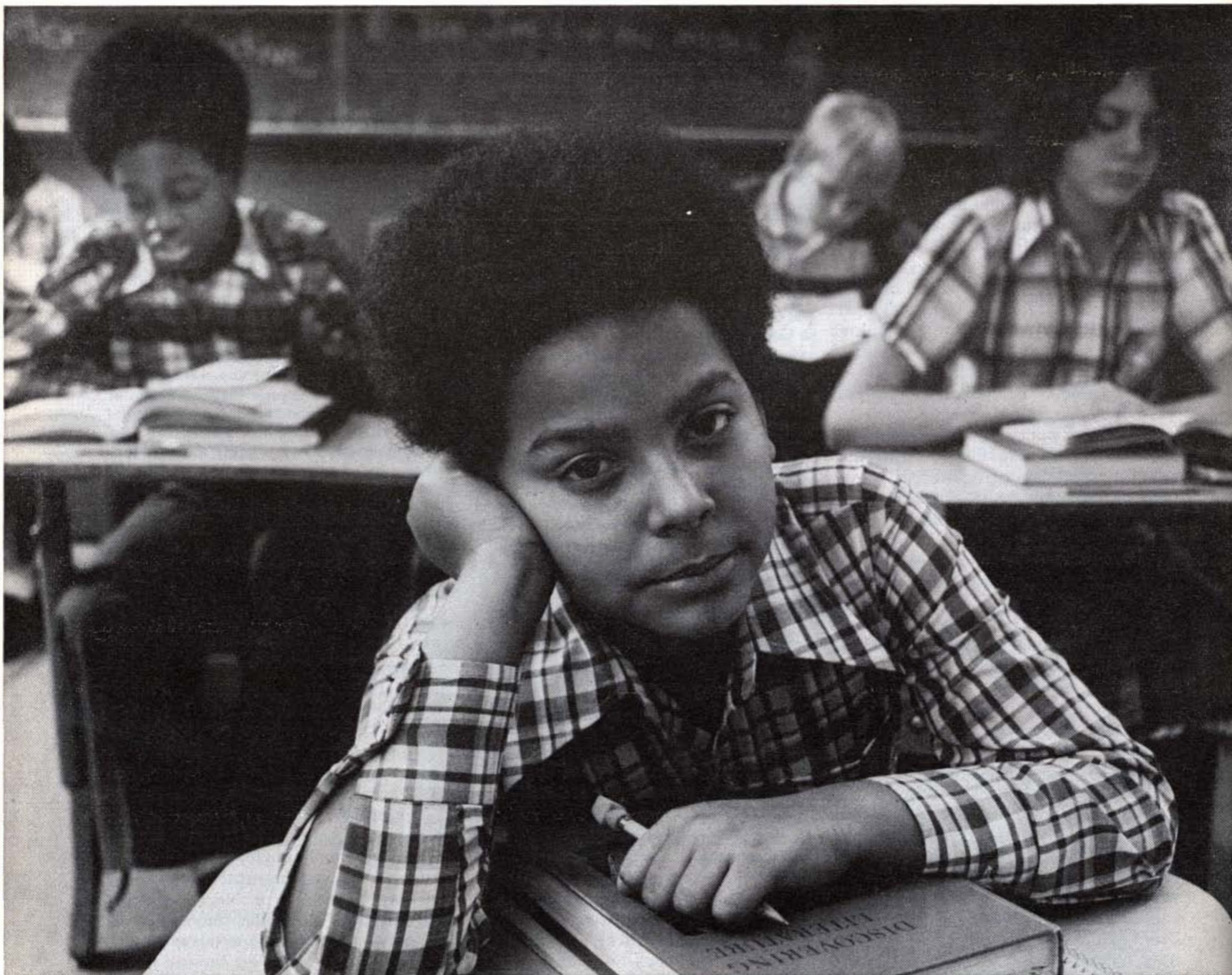
Any real number x is algebraic if it satisfies an equation of the form $a_nx^n + a_{n-1}x^{n-1} + \ldots a_1x + a_0 = 0$, where $a_1, a_2, \ldots a_n$ are integers. (For example, the rational numbers are those real numbers that can be expressed in the form p/q, where p and q are integers and q is not zero; thus any rational number is algebraic because it satisfies the equation $qx - p = 0$ for some values of p and q.) Cantor proved that numbers that do not satisfy this requirement—transcendental numbers—exist, and that there are in fact infinitely many of them, by first showing that the set of algebraic numbers is countably infinite. In other words, he proved that those numbers can be put into a one-to-one correspondence with the natural numbers. As we have seen, the set of real numbers is not countably infinite, and so there must be some "leftover" real numbers that are not algebraic. Now assume that those transcendental numbers are countable: either finite or countably infinite. It is not difficult to show that a set formed by combining two countable sets is itself countable. (This presents no problem if one or both of the sets are finite. If both sets are infinite, then to show that the combined set is countable it suffices to put one of the sets in one-to-one correspondence with the odd numbers and the other in one-to-one correspondence with the even numbers.) Hence if the assumption were true, then the set of real numbers (which is made up of the algebraic numbers and the nonalgebraic numbers) would be countable—a false conclusion. Therefore the assumption that the transcendental numbers are countable is false. They must be uncountably infinite.

Cantor's proof makes use of the classic method of reductio ad absurdum: proving a proposition by showing that its denial leads to a contradiction. To understand the outraged reaction of mathematicians such as Kronecker to the proof it is necessary to understand that according to mathematical tradition there could be no existence without construction. Or to be more precise, in order to establish that a mathematical object existed it was considered necessary to specify a procedure by which the object could, at least in principle, be constructed. This procedure did not have to be implemented, nor did it have to be implementable in any "reasonable" amount of time. The only stipulations were that the procedure must terminate after a finite number of steps and that at the end of each step there must be no question as to how to proceed to the next. Cantor's proof that the set of transcendental numbers was infinite failed rather spectacularly to meet these requirements, however, in that it did not give rise to one example of a transcendental number, let alone an infinity of them. (It is interesting to note that not everyone considered even the traditional definition of existence to be sufficiently stringent. One mathematician reportedly spent 10 years constructing with only a ruler and a compass a regular polygon of 65,537 sides, an object whose existence had been established by Carl Friedrich Gauss.)

Cantor's proof was the first important example of what has come to be called a pure-existence proof. Giving not the slightest hint of how even a single transcendental number might be found, it established the existence of such numbers by proving that it would be contradictory for them not to exist. Once again the basic issue is infinity. A proof by reductio ad absurdum that establishes the existence of an object in a finite set is perfectly acceptable to any mathematician; one can always in principle produce the object by checking through all the members of the set. But the same is not true for, say, the transcendental numbers, which belong to the infinite set of real numbers. For this reason many mathematicians rejected Cantor's proof completely, objecting that a contradiction was no substitute for a tangible example. This controversy was far from being resolved when in 1889 David Hilbert published an even more contentious pure-existence proof.

The proof concerned a problem that had been formulated by the German mathematician Paul Gordon. One of the outstanding mathematical questions of the day, Gordon's problem proposed the existence of certain (finite) sets of objects with special properties in a number of situations. Hilbert showed by reductio ad absurdum that the sets under consideration existed in every case. Cantor had proved the existence of a set of objects for which at least some examples were known; certain transcendental numbers had been discovered as much as 30 years earlier. Now Hilbert was maintaining that he had established the existence of objects no one had ever seen and no one had the vaguest idea how to construct. "This is not mathematics," said Gordon. "It is theology."

In Hilbert's hands, however, the pure-existence proof soon became a powerful tool, and as the application of set theory began to yield important results in mathematics, the violent objections to its methods began to subside. "Theology," Gordon came to concede, "also has its merits." After Kronecker's death in 1891 the opposition to nonconstructive methods faltered, until in 1907 it found a new voice in the Dutch mathematician L. E. J. Brouwer. Mathematics would become meaningless, Brouwer argued, unless a clear distinction was made between real, or constructive, existence and pure existence. These two notions had become identified through



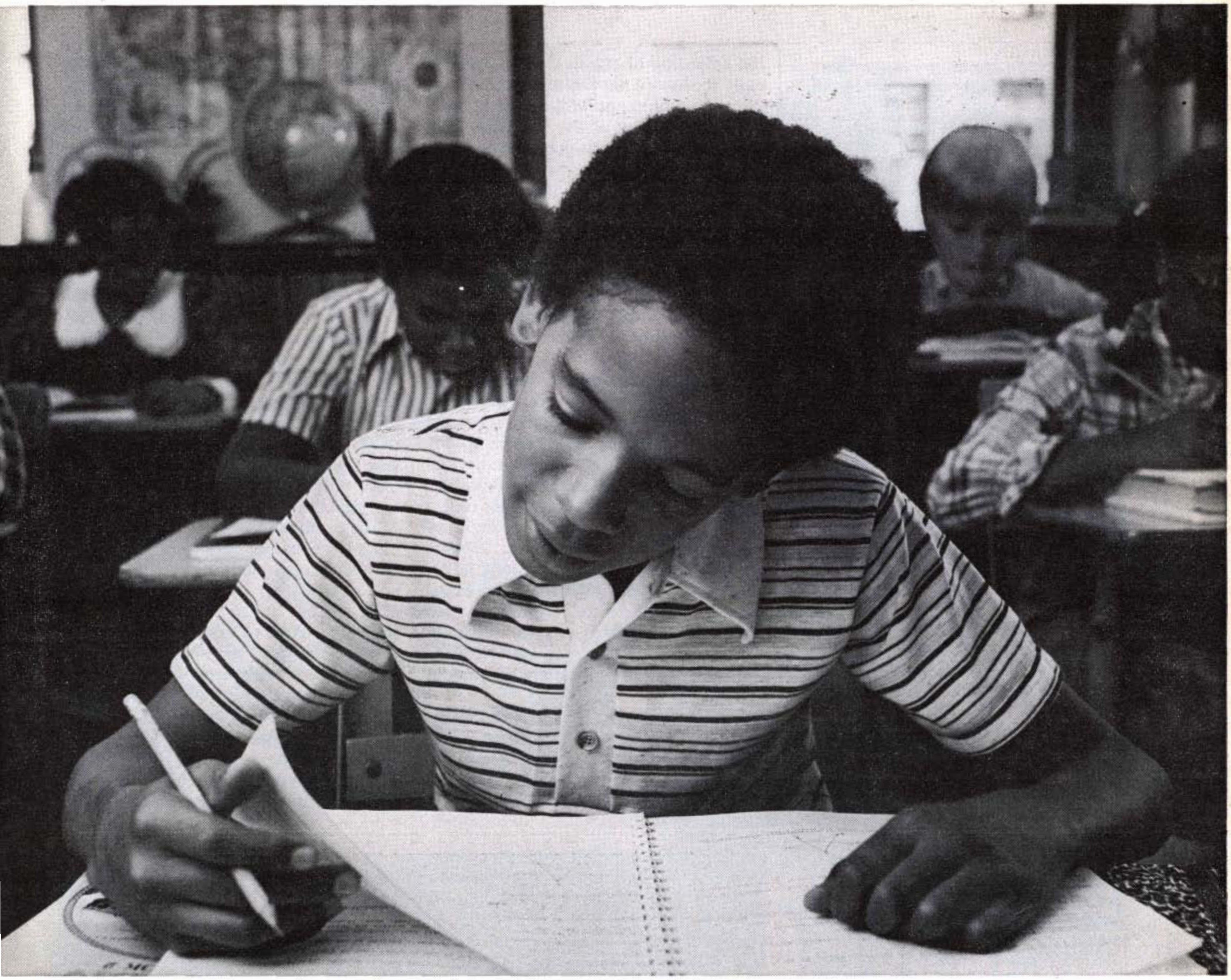
Why Johnny can't read . . .

The failure of so many of our children to learn the basics is so widespread in America today that there are thousands like Johnny who graduate from school without even being able to read.

But Jimmy and many others like him in the New Orleans Elementary Schools are learning how to read and a lot of other skills with the help of Mastery Learning Systems from Westinghouse.

Developed by educators and applied by Westinghouse engineers and systems specialists, the Mastery Learning System employs an effective combination of teaching aids, techniques and strategies – whatever is required to stimulate interest and help students learn the basics.

In addition to testing and measuring step-by-step progress, a computer system helps the teacher manage the class efficiently and inform the students how much they learn each day rather than how much they didn't learn.



and Jimmy can

Most important, Mastery Learning instills a sense of pride when the students realize they can succeed.

The system works. It works so well in their elementary schools that New Orleans is adopting Westinghouse Mastery Learning Centers™ for use in their secondary schools as well.

Mastery Learning Centers.™ Just another way Westinghouse directs its energy into helping America.

Westinghouse

A powerful part of your life

the extension of classical logic—specifically the law of the excluded middle—to apply to infinite, as well as finite, sets.

Aristotle's law of the excluded middle states that every proposition is either true or false, and without it there can be no proof by contradiction. For example, in Cantor's proof that there are transcendental numbers the assumption that all numbers are algebraic is shown to lead to a contradiction; therefore by the law of the excluded middle there must be transcendental, or nonalgebraic, numbers. Aristotelian logic is so much a part of daily living that questioning its validity in any situation may seem silly. Most people feel they know the law of the excluded middle is correct because it is continually reaffirmed by their daily experience. For example, given a bag filled with a number of marbles of different colors, no one would object to being told that either there is a red marble in the bag or there is not. In other words, it is not necessary to look in the bag to verify that one of the two assertions is true. And to determine which one is true it suffices to check through the contents of the bag; after a finite length of time the question will be decided.

The situation is quite different when the law of the excluded middle is applied to an infinite set such as a bag containing an infinity of colored marbles. In this case whether the marbles are examined one by one or 1,000 at a time it will never be possible to examine all of them. Therefore although at some point in the examination it may be possible to state that the bag does contain a red marble, it will never be possible to state that the bag does not contain a red marble. This is the fix implied in the notion of infinity: No matter how much of the set has been examined there will always be more to come. The only way to know that the bag does not contain a red marble is to "see" the infinity of marbles all at once.

Since no actual infinities have been found in nature, mathematicians have no real, intuitive knowledge of them. Hence to apply the law of the excluded middle to infinite sets they must extrapolate from their experience with finite sets. This extrapolation seems particularly difficult to justify in view of the fact that the Greeks who devised the law of the excluded middle were highly suspicious of actual infinities. Indeed, as Brouwer concluded, the results of such an extrapolation from finite experience need not be "constructively" true, that is, it is possible to prove the pure existence of objects that cannot be constructed [*see illustration on page 160*]. Brouwer argued that by according such "ideal" objects the same validity as real, or finitely constructible, objects, mathematics lost its certainty. Ideal objects would be used to create more ideal objects, and soon there would be no way

Proposition: There are infinitely many prime numbers.

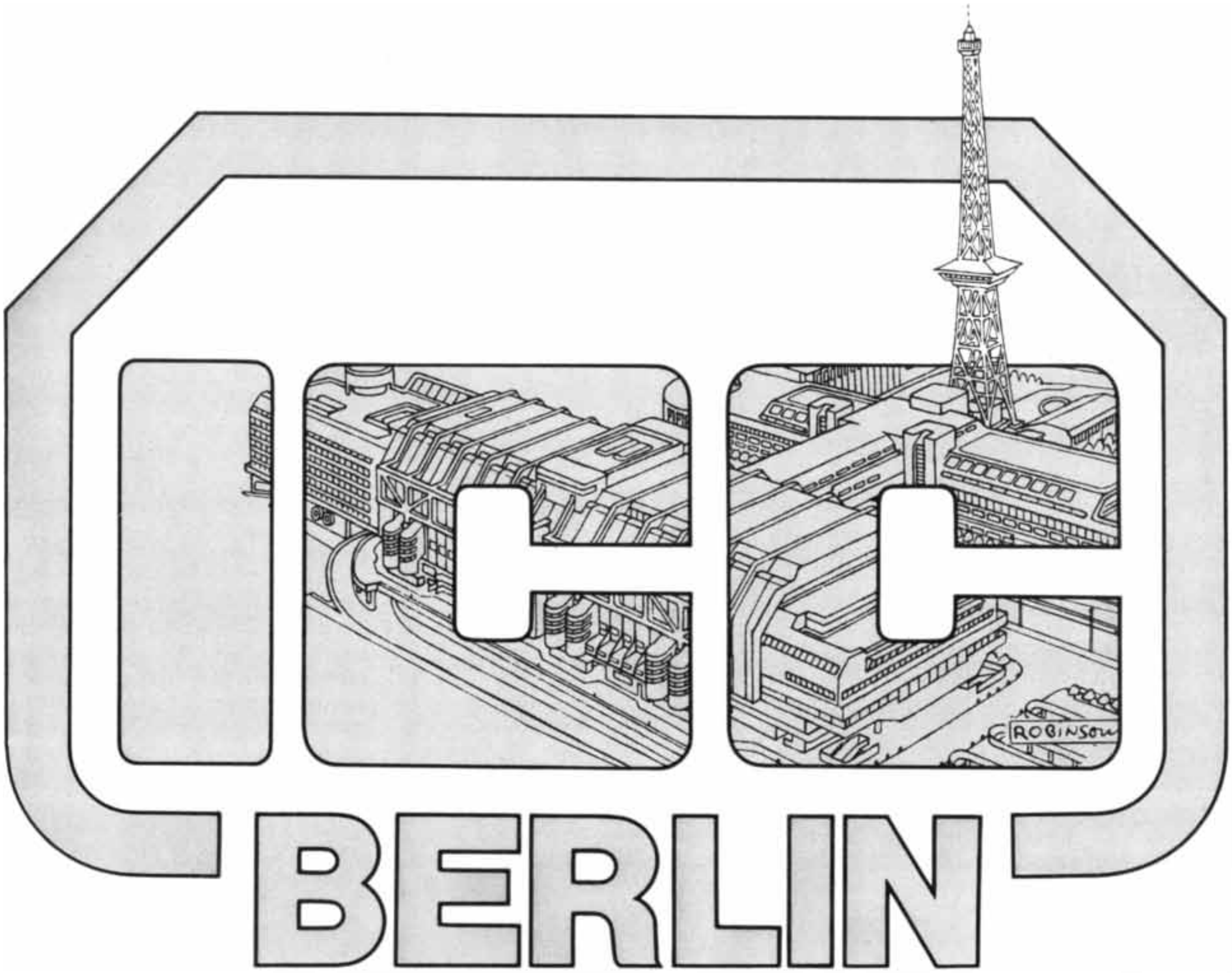
Euclid's Proof

1. Let $p_1, p_2, \ldots p_m$ be any finite set of prime numbers.
2. Consider the number $p_1 \times p_2 \times \ldots p_m + 1$, which cannot be divided evenly by any of the primes $p_1, p_2, \ldots p_m$. This number either is prime itself or it is divisible by some other prime p.
3. In either case there must be a prime that is not part of the set $p_1, p_2, \ldots p_m$. Hence the set of primes is infinite.

Modern Proof

1. Assume that the proposition is false, that is, there are only finitely many prime numbers $p_1, p_2, \ldots p_m$.
2. Consider the number $p_1 \times p_2 \times \ldots p_m + 1$, which cannot be divided evenly by any of the primes $p_1, p_2, \ldots p_m$. This number either is prime itself or it is divisible by some other prime p.
3. In either case there must be a prime that is not part of the set $p_1, p_2, \ldots p_m$. This result contradicts the assumption that the set of primes is finite, and so the assumption must be incorrect. Hence (by the law of the excluded middle) the set of primes is infinite.

CONSTRUCTIVE AND NONCONSTRUCTIVE PROOFS differ in their use of infinity, as can be seen by comparing Euclid's original proof of the proposition that there exist infinitely many prime numbers with the standard modern proof of the same assertion. (A prime number is a positive integer that can be divided only by 1 and itself.) Euclid's constructive proof (*left*) demonstrates that for any finite group of prime numbers there is a larger group; the modern, nonconstructive proof (*right*) demonstrates the existence of an actual infinity of prime numbers. The nonconstructive proof employs the technique of reductio ad absurdum: proving a proposition by showing that its denial leads to a contradiction. Proofs by contradiction rely on Aristotle's law of the excluded middle, which states that any proposition is either true or false. Constructive mathematicians object to the application of this law to infinite sets on the grounds that it may never be possible to decide which alternative is true (*see illustration on page 160*).


ICC
BERLIN
ROBINSON

to know whether the subject had any real meaning.

Brouwer's critique came at a time when set theory was particularly vulnerable. After 1900 the reorganization of mathematics on the basis of sets had resulted in the formulation of a number of unresolved paradoxes. For example, Cantor himself observed that since for every set there is a set of higher cardinality, there must be a set V that has higher cardinality than the set of all sets U, which includes everything; on the other hand, because U does include everything the elements of V must be elements of U, and so the cardinality of U must be equal to or greater than that of V. Paradoxes such as this one indicated that there was some defect in the set-theoretic foundations of mathematics, and it was Brouwer's assertion that the defect was due to the introduction of ideal objects.

Brouwer believed the crisis of the paradoxes would have to be resolved by purging mathematics of ideal objects and arguments, that is, by completely rebuilding mathematics along strictly constructive lines. Because this program, called intuitionism, would have wiped out most of contemporary mathematics, even mathematicians sympathetic to Brouwer's criticisms were unwilling to adopt it. They were far more amenable to a plan devised by Hilbert not for purging mathematics from within but for justifying it from without. In this way mathematicians would not, as Hilbert put it, be driven from the paradise Cantor had created for them.

Hilbert believed that although ideal objects might not have meaning in themselves, meaningful objects and theorems could be derived from them. In his view certainty could be restored to mathematics by considering it as a purely formal system: a game whose interest lies not in the objects with which it deals but in its rules and the relations they create among the objects. Mathematics would be shown to have meaning, Hilbert maintained, when the system was proved to be consistent, or free from contradictions.

Hilbert's program, generally referred to as formalism, called first for the representation of mathematics as a formal axiomatic system. It was hoped that by transforming the statements of mathematics into strings of meaningless symbols to be combined according to the rules of logic, whatever unavowed principles of reasoning had given rise to the paradoxes would be revealed. More important, once mathematics was completely formalized one could undertake to demonstrate that it was consistent. For this purpose Hilbert introduced a new theory called proof theory, or metamathematics, in which meaningful statements about the meaningless signs and configurations of the axiomatic sys-

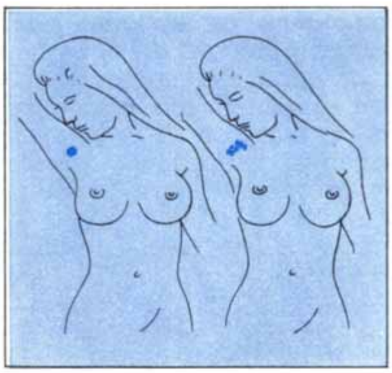
You can wait years for new developments to appear in standard texts, spend hundreds of hours extracting them from journals, or you can subscribe to SCIENTIFIC AMERICAN *Medicine*.

SCIENTIFIC AMERICAN ***Medicine*** is the clinician's answer to the diffuseness of the periodicals and the prepublication obsolescence of the textbook.

Leading scholar-practitioners from Harvard and Stanford have created this 2,000-page innovative union of electronic and publishing technology. The editors are Edward Rubenstein, M.D., F.A.C.P., and Daniel D. Federman, M.D., F.A.C.P.

The display-retrieval element in this sophisticated medical-information system is a loose-leaf bound text. Each month as authors update their contributions, revisions are entered on the magnetic tape on which the text and index are stored. The tape drives high-speed phototypesetting equipment so that subscribers receive about eight new chapters and a new index every four weeks; a bulletin highlights the new developments.

—Adjuvant chemotherapy for breast cancer utilizing CMF enhances survival among premenopausal women with axillary lymph node involvement.



—Posttransfusion hepatitis is usually caused by "non-A, non-B" viruses.

—Lidocaine fails to control ventricular arrhythmias in as many as 20 percent of acute M.I. patients. Procainamide, given by a bolus and continuous infusion method, is effective in most of these cases.

—Poorly ventilated recreational vehicles are a new, significant carbon monoxide poisoning hazard.

—Cerebral embolism, producing both transient ischemic attacks and permanent strokes, is a complication of mitral valve prolapse.

—Spectinomycin is effective for disseminated gonococcal disease caused by penicillinase-producing organisms.

—Adenine arabinoside has been licensed for treatment of herpes simplex encephalitis.

New material replaces old material in the steady-state text, so that the information is there, up-to-date, at your fingertips.

A CME program of eight patient management problems offered over a 12-month period is available at no extra cost. Authorities from first-line institutions create the problems under the direction of Stanford's Office of Postgraduate Medical Education and the Educational Testing Service.

As an organization accredited for continuing medical education, the Stanford University School of Medicine designates this continuing medical education activity as meeting the criteria for 32 credit hours in Category 1 for Educational Materials for the Physician's Recognition Award of the American Medical Association provided it has been completed according to instructions.

This program is approved by the American Academy of Family Physicians for 32 Elective Hours of CME credit.

Trial Offer

Please send us the card and you will receive over a period of 60 days at no cost, the two-volume text and two monthly updates. You may also take a CME test for credit. At the end of 60 days, if you decide to continue the subscription, we will bill you for $150 for the full 12 months; otherwise just return the volumes to us.

We urge you to subscribe now. We anticipate a price increase in January 1980 to $218 for the first year and $165 for the annual renewal as of January 1981. If you subscribe in 1979 for $150, you will be able to renew next year at $150, a saving of $83 over two years.

So please mail the card today and let us take the hassle out of keeping up.

SCIENTIFIC AMERICAN MEDICINE
415 Madison Avenue, New York, N.Y. 10017

Please enroll me as a subscriber to SCIENTIFIC AMERICAN *Medicine*. On receipt of this coupon you will send me the advanced two-volume text described in your announcement and update it regularly by sending me new monthly subsections. I understand that the annual subscription of $150 for SCIENTIFIC AMERICAN *Medicine* is tax-deductible. If I am not entirely satisfied, I may cancel at any time during the first 60 days, returning all materials for a complete refund.

☐ Please enter my subscription for SCIENTIFIC AMERICAN *Medicine*
☐ I shall also enroll in the CME Program
☐ I enclose a check made out to SCIENTIFIC AMERICAN *Medicine* for $150
☐ Please bill me

Please add sales tax for California, Illinois, Massachusetts, Michigan, Ohio and New York

Name ______________________

MD Specialty ______________________

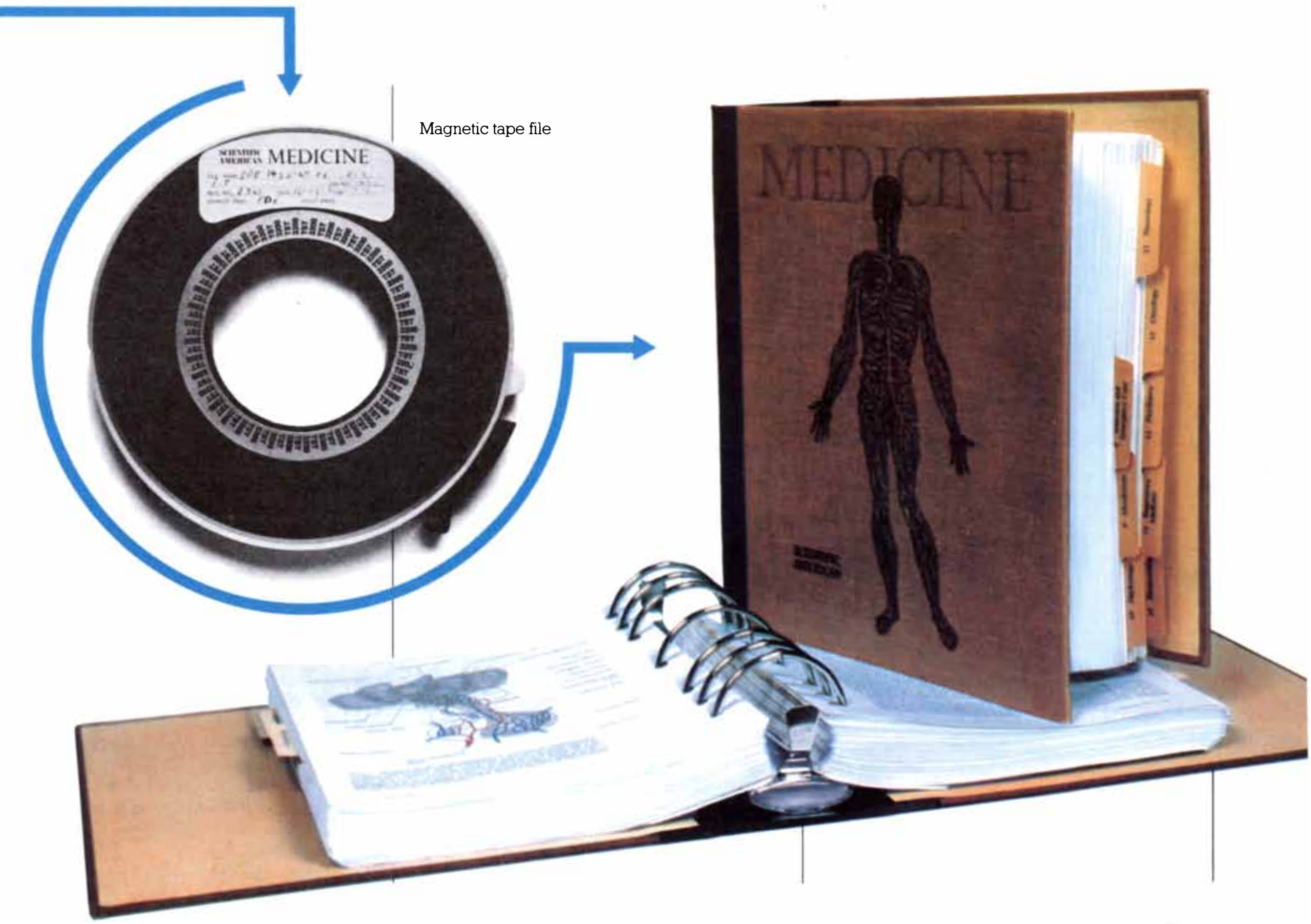
Address ______________________

City ______________________

State ______________ Zip ________

Signature ______________________

Subscribers outside of the U.S. and the possessions will be charged extra for shipping and handling; updates will be sent by surface routes unless air mail delivery is requested. 1D



Magnetic tape file

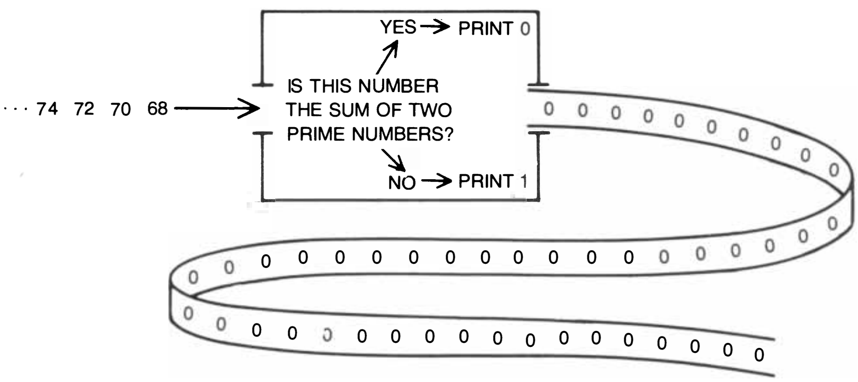
tem could be formulated. The arguments of metamathematics were required to be completely constructive, indeed finitistic, in nature. Hence there would be no questioning of the validity of a metamathematical proof of the consistency of mathematics.

Brouwer did not accept the proposition that proving mathematics to be consistent would restore its meaning, and he spent most of his life trying to show that mathematics could be done constructively. His program never gained much support, however, not least because the mathematical efforts of the intuitionists tended to be awkward and contrived. And when Brouwer attempted to rid mathematics of ideal notions while preserving certain of its long-established aspects, he introduced ideas that seemed to many mathematicians to be just as idealistic as those he was eliminating.

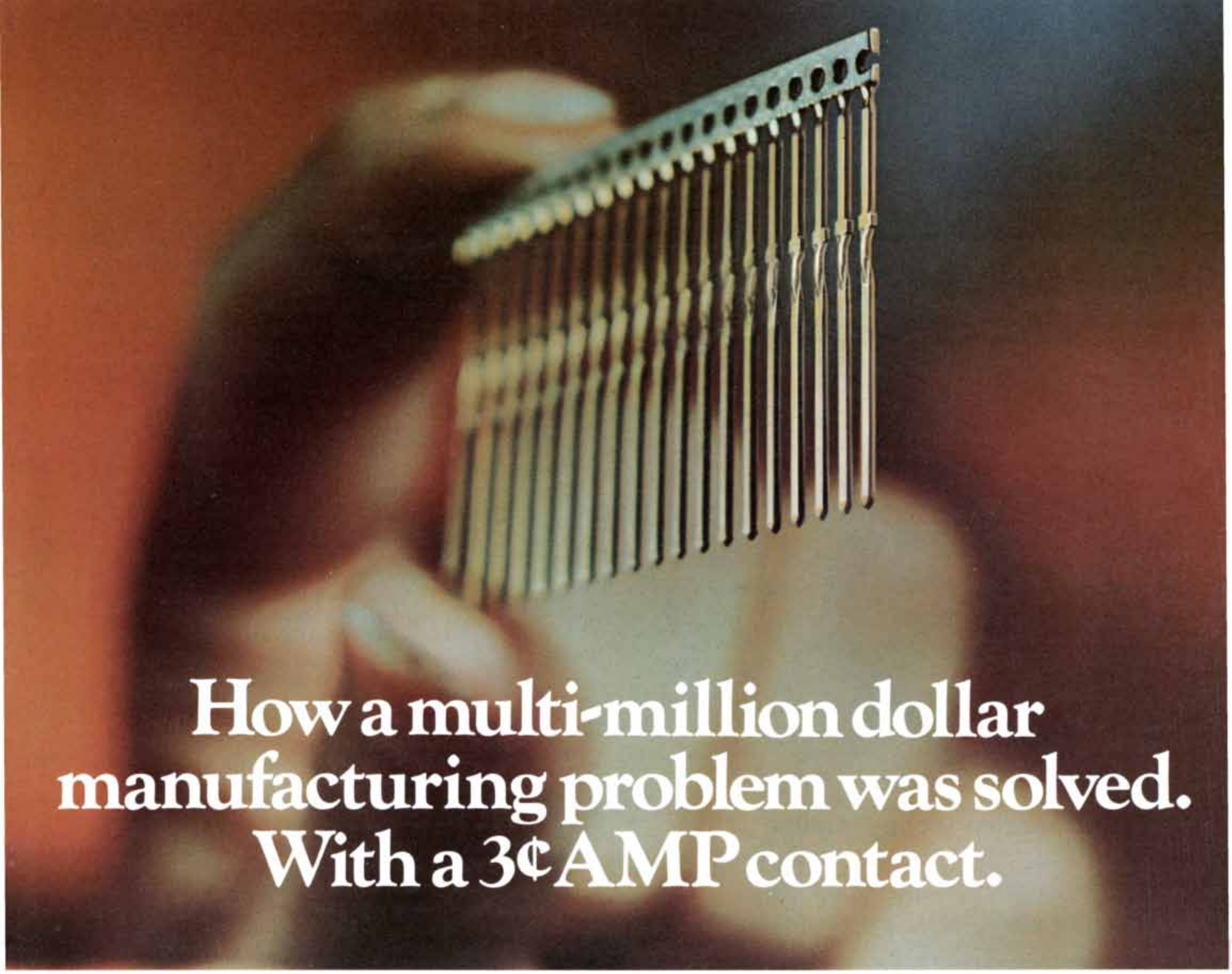
Ironically, it was not a great faith in mathematics that made mathematicians reject intuitionism but rather a lack of such faith. Although many feared their discipline was indeed losing real meaning, it appeared that the only way to keep the subject alive was to accept some ideal notions. As Hilbert put it: "Compared with the immense expanse of modern mathematics, what would the wretched remnant mean, a few isolated results, incomplete and unrelated, that the intuitionists have obtained." To most mathematicians Hilbert's formalism offered an attractive compromise. If mathematics could be shown to contain no contradictions, then at least it would be assured that if two people (or one person on two different occasions) played the game of mathematics, they would not come up with two different correct answers.

Then in 1931 Kurt Gödel showed that Hilbert's formalist scheme was doomed to failure. Gödel showed that it is not possible using metamathematics to prove that any formal axiomatic system rich enough to include basic arithmetic is consistent; the best that can be hoped for is to know that such a system is inconsistent. By this time, however, set theory and formalism were so much a part of mathematics that they could no longer be rejected simply because the fundamental goal of formalism had proved unobtainable.

Over the past half century formalism has become an integral part of mathematical thinking and intuitionism has come to be viewed as a quaint mathematical curiosity. Today most mathematicians choose the clean elegance of pure-existence proofs even when constructive proofs of the same results are available, and the certainty of such mathematics is rarely questioned. A few issues concerning set theory continue to be debated (in particular the use of the axiom of choice, which is discussed below), but for the most part the controversy over the set-theoretic methods of Hilbert and Cantor has subsided. Hence



GOLDBACH'S CONJECTURE asserts that every even integer is the sum of two prime numbers, and although this famous proposition has never been proved, there is an efficient procedure for determining whether any particular even number is the sum of two primes. Hence it would be a simple task to program a computer to check each even number in turn and print out a 0 if the number is the sum of two primes and a 1 if it is not, as is shown here. Because the set of even numbers is infinite, the string of 0's and (perhaps) 1's generated would also be infinite. Is it correct to apply Aristotle's law of the excluded middle to this infinite string and say that it either does include a 1 or does not? The problem is that it may never be possible to decide which of the two alternatives is correct. If a 1 is printed out, one can say conclusively that there is a 1 in the string (and the conjecture is false), but no matter how long the computer runs, as long as no 1 is printed out one can never say that there is no 1 in the string (and the conjecture is true). In constructive mathematics it is possible to assert that the string either does include a 1 or does not only if a finite procedure can be specified for deciding which alternative is true. Constructive mathematicians believe every problem in mathematics can be reduced to a numerical problem such as Goldbach's conjecture and every numerical question can be reduced to a problem of deciding whether or not a 1 appears in a sequence of 0's and 1's. Therefore if it were possible to prove constructively that every sequence of 0's and 1's either does include a 1 or does not, all problems in mathematics would be solved, since proof would have to include a finite procedure for determining which alternative is true. Most mathematicians, however, believe no such procedure exists, so that no all-embracing constructivist proof can be found.



Many companies don't call AMP until they need electrical connectors. But why wait till then? If you involve us early enough, we can help you save money. And not just on connectors.

Here's an example:

Printed circuit boards solved a lot of problems for the electronics industry. And created others.

The biggest problem was connecting the thousands of circuits.

Using solder was slow, expensive and unreliable. So we found a better way—a solderless connector with our unique ACTION PIN contact.

Because 10,000 contacts can be inserted in an hour, production rates increased while costs decreased. And it's proved so reliable it was adopted as a standard in the major telephone system.

Can we produce the same kind of savings for you?

If your company manufactures electrical/electronic products, the answer is probably yes. But it's important for you to call us in early.

That's when our experience in developing connectors, switches and application equipment for hundreds of industries worldwide can help you the most.

Early Involvement.

At AMP, we call this approach to solving our customers' problems "Early Involvement." It's our better way. And it's what makes a 3¢ AMP contact worth so much more than 3¢.

Ask for a copy of our brochure "AMP Has A Better Way." Call (717) 564-0100, Ext. 8420.

Or write to AMP Incorporated, Harrisburg, PA 17105.

AMP

AMP has a better way.

Barcelona • Brussels • Buenos Aires • Frankfurt • Harrisburg • Helsinki • s-Hertogenbosch • London • Luzern • Mexico City • Montreal
Paris • San Juan • Sao Paulo • Stockholm • Sydney • Turin • Toronto • Tokyo

it may seem all the more remarkable that a revival of constructive mathematics is now under way.

The new interest dates from the publication in 1967 of Errett Bishop's book *The Foundations of Constructive Mathematics.* Although the efforts of Bishop's more polemical predecessors generally seemed clumsy and contrived, Bishop's work demonstrates that the results obtained by constructive methods can be just as beautiful and useful as those obtained by formalist methods. Brouwer wanted to build certainty into mathematics at its foundations, starting with an intuitive system of axioms for arithmetic and working up, but Bishop's plan for constructivizing mathematics has a different orientation. He is concerned not with founding mathematics but with doing it constructively, and so he begins with the standard rules of arithmetic and the rational numbers and tries to proceed "realistically" from there. In Bishop's view certainty can be achieved by avoiding idealistic notions and continually checking the objects and theorems generated for intuitive meaning. His results show that a great deal of valuable mathematics can be effectively and elegantly developed in this straightforward manner.

For example, rather than rejecting Cantorian set theory, as Brouwer did, Bishop modifies the theory to give it constructive validity. In particular he insists that to define a set it is not sufficient to simply give a rule by which membership in the set can be decided. He requires that procedures also be specified for actually constructing a member of the set and for demonstrating that two members of the set are distinct. An interesting result of these modifications is that they make one of the most disputed postulates of Cantorian set theory—the axiom of choice—appear to be completely acceptable.

Cantor's axiom of choice states that for any collection of nonempty sets (sets with at least one member) it is always possible to choose a representative from each set. In other words, it is possible to give a rule for selection such that if two people go through the collection following the rule, they will choose the same representatives. To understand why this axiom, which is employed throughout mathematics, is controversial consider the following example proposed by Bertrand Russell.

For any infinite collection of pairs of shoes the rule "Choose the left shoe from each pair" satisfies the axiom of choice, but what rule will serve with an infinite collection of pairs of socks? If there were only a finite number of pairs of socks, it would be possible to go through the collection and mark one sock in each pair so that the rule "Choose the marked sock" would suffice. Since there are infinitely many pairs, however, there is no way to mark

"We weren't seriously thinking of replacing our stereo. Then we heard Yamaha."

"There was a noticeable difference in sound quality. Enough of a difference to convince us to own Yamaha now."

For almost a century, the Yamaha name has graced the best musical instruments in the world. It is also synonymous with audio components that set unmatched industry standards.

Take the components shown—the CR-640 receiver, the YP-B2 turntable, and the NS-244 speakers. Each is built with the most advanced technology and highest quality craftsmanship. Their visual beauty is rich and warm. Their sound is distinctly natural and truly adds to the enjoyment you expect from listening to music. And the price is surprisingly affordable.

If you already have a stereo, and it's not Yamaha, you owe it to yourself to give Yamaha systems and components a listen. Now is the time to own Yamaha.

Plan to visit your Yamaha Audio Specialty Dealer this weekend and discover the difference for yourself. He's listed in the Yellow Pages.

YAMAHA

Yamaha International Corporation,
P.O. Box 6600, Buena Park, CA 90622.
Also available in Canada.

Free for the listening!

The Session II album, valued at more than $15.

Hear today's hot studio musicians play the sounds of tomorrow on this lavishly produced album. You can't buy the Session II record, but you can get it free with a demonstration of this Yamaha CR-640 component system at your Yamaha dealer.

Recorded with state-of-the-art techniques for astonishingly live ambience, Session II is destined to become a Collector's Item. Quantities are limited, so don't delay! See the Yellow Pages for the name of your Yamaha Audio Specialty Dealer. Offer good thru Dec. 15 at participating dealers. Offer limited to one per household.

HOME STEREO II ENERGY EFFICIENT ENTERTAINMENT

YAMAHA
Nippon Gakki Co., Ltd., Japan

a sock in every pair ahead of time, and so there appears to be no such rule for selection. Moreover, the axiom of choice can also present problems when it is applied to a collection consisting of a finite number of sets or even of a single set. For example, Fermat's last theorem is a famous unproved proposition in number theory. A set can be defined to consist of 0 and 1 if Fermat's theorem is true or 2 and 3 if the theorem is false. In this case the rule "Choose the smallest number in the set" specifies a unique representative, but until the theorem is proved there is no way to identify an actual number representing the set.

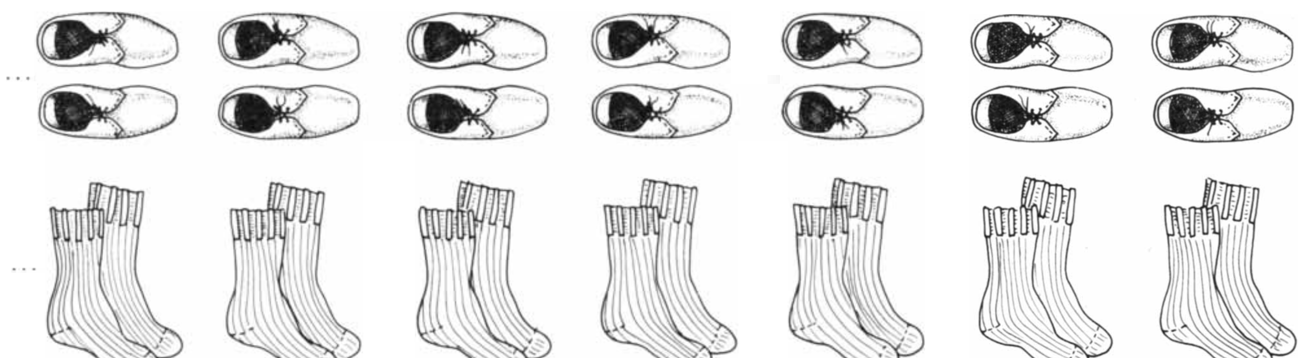
In Bishop's constructive revision of set theory, however, the axiom of choice appears to be noncontentious, because it is stipulated that to call a set nonempty one must show that the construction method specified in the definition of the set actually generates a member of the set. In constructive theory, then, each nonempty set in a collection can be assumed to have a distinguished member, namely the one that had to be constructed to prove the set nonempty. The axiom of choice can always be satisfied by choosing the distinguished member from each set in the collection.

To understand the ramifications of Bishop's constructive set theory consider the set of real numbers. In order to define this set constructively it is necessary to specify a method for constructing the real numbers in terms of the intuitively valid rational numbers. In other words, it is necessary to construct the irrational numbers such as $\sqrt{2}$ and e from the rationals. One might assume that such a number could be considered to be constructed if there were a procedure for computing its decimal expansion to any desired decimal place. This definition of construction is not sufficient, however, because the set of numbers obtained is not closed under the basic operations of arithmetic, as the real numbers must be. In other words, with this definition the sum of two real numbers is not necessarily a real number, the product of two real numbers is not necessarily a real number, and so on. The following example shows how it is possible to construct two numbers a and b in this manner and yet not be able to construct their sum.

It is not known whether or not the sequence 0123456789 appears in the decimal expansion of the irrational number π. Moreover, if the sequence does appear, say, beginning with the kth decimal place of π, then it is not known whether the "critical number" k is even or odd. There is a purely mechanical procedure, however, for determining whether a particular number n is indeed the critical number of π: one simply computes π to the $n + 9$th decimal place. As a result one can define a to be the number whose decimal expansion begins .333... and continues the same way unless some odd number n turns out to be the critical number of π; in that case 4 is placed in the nth decimal place and all decimal places to the right. Similarly, b can be defined as the number that begins .666... and continues the same way unless some even number n turns out to be the critical number of π; in that case 5 is placed in the nth decimal place and all decimal places to the right.

Note that if π has no critical number, then a equals 1/3, b equals 2/3 and $a + b$ equals 1. On the other hand, if the critical number of π exists and is even, then a is greater than 1/3, b equals 2/3 and $a + b$ is greater than 1, whereas if the critical number of π exists and is odd, then a equals 1/3, b is less than 2/3 and $a + b$ is less than 1. Now try to begin writing down the decimal expansion of $a + b$. If the sum begins with 1, then it cannot be less than 1 and the critical number of π, if it exists, must be even; if the sum begins with .9, then it cannot be greater than 1 and the critical number of π, if it exists, must be odd. Since no one has any idea how to show that the critical number of π, if it exists, is either odd or even, it is not possible to write down even the first digit of the decimal expansion of $a + b$. As this example demonstrates, defining numbers by specifying how to compute their decimal expansion is too naive an idea, by nonconstructive standards as well as constructive ones.

How, then, can the set of real numbers be constructed? A real number is determined (in constructive or nonconstructive terms) when it can be approximated arbitrarily closely by rational numbers. For instance, the decimal expansion of a number x to, say, n decimal places differs from the number by less than $1/10^n$. (Consider the familiar decimal expansion of π, which begins 3.141592...; taking the expansion to the fourth decimal place one obtains 3.1415, a number that differs from π by .00009..., which is less than .0001, or $1/10^4$.) By taking a large enough value of n one can always obtain in this way a rational number that is arbitrarily close to x. As we have seen, using decimal expansions to represent real numbers is not always a rich enough notion. There are, however, many other methods for obtaining rational approximations of real numbers. Considering a real number as being defined once it has been approximated by any such method (not just decimal expansions) provides a perfectly adequate system. In constructive theory this means that a real number is



AXIOM OF CHOICE, one of the most widely used and widely disputed tenets of modern set theory, states that for every collection of nonempty sets it is possible to choose a representative from each set. In other words, given a collection such as one of those shown here it should be possible to specify a rule for choosing representatives so that two people going through the collection will make the same choices. For instance, for the infinite collection of pairs of shoes shown at the top the rule "Choose the right shoe from each pair" will suffice. There seems to be no such rule, however, that will serve for the infinite collection of pairs of socks shown at the bottom. This example, due to Bertrand Russell, involves infinite collections of sets, but similar problems can arise with finite collections. In Bishop's version of set theory these problems appear to be eliminated because to define a set constructively it is necessary to give not only a rule by which membership in the set can be decided, as is required in traditional set theory, but also a (finite) procedure for constructing a member of the set. To show that a set is nonempty, then, it is necessary to show that the method of construction actually gives rise to a member of the set. As a result it is always possible to assume that each set in a collection of nonempty sets contains a distinguished member, namely the one that had to be constructed to show that the set was nonempty. In this way the axiom of choice can always be satisfied.

considered to be constructed when a (finite) procedure has been specified for producing a rational approximation to within any given degree. In other words, it must be possible to obtain in a finite number of steps a rational number that differs from the real number in question by less than that degree. For example, to approximate $a + b$, the number whose decimal expansion cannot be written down, to within $1/10^n$ one simply computes π to $n + 9$ decimal places. If this computation shows that the critical number of π is less than or equal to $n + 1$, $a + b$ can be computed exactly to $n + 1$ decimal places. Otherwise a will look like $.333\ldots$ and b will look like $.666\ldots$ up to at least the $n + 1$th decimal place, and so the rational number 1 will differ from $a + b$ by less than $1/10^n$.

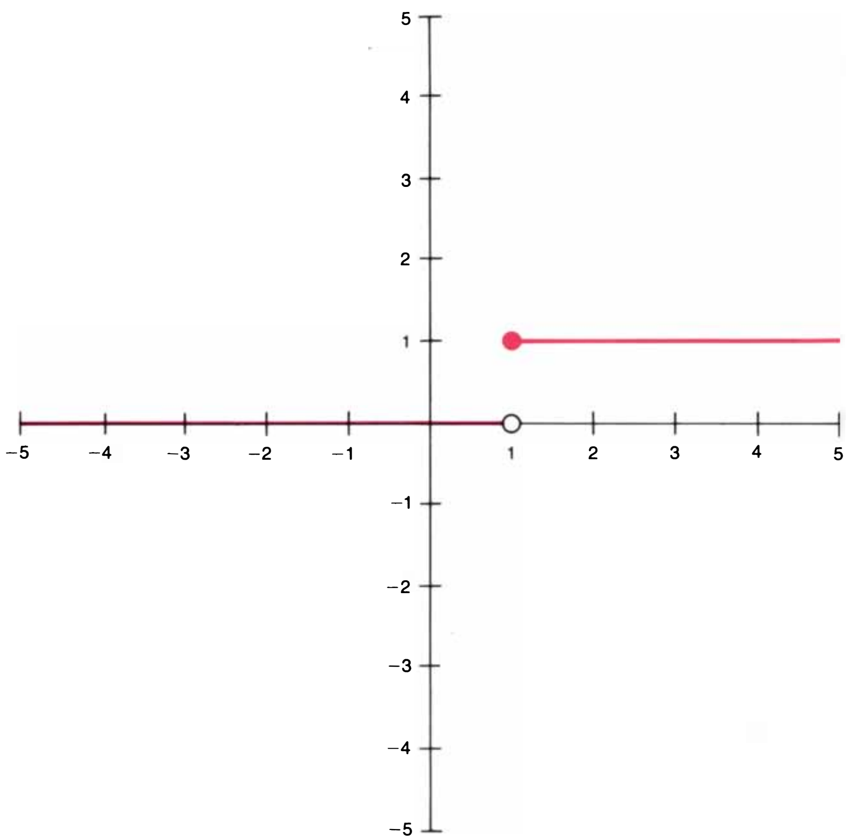
More generally the numbers constructed by Bishop's definition obey all the rules of standard arithmetic, that is, for any numbers x and y that can be constructed by this definition it will be possible to construct their sum $x + y$, their difference $x - y$ and so on. To see that $x + y$ can be approximated arbitrarily closely, or to within $1/10^n$, note that since x and y can be constructed, there must be numbers r and s that differ respectively from x and y by less than $1/10^{n+1}$. Therefore the sum $r + s$ differs from $x + y$ by less than $2/10^{n+1}$, which is less than $1/10^n$, and so $r + s$ is a suitable approximation.

Now that a proper method for constructing the real numbers has been established, only a procedure for deciding when two real numbers are distinct is needed to complete the definition of the set of real numbers. That can be done quite simply, by calling two real numbers x and y different if it is possible to construct a rational number r that lies between them. In other words, either x will be greater than r and r will be greater than y or x will be less than r and r will be less than y.

A curious consequence of Bishop's constructive notion of a set is that it is not possible to consider the set of real numbers as being made up of two nonempty sets with no members in common, a formulation that is frequently used in contemporary mathematics. For example, in nonconstructive set theory the set G consisting of all real numbers less than 1 and the set D consisting of all real numbers greater than or equal to 1 can be considered to be the whole of the set of real numbers. The same is not true constructively, however, because it is not possible to assign certain real numbers such as the number $a + b$, described above, to one set or the other. And one cannot say $a + b$ belongs either to G or to D and therefore G and D make up the real numbers, because that requires applying the law of the excluded middle to the infinite set of real numbers. (There is always the chance that it may someday be possible to place a number such as $a + b$ that is defined with respect to an unsolved problem in arithmetic in one of the sets G and D. It seems safe to assume, however, that there will always be unsolved problems in arithmetic, so that it will always be possible to define similar numbers.)

This restriction on the makeup of the set of real numbers is reflected in the types of real-number functions that can be defined constructively. In nonconstructive set theory a function f from set S to set T is a rule assigning to each member s of S a member t of T, that is, $f(s)$ equals t. For example, the rule that assigns to each person that person's mother can be thought of as a function from the set of all people to the set of all women. As this example indicates, each member of S must be assigned to one and only one member of T but not all members of T must be assigned to some member of S. (In other words, every person has a mother but not every woman is necessarily a mother.) In constructive set theory a function from set S to set T is still a rule that assigns to each member of S a member of T, but it is also required that the rule be effectively computable, that is, for any particular member s of S it must be possible to construct the member $f(s)$, or t, to which it is assigned. Now consider the function from the set of real numbers to itself specified by the following rule: $f(r)$ equals 0 if r is less than 1, and $f(r)$ is 1 if r is greater than or equal to 1. Functions of this type—called step functions because of the appearance of their graph—are widely used in contemporary mathematics [*see illustration below*]. They are not well defined constructively, however, because it is not possible to decide whether numbers such as $a + b$ are in the set of numbers less than 1 (G) or in the set of numbers greater than or equal to 1 (D), and so it is not possible to decide whether 0 or 1 should be assigned to them. On the other hand, the function considered as being defined not from the set of all real numbers to itself but from the set consisting of G and D to the set of real numbers is perfectly acceptable.

The nonconstructive step function defined on all the real numbers is one



STEP FUNCTION (*color lines*) is a widely used function on the real numbers that cannot be defined constructively. In nonconstructive mathematics a function on the real numbers is a rule that assigns to each real number r another real number $f(r)$. For example, the rule that defines the step function is: If r is less than 1, then $f(r)$ equals 0, and if r is greater than or equal to 1, then $f(r)$ equals 1. In some cases, however, it is possible to define a real number and yet not be able to determine in a finite number of steps whether it is less than 1 or greater than or equal to 1. As a result there is no way to determine whether a 1 or a 0 should be assigned to the number. In Bishop's constructive mathematics it is required that the rule defining a function on real numbers be effectively computable. In other words, for each r it must be possible to construct $f(r)$ in finite number of steps. Hence the step function cannot be defined constructively on real numbers.

Getting more out of coal takes skills that go beyond coal.

Coal represents 85 percent of America's energy reserves, but only 18 percent of our energy consumption. So, as the nation becomes dangerously dependent on foreign oil, it is a major challenge to overcome the barriers that block the greater use of coal.

As an oil company that is also a leading coal producer, Conoco is helping meet this challenge—often by applying oil technology to coal.

For instance, because America runs primarily on petroleum, we're working on ways to change coal into synthetic gas and oil.

And we've developed ways to reduce pollution when coal is burned.

At a time when some people in government are trying to limit oil company participation in the coal business, Conoco's efforts to reduce dependence on foreign oil demonstrate a very important point:

America will get more *out* of coal if any company that wants to can put its skills *into* coal.

CONOCO

Doing more with energy.

Conoco, Stamford, CT 06904. © 1979

of the simplest examples of a discontinuous function. Intuitively speaking, a function from the set of real numbers to itself is continuous if the value of $f(r)$ varies continuously, or without jumps. (Heuristically, the graph of such a function can be drawn without lifting a pencil from a piece of paper.) As r varies through 1, the value of the step function $f(r)$ jumps from 0 to 1, and so the function is discontinuous. Since the real numbers are defined in terms of the rational numbers close to them, the value a constructively defined function assigns to a given real number will depend on the values it assigns to nearby rationals. Hence it appears unlikely that it will be possible to constructively define a function on the whole of the real numbers that is not continuous. (Although this observation, which was true of Brouwer's construction of the real numbers, has not been proved for Bishop's, no constructively defined discontinuous function has ever been found.)

Discontinuous functions arise constantly in mathematics, particularly in its applied branches, and so the prospect that only continuous functions can be defined in constructive theory is not pleasing. On the other hand, even in nonconstructive theory it is not possible to apply a discontinuous function to all real numbers. One cannot determine the value of, say, the step function for a given real number until it has been determined which of the sets G or D the number is in. Hence in practical terms this characteristic of constructive functions is less troubling than it might first appear to be.

Because the requirements of constructive mathematics are more stringent than those of nonconstructive mathematics every theorem that is true constructively is also true nonconstructively. (It is possible, however, that two theorems that are distinct constructively might have the same nonconstructive interpretation.) Therefore constructive mathematics could be considered to be simply a branch of modern mathematics. Alternatively, every nonconstructive theorem can be made constructive by adding "assuming the law of the excluded middle" to its statement. Or nonconstructive proofs can be reinterpreted constructively. For example, consider the constructive meaning of a nonconstructive proof that there exist irrational numbers a and b such that the number a^b is rational.

The proposition is proved nonconstructively as follows. The number $\sqrt{2}$ is irrational, and the number $\sqrt{2}^{\sqrt{2}}$ is (by the law of the excluded middle) either rational or irrational. If $\sqrt{2}^{\sqrt{2}}$ is rational, then obviously the proposition is proved by the example in which a equals $\sqrt{2}$ and b equals $\sqrt{2}$. On the other hand, if $\sqrt{2}^{\sqrt{2}}$ is irrational, then the proposition is proved by the example in which a equals $\sqrt{2}^{\sqrt{2}}$ and b equals $\sqrt{2}$, because

$(\sqrt{2}^{\sqrt{2}})^{\sqrt{2}}$ equals $\sqrt{2}^{\sqrt{2}\sqrt{2}}$, which equals $\sqrt{2}^2$, or 2. (This is a particularly interesting proof because unlike many pure-existence proofs it appears to come close to generating an example. To actually obtain the example, however, it would be necessary to know whether $\sqrt{2}^{\sqrt{2}}$ is indeed rational or irrational.)

What does the proof show constructively? To begin with, it indicates that trying to prove the proposition that for any irrational numbers a and b, a^b is irrational is a waste of time. One can show constructively that this task is hopeless by assuming that the proposition is true. In that case $\sqrt{2}^{\sqrt{2}}$ must be irrational, and by the same reasoning $(\sqrt{2}^{\sqrt{2}})^{\sqrt{2}}$, or 2, must be irrational, an obvious contradiction. Hence the proof shows constructively that it is not the case that for all irrational numbers a and b, a^b is irrational.

Bishop predicted that once the advantages of his program became clear, modern mathematics as it is currently practiced would cease to exist as an independent field, becoming instead a branch of constructive mathematics. *The Foundations of Constructive Analysis* is now out of print, however, and in the 12 years since its publication there has been little sign that such a transformation is taking place. Although the number of mathematicians practicing constructive mathematics has been slowly increasing, for the most part mathematicians have seen little reason to change the fundamental ways in which they work. In spite of the unresolved problems at its foundations modern mathematics is extremely successful, and it is not even clear that the constructivists' charge that the subject lacks real meaning is valid. After all, the utility of mathematics in the real world is undeniable. To many applied mathematicians this issue that Albert Einstein called a "frog-and-mouse battle among the mathematicians" seems totally irrelevant. Moreover, the major problems that determine the development of mathematics are usually highly intuitive, so that the subject never drifts too far from "reality." The widely held belief that abstract mathematics is more than a meaningless game would appear to be well founded.

Even if there is never to be a constructive revolution, however, the future for constructive mathematics as a vital mathematical discipline seems secure. The current surge in the production of constructive results has shown that in some cases there are substantial advantages to tackling problems from this point of view. Not only can new and sometimes deeper insights be gained but also the resulting theorems are generally more useful. For instance, when the existence of a certain number has been proved constructively, that number can be computed, often in practice as well as in theory. Bishop has pointed out that modern probability theory includes few general methods for actually computing probabilities, and so now a constructive development of the subject is being carried out, most notably by Bishop's former student Y. K. Chan. Constructive mathematics could surely be profitably employed in many other areas of applied mathematics where it is ultimately more important to provide solutions than simply to show they exist.

The majority of mathematicians are now the product of several generations of formalist training, and that places them in a mental "trap" from which it is difficult to view mathematics objectively. From within this trap mathematical certainty can be called into question only by the discovery of unavoidable inconsistencies inherent in the formal axiomatic system of mathematics. And it is generally believed that although such inconsistencies may exist, they can be corrected by making minor modifications of axioms. Hence the question of whether abstract mathematics has real meaning cannot even be formulated. Mathematicians who find themselves doubting the validity of their subject are generally expected to follow the advice the 18th-century French philosopher Jean d'Alembert gave to those who questioned the calculus: "Go on and faith will return."

The formalist point of view, which allows criticisms of mathematics only in its own terms, has led to excesses on both sides of the constructivist controversy. In particular the unfounded belief that constructivism is a cancer that would destroy mathematics has prevented the systematic treatment of tangible existence. It may be impossible to convert every pure-existence proof into a constructive proof, but often it may be feasible to start from a point where tangible existence has already been established and build from there. On the other hand, it is misleading for constructivists to categorically reject pure existence, which does imply the constructively valid notion that it would be contradictory for something not to exist.

It is generally supposed that mathematical ability comes from the human mind's capacity to blend intuition and reason. The vitality of mathematics springs mainly from intuition, however, which formalism tends to devalue. Indeed, in the finished product of mathematics it is not uncommon to find that all signs of intuition have been removed. A proper consideration of constructivism could restore a more appropriate balance between intuition and reason. Moreover, to believe that mathematical truth exists independently of man's knowledge of it is an act of faith that most mathematicians enter into unconsciously. The constructivists are beginning to show that much of mathematics can be done without such an assumption.

THE AMATEUR SCIENTIST

The mysterious "rattleback": a stone that spins in one direction and then reverses

by Jearl Walker

If you spin a certain type of stone in the "wrong" direction, it will quickly stop, rattle up and down for a few seconds and then spin in the opposite direction. Going in the "right" direction, however, it will usually spin stably. The stone is apparently biased toward one direction of spin. It will even develop a spin in that direction if you just tap one end downward. The rocking of the stone caused by the tap is quickly converted into a spin.

These curious stones were originally called celts because their behavior was discovered by archaeologists studying the prehistoric axes and adzes so named. I imagine the spin reversal was first noticed by an archaeologist idly spinning one of the celts on a table. My introduction to the stones came some years ago when I read Harold Crabtree's delightful book *An Elementary Treatment of the Spinning Tops and Gyroscopes.* Among the many spinning tops, hoops and other toys he described were the mysterious celts, but he did not indicate why they exhibited a bias for one direction of spin. Recently two people have rekindled my interest in the stones. Nicholas A. Wheeler of Reed College has been considering doing an analytical investigation of them. (If he does one, it will be to my knowledge the first study of the stones since the 19th century.) A. D. Moore of the University of Michigan has sent me several of the hundreds of celts he has made out of dental stone, spoons and other materials. Moore has dubbed his stones rattlebacks, a term I shall employ here.

Most rattlebacks have a smooth, ellipsoidal bottom. The top can be flat, hollowed out or ellipsoidal. The spin bias apparently results from the shape of the bottom and the distribution of the mass with respect to the axis of spin. The basic rattleback in Moore's collection has a flat rectangular top. The key feature of the design is that the long axis of the ellipsoid is aligned at an angle of from five to 10 degrees to the long axis of the rectangular top. This skewing (and perhaps some subtle shaping of the ellipsoid) introduces the bias to the spin of the stone but in a way that may be difficult to see in detail.

Most of the rattlebacks Moore has sent me spin clockwise as seen from above. When I spin such a rattleback clockwise, it rotates rather smoothly until friction eventually stops it. When I spin it counterclockwise, it first turns through a few revolutions but then stops as its ends begin to rock up and down. Soon it starts to spin clockwise until friction halts it. If I tap the end of the stone when it is stationary, the ends chatter up and down for a few seconds and then the stone begins to turn clockwise. Again it is friction with the table that eventually stops the stone.

Some of the rattlebacks display a second reversal, from clockwise to counterclockwise, near the end of the spinning just as the friction is about to remove the last of the stone's energy. The second reversal is usually not as strong as the first and does not display the same type of vertical oscillation. Instead the stone rocks from side to side rather than longitudinally. A few of the rattlebacks display strong reversals with either direction of spin.

Moore forms most of his rattlebacks with dental stone, which he polishes and varnishes. Once he has made a good one (a feat that requires both experimentation and luck) he makes a mold of it for duplication. Some of his rattlebacks are made from the bowl of a spoon that he has detached from its handle. The bowl is glued to a flat rectangular piece of metal or Plexiglas with the long axis of the bowl's ellipsoid at an angle to the long axis of the rectangle. Pennies are glued on top of some of the rectangular pieces to add weight and to change the rattleback's moment of inertia.

One type of rattleback looks like half an egg. Across the flat top Moore tapes a brass rod at an angle to the long axis. Another type is similar to the basic design except that the inside is hollowed out to resemble a boat. Still another type has central terraces built on top of a rattleback of the basic design.

Moore's favorite rattleback must certainly be the half egg. He once engaged in a contest with C. L. Stong, who conducted this department for almost two decades, to determine which of them could get the most back turns out of a rattleback. Moore won with the half egg, which made more than 15 back turns before it stopped.

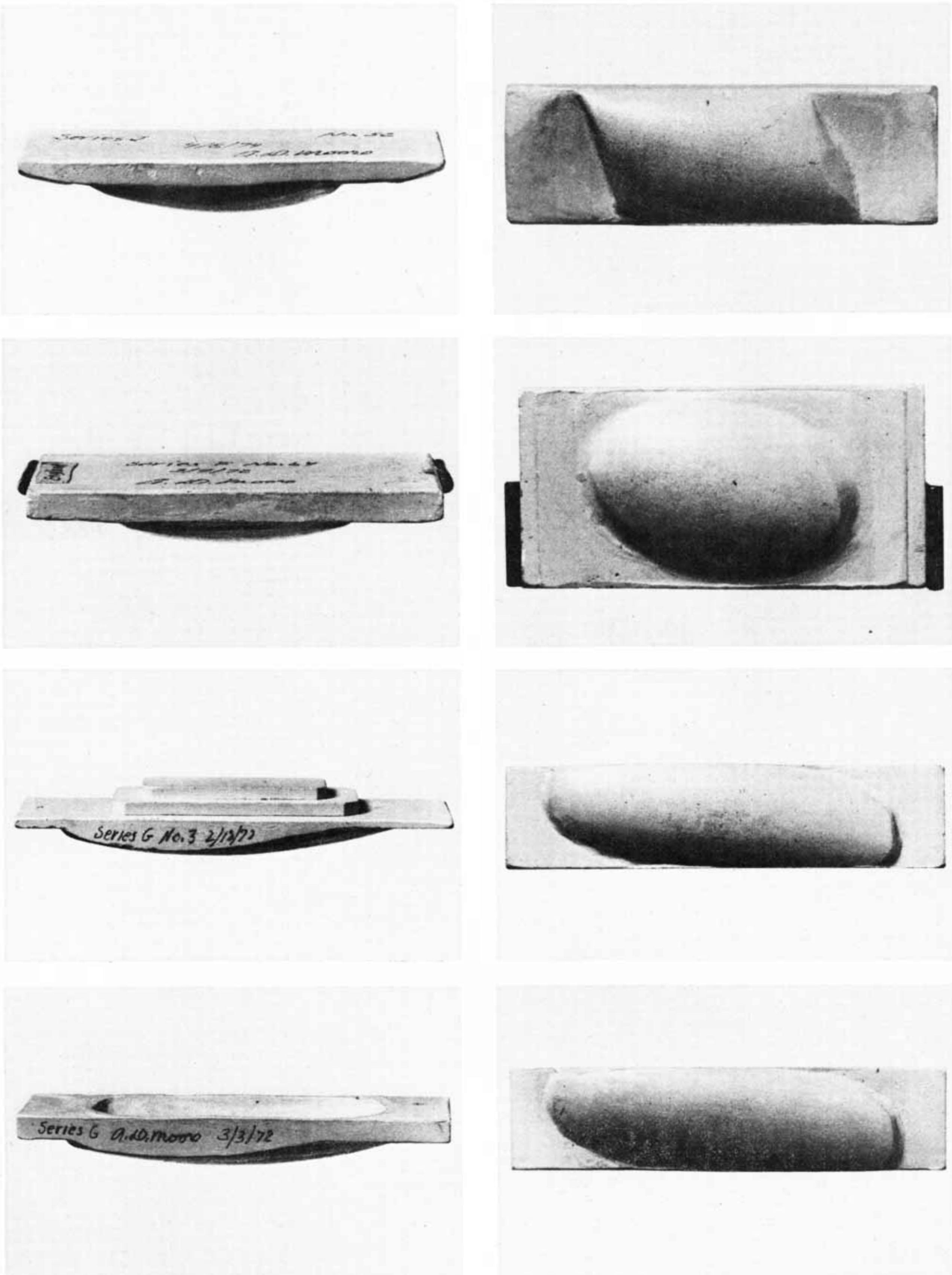
When I examined the rattlebacks, I found that the basic kind exhibited the strongest reversal (counterclockwise to clockwise), which is followed by a slight second reversal as the stone nears the end of its energy. The rattleback with the pennies and the one with the terraces performed about as well as the basic type. The half egg did not spin well, tending to wobble so much that the ends of the brass rod scraped the table and thus drained energy from the egg. Moore is not sure why this former winner now does so poorly, but he suspects that a hairline crack on the bottom surface may be the cause.

The most fascinating rattleback is the boatlike one because it displays strong reversals in each direction of spin. When I spin it counterclockwise, it rotates for part of a revolution, stops, wobbles briskly up and down and then reverses for several revolutions. Then it stops spinning again, rocks from side to side and begins to spin counterclockwise once more. The performance continues until the stone runs out of energy.

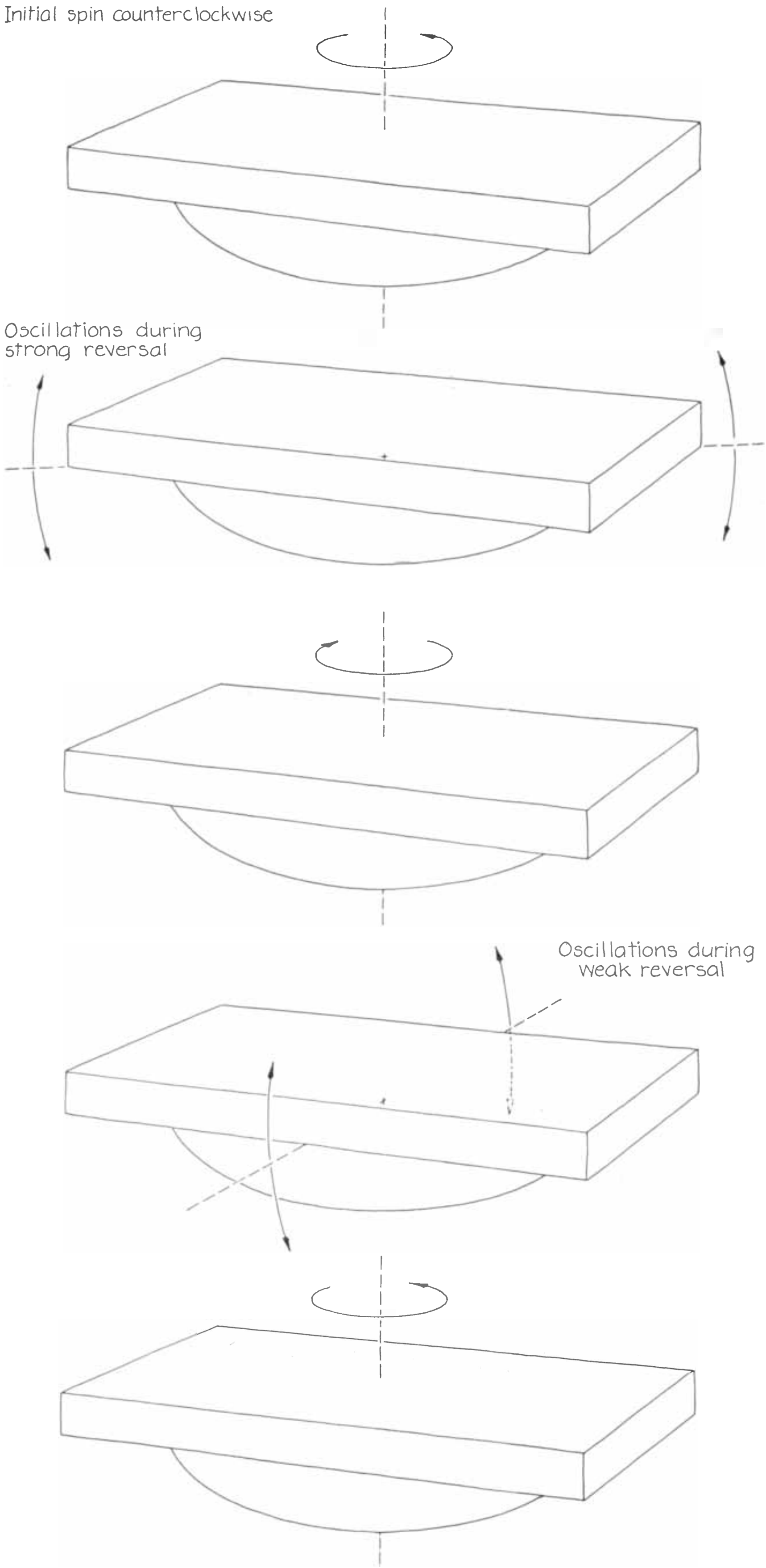
What causes the reversals of spin? Although a complete mathematical explanation is difficult, careful observation reveals a few clues. When a rattleback of the basic design begins to reverse from a counterclockwise rotation, its longitudinal oscillations appear to be approximately around the short axis of the ellipsoid because the points marked *C* and *D* in the illustration on page 176 vibrate up and down more than the points marked *A* and *F*. (I think the oscillation is actually around another important axis, one of the stone's principal axes that is almost parallel to the short axis of the ellipsoid.) Somehow the counterclockwise rotation excites an instability of this kind.

When the rattleback executes its second and milder spin reversal, it appears to oscillate approximately around the long axis of the ellipsoid, sending points *G* and *H* up and down with more amplitude than other points. (Again, theoretically the actual axis is a principal axis closely aligned with the long axis.) Apparently the clockwise spin excites this type of instability.

I can easily set up either type of oscillation in an initially stationary stone by pressing down on an appropriate part of the stone and then letting go. When I do this at point *A*, not much happens. The



Paired views of the side (**left**) *and bottom* (**right**) *of rattlebacks made by A. D. Moore*



A rattleback that exhibits a strong reversal and a weak one

rattleback oscillates for a short time and may turn counterclockwise slightly. At *B* a similar action results in more pronounced oscillations and a clockwise rotation, but the performance is still far from dramatic. Livelier effects are obtained with a press and release at point *C;* the longitudinal oscillations are vigorous and the stone is quickly propelled into a clockwise rotation. The oscillations are identical with the ones I see when the spin reverses from counterclockwise to clockwise.

When I press point *G* and release it, the stone rocks from side to side and soon begins to spin counterclockwise. Thus when I excite the vertical oscillations at *G,* I am duplicating the second type of spin reversal (of the initial clockwise rotation). Similar results are obtained with a press and release at other points between *G* and *A.*

The transverse oscillations are much slower than the longitudinal ones, but both types are so fast that I cannot follow the motion easily. Moreover, once the oscillations start the stone begins to turn and my perspective is altered. To follow the oscillations better I tried tapping a rattleback I had placed on a strip of tape that had its sticky side up. The tape kept the stone from rotating, but of course it may also have interfered with the stone's natural motion. Although the stone seemed to oscillate primarily around a single axis, it also gave the appearance of rolling out of the primary plane of oscillation. I do not know if such a feature is essential to the reversals of spin.

The first scientific investigation of spin-reversing stones was made by G. T. Walker (no relation) in 1896. Some of his results can still be found in reprints of old books dealing with rotational mechanics. Usually the description of the stones is put in the chapter on asymmetrical tops, that is, tops that are not symmetrical around the vertical axis. Theoretical discussions of symmetrical tops are challenging enough, but the theory of asymmetrical tops might be called mind-spinning.

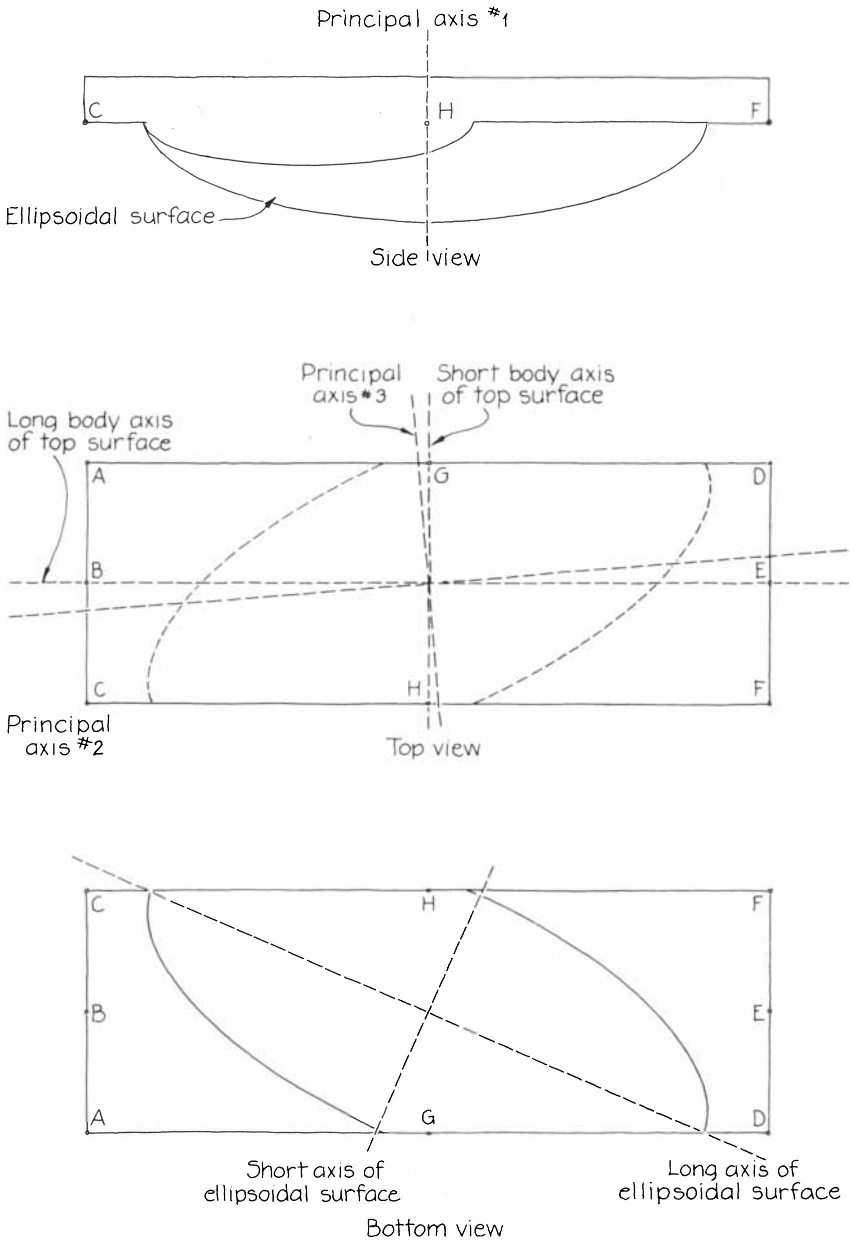
A successful rattleback has several key features. One has to do with the misalignment of the stone's principal axes and the axes of the ellipsoid on the bottom. A principal axis of an object is an axis about which the object could be rotated freely with no rotation about either of the other two principal axes. The three mutually perpendicular axes are usually found along the axes of symmetry of many common objects. If the ellipsoid were aligned with the rectangular top in Moore's basic rattleback, the principal axes would be easy to find because they would coincide with the axes of symmetry. One axis would be vertical and the other two would be parallel to the short and long axes of the ellipsoid.

They would all cross at the center of mass of the stone.

A good rattleback's ellipsoidal bottom is not aligned with the rectangular top. The vertical axis is still a principal axis, but the other two axes are now shifted around the vertical one in the direction in which the ellipsoidal bottom has been shifted out of alignment with the rectangular top. One of these principal axes, call it axis 2, is somewhere between the long body axis of the rectangular top and the long axis of the ellipsoid. The other one, axis 3, is somewhere between the corresponding two short axes and is perpendicular to axis 2. Their exact locations depend on the relative mass distributions of the rectangular top and the ellipsoid. The misalignment of the ellipsoid from the principal axes is an essential feature in the reversal of spin.

A second important feature of a well-trimmed rattleback is that the radius of curvature is different along the two axes of the ellipsoid. (A large radius of curvature corresponds to a surface with a small amount of curvature.) If the bottom were perfectly spherical, the rattleback would not reverse its spin.

It is also important that the mass distribution of the stone be different for the two principal axes. The function giving the distribution of mass with respect to a particular axis is called the moment of inertia. Consider principal axis 3. When

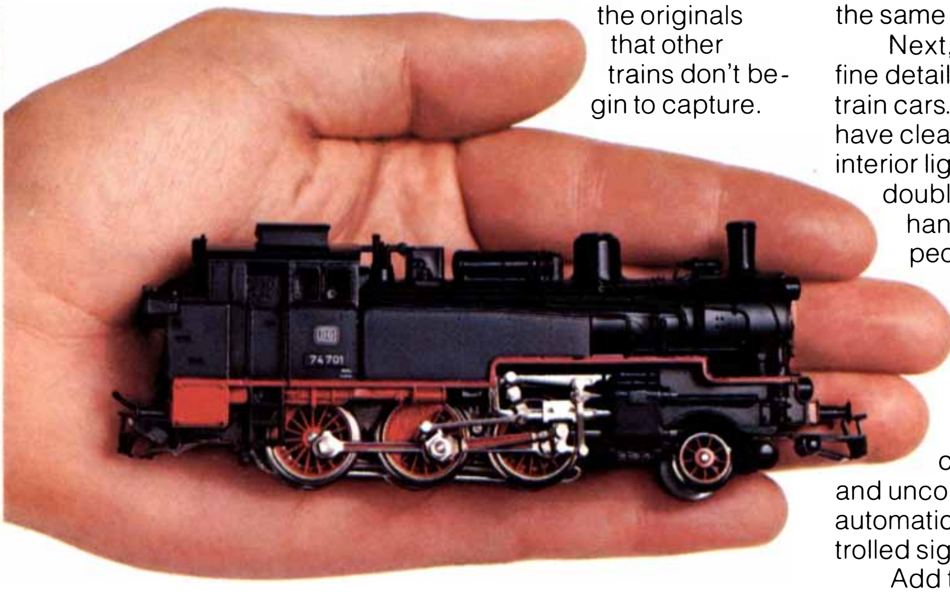


The axes of a rattleback biased to rotate clockwise

Basic Training

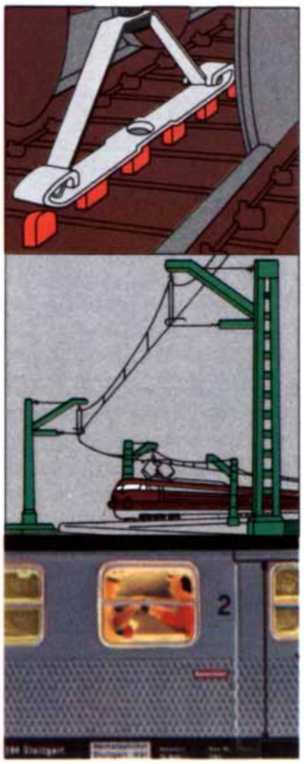
First of all, there's a really big difference between Marklin and the other more ordinary HO trains.

Our trains have a precision and faithfulness to the originals that other trains don't begin to capture.



Take our HO gauge locomotives for example. We make over forty-five different locomotives and each one is a scaled-down replica of an actual train engine.

Detailed down to the bolts on the boiler, we put an incredible amount of precision engineering right in the palm of your hand.



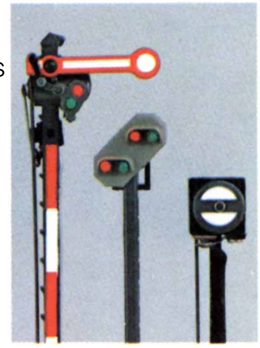
Three unusual features make the Marklin system more functional (and therefore, more enjoyable) than a standard HO train set.

1. Snap together tracks that stay together during operation.
2. Color-coded wiring that simplifies setting up.
3. Deep wheel flanges that do a better job of keeping the train on its tracks.

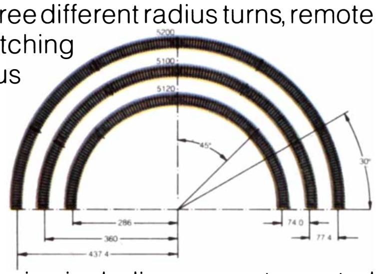
Another important feature in a model train is reliability. That's why we have developed a center conductor power system that is more reliable and more flexible.

There's also a catenary power system that can be used as a second circuit, thereby permitting two locomotives to be operated independently on the same track at the same time.

Next, take a look at the fine details on the individual train cars. Our passenger cars have clear plastic windows, interior lighting, single and double seats and even hand painted "little people." Many of our freight cars have working doors or other functional features.



When it comes to track layout, you'll want at least three different radius turns, remote control switching and uncoupling, plus automatically controlled signals.



Add to this a variety of graded crossings, bridges, lighting and a wide selection of accessories including a remote control motor-driven turntable and roundhouse with automatically closing doors.

About the only limitations are your imagination and perhaps your pocketbook. Our trains are not inexpensive, but then craftsmanship, precision and excellence seldom are.

So visit your hobby store or finer department store; or send $3 for a catalog to: Marklin Catalog,

Reeves International, 1107 Broadway, N.Y., N.Y. 10010. On the West Coast contact Boyd Models, Bldg. 1, 1835 Whittier Ave., Costa Mesa, Ca. 92627.

märklin

This is a reality. Few airlines fly so high.

And, since a picture is worth a thousand words, we'll limit ourselves to offering the first 17 of the 108 airlines IATA (International Air Transport Association) worldwide ranking in order of importance in number of passengers flown:*

1.................................. United Airlines,U. S. A.
2................................. Eastern Airlines,U. S. A.
3..................................... Delta Airlines,U. S. A.
4............................. American Airlines,U. S. A.
5.. TWA,U. S. A.
6.......... British Airways (Merger of Bea and Boac),U. K.
7.. IBERIA,Spain.
8.. Allegheny,U. S. A.
9.................................... Japan Airlines,Japan.
10.................................... Lufthansa,Germany.
11.................................... Air Canada,Canada.
12.. Braniff,U. S. A.
13...................................... Air France,France.
14.. Pan Am,U. S. A.
15..................................... SAS,Scandinavian.
16.. Alitalia,Italy.
17.. National,U. S. A.

Facts speak for themselves.

This is Iberia today.
But we want to be better.

*Ranking 1977

IBERIA
INTERNATIONAL AIRLINES OF SPAIN

MORE THAN 50 YEARS MAKING FRIENDS.

one of Moore's rattlebacks oscillates up and down around that axis, as it does during the strong type of spin reversal, the moment of inertia for the oscillation is relatively large because some of the mass is at a relatively large distance from the axis. Suppose the stone oscillates around principal axis 2, as it does during the second, weaker type of spin reversal. The moment of inertia for that oscillation is relatively small because most of the mass is relatively close to the axis. This difference in the moment of inertia of the two horizontal principal axes is essential to the success of a rattleback.

In sum, Moore's basic rattleback has the following key features. The long axis of the ellipsoid (the one with the large radius of curvature) is shifted counterclockwise (as seen from above) from principal axis 2, which has a small moment of inertia. This arrangement biases the rattleback for a strong spin reversal from counterclockwise to clockwise and a weaker spin reversal in the other direction. What would happen if the long axis of the ellipsoid were shifted in the other direction? Would the strong and weak spin reversals be in the opposite directions? Although G. T. Walker's theoretical work predicts the behavior, I shall leave the questions for you to answer with suitably carved rattlebacks.

How do the key features of Moore's rattleback cause the stone to reverse spin? Primarily they make it unstable against small perturbations of the spin around the vertical axis. If the bottom were spherical or the ellipsoid were aligned with the principal axes, any small perturbation would have a small effect on the spin and would not cause noticeable oscillations. With a properly shaped stone, however, small perturbations from the initial spinning of the stone or from the tabletop generate oscillations that rapidly grow in amplitude. Which of two types of instability is excited depends on the direction of spin and the design of the stone.

Suppose a rattleback of the basic shape is spun counterclockwise. Because of its design any small perturbation causing an oscillation around principal axis 3 is rapidly enhanced. The frictional forces acting on the stone during the oscillation stop the spin and then initiate an opposite one. Once the reverse spin has begun, the friction forces stop the oscillation. Hence a spin in the counterclockwise direction allows one type of instability—the oscillation around principal axis 3—to grow large quickly, and then the frictional forces accompanying the oscillations reverse the spin.

If the stone is initially spinning clockwise, another instability is encouraged, one that enhances oscillation around the other horizontal principal axis. Again

HP measurement and computer advances

For designers of microcomputer-based products, HP's new 64000 System is a very logical development.

Hewlett-Packard's new logic development system simplifies logic design, development, and analysis for all products built around a microcomputer. It can help move a design rapidly from definition to production with minimum cost and maximum design team performance.

As microcomputers increase in power and versatility—and get smaller and cheaper at the same time—they are finding their way into a widening range of products as intelligent controllers. And this "microcomputer revolution" is growing at such a rate that few product designers can afford to ignore the possibility that the inclusion of a microcomputer might improve their product.

The HP 64000 Microcomputer Development System is, essentially, a special-purpose computer system dedicated to developing source code, downloading the code to a microcomputer, and debugging the resulting system. It supports the major microcomputers currently available, and is designed to anticipate future microcomputer developments such as 16- and 32-bit devices. It handles conventional software tasks including editing, assembly, and compiling, as well as microcomputer emulation and logic analysis.

The HP 64000's modular architecture lets you choose the system components you need to develop microcomputer software programs for your application: from one to six development stations, from 20 to 120 megabytes of storage on a high-speed hard disc, and one of several printers.

Each development station shares the disc memory, giving each user access to current programs. Thus the design team can work simultaneously on hardware, software, and interactive problems, with resulting time and system cost savings. Each station contains a host processor with 64K bytes of user memory, I/O control, and display control.

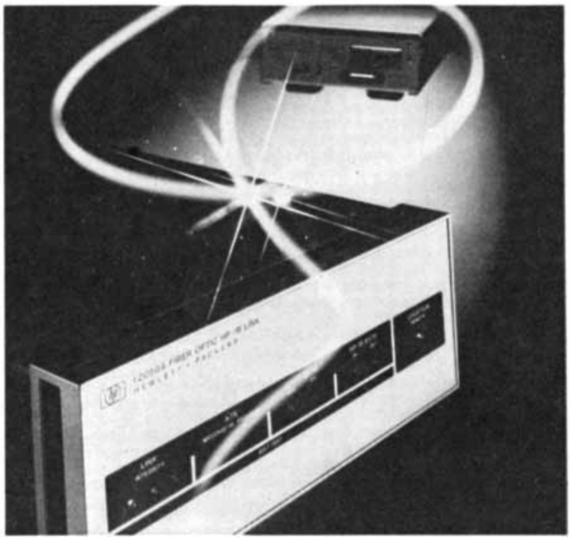
Real-time emulation of microcomputers from different manufacturers at full specified processor speed does not alter the performance of the system under test, and isolates potential time-dependent problems.

An optional, built-in logic analyzer gives a real-time view of emulator-bus activity to identify and solve problems that show up when your system operates at speed. As a pioneer in logic state analysis, HP offers advanced, highly practical debugging and troubleshooting capability.

Price of a completely configured system starts at $35,000*.

extend your possibilities.

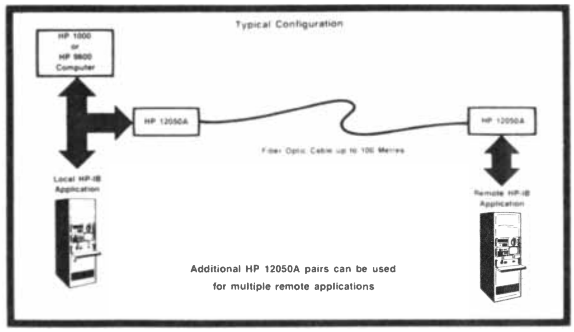
Fiber optics technology now speeds and protects data communication via the HP interface bus.



The new Fiber Optic HP-IB Link provides a high-speed, highly reliable extension to the Hewlett-Packard Interface Bus (HP-IB). The Link joins instrument clusters and computers up to 100 metres apart, and transmits data by light impulses over fiber optic cable at up to 20k bytes per second. Because the cable is immune to electromagnetic interference, it is especially well suited to industrial environments.

The HP 12050 Fiber Optic HP-IB Link combines HP's expertise in instrumentation interfacing with the advantages of fiber optics: broad bandwidth over long distances; lightweight, small-diameter, flexible cables; electrical isolation between connected units; and no externally radiated signal. The Link solves two big problems of remote interfacing: the data transfer rate is fast enough for computer-oriented real-time applications, and data is not lost or distorted by electromagnetic disturbances along the way.

The HP 12050 accepts bit-parallel signals from any HP-IB compatible computer/controller or instrument. It converts the signals to bit-serial light impulses; these are transmitted through a dual-channel fiber optic cable, reconverted to bit-parallel form by a second 12050, and sent on to instruments or computer through a standard HP-IB connector.



A pair of HP 12050's will link a cluster of as many as 14 remote instruments to a computer up to 100 metres distant. No special programming is required, and remote interrupts are detected within 100 microseconds for real-time response. For additional data integrity, a built-in silicon-on-sapphire microprocessor detects transmitted errors using a checksum technique and, if necessary, data can be retransmitted without computer intervention.

Since optical data transmission is not affected by electromagnetic interference, the cable needs no shielding near the rotating machinery, transformers, and switching gear likely to be found in an industrial environment. And since the cable carries no electrical charge, it is safe even in explosive environments. It is lightweight and easy to handle, but rugged enough to be installed by conventional cable-pulling techniques.

The price of each 12050A unit is $1950*; two units are needed per system. Dual channel fiber optic cable with connectors attached can be ordered in several lengths; price is about $8.50* a metre.

HEWLETT hp PACKARD

1504 Page Mill Road, Palo Alto, California 94304

For assistance call: East (301) 258-2000, West (213) 877-1282, Midwest (312) 255-9800, South (404) 955-1500, Canada (416) 678-9430

Mail to: Hewlett-Packard, 1504 Page Mill Road, Palo Alto, CA 94304
Please send me further information on

( ) HP 12050 fiber optic HP-IB link
( ) HP 64000 microcomputer development system

Name ____________

Company ____________

Address ____________

City ____________ State ________ Zip ________

*Domestic U.S. prices only.

00949

the design of the rattleback allows this instability to grow in amplitude exponentially. Frictional forces during this kind of oscillation cause the clockwise spin to stop and then reverse.

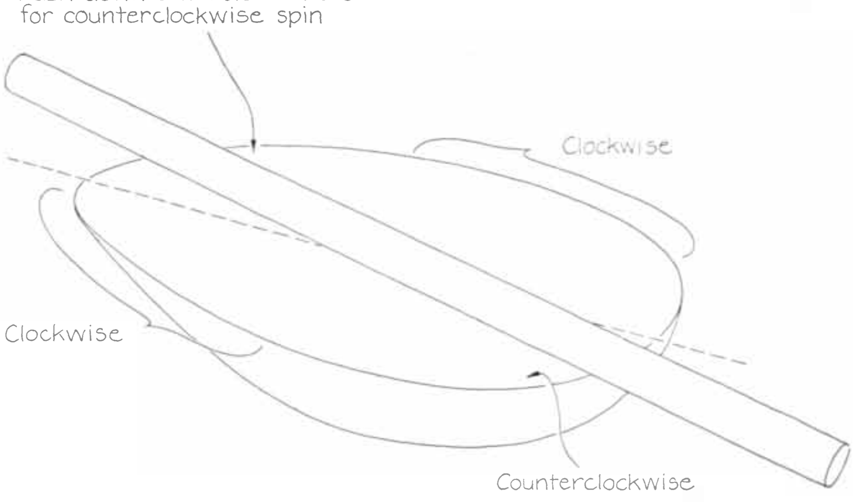
I wanted to investigate how the behavior of a basic rattleback would change if I altered one of the key design features. The only feature I could change easily was the relative moment of inertia for the two horizontal principal axes. To alter it I taped a pencil across the top of a rattleback, first parallel to the length of the rectangle and then parallel to the width. Each time I was careful to balance the pencil so that the rattleback would spin on the same area on the ellipsoid.

With the pencil lengthwise the rattleback behaved generally as it had before because the pencil merely increased the difference in the moments of inertia around the principal axes. With the pencil parallel to the width the stone did not reverse spin or show any of the characteristic instabilities. By placing the pencil across the width of the rectangle I was increasing the moment of inertia around principal axis 2 without much changing the moment of inertia around principal axis 3. Apparently I increased the one enough to make it comparable to the other. When the moments of inertia are much the same, the oscillations due to small perturbations do not grow exponentially with time and therefore remain small. In the absence of large oscillations around the principal axes the frictional forces from the tabletop cannot stop the rattleback and then reverse its spin.

A moment of inertia can be varied in a more controlled way by mounting long stove bolts on a rattleback. Extend a bolt in each direction along a principal axis and screw several nuts on the ends of the bolts. The moment of inertia is changed by moving the nuts closer to the center of the rattleback or farther away from it. (The stone must still be balanced.) With this arrangement I was again able to eliminate the spin reversals when I equalized the moments of inertia for the two horizontal principal axes. What would happen if one adjusted the nuts to make the moment of inertia around principal axis 2 larger than the one around principal axis 3? Would spin reversal appear again but with the directions of the strong and weak reversals interchanged?

Moore's half-egg rattleback has a brass rod by which the moments of inertia can be altered. (Repositioning the rod also shifts the principal axes somewhat.) When I tape the rod across the short width of the flat top, the egg spins stably in each direction. With the rod along the longer axis of the flat top instabilities and spin reversals appear. Normally the rod is positioned at a small angle to the longer axis of the flat top, as is shown in the illustration below. The half egg reverses weakly in each direction of spin. With the rod mounted in this way the half egg is also sensitive to a tap on its perimeter. The direction of the resulting spin depends on the location of the tap. The illustration indicates the possibilities.

G. T. Walker's analysis predicts that the frequency of oscillation is higher around principal axis 3 than it is around axis 2, a point easily verified with a basic rattleback. His equations also predict that neither instability will arise if the spin frequency is higher than the oscillation frequency. With Moore's rattlebacks my attempt to eliminate the instability for the strong reversal (the one for a counterclockwise spin) is impractical because the vigor of my initial twist inevitably introduces instabilities. I can easily spin the rattleback clockwise, however, with a spin frequency greater than the oscillation frequency around



The rotational possibilities with Moore's half egg principal axis 2. As far as I can tell the oscillation does not appear until friction has slowed the stone, presumably to a spin frequency lower than the oscillation frequency.

I played with Moore's rattlebacks on a smooth Formica top and a few other surfaces. I have not investigated the effects of the friction coefficient between the rattleback and the surface on which it is spun. If the friction is too high, the spinning should be eliminated so quickly that no spin reversal develops. If the friction is too low, the stone will reverse spin with difficulty because the reversal requires friction.

You can easily carve rattlebacks from dental stone or wood. Adjust the shape of the bottom until the rattleback works properly. You may be able to think of variations of the basic design. For example, you could investigate the effect of weight distribution by taping small weights or metal rods on the top of the rattleback. By making a number of rattlebacks you could determine how the orientation of the ellipsoid on the bottom affects the spin reversal. If you are out to win a contest for the number of spin reversals of a rattleback, look for an optimum angle between the long body axis and the long axis of the ellipsoid. You might also try to build a rattleback big enough to ride on. If you could make one with a vigorous spin reversal, riding it would be like riding a bucking bronco.

I turn now to the Tippe Top, which also reverses unexpectedly. Such a top usually has a spherical bottom and a central stem. It is spun by sharply rolling the stem through the fingers and dropping the top on a flat surface. The surprise is that the top spins on its spherical end for only a few seconds and then, turning upside down, spins on its stem. The motion appears to violate the law of the conservation of energy because the top seems to raise its center of mass (which is in the spherical section) without outside help.

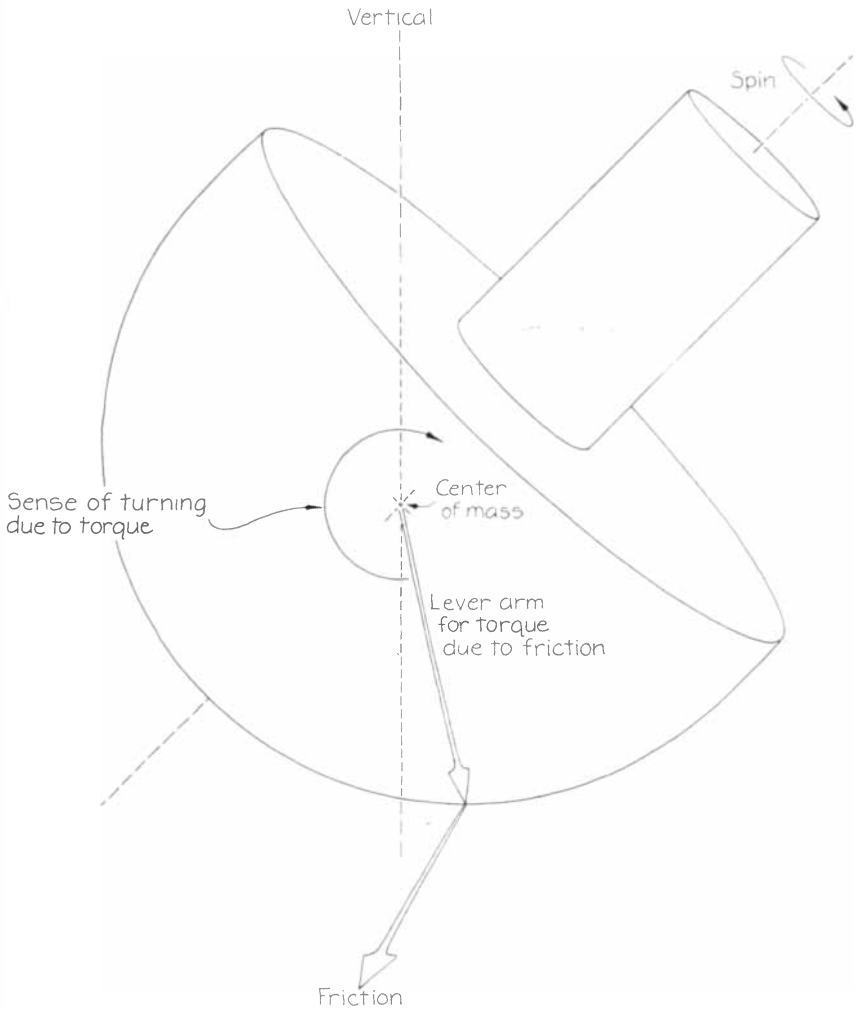
The top has long fascinated observers, including several distinguished physicists and mathematicians. In a recent paper Richard J. Cohen of the Massachusetts Institute of Technology describes how William Thomson (the eminent physicist better known as Lord Kelvin) spent his time spinning smooth stones on the beach instead of preparing for his mathematical examination at the University of Cambridge. Later Niels Bohr, who developed the first modern model of the hydrogen atom, became similarly entranced with the mechanics of the Tippe Top.

The surprising inversion of the top arises from friction between the top and the surface on which it spins. Suppose it spins with its stem tilted away from the vertical axis, as is shown in the illustration below. Because it is spinning it develops a sliding friction force (perpendicular to the plane of the page in the illustration). This force creates a torque about the center of mass, causing the top to invert.

You could make a Tippe Top by cutting off part of a solid rubber ball and inserting a bolt in the flat surface. A typical school ring with a smooth stone displays the same type of inversion. Spin the ring on the stone. A hard-boiled egg that is spun flat will rise to spin on one end. Indeed, spinning an egg is one way to ascertain whether it is hard-boiled; a fresh egg will not stand up because of the sloshing of its internal liquid.

At this year's International Science and Engineering Fair in San Antonio, Tex., several projects dealt with topics discussed recently in this department. In the physics section David Kirk of Edmond, Okla., sought to determine whether initially hot water might under some circumstances freeze faster than initially warm water. He found, as I did, a great deal of variation, but his results seemed to indicate that with open-top containers the effect can be demonstrated. Joella Carr of Winchester, Tenn., studied the floating of water drops above the surface of a body of water, confirming that vibration of the bulk water prolongs the lifetime of the drops for 10 minutes or so.

In the chemistry division two students investigated oscillating chemical reactions. Helen Runnels of Baton Rouge, La., studied the periodicity of the Belousov reaction, finding that by increasing the cerium concentration from 0 to 10^{-4} molar she was able to increase the period of the reaction threefold. She also found that contaminations of chloride, bromide or iodide stopped the oscillation. Her indicator was nitroferroin. William Crooke of Pensacola, Fla., investigated the chemical waves in a similar reaction. A booklet containing abstracts of all the projects at the fair can be obtained for $2.50 from Science Service, 1719 N Street NW, Washington, D.C. 20036.



What happens as a spinning Tippe Top inverts

SCIENCE/SCOPE

Highly complex microcircuitry soon may be mass produced with a technique being pioneered at Hughes. The approach, called ion beam lithography, has application toward making very large-scale integrated circuits (VLSI's) having circuit lines as narrow as 0.1 micrometer, about 4 millionths of an inch. These minute dimensions have been possible only by tedious, painstaking methods that use an electron beam to draw circuitry on a wafer. Ion beam lithography, however, is faster and less costly because it uses a collimated beam of protons to "photograph" circuit patterns from a mask onto a whole chip.

Information for a unique international weather study is being gathered from instruments on hot air balloons, aircraft, ships, buoys, and five synchronous-orbit weather satellites -- three operated by the U.S., one by Japan, and one by Europe. The purpose of the Global Weather Experiment, as the year-long study is known, is to help improve weather forecasts and monitoring methods. Hughes built one of the spacecraft, Japan's Geostationary Meteorological Satellite. A subsidiary, the Santa Barbara Research Center, supplied the prime instrument aboard the U.S. and Japanese craft. This device, a visible-infrared spin-scan radiometer, provides detailed pictures of cloud cover day or night.

Patients can now receive lower doses of radiation while undergoing fluoroscopic and other X-ray examinations, thanks to a device that makes snapshots of television pictures. The equipment, called the Hughes Anaram 80™ digital picture processor, ties into the TV portion of an X-ray system. It creates a picture from a brief burst of radiation, so patients avoid lengthy exposure to X-rays. The system, in use in hospitals nationwide, can display 60 images per second, freeze one picture for an hour, enhance obscured detail, and display four different pictures simultaneously for comparative analysis. The processor's other uses include ultrasonography, non-destructive electronics testing, radar, sonar, graphics terminals, and image transmission.

Hughes is seeking engineers to develop advanced systems and components for the following satellites: GOES D, E, and F for NASA; Anik C for Canada; GMS II for Japan; LEASAT for the U.S. Navy; and SBS, the next-generation domestic system. Fields with immediate openings include advanced communications, scientific and engineering programming, systems test & evaluation, microwave & RF design, structural design, spacecraft stress analysis, power systems design, data processing systems, control electronics computers, operating systems, and network systems. Please send your resume to T.W. Royston, Dept. SE109, Hughes Space & Communications Group, P.O. Box 92919, Los Angeles, CA 90009.

The day when satellites will broadcast television directly into the home has drawn closer with the development of a receiver that can be mass produced at low cost. The unit, delivered by Hughes to NASA's Goddard Space Flight Center, is used with an antenna and amplifier to tune in TV signals from any particular transponder of a broadcast satellite. It serves as a prototype for inexpensive models to be developed in the next decade. Rooftop terminals are expected to become commonplace in many countries by the mid-1980s.

Creating a new world with electronics

HUGHES

HUGHES AIRCRAFT COMPANY
CULVER CITY, CALIFORNIA 90230

BIBLIOGRAPHY

Readers interested in further explanation of the subjects covered by the articles in this issue may find the following lists of publications helpful.

MATHEMATICAL GAMES

DECOMPOSITION OF A CUBE INTO SMALLER CUBES. Christoph Meier in *The American Mathematical Monthly*, Vol. 81, No. 6, pages 630–631; June–July, 1974.

ON PACKING SQUARES WITH EQUAL SQUARES. P. Erdös and R. L. Graham in *Journal of Combinatorial Theory, Series A*, Vol. 19, No. 1, pages 119–123; July, 1975.

INEFFICIENCY IN PACKING SQUARES WITH UNIT SQUARES. L. F. Roth and R. C. Vaughan in *Journal of Combinatorial Theory, Series A*, Vol. 24, No. 2, pages 170–186; March, 1978.

ACID RAIN

IMPACT OF ACID PRECIPITATION ON FOREST AND FRESHWATER ECOSYSTEMS IN NORWAY. Edited by F. H. Braekke. SNSF Project, Research Report FR/76, Oslo, Norway, 1976.

PROCEEDINGS OF THE FIRST INTERNATIONAL SYMPOSIUM ON ACID PRECIPITATION AND THE FOREST ECOSYSTEM. Edited by Thomas A. Seliga and Leon S. Dochinger. USDA Forest Service General Technical Report, NE-23, 1976.

THE ACIDITY PROBLEM—AN OUTLINE OF CONCEPTS. Svante Odén in *Water, Air and Soil Pollution*, Vol. 6, Nos. 2, 3, 4, pages 137–166; September/October/November, 1976.

LINEAR-CHAIN CONDUCTORS

CHEMISTRY AND PHYSICS OF ONE-DIMENSIONAL METALS. Edited by Heines J. Keller. NATO Advanced Study Institutes Series, Vol. 25B. Plenum Press, 1977.

SYNTHESIS AND PROPERTIES OF LOW-DIMENSIONAL MATERIALS. Edited by Joel S. Miller and Arthur J. Epstein in *Annals of the New York Academy of Sciences*, Vol. 313; September 29, 1978.

QUASI-ONE-DIMENSIONAL CONDUCTORS. Edited by S. Barišič, A. Bjeliś, J. R. Cooper and B. Leontič in *Lecture Notes in Physics*, Vols. 95 and 96. Springer-Verlag, 1979.

THE ACOUSTIC MICROSCOPE

ACOUSTIC MICROSCOPY: BIOMEDICAL APPLICATIONS. Ross A. Lemons and Calvin F. Quate in *Science*, Vol. 188, No. 4191, pages 905–911; May 30, 1975.

ACOUSTIC IMAGING: CAMERAS, MICROSCOPES, PHASED ARRAYS, AND HOLOGRAPHIC SYSTEMS. Edited by Glen Wade. Plenum Press, 1976.

ACOUSTIC MICROSCOPY AT OPTICAL WAVELENGTHS. V. Jipson and C. F. Quate in *Applied Physics Letters*, Vol. 32, No. 12, pages 789–791; June 15, 1978.

AN ANGULAR-SPECTRUM APPROACH TO CONTRAST IN REFLECTION ACOUSTIC MICROSCOPY. Abdullah Atalar in *Journal of Applied Physics*, Vol. 49, No. 10, pages 5130–5139; October, 1978.

LEGIONELLOSIS

DEMONSTRATION OF THE AGENT OF LEGIONNAIRES' DISEASE IN TISSUE. Francis W. Chandler, Martin D. Hicklin and John A. Blackmon in *The New England Journal of Medicine*, Vol. 297, No. 22, pages 1218–1220; December 1, 1977.

LEGIONNAIRES' DISEASE: DESCRIPTION OF AN EPIDEMIC OF PNEUMONIA. David W. Fraser, Theodore F. Tsai, Walter Orenstein, William E. Parkin, H. James Beecham, Robert G. Sharrar, John Harris, George F. Mallison, Stanley M. Martin, Joseph E. McDade, Charles Shepard, Philip S. Brachman and the Field Investigation Team in *The New England Journal of Medicine*, Vol. 297, No. 22, pages 1189–1197; December 1, 1977.

LEGIONNAIRES' DISEASE: ISOLATION OF A BACTERIUM AND DEMONSTRATION OF ITS ROLE IN OTHER RESPIRATORY DISEASE. Joseph E. McDade, Charles Shepard, David W. Fraser, Theodore F. Tsai, Martha A. Redus, Walter R. Dowdle and the Laboratory Investigation Team in *The New England Journal of Medicine*, Vol. 297, No. 22, pages 1197–1203; December 1, 1977.

PONTIAC FEVER: AN EPIDEMIC OF UNKNOWN ETIOLOGY IN A HEALTH DEPARTMENT. I. CLINICAL AND EPIDEMIOLOGIC ASPECTS. Thomas H. Glick, Michael B. Gregg, Bernard Berman, George Mallison, Wallace W. Rhodes, Jr., and Ira Kassanoff in *American Journal of Epidemiology*, Vol. 107, No. 2, pages 149–160; February, 1978.

THE PHOTOSYNTHETIC MEMBRANE

CHLOROPLAST MEMBRANE ORGANIZATION AT THE SUPRAMOLECULAR LEVEL AND ITS FUNCTIONAL IMPLICATIONS. L. Andrew Staehelin, Paul A. Armond and Kenneth R. Miller in *Brookhaven Symposia in Biology*, Vol. 28, pages 278–315; 1976.

A PARTICLE SPANNING THE PHOTOSYNTHETIC MEMBRANE. Kenneth R. Miller in *Journal of Ultrastructure Research*, Vol. 54, No. 1, pages 159–167; January, 1976.

DYNAMIC STRUCTURAL FEATURES OF CHLOROPLAST LAMELLAE. Charles J. Arntzen in *Current Topics in Bioenergetics*, Vol. 8, pages 111–160; 1978.

THE LAVA LAKES OF KILAUEA

PETROLOGY OF THE KILAUEA IKI LAVA LAKE, HAWAII. Donald H. Richter and James G. Moore. Geological Survey Professional Paper 537-B, 1966.

THE ERUPTION OF AUGUST 1963 AND THE FORMATION OF ALAE LAVA LAKE, HAWAII. Dallas L. Peck and W. T. Kinoshita. Geological Survey Professional Paper 935-A, 1976.

THE EVOLUTION OF THE CROCODILIANS

THE CROCODILIAN SKULL IN HISTORICAL PERSPECTIVE. Wann Langston, Jr., in *Biology of the Reptilia: Vol. 4*, edited by Carl Gans and Thomas S. Parsons. Academic Press, 1973.

CROCODYLIA. Rodney Steel in *Handbuch der Paläoherpetologie/Encyclopedia of Paleoherpetology: Part 16*, edited by Oskar Kohn. Gustav Fischer Verlag, 1973.

THE MORPHOLOGY AND RELATIONSHIPS OF A CROCODILIAN, *ORTHOSUCHUS STORMBERGI*, FROM THE UPPER TRIASSIC OF LESOTHO. Diane S. Nash in *Annals of the South African Museum*, Vol. 67, Part 7, pages 227–329; 1975.

CONSTRUCTIVE MATHEMATICS

THE FOUNDATIONS OF CONSTRUCTIVE ANALYSIS. Errett Bishop. McGraw-Hill Book Company, 1967.

WHAT IS A REAL NUMBER? John Myhill in *The American Mathematical Monthly*, Vol. 79, No. 7, pages 748–754; August–September, 1972.

A RENAISSANCE OF EMPIRICISM IN THE RECENT PHILOSOPHY OF MATHEMATICS. Imre Lakatos in *Mathematics, Science and Epistemology*. Cambridge University Press, 1978.

THE AMATEUR SCIENTIST

AN ELEMENTARY TREATMENT OF THE THEORY OF SPINNING TOPS AND GYROSCOPIC MOTION. Harold Crabtree. Longmans, Green and Co., 1909.

A TREATISE ON GYROSTATICS AND ROTATIONAL MOTION: THEORY AND APPLICATION. Andrew Gray. Dover Publications, Inc., 1959.

MECHANICS: LECTURES ON THEORETICAL PHYSICS, VOL. 1. Arnold Sommerfeld. Academic Press, 1964.

THE TIPPE TOP REVISITED. Richard J. Cohen in *American Journal of Physics*, Vol. 45, No. 1, pages 12–17; January, 1977.

Finding new ways to

Lockheed knows how.

"Biggest advance since the autopilot."

That's one widely held opinion of an exclusive fuel-saving system now being installed in many L-1011 TriStars.

Called the Lockheed Flight Management System, it senses altitude, speed, and other flight factors. Then it can act automatically to control the engine throttles for best fuel-use efficiency at all times during flight—takeoff to landing.

Flight management control panel.

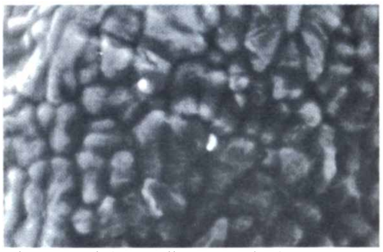
Though a new feature, the system isn't just for new planes; it can easily be added to existing L-1011s. And the savings, over the life of each plane, could total millions of dollars.

Making aircraft lighter.

Since the 1930s, most aircraft have been structured of strong, lightweight aluminum. Future generations of aircraft may also be made mostly of aluminum, but of a very different nature.

A new technology being explored by Lockheed uses powdered aluminum, alloyed with such elements as iron, lithium or cobalt, to create lighter, stronger and stiffer metals than can be achieved with conventional aluminum technology.

Results indicate this 'new aluminum' could reduce present-day jetliner weight by 10%. That in turn means savings of millions of gallons of fuel over each aircraft's life, a big benefit in energy-short years.



Aluminum lithium alloy magnified 1,000 times.

For this new-material process, Lockheed scientists even had to employ astonishing techniques for turning the powdered aluminum and alloys into superior metal. After heating and melting, for example, the mixture is rapidly cooled a *million times faster* than the cooling rate used with standard aluminum ingots.

Down the road: liquid hydrogen.

Lockheed scientists are deep into a NASA-sponsored study of a future aircraft, powered by a clean-burning liquid hydrogen, that could cruise at a flashing hypersonic speed of about 4,000 mph, or six times the speed of sound.

Because liquid hydrogen can be made from coal or even water, its future use in aircraft could eliminate the need for scarce, expensive petroleum.

The aircraft would cruise at altitudes as high as 120,000 feet and have ocean-spanning range. Its 200 passengers

save fuel in flight.

Design for liquid hydrogen fueled aircraft.

could be whisked from New York to London in under two hours, and could make the Los Angeles/Tokyo flight in a mere two hours, 18 minutes—takeoff to landing.

Fine-tuning the flight plan.

When a pilot calls Lockheed's JetPlan, he's in touch with a computerized encyclopedia of information affecting his proposed flight. And the benefit he gets is eye-opening economy—savings of both time and fuel.

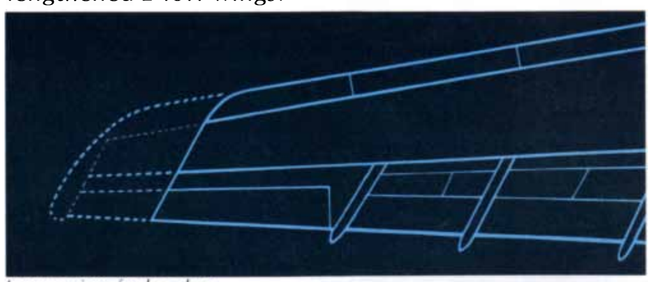
To tap a constantly updated reservoir of data, a pilot makes preflight contact via telephone and a small printer terminal. In minutes, he gets back his most efficient flight plan. Given origin, destination, and type of aircraft, JetPlan reckons its answer from such factors as route analysis, current weather, recommended fuel load, cruise altitude, speed and shortest route.

How well does it work? Over a recent 12-month test period, U.S. military and corporate aviation saved more than 30 million gallons of fuel, or better than $1.5 million. But that's just the beginning. The number of JetPlan users and the savings are both growing steadily.

Stretching wings.

Longer wings mean reduced drag and consequently a substantial cut in fuel use. But among big wide-bodied jetliners, only the Lockheed L-1011 TriStar can add nine feet of wingspan without costly structural redesign.

That's because of the L-1011's advanced technology, in this instance an exclusive system of Active Control ailerons. Controlled by computers, these ailerons will act automatically to reduce the structural loads on the lengthened L-1011 wings.



Longer wings for less drag.

The longer wings and Active Control ailerons will mean a healthy savings of fuel when they enter airline service in 1980. And they'll also give the L-1011 a smoother ride through gusts and turbulence.

When it comes to devising farseeing programs to meet the growing fuel crisis, Lockheed knows how.

Lockheed