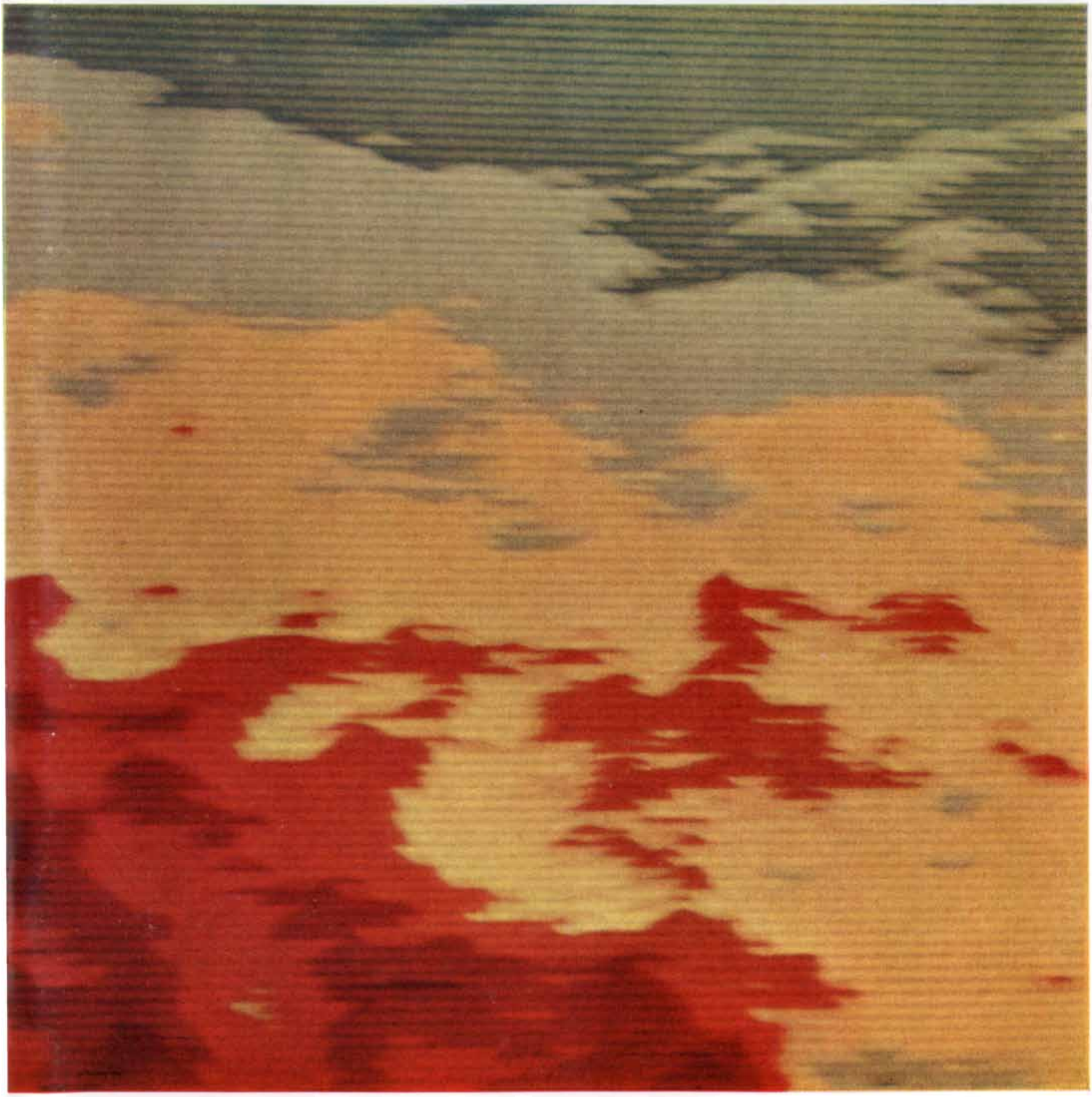
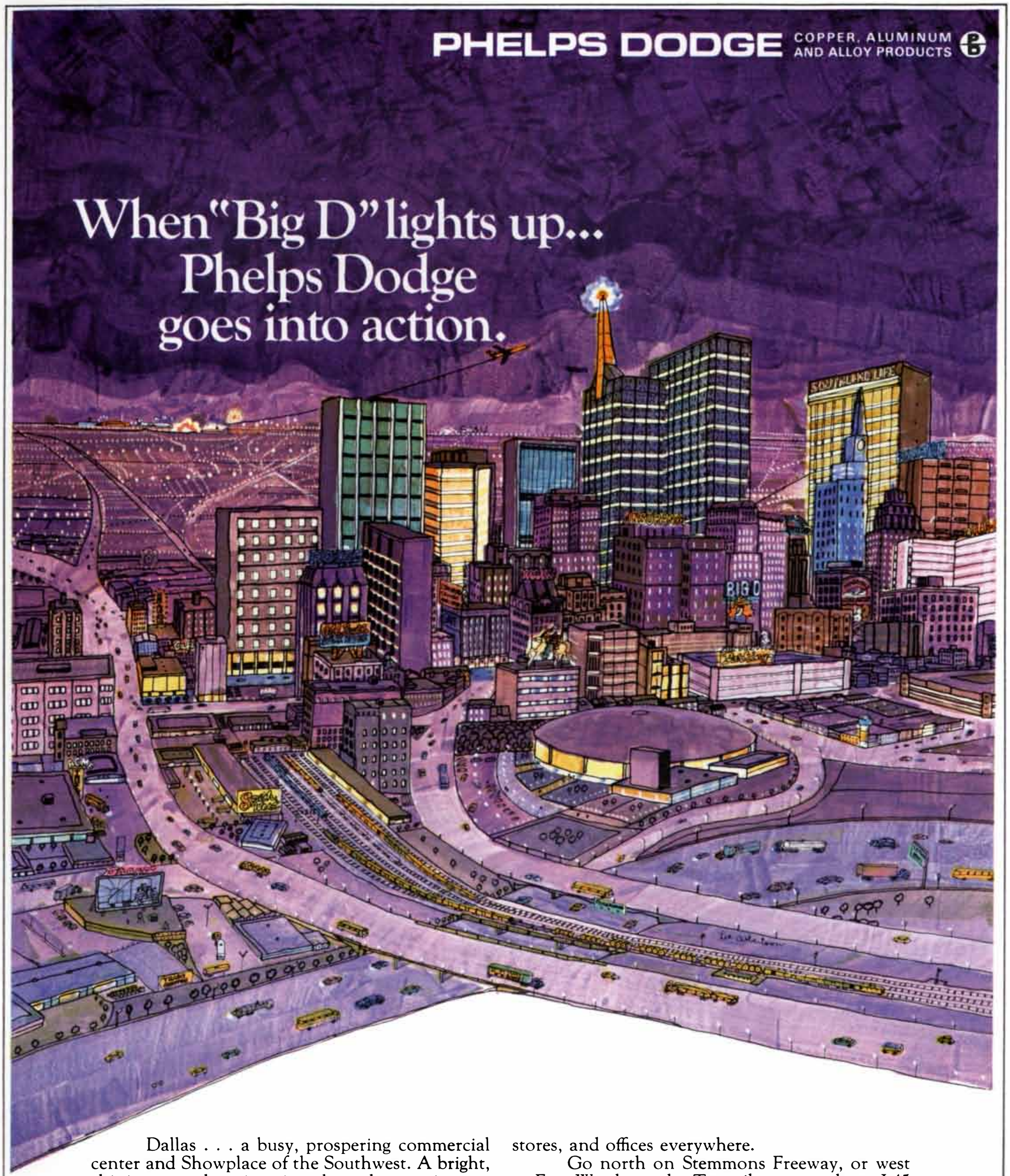
SCIENTIFIC AMERICAN



THERMAL POLLUTION

SEVENTY-FIVE CENTS

March 1969



Dallas . . . a busy, prospering commercial center and Showplace of the Southwest. A bright, shining ever-changing city where the new is commonplace.

Look behind the splendor and the bright lights and you'll see that Dallas is also a Phelps Dodge city. Our condenser tubes are used at the generating plants of the Dallas Power and Light Company. Our 135-kv transmission cables and other high-voltage power cables distribute power throughout the city . . . and the transformers, coils and motors wound with our magnet wire make things happen . . . from the flashing signs downtown, to factories along the river . . . to homes, stores, and offices everywhere.

Go north on Stemmons Freeway, or west to Fort Worth on the Turnpike, or south on I-45 and Phelps Dodge buried lighting cables, telephone or coaxial CATV cables are following alongside. You'll also find our building wire and aluminum conduit . . . our plumbing, gas and refrigeration copper tubing at work everywhere. Many new buildings, like the Statler Hilton Hotel, use PD building wire and copper tubing exclusively.

We specialize in conductors of electricity, liquids, gases and heat made of copper, aluminum and alloys. Look closely, and you'll find Phelps Dodge products at work everywhere.

Return a Hertz car in 25 seconds or less.

Just write the mileage in the space provided on your rental envelope.

We know what it's like to stand in line.

We may not stand in line at Hertz counters, like you. But we've served our time in lines at airline counters and hotel counters, etc.

And it is this knowledge that has led us to the invention of the Express Check-in.

If you're charging one of our Fords or other new cars, all you do to return it is write the mileage in the space provided on the rental envelope–we'll check it for you later. Put the keys inside. Throw it on the counter or give it to an attendant and run.

The whole process takes about 25 seconds or less. Which is very important if your plane happens to be taking off in 25 seconds or less.

Hertz®

The biggest <u>should</u> do more. It's only right.

Established 1845 SCIENTIFIC AMERICAN March 1969 Volume 220 Number 3

ARTICLES

18 THERMAL POLLUTION AND AQUATIC LIFE, by John R. Clark
The disposal of industrial heat in natural waters creates ecological problems.

28 SUPERPLASTIC METALS, by H. W. Hayden, R. C. Gibson and J. H. Brophy
Certain alloys can be shaped by techniques usually limited to glass and plastics.

36 PHASES IN CELL DIFFERENTIATION, by Norman K. Wessells and William J. Rutter
Experiments with embryonic pancreas point to regulatory periods.

54 CONTINENTAL DRIFT AND EVOLUTION, by Björn Kurtén
The breakup of continents may have shaped the rise of reptiles and mammals.

66 BROWNIAN MOTION AND POTENTIAL THEORY, by Reuben Hersh and Richard J. Griego
These two branches of physics are mathematically akin.

78 THE ATMOSPHERES OF MARS AND VENUS, by Von R. Eshleman
Their similarities and differences are compared on the basis of three space probes.

92 PLAGUE TOXIN, by Solomon Kadis, Thomas C. Montie and Samuel J. Ajl
How does *Pasteurella pestis,* the bacillus of plague, kill its mammalian host?

104 THE FIRST ELECTRON TUBE, by George Shiers
It was the two-element tube invented in 1904 by John Ambrose Fleming.

DEPARTMENTS

6 LETTERS

8 50 AND 100 YEARS AGO

14 THE AUTHORS

46 SCIENCE AND THE CITIZEN

116 MATHEMATICAL GAMES

124 THE AMATEUR SCIENTIST

131 BOOKS

148 BIBLIOGRAPHY

PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. COPYRIGHT © 1969 BY SCIENTIFIC AMERICAN, INC. ALL RIGHTS RESERVED. 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: $8 PER YEAR.

THE COVER

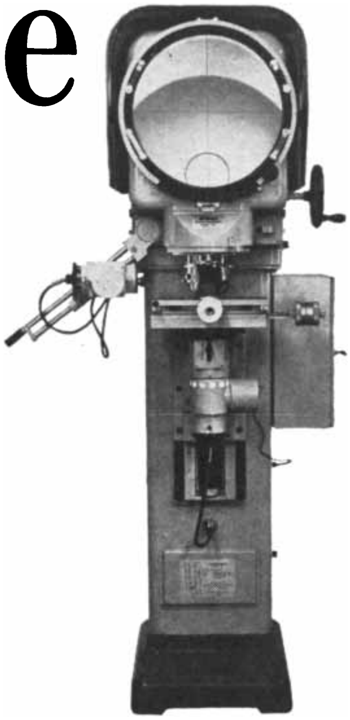
The pattern on the cover is part of a color thermogram, a representation in color of the temperature of an object. What this thermogram shows is a patch of the surface of the Connecticut River at a point where warm water from a power plant is discharged into the river. Heated effluent of this kind is considered an increasing threat to U.S. rivers and lakes (see "Thermal Pollution and Aquatic Life," page 18). The thermogram is made by an instrument that measures the infrared radiation emitted by any body because of the thermal agitation of its molecules. The output of the infrared meter drives a color modulator that translates temperature into color. Here the warmest water, at about 83 degrees Fahrenheit, is shown in a dark magenta. Increasingly cool water is shown, in three-degree steps, in red, orange, yellow, light green and dark green. The pattern is part of a larger thermogram showing how warm water spreads across the river at slack tide (see page 18).

THE ILLUSTRATIONS

Cover photograph by the Barnes Engineering Company

Page	Source
18–19	Barnes Engineering Company
20–21	U.S. Geological Survey
22–26	Joan Starwood
27	Tennessee Valley Authority (*top*), The Marley Company (*bottom*)
28–29	Stephen Manville
31–32	H. W. Hayden and L. S. Torressen, International Nickel Company, Inc.
33–34	Eric Mose
35	David Holt, Massachusetts Institute of Technology (*top*); H. W. Hayden, R. C. Gibson and J. H. Brophy, International Nickel Company, Inc. (*middle*); Eric Mose (*bottom*)
37	Norman K. Wessells
38–39	Bunji Tagawa
40	Bunji Tagawa (*top*), Norman K. Wessells (*bottom*)
41	Bunji Tagawa (*top*), Norman K. Wessells (*middle, bottom*)
42	Bunji Tagawa
43	Bunji Tagawa (*top*), Norman K. Wessells (*second from top*), Bunji Tagawa (*third from top*), Norman K. Wessells (*bottom*)
44	Bunji Tagawa
49	Lick Observatory
55	Enid Kotschnig and Thomas Prentiss
56–57	Thomas Prentiss
59	Enid Kotschnig
60–63	Enid Kotschnig and Thomas Prentiss
66	B. J. Alder and Thomas E. Wainwright
67–74	Allen Beechel
78	Robert B. Leighton, California Institute of Technology (*top*); Lowell Observatory (*bottom*)
80–87	Jim Egleson
92	Paul Weller
94	Solomon Kadis, Thomas C. Montie and Samuel J. Ajl
95	Sol Mednick
96	Thomas Prentiss
98	Solomon Kadis, Thomas C. Montie, Samuel J. Ajl and Chieko Yoshida
100	Thomas Prentiss
105	United Press International
106–107	Dan Todd
108	Edison National Historic Site
109	Gradan Randle
110	The Marconi Company Limited
112	Dan Todd (*top*), Thorn-AEI Radio Valves and Tubes Limited, London (*bottom*)
116–117	Chieko Yoshida
118	Thomas Prentiss
119–120	Chieko Yoshida
124–128	Roger Hayward

How do you measure a sour note optically?



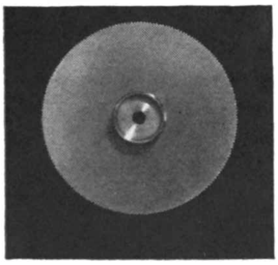
Ask the people at Hammond Organ

It may be difficult to see the connection between a Nikon Profile Projector and, let us say, Bach's G minor fugue—until you discover how a Hammond organ gets its voice.

Every note on the keyboard has a corresponding tone generator, the heart of which is a small, rotating disc, about the size of a silver dollar. There are from 86 to 91 of these tone generators in an organ, and as many discs, all spinning at a uniformly constant speed.

Were you to examine one of these discs, you would discover a series of identical, small, scallop-shaped projections around the rim. As the disc turns, these projections pass through a magnetic field, setting up a minute flow of alternating current, which is amplified and fed to loudspeakers, where it is converted into sound.

The number of projections around the rim determines the frequency or pitch; the shape and dimensions, the purity of the tone. Hammond design demands sine-wave, tonal purity. The harmonics and overtones required for timbre and musical character are obtained from the tones of other discs, added in various patterns of intensity.



Any deviation in the shape of a projection, any dimensional inaccuracy by as little as .0002″, will produce spurious, undesirable sounds—noise and distortion. Consequently, Hammond subjects these tone discs to rigid, 100% inspection.

A Nikon profile projector (such as shown) is used for this purpose. It is essentially a projection microscope. A part (in this case, a tone disc) is placed on the stage. It appears as a magnified image on the screen, where dimensions, shapes and contours can be precisely measured or checked against master charts.

Because of the degree of magnification and the absence of image distortion, repeat measurement accuracy is obtainable to within 0.0001″ (twice that required to meet Hammond specifications). Moreover, inspection now takes less than one quarter of the time previously required.

The Hammond story is only one example of Nikon optical equipment at work. It is the singular story of how a specific instrument has dealt with a specific inspection problem in quality control.

There are many other such stories, in science as well as in industry, in education and communications, in photography and the graphic arts, dealing with almost as many different types of Nikon instruments as there are ways in which optics can serve these endeavors.

If you are faced with a problem which seems to defy other methods of approach, let us show, you how

Nikon can make light work for you.

Instrument division / Nikon Inc., Garden City, N.Y. 11530 • Subsidiary of Ehrenreich Photo-Optical Industries, Inc.

LETTERS

Sirs:

I am writing concerning Karl H. Pribram's interesting article "The Neurophysiology of Remembering," published in your January issue. I wish to report the existence of a mathematical theory of learning, remembering and recall, having a plausible psychological, neurophysiological and anatomical interpretation, that has already given a rigorous form to Dr. Pribram's heuristic suggestion that neural "holograms" exist. A score of papers on this theory have already appeared or are in press, and a monograph applying the theory to learning phenomena was distributed from the Rockefeller Institute to more than 100 laboratories in 1965. The first rigorous equations for the theory were derived in 1959–1961 at Dartmouth College.

The theory, whose mathematical objects are called "embedding fields," has been used to study such diverse phenomena as bowing, anchoring, chunking, backward learning and reminiscence in serial learning; spatiotemporal masking and consolidation; influence of stimulus energy, experimental warm-up and prior learning on reaction time; generalizations of the Hartline-Ratliff equation for lateral inhibition; stimulus sampling; possible mechanisms of excitatory transmitter production, mobilization and release; learning of any number of essentially arbitrarily complicated motor sequences and reflexes by operant and respondant conditioning; the influences of particular anatomical cell distributions on the phenomena these cells can learn, remember and recall; etc.

STEPHEN GROSSBERG

Massachusetts Institute of Technology
Cambridge, Mass.

Scientific American, March, 1969; Vol. 220, No. 3. Published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017; Gerard Piel, president; Dennis Flanagan, vice-president; Donald H. Miller, Jr., vice-president and treasurer.

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

Advertising correspondence should be addressed to Allan Wittman, Advertising Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017.

Offprint correspondence and orders should be addressed to W. H. Freeman and Company, 660 Market Street, San Francisco, Calif. 94104. For each offprint ordered please enclose 20 cents.

Change of address (or other subscription correspondence) should be addressed to Circulation Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. Please notify us four weeks in advance of change. If possible, kindly furnish an address imprint from a recent issue. Send both old and new addresses or complete and mail Post Office Form 357B.

Please enter change of address ☐
Please enter subscription ☐
☐ 1 year, $8 ☐ 2 years, $15 ☐ 3 years, $21

NAME

NEW ADDRESS

OLD ADDRESS

Sirs:

After reading John Updike's charming verses "The Dance of the Solids" [SCIENTIFIC AMERICAN, January], it occurred to me that you might be interested in a similar effort of my own. It was written after reading the article "The Problem of the Quasi-stellar Objects," by Geoffrey Burbidge and Fred Hoyle [SCIENTIFIC AMERICAN; December, 1966].

TO A QUASAR

Twinkle, twinkle, little quasar,
Candidate for Occam's razor:
Are you near or are you far?
Are you nebula or star,
Emitting all that energy
Like any normal galaxy?

Is your message from the dark
Sent by positron or quark?
Spectrum lines, though rather faint,
Tell us only what you ain't.
What strange phenomenon's involved
In this enigma, yet unsolved?

Astronomers have taxed their wits;
There's no hypothesis that fits.
Does Hubble's constant not apply
To these queer objects in the sky?
Is your red shift due to gravitation,
Or to speeding from our observation?

Content with what we think we know,
We sometimes find it isn't so.
Perhaps the gods, somewhat amused,
Though tolerant of facts perused
With all our scientific lore,
Are telling us, "There's much, much more!"

STANLEY A. BELL

Laguna Hills, Calif.

CABOT CORPORATION
DEPT. SA-16, 125 HIGH ST., BOSTON, MASS. 02110
(617) 423-6000

LABORATORY REPORT NO. 16

NOW . . . FROM CABOT FLAME RESEARCH

alon® Al_2O_3

An oxide with exciting possibilities — What can it do for your product?

The unique properties of positive surface charge, large surface area, high purity and extremely small particle size make ALON a versatile and exciting new material with potential uses in many diverse industries. A fumed alumina of different crystalline forms consisting predominantly of the gamma modification, ALON is made by the hydrolysis of aluminum chloride in a flame process similar to that for making Cabot fumed silica, CAB-O-SIL®.

ALON is an inorganic oxide with a surface charge that remains positive up to a pH level of 9.1, the isoelectric point. Each gram of ALON has a surface area of 100 square meters. This produces high functional activity from relatively small quantities.

SUGGESTED USES FOR ALON

TEXTILES — As a wool spinning aid, ALON has been found to considerably lessen end breakage and "fly." ALON substantially increases yarn breaking strength, permits the use of shorter stocks and higher productivity with less waste. The ALON particles form a strong bond with the negatively charged wool fibers, making them less susceptible to pulling apart. The same function is under investigation in cotton and synthetic fiber spinning.

Its positive charge indicates ALON potential as an anti-static, anti-soiling and anti-pilling agent for synthetic fibers and blends. This same property may promote dyeability in certain hard-to-dye fibers.

PAPER — Paper machine down-time is reduced when ALON is used as a de-sliming agent. The positive surface charge helps control build-up and deposition of pitch on rollers, and keeps agglomerates from depositing on the paper thereby improving quality. The high purity and good electrical properties make ALON an effective filler in specialty electronic capacitance papers.

OTHER USES

FERRITES — ALON improves cold temperature operation of ferrites used as armature cores in fractional hp motors.

LIGHT TRANSMISSION AID — ALON is suggested as an additive to phosphor coatings in fluorescent or cathode ray tubes to improve the light output on aging and where high efficiency light transmission is desired on incandescent bulbs. COSMETICS — Hair sprays may be improved by ALON because the positive surface charge reacts with negatively charged hair fibers.

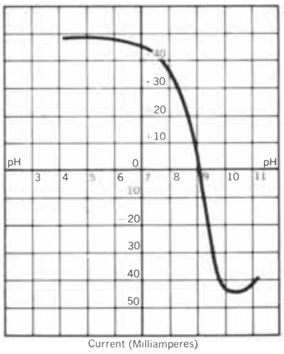
ELASTOMER REINFORCEMENT — ALON has demonstrated an ability to increase creep strength and reinforcement of adhesives and such elastomers as SBR, butyl, nitrile and neoprene rubbers. THICKENING AGENT — Its positive surface charge suggests ALON can be combined with CAB-O-SIL fumed silica to thicken polar liquids such as water. CERAMIC BODY — ALON can be used to improve strength and electrical properties. AIR POLLUTION CONTROL — ALON is being considered as an additive in fuel oils to reduce contaminant emission and minimize scale build-up on furnace firewalls and boiler tubes. HYDROPHILIC AGENT — ALON can be applied to normally hydrophobic surfaces such as plastic films and fibers, photographic films and offset printing plates to improve wettability and eliminate water spotting.

Typical Properties of ALON

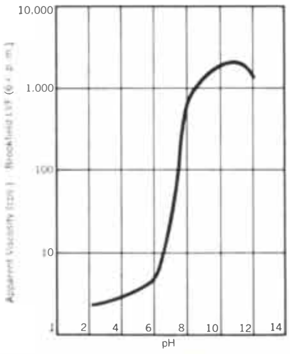
Color and form	White powder
X-ray structure	90% gamma
Alumina content*	99% minimum
Moisture content	2.5%
Ignition loss	4.5%
Metallic oxides (other than Al_2O_3)	1% maximum
Average particle diameter	0.03 micron
Surface area	100 square meters/gram
pH (10% aqueous suspension)	4.4
Specific gravity	3.6
Loose bulk density	4 pounds/cu. ft.

*Excludes physically and chemically combined water

Electrophoretic properties of ALON
Streaming currents of 2% dispersions of ALON in 0.01N sodium chloride solution are plotted as a function of pH.



Viscosity characteristics of ALON (10%, aqueous)



FREE

Send for general brochure containing ideas for potential uses of ALON

CABOT CORPORATION
Dept. SA-16, 125 HIGH ST., BOSTON, MASS. 02110
Please send:
☐ General brochure ☐ Textile applications brochure ☐ Sample
☐ Please call me for appointment. Tel: ____________
I am especially interested in ____________

Name ____________ Title ____________
Company ____________ Dept. ____________
Address ____________
City & State ____________ Zip ____________
Note: To insure receipt of material, please print or write clearly

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

MARCH, 1919: "The total solar eclipse of May 29, 1919, the path of which crosses South America and Africa, will be observed by two British expeditions sent out by the Joint Permanent Eclipse Committee of the Royal Society and the Royal Astronomical Society. Messrs. Crommelin and Davidson of Greenwich Observatory will occupy a station at Sobral in the Brazilian state of Ceará, while Professor Eddington and Mr. Cottingham will observe on the island of Principe, 110 miles from the coast of Africa. This eclipse will be notable, not only because of the long period of totality (six minutes 50 seconds in the mid-Atlantic and more than five minutes at the land stations) but also because of the location of the sun in a rich field of stars, offering a favorable opportunity for testing Einstein's theory of relativity, according to which rays coming from stars close to the sun's limb should undergo a certain deflection."

"It was inevitable that the close of the war would see a revival of the discussion of the proposed tunnel under the English Channel, but there is this difference between the project in pre-war and post-war days. Today it has behind it all the driving force derived from the urgent need for such a tunnel that has been revealed during the 4½ years of the great war. League of Nations or no League of Nations, it is accepted on both sides of the Channel that the construction of the tunnel would greatly strengthen the alliance between the French and British nations. In all the years through which the agitation has been carried on, more or less intermittently, for the construction of this work, the scheme was handicapped, fatally handicapped, by the reluctance of the British people to permit the construction of a work which they feared would destroy their absolutely insular position. The objection was largely a sentimental one, and it has been completely obliterated by the bonds of good feeling that have been forged by the war. Furthermore, the idea that the tunnel would expose Great Britain to military invasion was largely a myth, for it would be a very simple matter in these days of high explosives and highly developed fuses and electric connections to flood the lower reaches of the tunnel at a moment's notice."

"It was only a short time ago that we could scarcely believe our eyes when we read that an automobile had made a speed of 120 miles an hour, or a mile in half a minute. With the advent of the airplane we grew quite accustomed to think of travel at speeds of 125 to 150 miles per hour, and so when we learned that Ralph de Palma, racing at Ormond Beach on Lincoln's Birthday, had made a mile in 24.04 seconds, we were not half as astonished as we really should have been. This figures out to nearly 150 miles per hour, or 149.8, to be exact. Few airplanes have made as high a speed as this. The car with which the record was smashed has a twin-six aviation-engine power plant."

SCIENTIFIC AMERICAN

MARCH, 1869: "Hon. Caleb Cushing has returned from the capital of Colombia, the most northern of the South American republics, whither he was sent by the Department of State, and the draft of a treaty he there negotiated for the right of way, etc., of a ship canal across the Isthmus of Darien, or Panama, is now before the Senate for ratification. We are assured that upon ratification of the treaty measures will at once be taken to begin what will, when completed, be the greatest work in importance of this century."

"Notwithstanding the failures which have uniformly attended the attempts to construct a useful flying machine, and the emphatic negative given by a large number of scientific writers to the question of whether a flying machine is a mechanical possibility, the belief in the ultimate accomplishment of flight by means of human devices has never lacked adherents among the learned and the unlearned. The organization of the Aeronautical Society, which gave its first exhibition at London last June, is evidence that the belief is gaining rather than losing ground. The council of the society recently voted their £100 prize to Mr. Stringfellow for a steam engine weighing only 16 pounds and able to work to one horse-power; whether or not it ever becomes the motive power

Report from

BELL LABORATORIES

A New Use for the Sun

Radio frequencies in the range of 10 to 40 GHz are of interest for satellite communications. But signals at these frequencies—including millimeter wavelengths at 30 to 40 GHz—are attenuated by rain, snow, fog, and other weather conditions. To study these effects, we needed a source of millimeter waves in the sky. What could we use?

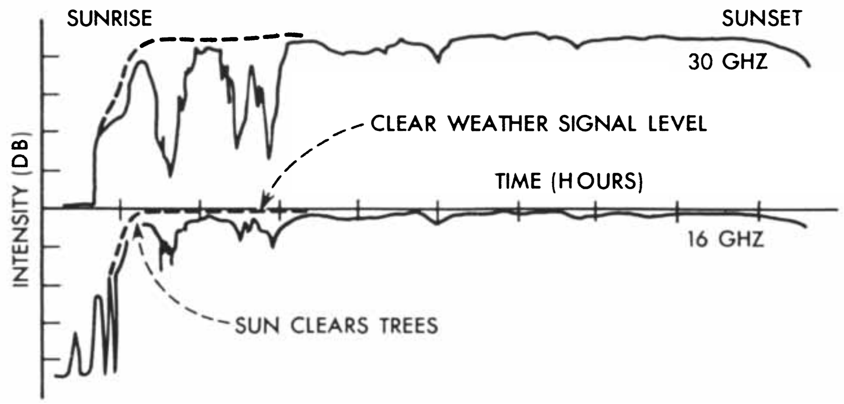
Putting up a research satellite would be difficult and expensive. But fortunately a millimeter-wave transmitter in the sky already exists ... the sun. By tracking the sun and correlating the signal losses with atmospheric effects, we are gaining valuable insight into the problems and possibilities of maintaining reliable communications at millimeter-wave frequencies in all kinds of weather.

Our apparatus includes a steerable, plane metal mirror that reflects solar energy into a stationary horn-reflector antenna (photo). The signals are processed and recorded in a temperature-controlled equipment house. The system automatically follows the sun on its daily path across the sky by a clock mechanism which can be set for a full week's automatic observation.

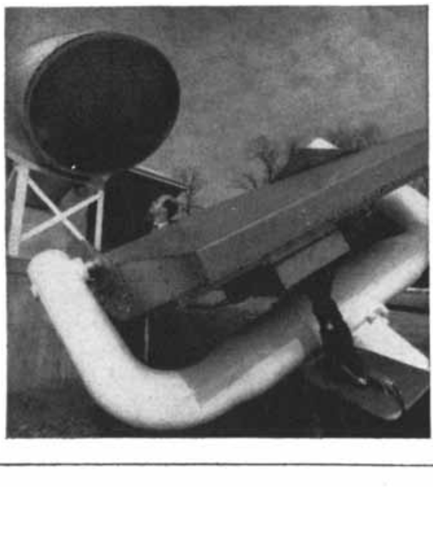
The sun emits radio signals at a great many frequencies, but the sun-tracker is tuned only to signals at 16 and 30 GHz. The received energy has two significant components: one due to the sun and one due to the atmosphere (which attenuates solar energy but also radiates energy of its own). To allow for the atmospheric component, we tilt our mirror away from the sun once each second, thus getting a sky-only reading. We subtract this from the sun/sky total and plot the difference. The equipment responds to and records signal changes as rapid as 30 dB in 15 seconds.

The graphs (left) indicate that there are periods when rain will attenuate the received signal by 30 to 40 dB. Such attenuation would seriously impair communications signals from satellites. The problem can be solved by spacing several ground terminals far enough apart that at least one of them would always have a clear path to the satellite. This would be done at both receiving and transmitting terminals.

Data from the sun-tracking experiments will help us determine how many terminals might be needed for communication systems operating at high frequencies, and how these terminals might be spaced to achieve maximum efficiency and economy.



Intensity of solar radiation at two frequencies, recently observed at Bell Laboratories' installation at Holmdel, N.J. Graphs show one 9-hour period out of months of study. Note attenuation due to rain early in the day.



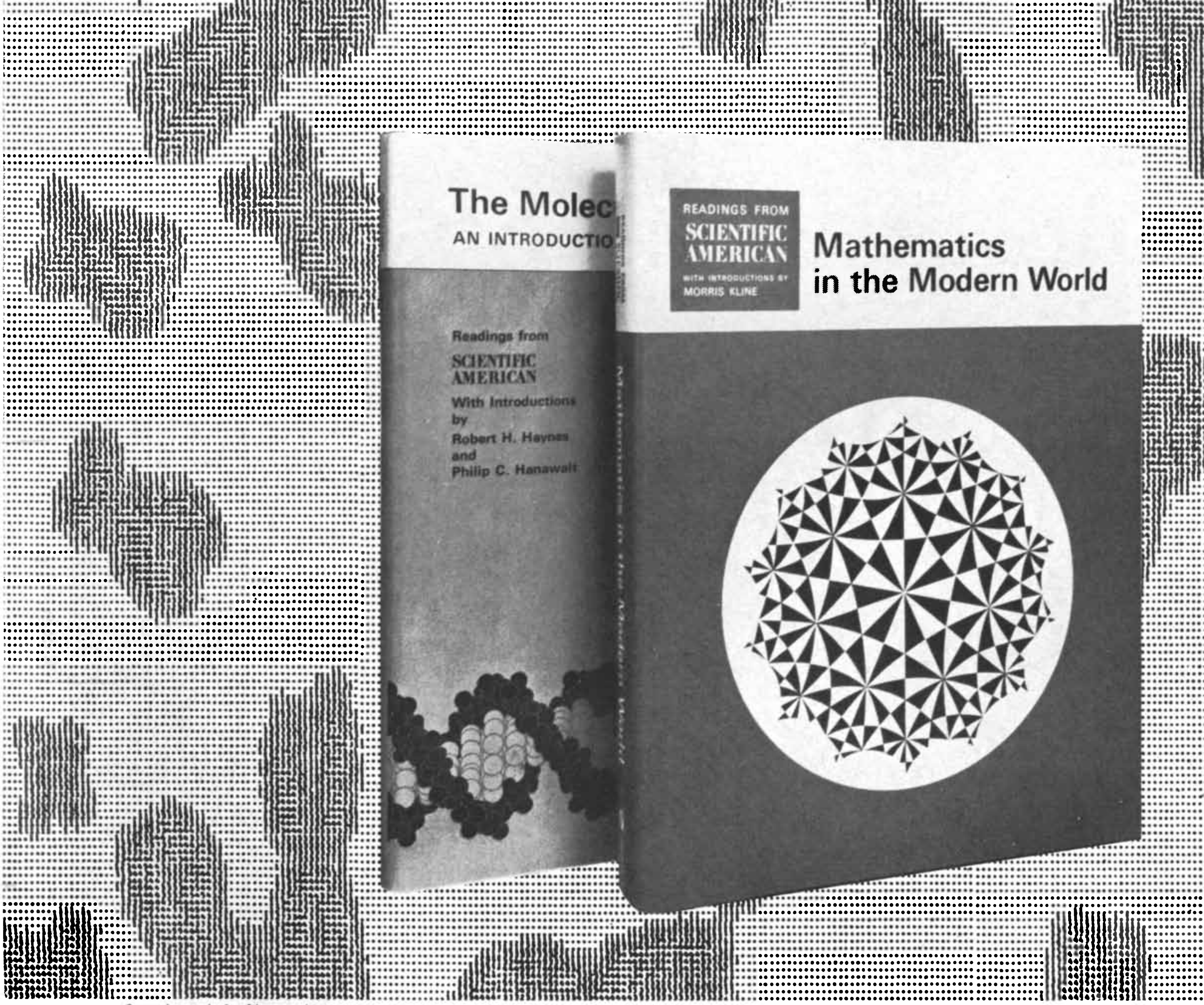
Bell Laboratories' station for studying atmospheric effects on solar radio energy. The flat mirror has two axes. One is parallel to the earth's polar axis—for daily rotation. The other is perpendicular—for seasonal variation. Experiment designer R. W. Wilson checks the horn antenna into which signals are reflected.

for flight, it would seem from its ingenuity to be well worth the reward. This engine appears to demonstrate the possibility of making engines light enough and powerful enough for the purposes of flight. The American wild goose frequently weighs more than this entire machine, boiler, propeller and all, and the power exerted by this bird in flight must be vastly less than that achieved by the engine."

"It must be acknowledged that the applications of aluminum in the arts are not so numerous as was at first predicted, and its manufacture, as compared with other metals, can at the present time hardly be called a metallurgical one. As soon as the metal is required in large quantities, however, some method will be devised for producing it at a cheap rate, and when that time arrives we shall not have to fit out expeditions to go and search for the ore in remote regions but can dig for it under our feet, nearly everywhere, and make a mine of every stone quarry. So abundant is this metal that it is safe to predict that the day is not far distant when our houses may be built of it instead of bricks, and we shall use it for many purposes now unknown."

"A company has been organized in London to tunnel from the Post Office to the marble arch entrance on Hyde Park. The trains of the proposed line are to be drawn by wire ropes from fixed engines at each end, so that the air of the tunnel will not be poisoned by the smoke and vapors of the locomotives; since there can be no collisions, trains will start every two minutes. In the opinion of competent engineers the substitution of locomotives for ropes was a mistake, whether regarded from the scientific or the economic point of view. Improved means of communication in cities is one of the greatest necessities of the day. We want, if possible, to get rid of the surface roads."

"Each ant in an anthill knows its companions. Mr. Darwin several times carried ants from one hill to another, inhabited apparently by tens of thousands of ants, but the strangers were invariably detected and killed. Thinking that there might be a family odor by which they were recognized, he put some ants from a very large nest into a bottle strongly perfumed with asafoetida and restored them after 24 hours. At first they were threatened by their companions, but they were soon recognized and allowed to pass."



Computer analysis of human chromosomes

So quiet, so smooth, that passengers don't even know there are engines on the plane.

It works on and on, with little care and only occasional feeding.

THE AUTHORS

JOHN R. CLARK ("Thermal Pollution and Aquatic Life") is assistant director of the Sandy Hook Marine Laboratory of the U.S. Bureau of Sport Fisheries and Wildlife. "My forte," he writes, "has truly been sea work; I've probably spent 1,000 days at sea in the past 20 years, and before that in Seattle I worked summers as a commercial fisherman to get through college." Clark was graduated from the University of Washington in 1949 with a degree in fisheries science. For 10 years after that he worked at the Fisheries Laboratory of the U.S. Bureau of Commercial Fisheries in Woods Hole, Mass. Clark was a founder of the American Littoral Society and is now its president. He wishes to note that his article expresses his personal views and is not to be regarded as an official statement of the U.S. Department of the Interior.

H. W. HAYDEN, R. C. GIBSON and J. H. BROPHY ("Superplastic Metals") are at the Paul D. Merica Research Laboratory of the International Nickel Company. Hayden, a research associate in the nonferrous group at the laboratory, was graduated from the Massachusetts Institute of Technology in 1960 and received his D.Sc. there in 1963. His outside interests, which include skiing, bridge, music and debating, have brought him an unusual distinction: he is an honorary parolee of Walpole State Prison in Massachusetts, because as a member of the M.I.T. debate team he participated in a debate with a prison team. Gibson, who is supervisor of the nickel alloys research section at the laboratory, was graduated from Lehigh University and obtained a master's degree in metallurgy at the Stevens Institute of Technology. Brophy, who is group leader of the nonferrous metallurgy group at the laboratory, received his bachelor's, master's and doctor's degrees from the University of Michigan. From 1958 to 1963 he was assistant professor of metallurgy at M.I.T.

NORMAN K. WESSELLS and WILLIAM J. RUTTER ("Phases in Cell Differentiation") are respectively associate professor of biological sciences at Stanford University and professor of biochemistry and genetics at the University of Washington. Wessells did both his undergraduate and his graduate work at Yale University and began teaching at Stanford in 1962. Rutter was graduated from Harvard College in 1949, went to the University of Utah for his master's degree and received his Ph.D. from the University of Illinois, where he taught from 1955 to 1965. Later this year he will become chairman of the department of biochemistry and biophysics at the San Francisco Medical Center of the University of California. The authors wish to thank a number of associates in the work they describe: Thomas G. Sanders, William R. Clark, William S. Bradshaw, John D. Kemp, William D. Ball, Leslie C. Brock, Julia H. Cohen and Jean Evans.

BJÖRN KURTÉN ("Continental Drift and Evolution") is a lecturer in paleontology at the University of Helsinki, where he obtained his Ph.D. in 1954. He has written a number of articles dealing with such subjects as fossil carnivores, dating of early man, late Tertiary and Quaternary stratigraphy, and aspects of population dynamics, evolutionary theory and paleobiogeography. His recent books include *Pleistocene Mammals of Europe* and *The Age of the Dinosaurs*.

REUBEN HERSH and RICHARD J. GRIEGO ("Brownian Motion and Potential Theory") are in the department of mathematics and statistics at the University of New Mexico; Hersh is professor and Griego assistant professor. Hersh was graduated from Harvard College with a degree in English literature. After four years as an office boy and editorial assistant with SCIENTIFIC AMERICAN and four years as a machinist he obtained his Ph.D. in mathematics from New York University. He has taught at New York University, Fairleigh Dickinson University and Stanford University. Griego was graduated from the University of New Mexico in 1961 and received his Ph.D. from the University of Illinois four years later. He taught for a year at the University of California at Riverside before taking his present position. In addition to his academic work at the University of New Mexico he is director of an enrichment program that seeks, with funds from the U.S. Office of Economic Opportunity, to help incoming freshmen from disadvantaged backgrounds.

VON R. ESHLEMAN ("The Atmospheres of Mars and Venus") is professor of electrical engineering at Stanford University and codirector of the Center for Radar Astronomy, a joint group of research workers from the university and the Stanford Research Institute. He was graduated from George Washington University and received advanced degrees in electrical engineering at Stanford in 1950 and 1952. Since then he has been at Stanford as a researcher and teacher, heading a number of projects related to radio propagation and radar studies of the upper atmosphere, the sun, the moon, the planets and interplanetary space. He is a member of the Lunar and Planetary Missions Board of the National Aeronautics and Space Administration and a consultant to other Government and industrial groups.

SOLOMON KADIS, THOMAS C. MONTIE and SAMUEL J. AJL ("Plague Toxin") are associated with the Albert Einstein Medical Center in Philadelphia. Kadis, who received his bachelor's degree at St. John's College in 1950, his master's degree at the University of Virginia in 1951 and his Ph.D. from Vanderbilt University in 1957, has been an associate member of the Einstein Medical Center since 1963. Montie, who is also an associate member, was graduated from Oberlin College in 1956 and obtained his master's degree and his Ph.D. at the University of Maryland. Ajl is director of research at the Einstein Center. Born in Poland, he came to the U.S. in 1939 at the age of 16. He was graduated from Brooklyn College in 1945 and received his Ph.D. from Iowa State College in 1949. His association with the Einstein Center began in 1960 after he had worked at the Washington University School of Medicine, the Walter Reed Army Medical Center and the National Science Foundation. Ajl also serves as professor of biology and microbiology at the Temple University School of Medicine.

GEORGE SHIERS ("The First Electron Tube") is a free-lance writer who also teaches in the adult education division of the Santa Barbara City College and in the extension service of the University of California. He recalls that as a boy in England he and "dozens of other boys in their early teens made triodes largely by hand during the early 1920's while employed by a pioneer tube manufacturer in London." Shiers is a senior member of the Institute of Electrical and Electronics Engineers. His article is based on material gathered for a book to be published by Prentice-Hall on the history of electronics.

SALVADOR E. LURIA, who in this issue reviews *Biology*, by Helena Curtis, is Sedgwick Professor of Biology at the Massachusetts Institute of Technology.

An aircraft carrier may be a very big ship but it's also a very small airport.

Over the past few years the Navy's planes have grown more and more complex. A lot more maintenance checks had to be made—and a lot more men and equipment were needed to make them.

All of which took more time and more space. The trouble is on an aircraft carrier you never have enough of either.

The advantage of UNIVAC® computer systems is they save on both.

The Navy worked with PRD Electronics to develop a computer-based Automatic Test System that would seek and trouble shoot equipment by zipping impulses through a plane's electronic package.

The system is called VAST—for Versatile Avionic Shop Test.

VAST will do routine aircraft maintenance in a fraction of the time taken by the equipment it replaces.

It will also take less than half the space.

It will cut down on the men needed by twenty-five percent.

VAST is easier to use so it will be easier to train men to use it.

And it can be shared by six different repair crews at the same time.

Univac systems are at work in many fields. In industry, science, education and government.

On five continents.

And the seven seas.

UNIVAC

Univac is saving a lot of people a lot of time.

SPERRY RAND

When it comes to expanding facilities, some airports are at a disadvantage.

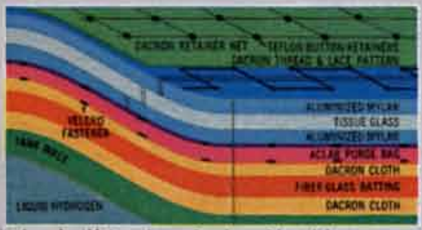
So the Navy uses computer systems to keep its planes shipshape.

Achieving critical cooling ...anywhere

• Multilayered "blanket" of ultra-thin sheets can insulate cryogenic propellants for long space trips. • Turbine fog is turned from nuisance to cooling agent in new air-cycle system. • Solid cryogens in tandem will work to refrigerate infrared detectors over year-long space missions. • Wind tunnel and computer help to solve unique high-heat problems for USAF's Mach 3+ SR-71.

The interplay of heating and cooling vitally affects the performance—and the life—of materials, fuels, instruments, and controls. Some of Lockheed's advanced ways of beating the heat with insulation and cooling techniques are briefly described here.

Thinning down a heat barrier. As space missions grow longer in time, the effectiveness of protective insulation for liquid cryogen propellants will become ever more crucial. Yet such insulation systems, while being highly reliable, should be ultra-thin and ultra-light so that they consume little of the spacecraft's precious allotments of volume and weight.



"Blanket" for insulating liquid cryogen propellants.

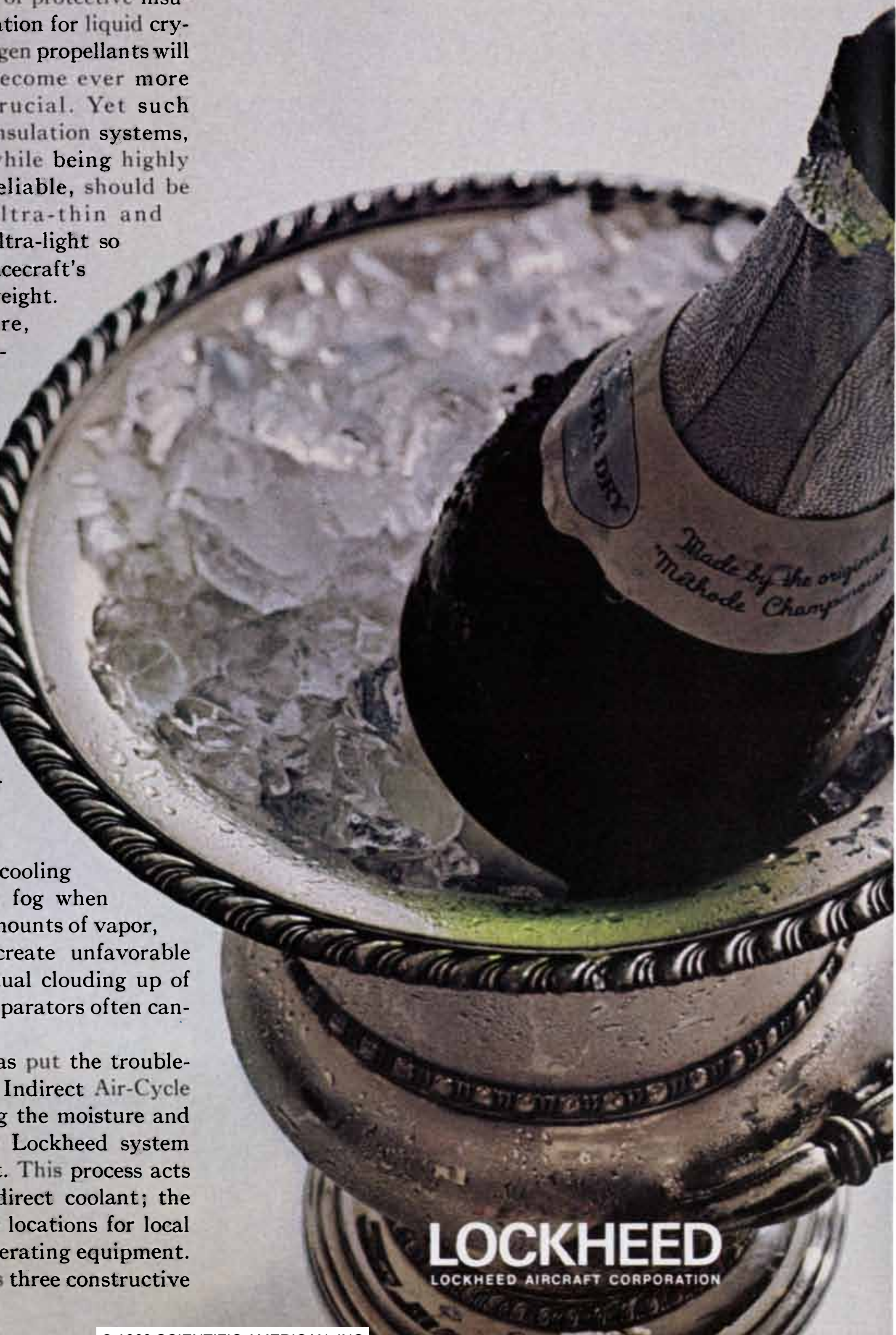
For exotic applications, therefore, Lockheed has developed an exceptionally thin, light, high-performance laminated system in which a multiple number of 0.00015-in.-thick aluminized plastic radiation shields and 0.0006-in.-thick fiber-glass paper separators are sandwiched in layers.

The resulting multilayered "blanket" has proved to be the most formidable heat barrier yet devised for shielding hypercold propellants like liquid hydrogen, oxygen and fluorine. And it promises to see many other space uses; for example, keeping instruments, optics, sensors, antennas and other equipment at optimum temperatures.

Harnessing an unwelcome fog. Present-day aircraft using air-cycle cooling systems can be plagued by internal fog when operating in humid climates. Dense amounts of vapor, discharged from cooling turbines, create unfavorable humidity for instruments and an actual clouding up of cabins. Unfortunately, so-called fog separators often cannot relieve the condition.

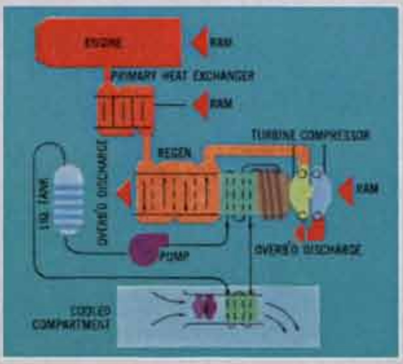
To beat this problem, Lockheed has put the trouble-making fog to good use in a unique Indirect Air-Cycle Cooling System. Instead of extracting the moisture and draining it overboard as waste, the Lockheed system intercepts the fog and reevaporates it. This process acts to reduce the temperature of an indirect coolant; the coolant, in turn, is plumbed to other locations for local cooling of occupied spaces or heat-generating equipment.

The Lockheed system thus performs three constructive



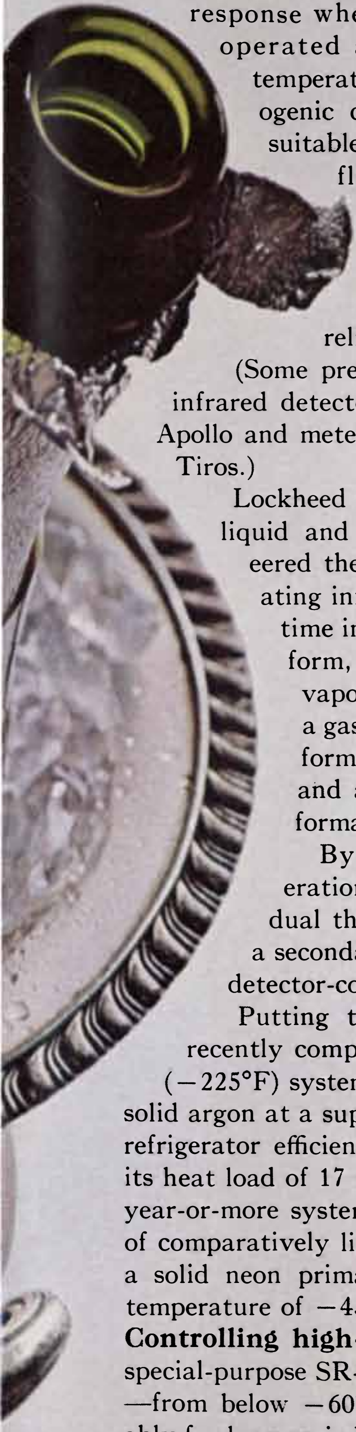
LOCKHEED
LOCKHEED AIRCRAFT CORPORATION

operations: It blocks visible fog from the cabin, utilizes the vapor's inherent cooling capacity, and prevents subfreezing fog from icing up the elements downstream of the turbine.



System uses turbine-generated fog to help cool remote areas of aircraft.

Refrigerating the refrigerants. Infrared detectors, critical sensors for spectral measurement and scanning devices used in space, show their best sensitivity and response when operated at temperatures below −279°F. This makes cryogenic cooling a necessity. But devising a suitable refrigeration system for long-time flights means solving a many-sided problem: The system must have low weight, volume, and power consumption...and the volatile, delicate cryogens must last and work reliably throughout the mission.



(Some present, typical applications for cooled infrared detectors are on such major programs as Apollo and meteorological satellites like Nimbus and Tiros.)

Lockheed scientists, after extensive research on liquid and solid cryogen characteristics, pioneered the use of cryogenic solids for refrigerating infrared detectors during long spans of time in space. The reason is basic: In liquid form, a cryogen has a given latent heat of vaporization—or phase transformation to a gaseous state. The same cryogen in solid form, however, has much higher density and a higher latent heat of phase transformation.

By applying solids in an advanced refrigeration concept, Lockheed has developed a dual thermal-protection technique in which a secondary solid cryogen guards the primary detector-cooler.

Putting these principles to work, Lockheed recently completed a solid-cryogen carbon dioxide (−225°F) system which acts as thermal protector for solid argon at a supercold temperature of −369°F. This refrigerator efficiently cools an infrared detector, with its heat load of 17 milliwatts, for a full year. And other year-or-more systems under development include those of comparatively light weight and low volume that use a solid neon primary coolant at a near-absolute-zero temperature of −432°F.

Controlling high-speed heats. The U.S. Air Force's special-purpose SR-71 endures harsh temperature cycles—from below −60°F to above +630°F—and flies reliably for long periods at searing Mach 3+ speeds. Every material and component in the aircraft must withstand environments hotter by far than ever experienced before in steady-state flight. And when Lockheed's Skunk Works started the SR-71 design, many of the complex thermodynamic problems had no precedents.

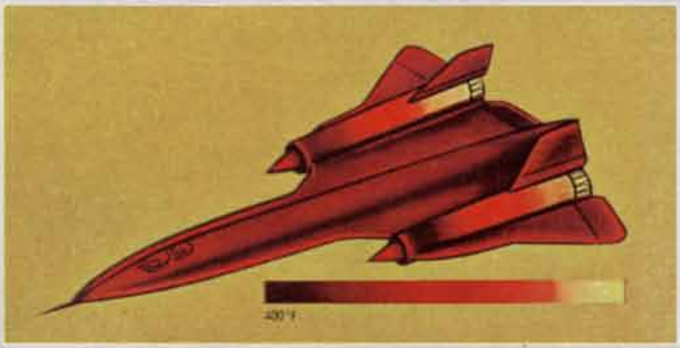
All aspects of internal and external thermodynamics were probed in the developmental studies—every tool was used from the straight-forward experimental approach in the wind tunnel to the complex analytical approach utilizing the maximum capability of the modern computer.

Wind tunnel coefficients helped project the external environment: High skin temperatures and stagnation ram air above 630°F. Because temperatures would be above aluminum alloy limits, titanium alloys were used for all skins and structure. A special high-emissivity black paint was applied to increase radiative heat loss and to help reduce skin temperatures. And, since the external skin is also the fuel tank skin, Lockheed created a unique tank sealant that would tolerate both severe flexing in flight and severe heat levels.

Accurate prediction of fuel temperature rise during high-speed flight was a must. Not only would the hydraulic and constant-speed drive systems be cooled by fuel but, for the first known time in an aircraft, fuel would serve as a direct heat sink for environmental control. The cooling performance gained through this bold approach provides a habitable cockpit atmosphere.

In the fuel tanks themselves, all equipment—pumps, vents, valves, switches, and wires—had to survive high-temperature operation as the tanks emptied. Furthermore, to prevent spontaneous ignition of vapor in the tanks, Lockheed also developed a fail-safe nitrogen inerting and pressurization system.

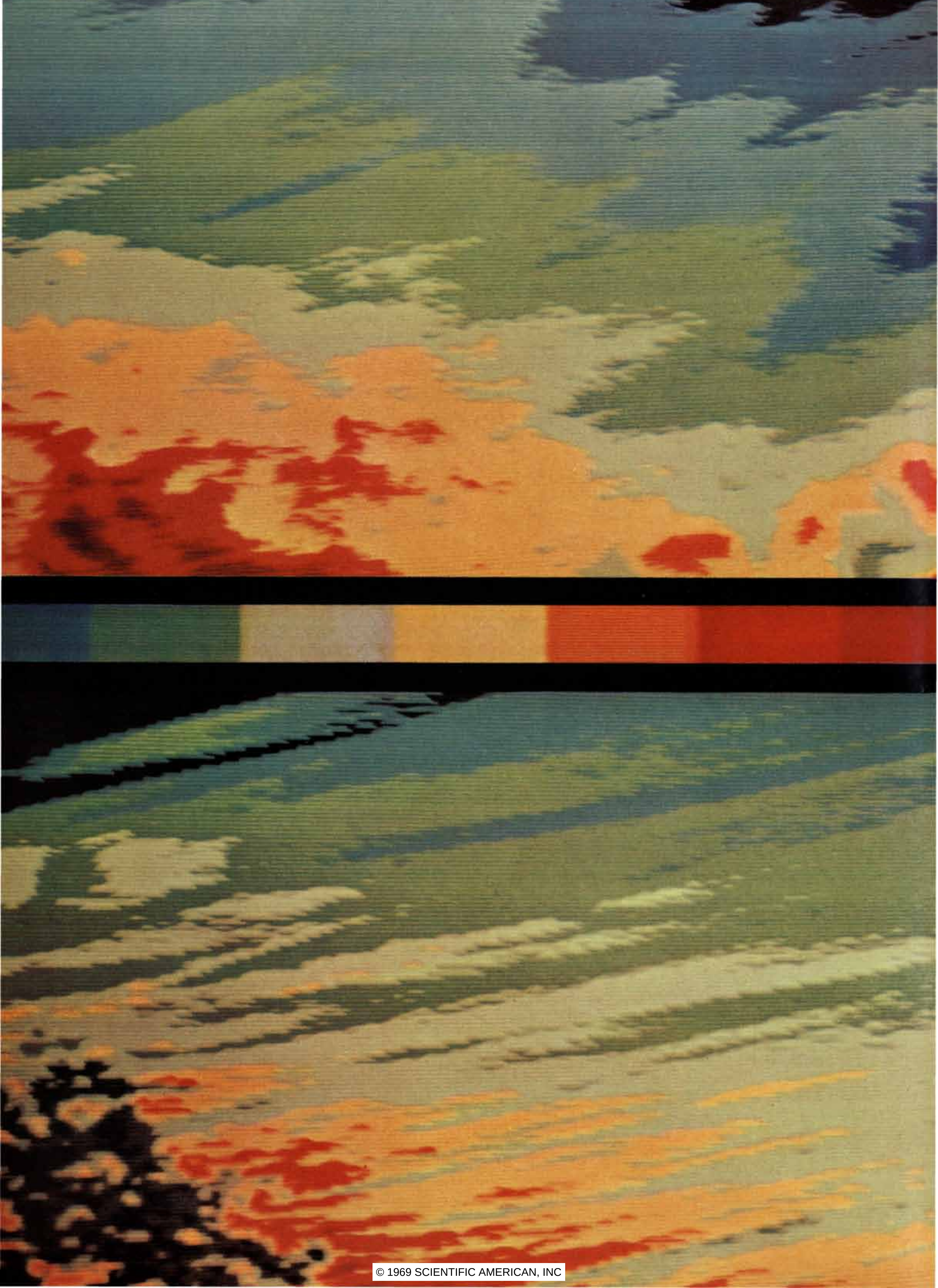
The hottest spots are the engine, its inlet system, and the nacelle. The engine inlet system is exposed entirely


Temperature profile of the SR-71 at Mach 3+ speeds.

to stagnation air temperatures above 630°F. Within the nacelle, the engine generates even higher temperatures. In addition to using all external radiation available, the engine is cooled sufficiently during prolonged missions by passing the inlet cowl's low-energy boundary layer air between the engine exterior and the nacelle skin.

The varied SR-71 cooling techniques have proved thoroughly effective. A typical example: Many high-precision components—servo valves, actuators, electrical connectors, and others—operate reliably in the engine inlet and nacelle at temperatures that remain, during long flights, far above normally destructive values.

The activities mentioned here are only a few of Lockheed's projects in environmental controls. If you are an engineer or scientist interested in this field of work, Lockheed invites your inquiry. Write: K. R. Kiddoo, Lockheed Aircraft Corporation, Burbank, California 91503. An equal opportunity employer.



Established 1845 SCIENTIFIC AMERICAN March 1969 Volume 220 Number 3

Thermal Pollution and Aquatic Life

The increasing use of river and lake waters for industrial cooling presents a real threat to fish and other organisms. To avoid an ecological crisis new ways must be found to get rid of waste heat

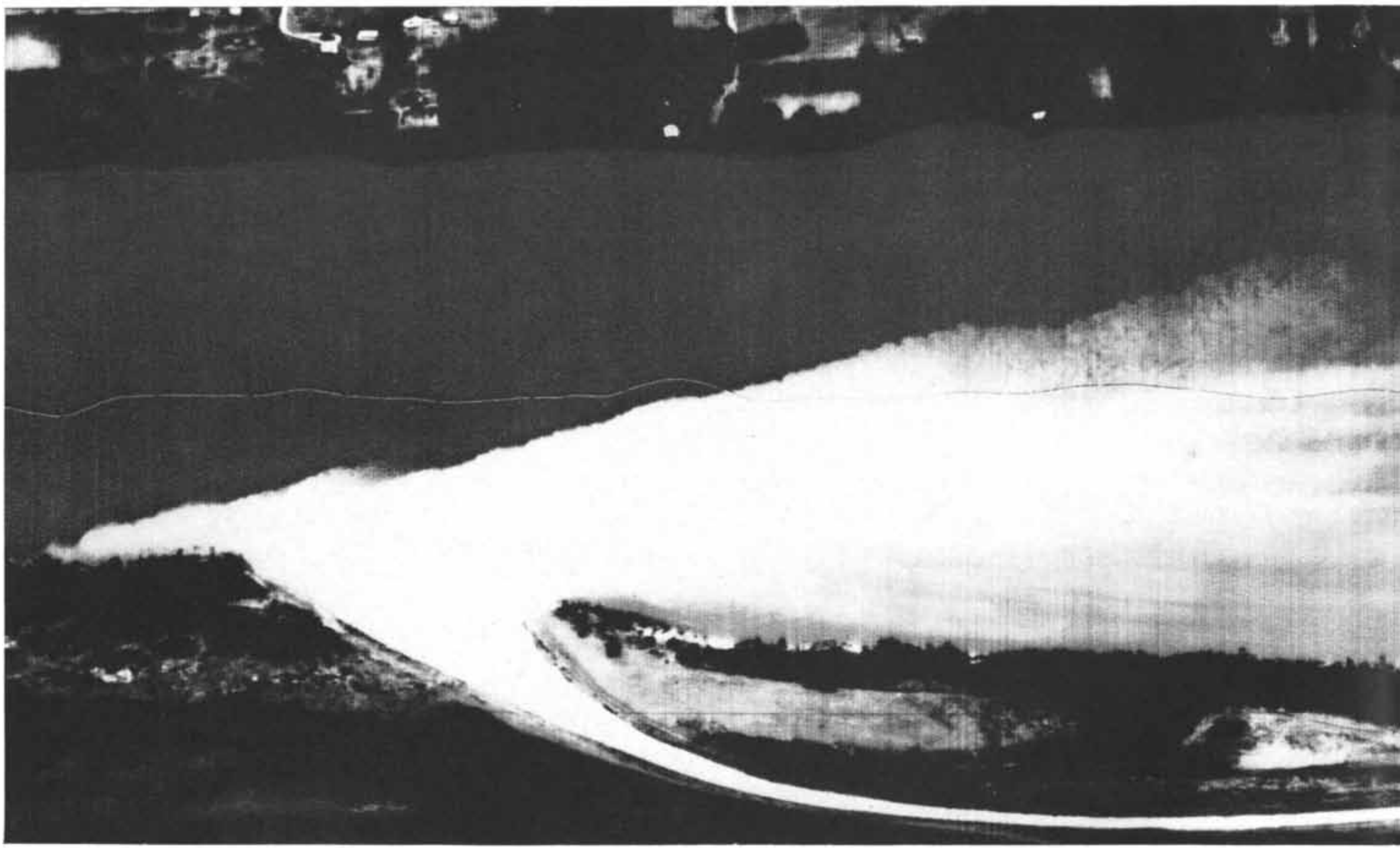
by John R. Clark

Ecologists consider temperature the primary control of life on earth, and fish, which as cold-blooded animals are unable to regulate their body temperature, are particularly sensitive to changes in the thermal environment. Each aquatic species becomes adapted to the seasonal variations in temperature of the water in which it lives, but it cannot adjust to the shock of abnormally abrupt change. For this reason there is growing concern among ecologists about the heating of aquatic habitats by man's activities. In the U.S. it appears that the use of river, lake and estuarine waters for industrial cooling purposes may become so extensive in future decades as to pose a considerable threat to fish and to aquatic life in general. Because of the potential hazard to life and to the balance of nature, the discharge of waste heat into the natural waters is coming to be called thermal pollution.

HEATED EFFLUENT from a power plant on the Connecticut River is shown in color thermograms (*opposite page*), in which different temperatures are represented by different hues. At the site (*above*) three large pipes discharge heated water that spreads across the river at slack tide (*top thermogram*) and tends to flow downriver at ebb tide (*bottom*). An infrared camera made by the Barnes Engineering Company scans the scene and measures the infrared radiation associated with the temperature at each point in the scene (350 points on each of 180 horizontal lines). Output from an infrared radiometer drives a color modulator, thus changing the color of a beam of light that is scanned across a color film in synchrony with the scanning of the scene. Here the coolest water (*black*) is at 59 degrees Fahrenheit; increasingly warm areas are shown, in three-degree steps, in blue, light blue, green, light green, yellow, orange, red and magenta. The effluent temperature was 87 degrees. A tree (*dark object*) appears in lower thermogram because the camera was moved.

The principal contributor of this heat is the electric-power industry. In 1968 the cooling of steam condensers in generating plants accounted for about three-quarters of the total of 60,000 billion gallons of water used in the U.S. for industrial cooling. The present rate of heat discharge is not yet of great consequence except in some local situations; what has aroused ecologists is the ninefold expansion of electric-power production that is in prospect for the coming years with the increasing construction of large generating plants fueled by nuclear energy. They waste 60 percent more energy than fossil-fuel plants, and this energy is released as heat in condenser-cooling water. It is estimated that within 30 years the electric-power industry will be producing nearly two million megawatts of electricity, which will require the disposal of about 20 million billion B.T.U.'s of waste heat per day. To carry off that heat by way of natural waters would call for a flow through power plants amounting to about a third of the average daily freshwater runoff in the U.S.

The Federal Water Pollution Control



NUCLEAR POWER PLANT at Haddam on the Connecticut River empties up to 370,000 gallons of coolant water a minute through a discharge canal (*bottom*) into the river. In this aerial thermogram, made by HRB-Singer, Inc., for the U.S. Geological Survey's

Administration has declared that waters above 93 degrees Fahrenheit are essentially uninhabitable for all fishes in the U.S. except certain southern species. Many U.S. rivers already reach a temperature of 90 degrees F. or more in summer through natural heating alone. Since the waste heat from a single power plant of the size planned for the future (some 1,000 megawatts) is expected to raise the temperature of a river carrying a flow of 3,000 cubic feet per second by 10 degrees, and since a number of industrial and power plants are likely to be constructed on the banks of a single river, it is obvious that many U.S. waters would become uninhabitable.

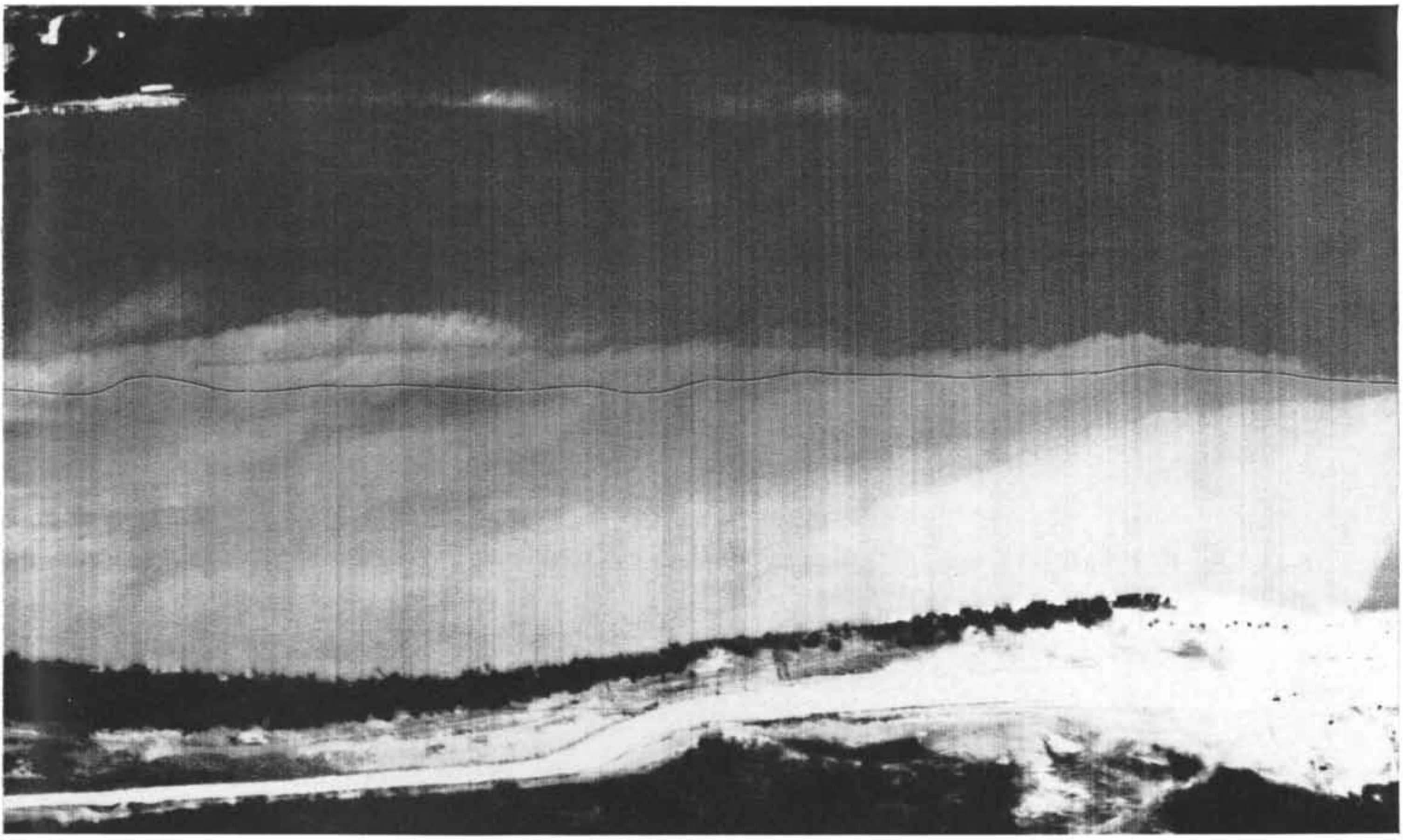
A great deal of detailed information is available on how temperature affects the life processes of animals that live in the water. Most of the effects stem from the impact of temperature on the rate of metabolism, which is speeded up by heat in accordance with the van't Hoff principle that the rate of chemical reaction increases with rising temperature. The acceleration varies considerably for particular biochemical reactions and in different temperature ranges, but generally speaking the metabolic rate doubles with each increase of 10 degrees Celsius (18 degrees F.).

Since a speedup of metabolism increases the animal's need for oxygen, the rate of respiration must rise. F. E. J. Fry of the University of Toronto, experimenting with fishes of the salmon family, found that active fish increased their oxygen consumption as much as fourfold as the temperature of the water was raised to the maximum at which they could survive. In the brown trout the rate of oxygen consumption rose steadily until the lethal temperature of 79 degrees F. was reached; in a species of lake trout, on the other hand, the rate rose to a maximum at about 60 degrees and then fell off as the lethal temperature of 77 degrees was approached. In both cases the fishes showed a marked rise in the basal rate of metabolism up to the lethal point.

The heart rate often serves as an index of metabolic or respiratory stress on the organism. Experiments with the crayfish (*Astacus*) showed that its heart rate increased from 30 beats per minute at a water temperature of 39 degrees F. to 125 beats per minute at 72 degrees and then slowed to a final 65 beats per minute as the water approached 95 degrees, the lethal temperature for this crustacean. The final decrease in heartbeat is evidence of the animal's weakening under the thermal stress.

At elevated temperatures a fish's respiratory difficulties are compounded by the fact that the hemoglobin of its blood has a reduced affinity for oxygen and therefore becomes less efficient in delivering oxygen to the tissues. The combination of increased need for oxygen and reduced efficiency in obtaining it at rising temperatures can put severe stress even on fishes that ordinarily are capable of living on a meager supply of oxygen. For example, the hardy carp, which at a water temperature of 33 degrees F. can survive on an oxygen concentration as low as half a milligram per liter of water, needs a minimum of 1.5 milligrams per liter when the temperature is raised to 95 degrees. Other fishes can exist on one to two milligrams at 39 degrees but need three to four milligrams merely to survive at 65 degrees and five milligrams for normal activity.

The temperature of the water has pronounced effects on appetite, digestion and growth in fish. Tracer experiments



Water Resources Division, temperature is represented by shades of gray. The hot effluent (*white*) is at about 93 degrees F.; ambient river temperature (*dark gray*) is 77 degrees. The line across the thermogram is a time marker for a series of absolute measurements.

with young carp, in which food was labeled with color, established that they digest food four times as rapidly at 79 degrees F. as they do at 50 degrees; whereas at 50 degrees the food took 18 hours to pass through the alimentary canal, at 79 degrees it took only four and a half hours.

The effects of temperature in regulating appetite and the conversion of the food into body weight can be used by hatcheries to maximize fish production in terms of weight. The food consumption of the brown trout, for example, is highest in the temperature range between 50 and 66 degrees. Within that range, however, the fish is so active that a comparatively large proportion of its food intake goes into merely maintaining its body functions. Maximal conversion of the food into a gain in weight occurs just below and just above that temperature range. A hatchery can therefore produce the greatest poundage of trout per pound of food by keeping the water temperature at just under 50 degrees or just over 66 degrees.

It is not surprising to find that the activity, or movement, of fish depends considerably on the water temperature. By and large aquatic animals tend to raise their swimming speed and to show more spontaneous movement as the temperature rises. In many fishes the temperature-dependent pattern of activity is rather complex. For instance, the sockeye salmon cruises twice as fast in water at 60 degrees as it does in water at 35 degrees, but above 60 degrees its speed declines. The brook trout shows somewhat more complicated behavior: it increases its spontaneous activity as the temperature rises from 40 to 48 degrees, becomes less active between 49 and 66 degrees and above 66 degrees again goes into a rising tempo of spontaneous movements up to the lethal temperature of 77 degrees. Laboratory tests show that a decrease in the trout's swimming speed potential at high temperatures affects its ability to feed. By 63 degrees trout have slowed down in pursuing minnows, and at 70 degrees they are almost incapable of catching the minnows.

That temperature plays a critical role in the reproduction of aquatic animals is well known. Some species of fish spawn during the fall, as temperatures drop; many more species, however, spawn in the spring. The rising temperature induces a seasonal development of their gonads and then, at a critical point, triggers the female's deposit of her eggs in the water. The triggering is particularly dramatic in estuarine shellfish (oysters and clams), which spawn within a few hours after the water temperature reaches the critical level. Temperature also exerts a precise control over the time it takes a fish's eggs to hatch. For example, fertilized eggs of the Atlantic salmon will hatch in 114 days in water at 36 degrees F. but the period is shortened to 90 days at 45 degrees; herring eggs hatch in 47 days at 32 degrees and in eight days at 58 degrees; trout eggs hatch in 165 days at 37 degrees and in 32 days at 54 degrees. Excessive temperatures, however, can prevent normal development of eggs. The Oregon Fish Commission has declared that a rise of 5.4 degrees in the Columbia River could be disastrous for the eggs of the Chinook salmon.

Grace E. Pickford of Yale University has observed that "there are critical temperatures above or below which fish will not reproduce." For instance, at a temperature of 72 degrees or higher the banded sunfish fails to develop eggs. In the case of the carp, temperatures in the range of 68 to 75 degrees prevent cell division in the eggs. The possum shrimp *Neomysis*, an inhabitant of estuaries, is blocked from laying eggs if the temperature rises above 45 degrees. There is also the curious case of the tiny crustacean *Gammarus*, which at temperatures above 46 degrees produces only female offspring.

Temperature affects the longevity of fish as well as their reproduction. D'Arcy Wentworth Thompson succinctly stated this general life principle in his classic *On Growth and Form:* "As the several stages of development are accelerated by warmth, so is the duration of each and all, and of life itself, proportionately curtailed. High temperature may lead to a short but exhausting spell of rapid growth, while the slower rate manifested at a lower temperature may be the best in the end." Thompson's principle has been verified in rather precise detail by experiments with aquatic crustaceans. These have shown, for example, that *Daphnia* can live for 108 days at 46 degrees F. but its lifetime at 82 degrees is 29 days; the water flea *Moina* has a lifetime of 14 days at 55 degrees, its optimal temperature for longevity, but only five days at 91 degrees.

Other effects of temperature on life processes are known. For example, a century ago the German zoologist Karl Möbius noted that mollusks living in cold waters grew more slowly but attained larger size than their cousins living in warmer waters. This has since been found to be true of many fishes and other water animals in their natural habitats.

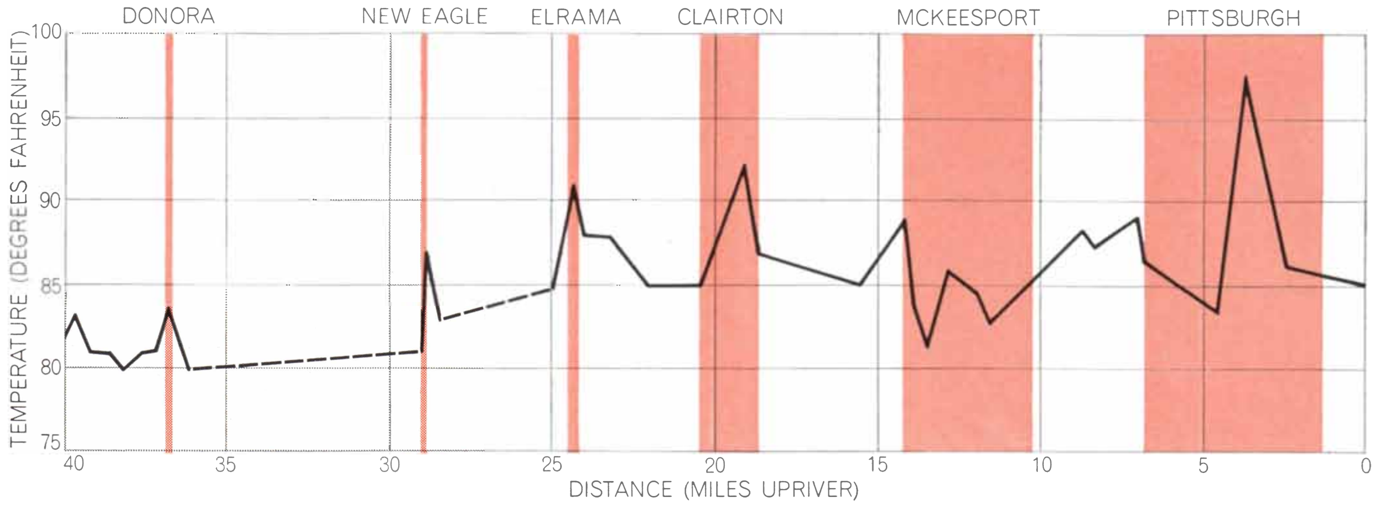
Fortunately fish are not entirely at the mercy of variations in the water temperature. By some process not yet understood they are able to acclimate themselves to a temperature shift if it is moderate and not too sudden. It has been found, for instance, that when the eggs of largemouth bass are suddenly transferred from water at 65 or 70 degrees to water at 85 degrees, 95 percent of the eggs perish, but if the eggs are acclimated by gradual raising of the temperature to 85 degrees over a period of 30 to 40 hours, 80 percent of the eggs will survive. Experiments with the possum shrimp have shown that the lethal temperature for this crustacean can be raised by as much as 24 degrees (to a high of 93 degrees) by acclimating it through a series of successively higher temperatures. As a general rule aquatic animals can acclimate to elevated temperatures faster than they can to lowered temperatures.

Allowing for maximum acclimation (which usually requires spreading the gradual rise of temperature over 20 days), the highest temperatures that most fishes of North America can tolerate range from about 77 to 97 degrees F. The direct cause of thermal death is not known in detail; various investigators have suggested that the final blow may be some effect of heat on the nervous system or the respiratory system, the coagulation of the cell protoplasm or the inactivation of enzymes.

Be that as it may, we need to be concerned not so much about the lethal temperature as about the temperatures that may be *unfavorable* to the fish. In the long run temperature levels that adversely affect the animals' metabolism, feeding, growth, reproduction and other vital functions may be as harmful to a fish population as outright heat death. Studies of the preferred, or optimal, temperature ranges for various fishes have been made in natural waters and in the laboratory [*see illustration on page 24*]. For adult fish observed in nature the preferred level is about 13 degrees F. below the lethal temperature on the average; in the laboratory, where the experimental subjects used (for convenience) were very young fish, the preferred level was 9.5 degrees below the lethal temperature. Evidently young fish need warmer waters than those that have reached maturity do.

The optimal temperature for any water habitat depends not only on the preferences of individual species but also on the well-being of the system as a whole. An ecological system in dynamic balance is like a finely tuned automobile engine, and damage to any component can disable or impair the efficiency of the entire mechanism. This means that if we are to expect a good harvest of fish, the temperature conditions in the water medium must strike a favorable balance for all the components (algae and other plants, small crustaceans, bait fishes and so on) that constitute the food chain producing the harvested fish. For example, above 68 degrees estuarine eelgrass does not reproduce. Above 90 degrees there is extensive loss of bottom life in rivers.

So far there have been few recorded instances of direct kills of fish by thermal pollution in U.S. waters. One recorded kill occurred in the summer of 1968 when a large number of menhaden acclimated to temperatures in the 80's became trapped in effluent water at 93 to 95 degrees during the testing of a new power plant on the Cape Cod Canal. A very large kill of striped bass occurred



WATER TEMPERATURES can become very high, particularly in summer, along rivers with concentrated industry. The chart shows the temperature of the Monongahela River, measured in August, along a 40-mile stretch upriver from its confluence with the Ohio.

in the winter and early spring of 1963 at the nuclear power plant at Indian Point on the Hudson River. In that instance the heat discharge from the plant was only a contributing factor. The wintering, dormant fish, attracted to the warm water issuing from the plant, became trapped in its structure for water intake, and they died by the thousands from fatigue and other stresses. (Under the right conditions, of course, thermal discharges benefit fishermen by attracting fish to discharge points, where they can be caught with a hook and line.)

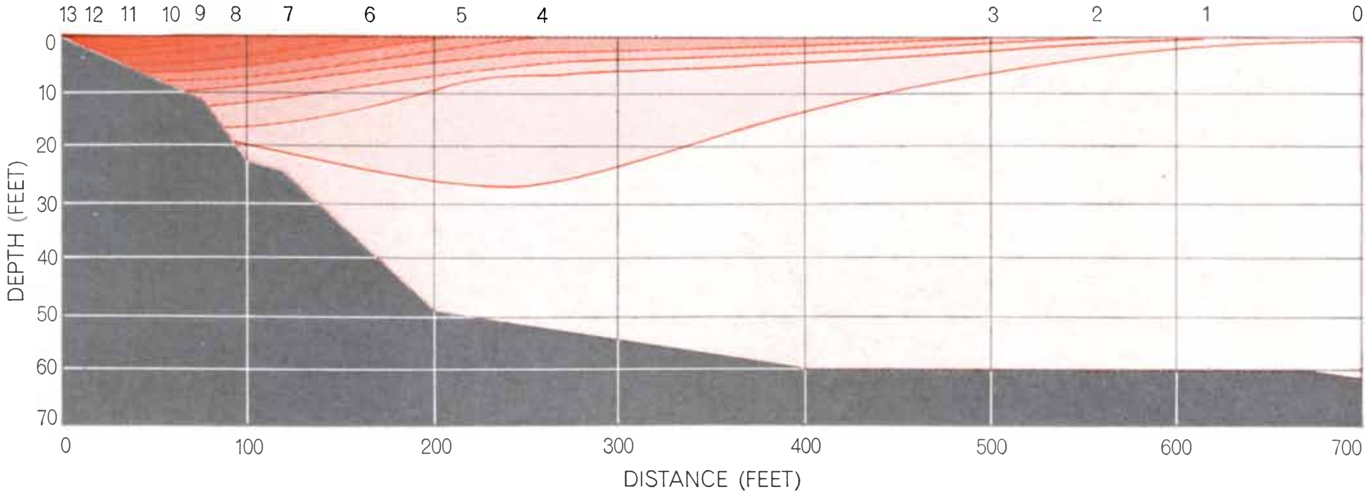
Although direct kills attributable to thermal pollution apparently have been rare, there are many known instances of deleterious effects on fish arising from natural summer heating in various U.S. waters. Pollution by sewage is often a contributing factor. At peak summer temperatures such waters frequently generate a great bloom of plankton that depletes the water of oxygen (by respiration while it lives and by decay after it dies). In estuaries algae proliferating in the warm water can clog the filtering apparatus of shellfish and cause their death. Jellyfish exploding into abundant growth make some estuarine waters unusable for bathing or other water sports, and the growth of bottom plants in warm waters commonly chokes shallow bays and lakes. The formation of hydrogen sulfide and other odorous substances is enhanced by summer temperatures. Along some of our coasts in summer "red tides" of dinoflagellates occasionally bloom in such profusion that they not only bother bathers but also may poison fish. And where both temperature and sewage concentrations are high a heavy toll of fish may be taken by proliferating microbes.

This wealth of evidence on the sensitivity of fish and the susceptibility of aquatic ecosystems to disruption by high temperatures explains the present concern of biologists about the impending large increase in thermal pollution. Already last fall 14 nuclear power plants, with a total capacity of 2,782 megawatts, were in operation in the U.S.; 39 more plants were under construction, and 47 others were in advanced planning stages. By the year 2000 nuclear plants are expected to be producing about 1.2 million megawatts, and the nation's total electricity output will be in the neighborhood of 1.8 million megawatts. As I have noted, the use of natural waters to cool the condensers would entail the heating of an amount of water equivalent to a third of the yearly freshwater runoff in the U.S.; during low-flow periods in summer the requirement would be 100 percent of the runoff. Obviously thermal pollution of the waters on such a scale is neither reasonable nor feasible. We must therefore look for more efficient and safer methods of dissipating the heat from power plants.

One might hope to use the heated water for some commercial purpose. Unfortunately, although dozens of schemes have been advanced, no practicable use has yet been found. Discharge water is not hot enough to heat buildings. The cost of transmission rules out piping it to farms for irrigation even if the remaining heat were enough to improve crop production. More promising is the idea of using waste heat in desalination plants to aid in the evaporation process, but this is still only an idea. There has also been talk of improving the efficiency of sewage treatment with waste heat from power plants. Sea farming may offer the best hope of someday providing a needed outlet for discharges from coastal power plants; pilot studies now in progress in Britain and the U.S. are showing better growth of fish and shellfish in heated waters than in normal waters, but no economically feasible scheme has yet emerged. It appears, therefore, that for many years ahead we shall have to dispose of waste heat to the environment.

The dissipation of heat can be facilitated in various ways by controlling the passage of the cooling water through the condensers. The prevailing practice is to pump the water (from a river, a lake or an estuary) once through the steam-condensing unit, which in a 500-megawatt plant may consist of 400 miles of one-inch copper tubing. The water emerging from the unit has been raised in temperature by an amount that varies from 10 to 30 degrees F., depending on the choice of manageable factors such as the rate of flow. This heated effluent is then discharged through a channel into the body of water from which it was taken. There the effluent, since it is warmer and consequently lighter than the receiving water, spreads in a plume over the surface and is carried off in the direction of the prevailing surface currents.

The ensuing dispersal of heat through the receiving water and into the atmosphere depends on a number of natural factors: the speed of the currents, the turbulence of the receiving water (which affects the rate of mixing of the effluent with it), the temperature difference between the water and the air, the humidity of the air and the speed and direction of the wind. The most variable and most important factor is wind: other things being equal, heat will be dissipated from



HOT-WATER "PLUME" that would result from an Indian Point nuclear power plant mixes with cooler Hudson River as shown by the one-degree contour lines. This section across the river shows temperature structure that was predicted by engineering studies.

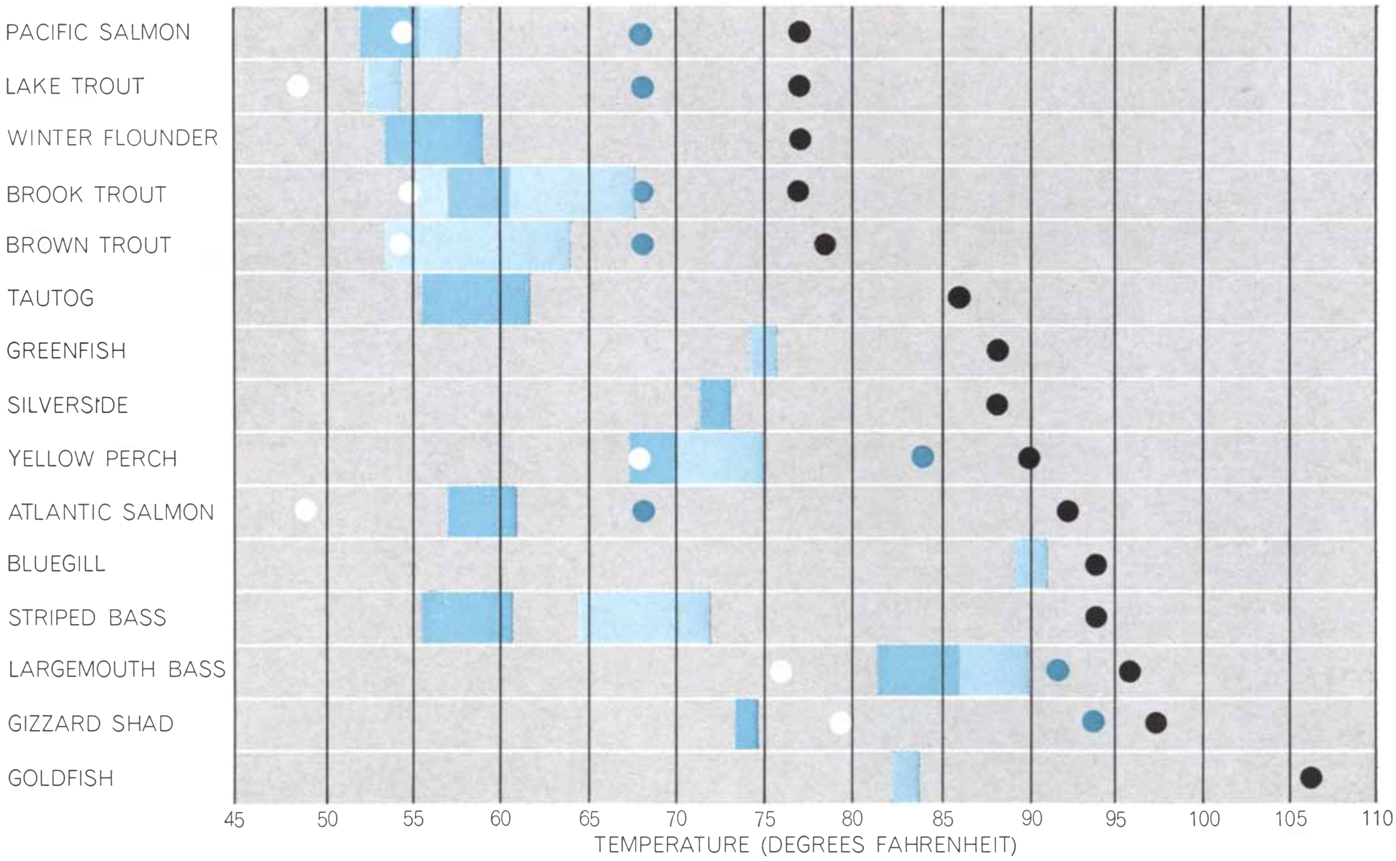
the water to the air by convection three times faster at a wind speed of 20 miles per hour than at a wind speed of five miles.

In regulating the rate of water flow through the condenser one has a choice between opposite strategies. By using a rapid rate of flow one can spread the heat through a comparatively large volume of cooling water and thus keep down the temperature of the effluent; conversely, with a slow rate of flow one can concentrate the heat in a smaller volume of coolant. If it is advantageous to obtain good mixing of the effluent with the receiving water, the effluent can be discharged at some depth in the water rather than at the surface. The physical and ecological nature of the body of water will determine which of these strategies is best in a given situation. Where the receiving body is a swift-flowing river, rapid flow through the condenser and dispersal of the low-temperature effluent in a narrow plume over the water surface may be the most effective way to dissipate the heat into the atmosphere. In the case of a still lake it may be best to use a slow flow through the condenser so that the comparatively small volume of effluent at a high temperature will be confined to a small area in the lake and still transfer its heat to the atmosphere rapidly because of the high temperature differential. And at a coastal site the best strategy may be to discharge an effluent of moderate temperature well offshore, below the ocean-water surface.

There are many waters, however, where no strategy of discharge will avail to make the water safe for aquatic life (and where manipulation of the discharge will also be insufficient to avoid dangerous thermal pollution), particularly where a number of industrial and power plants use the same body of water for cooling purposes. It therefore appears that we shall have to turn to extensive development of devices such as artificial lakes and cooling towers.

Designs for such lakes have already been drawn up and implemented for plants of moderate size. For the 1,000-megawatt power plant of the future a lake with a surface area of 1,000 to 2,000 acres would be required. (A 2,000-acre lake would be a mile wide and three miles long.) The recommended design calls for a lake only a few feet deep at one end and sloping to a depth of 50 feet at the other end. The water for cooling is drawn from about 30 feet below the surface at the deep end and is pumped through the plant at the rate of 500,000 gallons per minute, and the effluent, 20 degrees higher in temperature than the inflow, is discharged at the lake's shallow end. The size of the lake is based on a pumping rate through the plant of 2,000 acre-feet a day, so that all the water of the lake (averaging 15 feet in depth) is turned over every 15 days. Such a lake would dissipate heat to the air at a sufficient rate even in prolonged spells of unfavorable weather, such as high temperature and humidity with little wind.

Artificial cooling lakes need a steady inflow of water to replace evaporation and to prevent an excessive accumulation of dissolved material. This replenishment can be supplied by a small stream flowing into the lake. The lake itself can be built by damming a natural land basin. A lake complex constructed to serve not only for cooling but also for fishing might consist of two sections: the smaller one, in which the effluent is discharged, would be stocked with fishes tolerant to heat, and the water from this basin, having been cooled by exposure to



FISHES VARY WIDELY in temperature preference. Preferred temperature ranges are shown here for some species for which they have been determined in the field (*dark colored bands*) and, generally for younger fish, in the laboratory (*light color*). The chart also indicates the upper lethal limit (*black dot*) and upper limits recommended by the Federal Water Pollution Control Administration for satisfactory growth (*colored dot*) and spawning (*white dot*). Temperatures well below lethal limit can be in stress range.

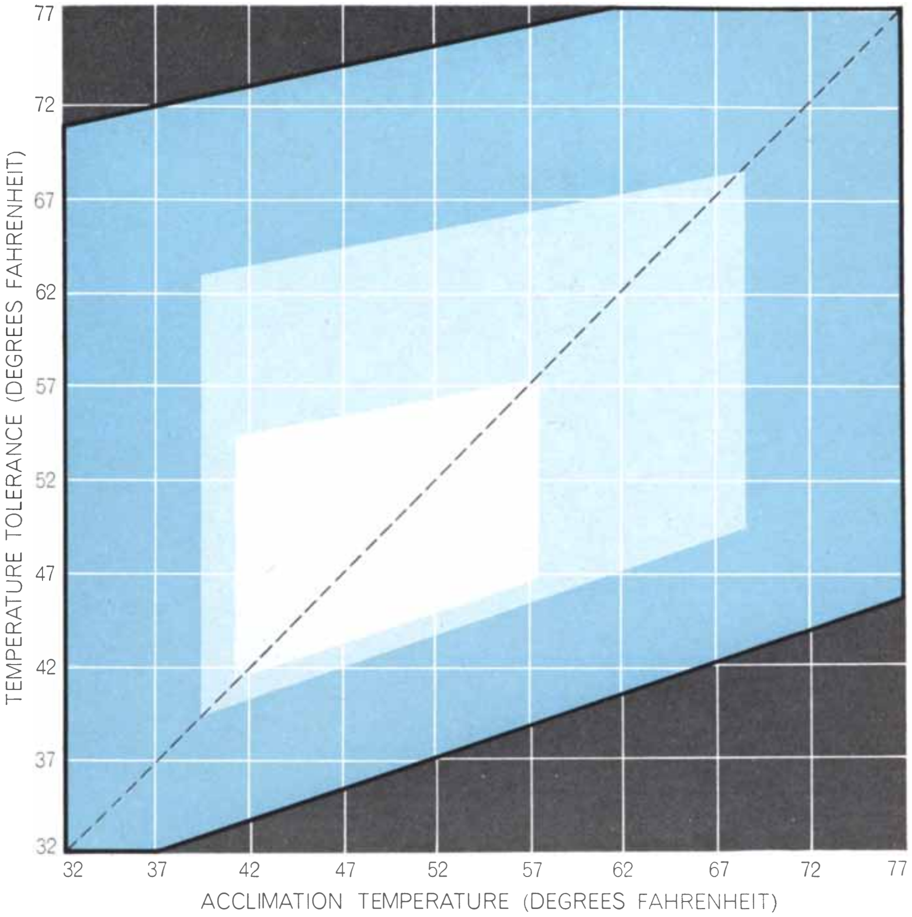
the air, would then flow into a second and larger lake where other fishes could thrive.

In Britain, where streams are small, water is scarce and appreciation of aquatic life is high, the favored artificial device for getting rid of waste heat from power plants has been the use of cooling towers. One class of these towers employs the principle of removing heat by evaporation. The heated effluent is discharged into the lower part of a high tower (300 to 450 feet) with sloping sides; as the water falls in a thin film over a series of baffles it is exposed to the air rising through the tower. Or the water can be sprayed into the tower as a mist that evaporates easily and cools quickly. In either case some of the water is lost to the atmosphere; most of it collects in a basin and is pumped into the waterway or recirculated to the condensers. The removal of heat through evaporation can cool the water by about 20 degrees F.

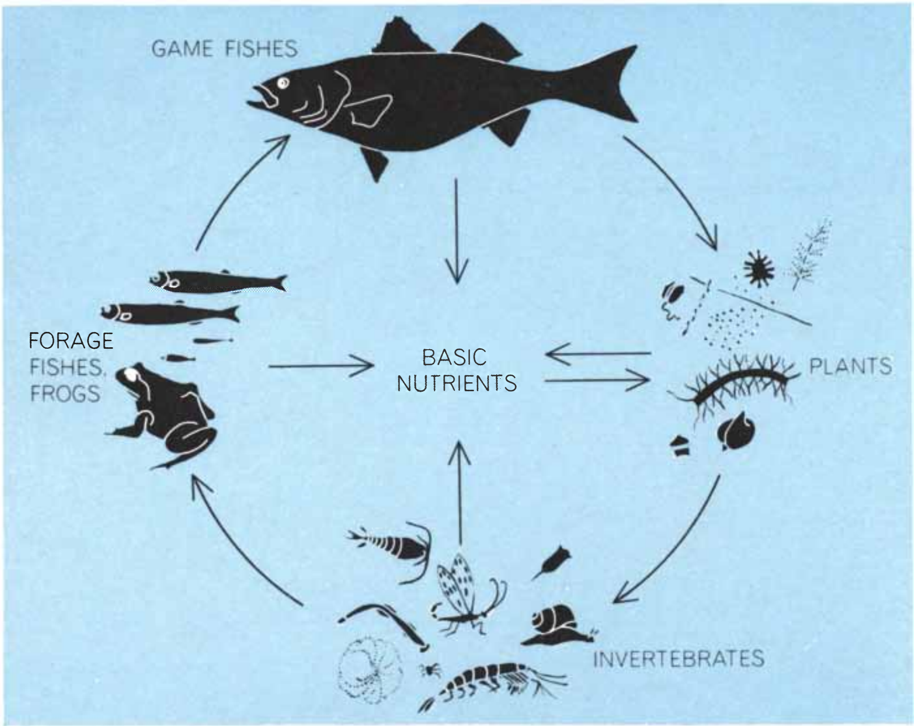
The main drawback of the evaporation scheme is the large amount of water vapor discharged into the atmosphere. The towers for a 1,000-megawatt power plant would eject some 20,000 to 25,000 gallons of evaporated water per minute—an amount that would be the equivalent of a daily rainfall of an inch on an area of two square miles. On cold days such a discharge could condense into a thick fog and ice over the area in the vicinity of the plant. The "wet" type of cooling tower may therefore be inappropriate in cold climates. It is also ruled out where salt water is used as the coolant: the salt spray ejected from a single large power plant could destroy vegetation and otherwise foul the environment over an area of 160 square miles.

A variation of the cooling-tower system avoids these problems. In this refinement, called the "dry tower" method, the heat is transferred from the cooling water, through a heat exchanger something like an enormous automobile radiator, directly to the air without evaporation. The "dry" system, however, is two and a half to three times as expensive to build as a "wet" system. In a proposed nuclear power plant to be built in Vernon, Vt., it is estimated that the costs of operation and amortization would be $2.1 million per year for a dry system and $800,000 per year for a wet system. For the consumer the relative costs would amount respectively to 2.6 and 1 percent of the bill for electricity.

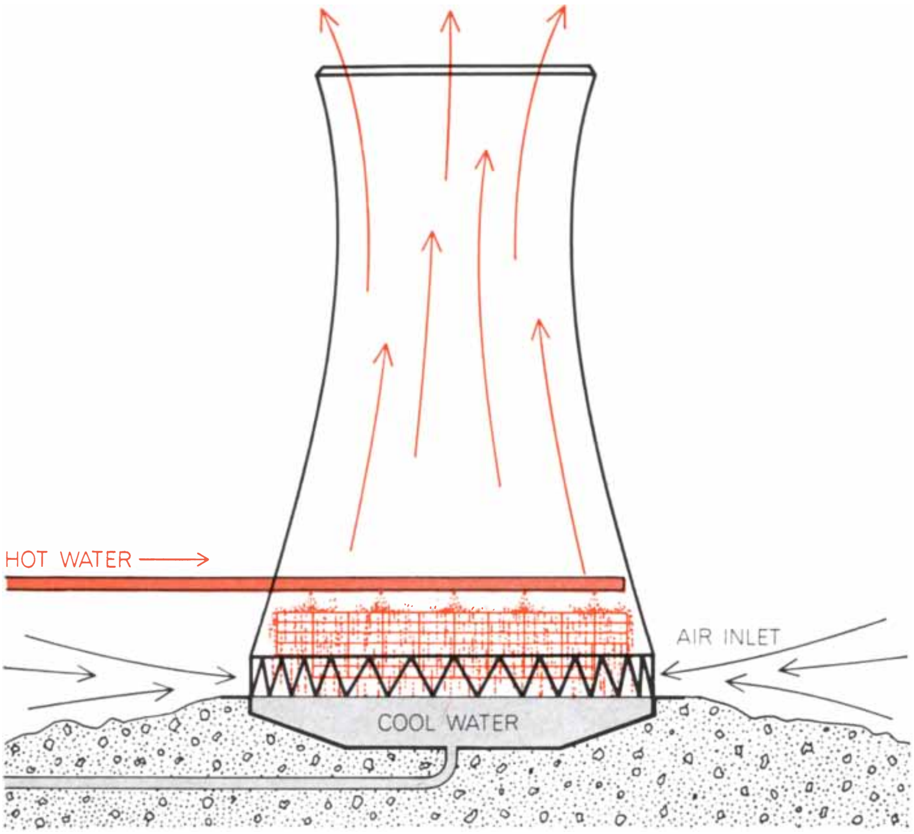
The public-utility industry, like other industries, is understandably reluctant to incur large extra expenses that add sub-



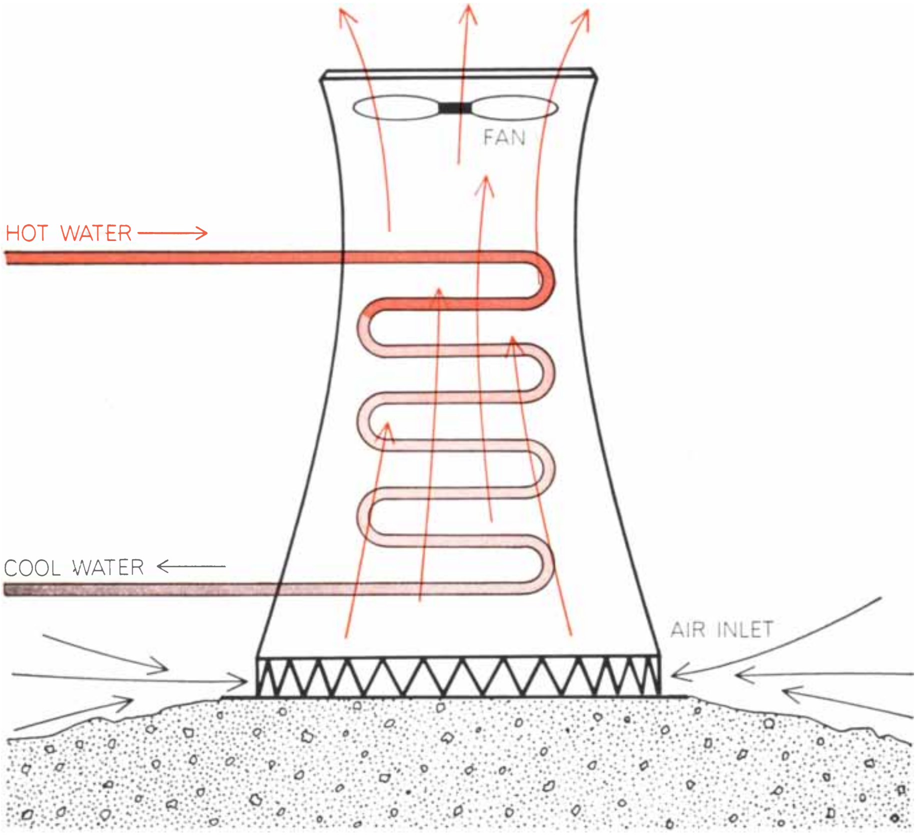
ACCLIMATION extends the temperature range in which a fish can thrive, but not indefinitely. For a young sockeye salmon acclimated as shown by the horizontal scale, spawning is inhibited outside the central (*light color*) zone and growth is poor outside the second zone (*medium color*); beyond the outer zone (*dark color*) lie the lethal temperatures.



AQUATIC ECOSYSTEM is even more sensitive to temperature variation than an individual fish. A single game-fish species, for example, depends on a food chain involving smaller fishes, invertebrates, plants and dissolved nutrients. Any change in the environment that seriously affects the proliferation of any link in the chain can affect the harvest of game fish.



COOLING TOWER is one device that can dissipate industrial heat without dumping it directly into rivers or lakes. This is a "wet," natural-draft, counterflow tower. Hot water from the plant is exposed to air moving up through the chimney-like tower. Heat is removed by evaporation. The cooled water is emptied into a waterway or recirculated through the plant. In cold areas water vapor discharged into the atmosphere can create a heavy fog.



"DRY" COOLING TOWER avoids evaporation. The hot water is channeled through tubing that is exposed to an air flow, and gives up its heat to the air without evaporating. In this mechanical-draft version air is moved through the tower by a fan. Dry towers are costly.

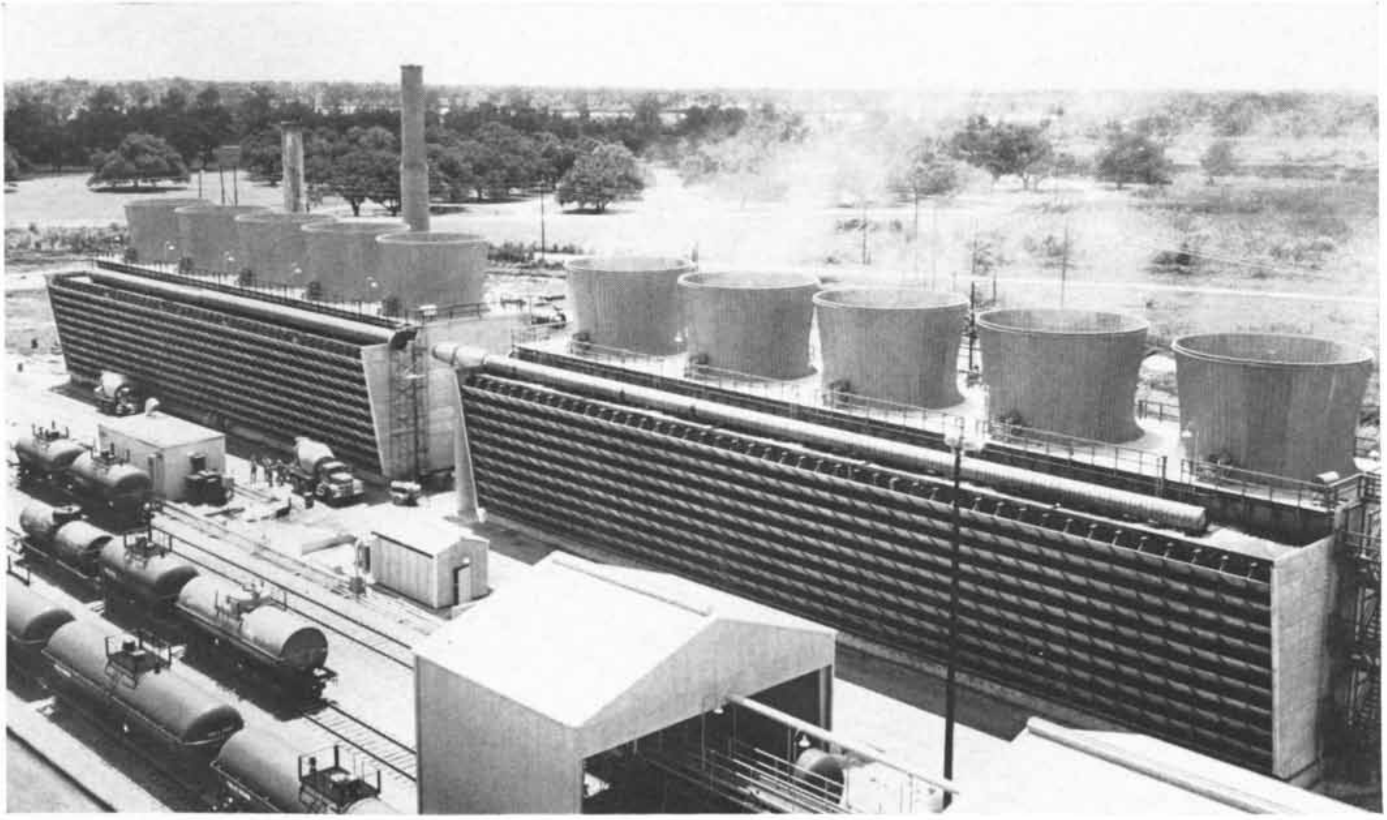
stantially to the cost of its product and services. There is a growing recognition, however, by industry, the public and the Government of the need for protecting the environment from pollution. The Federal Water Pollution Control Administration of the Department of the Interior, with help from a national advisory committee, last year established a provisional set of guidelines for water quality that includes control of thermal pollution. These guidelines specify maximum permissible water temperatures for individual species of fish and recommend limits for the heating of natural waters for cooling purposes. For example, they suggest that discharges that would raise the water temperature should be avoided entirely in the spawning grounds of cold-water fishes and that the limit for heating any stream, in the most favorable season, should be set at five degrees F. outside a permitted mixing zone. The size of such a zone is likely to be a major point of controversy. Biologists would limit it to a few hundred feet, but in one case a power group has advocated 10 miles.

It appears that the problem of thermal pollution will receive considerable attention in the 91st Congress. Senator Edward M. Kennedy has proposed that further licensing of nuclear power-plant construction be suspended until a thorough study of potential hazards, including pollution of the environment, has been made. Senator Edmund S. Muskie's Subcommittee on Air and Water Pollution of the Senate Committee on Public Works last year held hearings on thermal pollution in many parts of the country.

Thermal pollution of course needs to be considered in the context of the many other works of man that threaten the life and richness of our natural waters: the discharges of sewage and chemical wastes, dredging, diking, filling of wetlands and other interventions that are altering the nature, form and extent of the waters. The effects of any one of these factors might be tolerable, but the cumulative and synergistic action of all of them together seems likely to impoverish our environment drastically.

Temperature, Gordon Gunter of the Gulf Coast Research Laboratory has remarked, is "the most important single factor governing the occurrence and behavior of life." Fortunately thermal pollution has not yet reached the level of producing serious general damage; moreover, unlike many other forms of pollution, any excessive heating of the waters could be stopped in short order by appropriate corrective action.

POWER PLANT being completed by the Tennessee Valley Authority on the Green River in Kentucky will be the world's largest coal-fueled electric plant. Its three wet, natural-draft cooling towers, each 437 feet in height and 320 feet in diameter at ground level, will each have a capacity of 282,000 gallons of water a minute, which they will be able to cool through a range of 27.5 degrees.



INDUSTRIAL PLANTS of various kinds can use towers to cool process water. These two five-cell cooling towers were built by the Marley Company for a chemical plant. They are wet, mechanical-draft towers of the cross-flow type: a fan in each stack draws air in through the louvers, across films of falling water and then up. The towers cool 120,000 gallons a minute through a 20-degree range.

SUPERPLASTIC METALS

A new approach to the processing of metals makes it possible to form them by methods usually reserved for glass and plastics. This processing gives the metals a distinctive microstructure

by H. W. Hayden, R. C. Gibson and J. H. Brophy

The behavior of such substances as taffy and glass when they are heated to their softening point and gently pulled is conspicuously plastic: they can be stretched to many times their original length, and they retain the new shape after they have cooled. This property is not ordinarily shared by metals. A piece of metal under tension characteristically parts before it reaches twice its original length, unless it is squeezed as it deforms (as it is in the drawing of wire) to counteract its tendency to "neck down" and break. Recent work in metallurgy, however, has turned up a number of metal alloys that will behave like taffy or glass if certain significant steps are taken in their processing. The phenomenon is so remarkable in metals that it has come to be known as superplasticity.

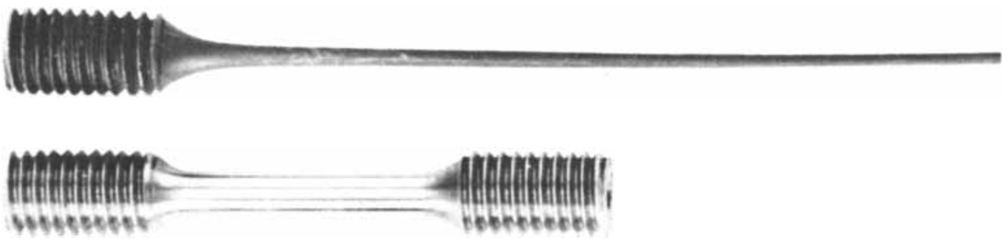
The implications of superplasticity in metals are far-reaching. Several groups of investigators have demonstrated on a laboratory scale that fabrication techniques formerly applied only to glass and plastics may now be applicable to superplastic metal alloys. For example, Walter A. Backofen and his colleagues at the Massachusetts Institute of Technology have shaped a zinc-aluminum alloy by gas pressure in an operation analogous to glassblowing [*see top illustration on page 35*]. D. S. Fields, Jr., of the International Business Machines Corporation has formed large and complicated shapes from superplastic alloys, producing parts that could never be fabricated as single pieces by conventional methods. Two companies in Britain, the British Leyland Motor Corporation and the Rio Tinto Zinc Company, have developed a superplastic zinc-aluminum alloy that can be formed into panels for automobile bodies by the conventional techniques for shaping plastics.

Superplasticity in metals was first observed by Claude E. Pearson of the University of Durham. In 1934, while testing the properties of certain lead-tin and bismuth-tin alloys, he found that samples could be pulled to 10 to 20 times their original length. For many years thereafter the phenomenon was regarded mainly as a laboratory curiosity. Only during the past five years, with the growing realization that superplasticity provides new opportunities for a highly versatile and potentially low-cost means of fabricating metals on a commercial scale, has there been any systematic search for superplastic alloys. A recent review listed 14 alloys in which superplasticity has been demonstrated.

Two types of superplasticity have been observed. One of them is called transformation superplasticity because it is characterized by allotropic transformations, meaning that under different conditions the alloys assume different crystal structures, much as carbon does in forming graphite and diamond. The alloys that exhibit this kind of superplasticity change from one crystal form to another at a fixed temperature or over a range of temperatures. If such materials are subjected to a load while the temperature is repeatedly cycled above and below the transformation temperature, they can be greatly stretched.

The second type of superplasticity is called isothermal superplasticity because the temperature need not be cycled for the alloys to be stretched. The only requirement is that the alloys be heated to about half their melting temperature (measured on an absolute basis such as the Kelvin scale rather than on an arbitrary one such as the Celsius and Fahrenheit scales). At such temperatures the alloys can deform superplastically when they are subjected to load. Most investigations of superplastic behavior have been concerned with systems that show this kind of superplasticity. The reason is that if superplasticity is to be applied to the commercial fabrication of metals, it would seem more feasible to have the equipment operating at a single temperature than to cycle it repeatedly over a range of temperatures.

Since isothermal superplasticity requires a temperature about half the absolute melting temperature of the alloy, superplasticity can occur over a wide range of temperatures, depending on the material. Alloys of lead and tin can be superplastic at room temperature. Alloys of aluminum and zinc require temperatures of several hundred degrees C. to achieve superplasticity. Although superplasticity has not yet been observed in alloys of tungsten, it can be expected (if it occurs at all) to require temperatures



SUPERPLASTICITY is demonstrated by a nickel-chromium-iron alloy. At bottom is a three-inch sample, threaded so that it can be held securely in a machine that puts the material

ranging from 1,500 to 2,000 degrees C.

The fact that superplasticity requires an elevated temperature suggests that the process is dependent on the diffusion of atoms in the crystalline structure of the material. A metal is made up of tiny crystalline "grains"; each grain is a single crystal in which the atoms are stacked in a regular array. The border between one such crystal and another (in which the atoms are arrayed in a different orientation) is called a grain boundary. When the atoms in a crystal are mobile enough, they can diffuse from one position in the crystal lattice to another. Such mobility arises from thermal agitation, although thermal agitation alone would cause only random motion. The application of a load tends to provide the more ordered movement that would seem to be essential to superplasticity.

Another characteristic of superplasticity is an unusual relation between stress and the rate of strain. (Stress is a force causing deformation and is expressed in terms such as pounds per square inch; strain rate is the speed of deformation and is expressed in terms such as inches per inch per minute.) The unusual relation is that a variation in applied stress causes at first a much greater change in the strain rate of a superplastic alloy than it does in a normal alloy. We have compared a nickel-based superplastic alloy with a normal one when both were put under tensile stress at a temperature of 1,000 degrees C. At stresses ranging from about 300 pounds per square inch to about 20,000 pounds per square inch the superplastic alloy stretched to more than 800 percent of its original length without breaking. How much farther it could have gone we do not know, because at 800 to 1,000 percent elongation it had reached the limit of travel in our testing machine. The normal alloy broke in every case before reaching as much as twice its original length. Its deformation rate was much less sensitive than that of the superplastic alloys to increases in stress. A superplastic alloy, however, is like taffy or "silly putty": if it is pulled too fast, it may break, but at a slow rate of pull it stretches a long way.

Superplastic alloys have a very fine grain structure. A case in point is a superplastic alloy of nickel, chromium and iron that we photographed after it had been under tensile stress for an hour at a temperature of 1,000 degrees C. and had elongated almost 1,000 percent; the diameter of the grains in the matrix phase (a nickel-rich solid solution) was about three microns. If a normal nickel-chromium-iron alloy were subjected to a similar test, it would elongate less than 100 percent, and the diameter of its grains would be between 25 and 2,500 microns.

Fine grain sizes are rather rare in metallurgy, particularly after a metal has been kept at a high temperature for a considerable time. Grain boundaries are high-energy imperfections in solids; hence when a solid is heated to an elevated temperature, it has a tendency to decrease the amount of excess energy represented by grain-boundary area. As a result the grains grow. The larger the grain size, the smaller the grain-boundary area per unit of volume. The secret of attaining superplasticity, then, seems to be finding ways of preventing the normal process of grain growth.

Most of the alloys that have been observed to behave superplastically are two-phase alloys, such as the nickel-chromium-iron alloy we have mentioned. In 1948 Clarence M. Zener, who was then working at the University of Chicago, demonstrated theoretically that the presence of a finely dispersed second phase (chromium-rich in our example) could inhibit the growth of grains in a major matrix phase (nickel-rich). It would be as difficult for large grains to form under such conditions as it would be to form large soap bubbles by blowing a film of soap through a very fine screen.

Thus a prerequisite to obtaining superplasticity is knowing how to create what we call a microduplex structure. The term reflects the fact that the structure must be fine-grained and have two phases. Such a structure can be created by a variety of methods. In the nickel-chromium-iron example the superplastic alloy and the normal one were of identical chemical composition. The difference between them lay in the different processing we gave them before subjecting them to tensile stress.

The methods that have been employed to produce superplasticity in various alloy systems utilize one or more of the familiar metallurgical processes of hot-working, cold-working, recrystallization and precipitation. All of them are included in the general term "thermomechanical processing." Let us describe the thermomechanical processing schedules followed to produce microduplex structures in three different alloy systems: lead-tin, zinc-aluminum and nickel-chromium-iron. The reader may find it helpful to refer to the three metallurgical phase diagrams [*page 33*] representing the three systems. (A phase diagram shows the effect of temperature and composition on the constitution of an alloy system, indicating whether the equilibrium state of a given composition at a given temperature is liquid, a mixture of a liquid phase and a solid phase, a single solid phase or a mixture of two or more solid phases.)

In the case of the lead-tin alloy the phase diagram shows that an alloy of 60 percent tin and 40 percent lead should be composed of two solid phases at all temperatures below the temperature at which a liquid is formed. This alloy can be given a microduplex structure by drastic extrusion. Extrusion is what happens when one squeezes toothpaste from a tube: a material with a large cross section is reduced to a smaller cross section by being pushed through a die. Extrusion has the effect of reducing the cross-sectional area of the grains in a metal. As a grain is squeezed its length is initially increased in accordance with the reduction of cross-sectional area; the volume of the grain remains constant. The energy put into the system by the operation of extrusion is dissipated to a certain extent by recrystallization, in which

under tensile stress. At top is a sample that has been heated and stretched; it is 13⅜ inches long. The effects of the heat have darkened the material. Superplasticity arises from a distinctive internal structure created in the material by preliminary processing steps.

an individual long, narrow and heavily strained grain is broken up into several short and narrow ones that are strain-free. Thus a microduplex structure is easily produced by a single drastic deformation.

In zinc-aluminum alloys phase transformations in the solid state play a more important role than they do in lead-tin alloys. The phase diagram for zinc-aluminum shows that when an alloy that contains 78 percent zinc and 22 percent aluminum is heated to a temperature above 275 degrees C., it will be transformed from a mixture of two solid phases into a single solid phase. If the alloy is heated to such a temperature and then quickly cooled to room temperature or lower (even to the temperature of liquid nitrogen), the single phase that existed at the high temperature will be momentarily retained, even though this is not the equilibrium state at the new temperature. Soon, however, since the tendency toward the equilibrium state is strong, the single-phase alloy will decompose into the equilibrium two-phase state, giving off heat. (This process of decomposition has become a classroom demonstration of an exothermic, or heat-releasing, reaction. A student is handed a sample of the alloy immediately after it is quenched from the high temperature; it is as cold as the temperature of the medium in which it was quenched. Then the decomposition reaction begins and the sample rapidly becomes warmer. Before long it is so hot that the student has to drop it.) The net result of the decomposition reaction is that an initially coarse-grained high-temperature phase decomposes into a very fine-grained mixture of two phases: a microduplex structure.

If, instead of following the quenching procedure, one slowly cools the alloy from above 275 degrees C. to a temperature below the equilibrium-transformation temperature, the resulting structure is a coarse-grained aluminum-rich phase with a dispersed zinc-rich precipitate distributed within the coarse grains. Mechanical testing demonstrates that the quenched and transformed material is superplastic, whereas the slowly cooled material is not. The contrast in the microstructures of the two materials is distinct [*see bottom illustration on page 35*].

We see thus far that the best processing schedule for producing a microduplex structure in a lead-tin alloy is quite different from the one for producing such a structure in a zinc-aluminum alloy. The procedure for the lead-tin material was based on deformation by hot-working; for the zinc-aluminum alloy the technique involved the appropriate heat treatment and cooling rates. Now let us look at a processing schedule (for a nickel-chromium-iron alloy) that involves a combination of deformation and the appropriate heating and cooling.

The phase diagram for the nickel-chromium-iron alloy shows that some mixtures of these metals may be heated to a quite high temperature (1,200 to 1,260 degrees C.), at which the equilibrium structure is a single-phase one. At lower temperatures (from room temperature to about 1,100 degrees C.) the equilibrium structure consists of two phases. One method of forming a microduplex structure in these alloys involves heating the material to the high temperature to form a single phase. The alloy is then hot-worked by such processes as forging or rolling. Hot-working is continued while the temperature of the workpiece decreases well into the two-phase region.

During the deformation and cooling there is a mutual occurrence of the two processes of recrystallization and precipitation. In recrystallization the stored energy in a heavily deformed large grain can be partly relieved by the formation of a number of smaller grains within the originally large one. In precipitation a second phase is formed from within a primary phase. In the superplastic nickel-chromium-iron alloys the second phase preferentially precipitates at the grain boundaries of the fine-grained primary phase. By its presence the second phase greatly retards the process of grain growth that would normally occur in the primary phase on reheating. This process has been successfully applied to the production of microduplex structures in several nickel-chromium-iron alloys in ingots ranging in weight from 30 pounds to 10 tons.

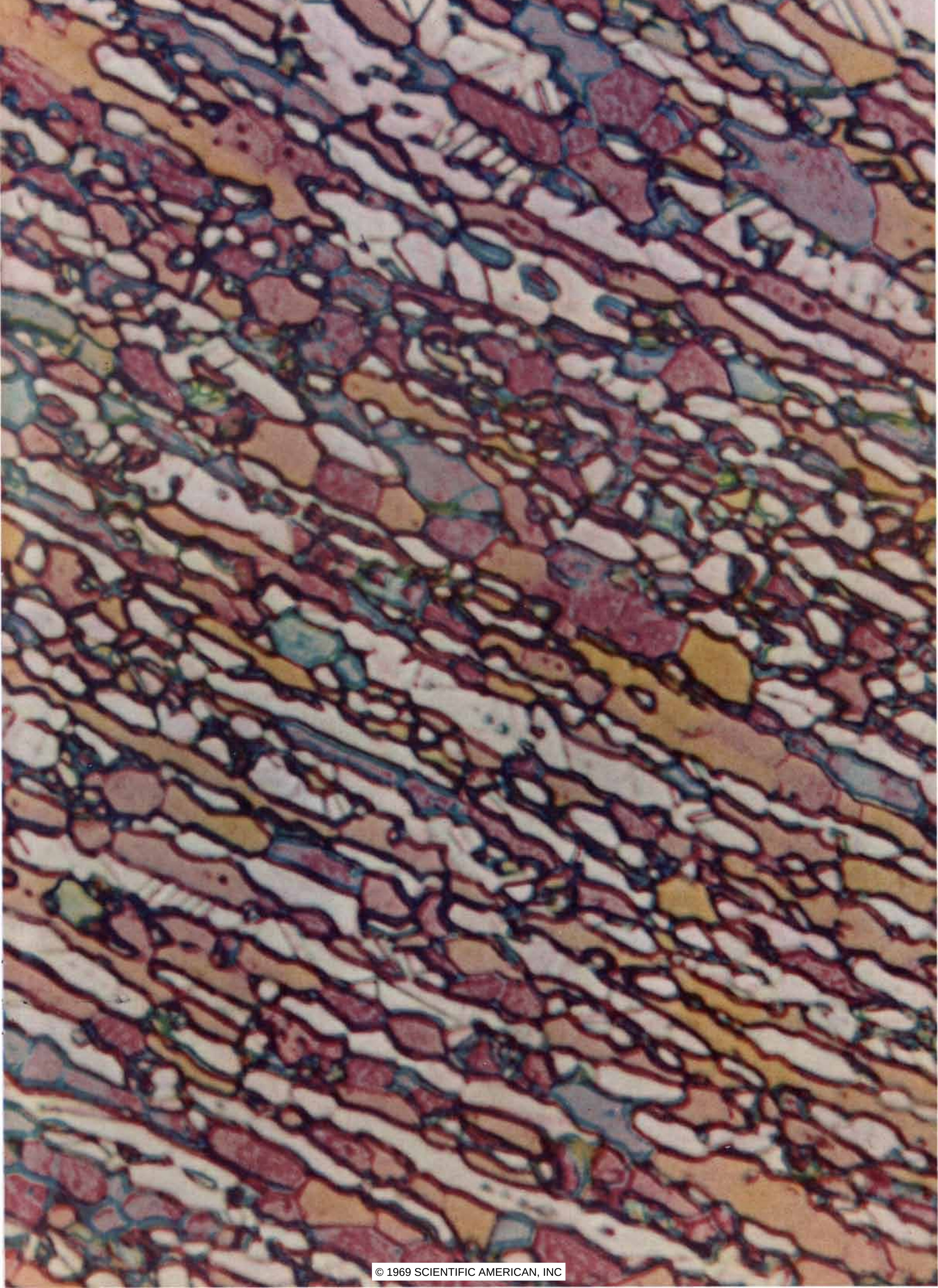
An alternative method for producing a microduplex structure in nickel-chromium-iron alloys involves heating the material to a temperature in the single-phase region and then quenching it to room temperature. In contrast to the zinc-aluminum alloys the high-temperature single phase does not spontaneously decompose into the equilibrium two-phase structure but is instead retained in a barely stable state. The quenched material is heavily cold-worked and then heated to a reasonably high temperature in the two-phase region; the temperature might be between 850 and 1,000 degrees C. Recrystallization and precipitation take place during the heating, with the creation of a microduplex structure. The crucial step is the cold-working. If it is omitted, there is not enough stored energy to cause recrystallization during the heating. Only precipitation occurs, so that the resulting structure consists of a coarse-grained matrix phase with most of the second phase distributed within the large grains.

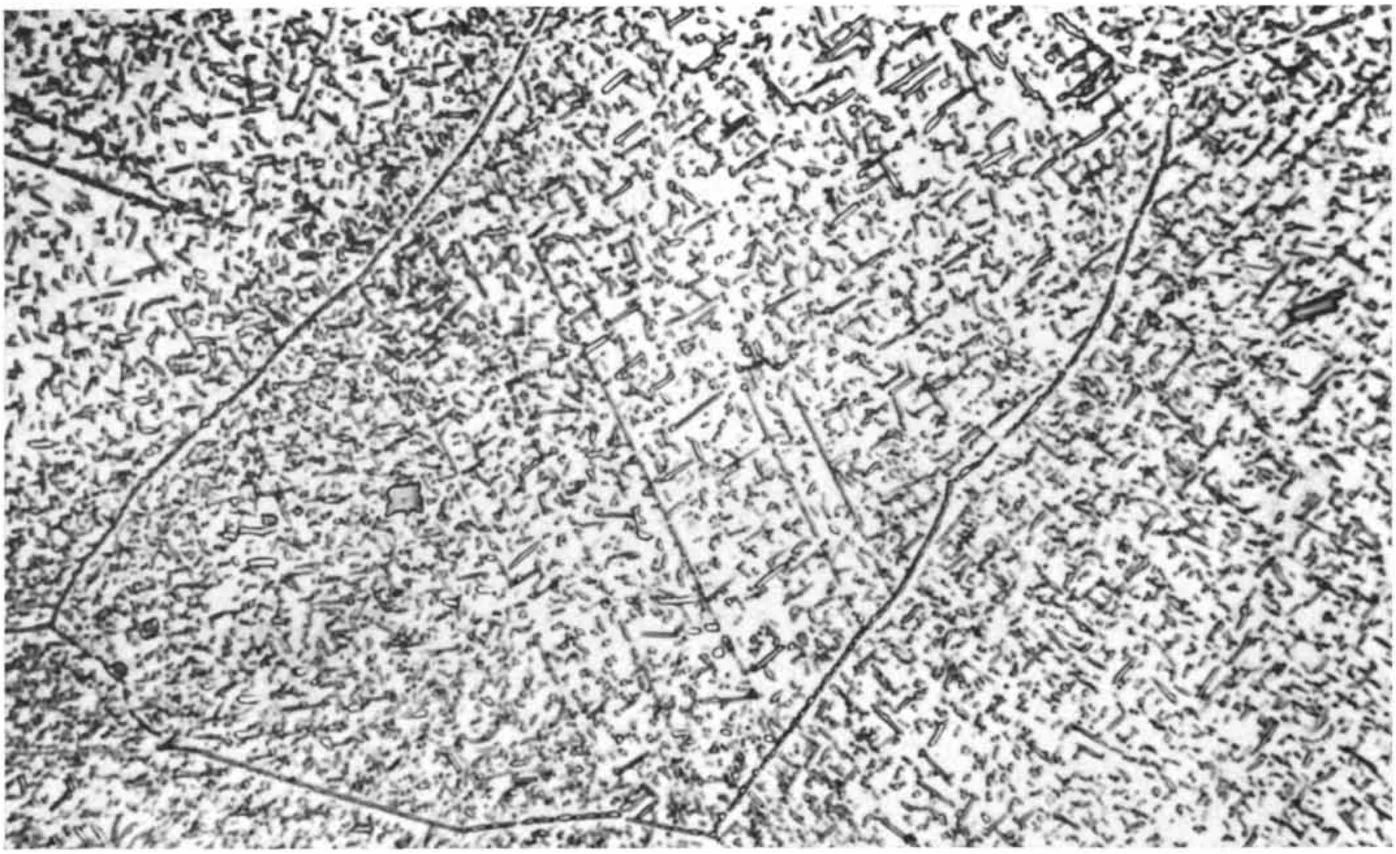
Many other procedures have been used to obtain microduplex structures, and doubtless still more of them will be worked out. Our main point here is that there are several ways. The central concept is that the processing steps have to be adapted to the particular thermomechanical responses of the given alloy system.

Knowing the procedures to follow in creating a microduplex structure is one thing; understanding the mechanisms that take place within the metal during superplastic deformation is quite another matter. Indeed, the answer to the seemingly simple question of what atomic mechanisms give rise to superplasticity is a subject of great metallurgical controversy. Several detailed explanations of the mechanisms have been proposed. Perhaps when the picture is complete it will reveal that different explanations apply to different superplastic alloys.

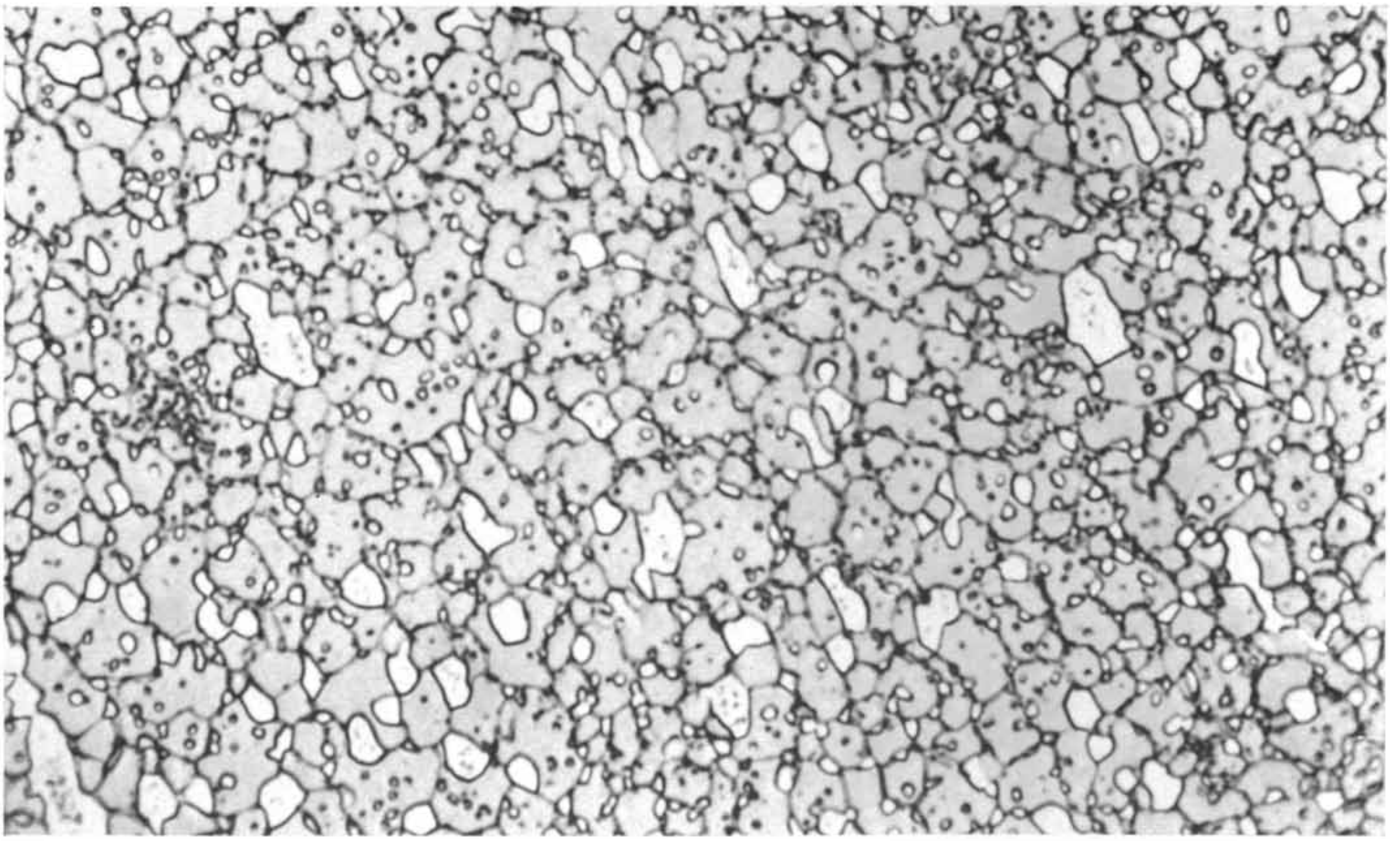
The presence of extremely fine grains means that any superplastic alloy has a larger amount of grain-boundary area per unit of volume than would be found in a coarse-grained alloy that was not superplastic. Grain boundaries can serve many important functions in high-temperature deformation; for example, they can act as sources and sinks for vacancies and dislocations, which are defects known to give rise to metallic deformation. Or, at very high temperatures, displacements between adjacent grains can take place by sliding along the grain boundaries. Hypotheses attempting to explain superplasticity must first account for the phenomenon's dependence on extremely fine grain sizes. Another matter to be accounted for is the distinctive rela-

INTERNAL STRUCTURE of a specimen of superplastic steel consisting of ferrite (*color*) and austenite (*white*) is emphasized by the colors in the photomicrograph on the opposite page. The specimen was prepared by an etching process that colored the ferrite phase but not the austenite phase. The dark lines outline grains, which are regions of similarly aligned atoms. Grains are here enlarged 3,000 diameters. A fine-grained, two-phase structure, which is called a microduplex structure, is characteristic of superplastic metals.





COARSE-GRAINED STRUCTURE of a nickel-iron-chromium alloy that is not superplastic is apparent in this photomicrograph at an enlargement of 2,700 diameters. A single grain is outlined in the center of the photograph by the thin, dark boundary. The matrix of the grain is nickel-rich; the finer particles inside the grain and the adjacent grains are a chromium-rich solid solution.



FINE-GRAINED STRUCTURE that is characteristic of superplastic metals appears in an alloy that is identical in chemical composition with the one shown in the photograph at the top of this page. The small grain size of the superplastic alloy results from the thermal and mechanical processing it received. Lighter regions are chromium-rich and the darker ones are nickel-rich. The dark lines in the nickel-rich matrix phase are the grain boundaries. In this photomicrograph the enlargement is 1,300 diameters.

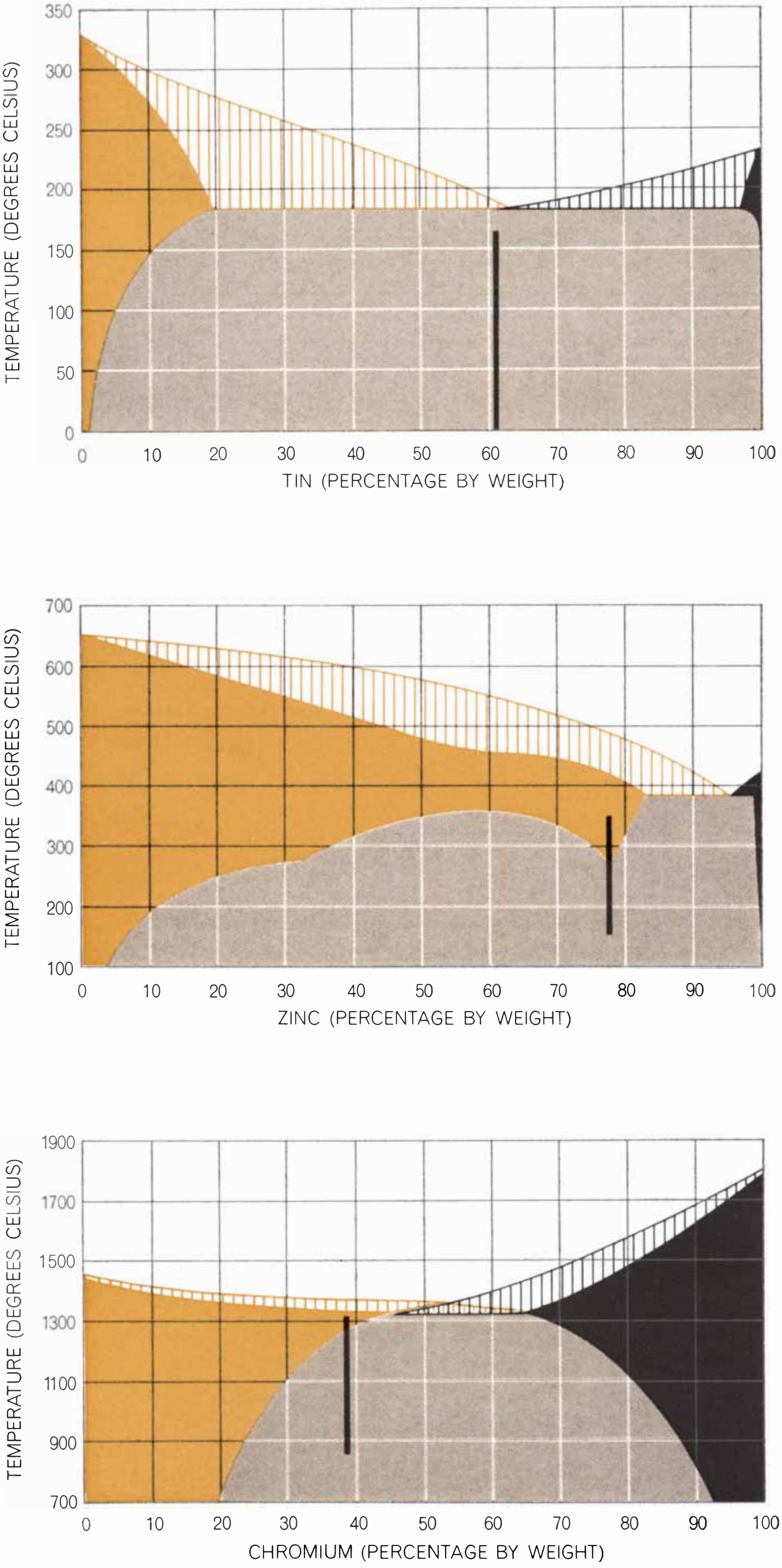
tion between stress and strain rate in superplastic metals.

Among the hypotheses that have been proposed are vacancy creep, dislocation creep and grain-boundary sliding [*see illustration on next page*]. A vacancy is the absence of an atom from a position where one would be expected in a crystal lattice; a dislocation is a line of disorder in the lattice caused by a missing plane of atoms or an extra plane. Vacancies and dislocations are known to exist in metals. Vacancy creep would involve deformations caused by the movement of atoms and vacancies. This basic mode of deformation was first proposed in 1950 by Conyers Herring of the Bell Telephone Laboratories and F. R. N. Nabarro of Bristol University. It recognizes that under a state of stress atoms tend to move from positions that are attempting to contract to positions that are attempting to extend.

Herring and Nabarro proposed that this diffusive movement takes place inside grains. Robert L. Coble of M.I.T. suggested later that the diffusion might occur through the grain boundary more readily than through the volume of the grain. In either case a fine grain size means shorter paths for diffusion than would be possible with large grains and hence a shorter period of time for an atom to move from one location to another. Both mechanisms would lead to faster rates of deformation at a fixed level of stress and so could account for superplasticity. With these mechanisms, however, stress should be directly proportional to strain rate, which is not usually the case in superplastic alloys.

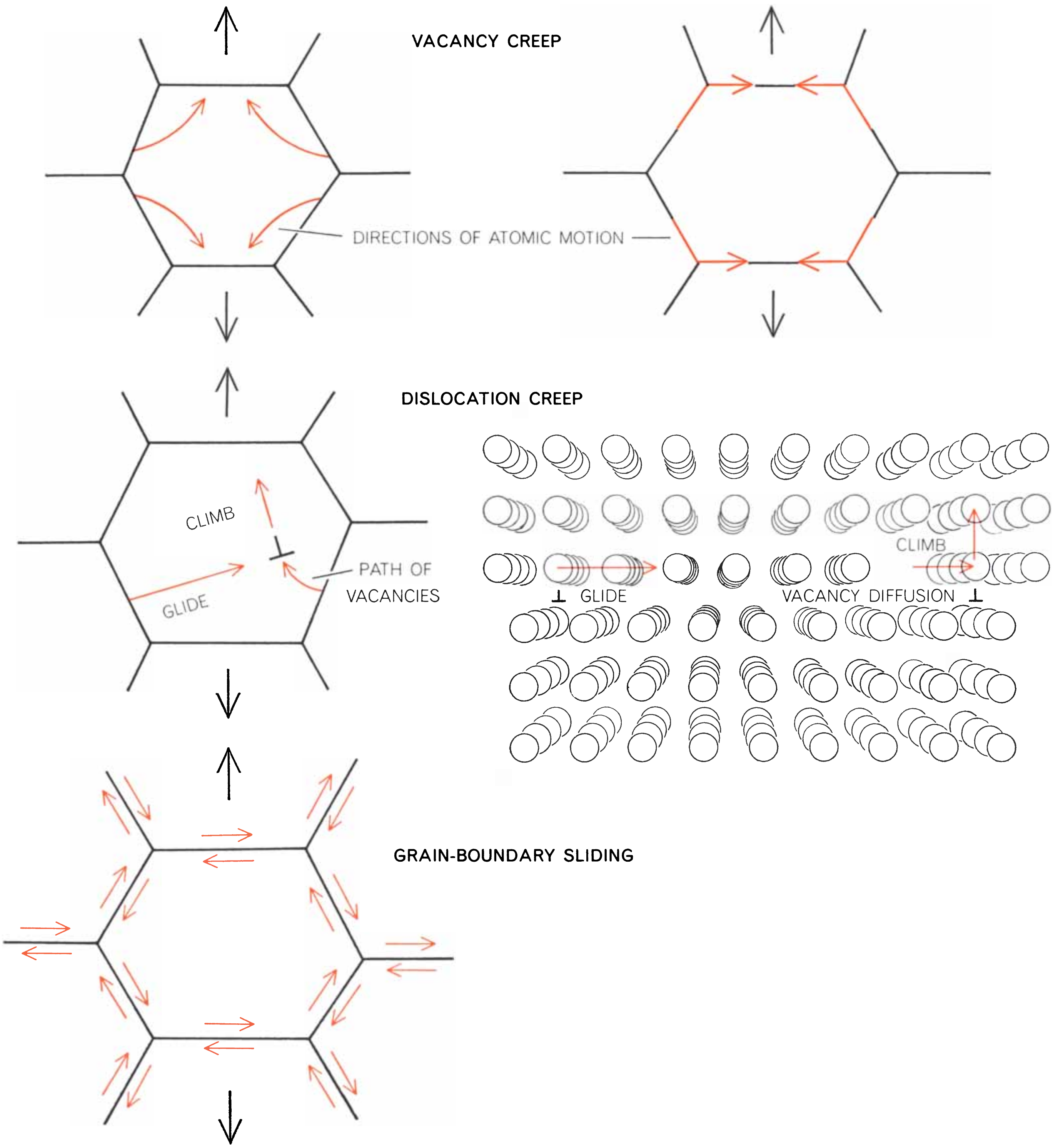
Our group has proposed the dislocation-creep mechanism on the basis of predictions made earlier by Johannes Weertman of Northwestern University. The mechanism involves the motion of atoms, vacancies and dislocations. Under conditions of normal high-temperature deformation (which is also called creep) in coarse-grained materials, dislocations usually become entangled, forming dislocation cells within each grain. Each dislocation cell consists of a region of almost perfect crystal bounded by a wall of tangled dislocations. The size of the cells is normally reciprocally related to the applied stress: the lower the stress, the larger the cells. Predictions of cell size in fine-grained materials suggested that even at rather high stresses the dislocation cells should be larger than grains as small as those in a microduplex alloy. Hence it is unlikely that dislocation cells would form in such materials.

Other predictions that we have made



PHASE DIAGRAMS show the states of various alloys at various temperatures. The alloys are lead-tin (*top*), zinc-aluminum (*middle*) and nickel-chromium-iron (*bottom*). Heavy black lines indicate the composition of superplastic alloys. Lead-tin phases include liquid (*white*), solid lead (*color*), solid tin (*black*) and various combinations of solid and liquid phases shown by mixtures of black, white and color. Other diagrams are read similarly.

on the basis of our experiments suggest another distinction between the behavior of dislocations in a coarse-grained material and their behavior in a fine-grained material. In coarse-grained materials dislocations can pile up on a single plane of slippage, whereas in fine-grained materials there would be little likelihood of pileups except at rather high levels of stress. We believe that the usual accumulation of dislocation cells and dislocation pileups found in conventional high-temperature deformation does not take place in the microduplex structure because the smallness of the grains allows no room for such dislocation phenomena. Since dislocation cells and pileups inhibit creep, their relative absence facilitates it and hence provides an explanation of superplasticity. Moreover, our calculations indicate that in dislocation creep, stress should be proportional to the square root of strain rate and to the grain size at a fixed strain rate. We have observed both relations in our experiments with superplastic nickel-chromium-iron alloys.



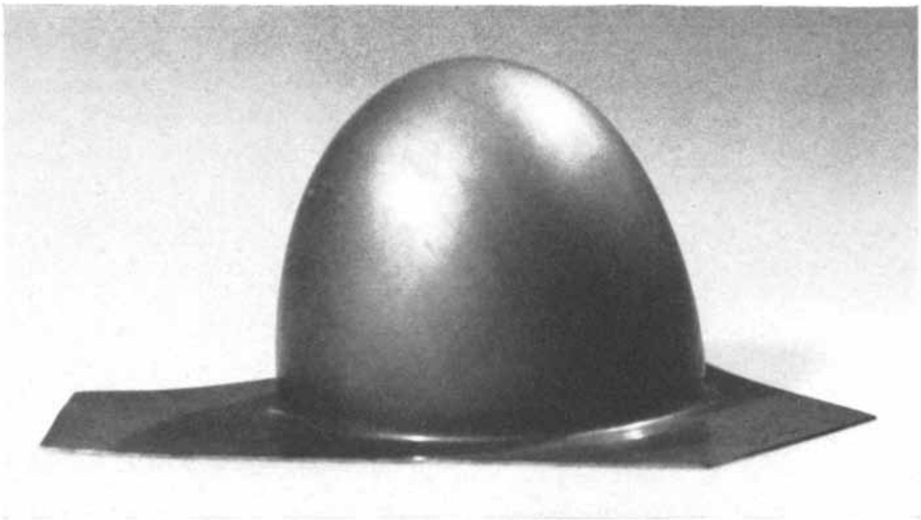
MECHANISMS OF SUPERPLASTICITY may include vacancy creep (*top*), dislocation creep (*middle*) and grain-boundary sliding (*bottom*). At top left is a single grain, together with parts of adjacent grains; the metal is under tensile stress indicated by the short vertical arrows. An atom can move into a vacancy, which is an empty place in a crystalline array of atoms; the movement leaves a vacancy, which is replaced by another atom, so that the vacancy creeps through the grain (*left*) or along its boundary (*right*). A similar mechanism occurs in dislocation creep; a dislocation is a line of disorder caused by a missing plane of atoms or an extra plane. Glide and climb movements are indicated. In grain-boundary sliding the crystal's grains slide past one another.

Grain-boundary sliding has been suggested by several investigators as a mechanism of superplasticity. It involves the displacement of one grain in relation to another. Recent evidence suggests the existence of grain-boundary dislocations, which can move by both gliding and climbing, as in the case of dislocations inside a grain. Calculations made before this evidence was obtained suggest that in grain-boundary sliding, stress should be proportional to strain rate. It remains to be seen if the recent recognition of grain-boundary dislocations will lead to new predictions for strain-rate sensitivity that are more in line with what is observed in superplastic alloys.

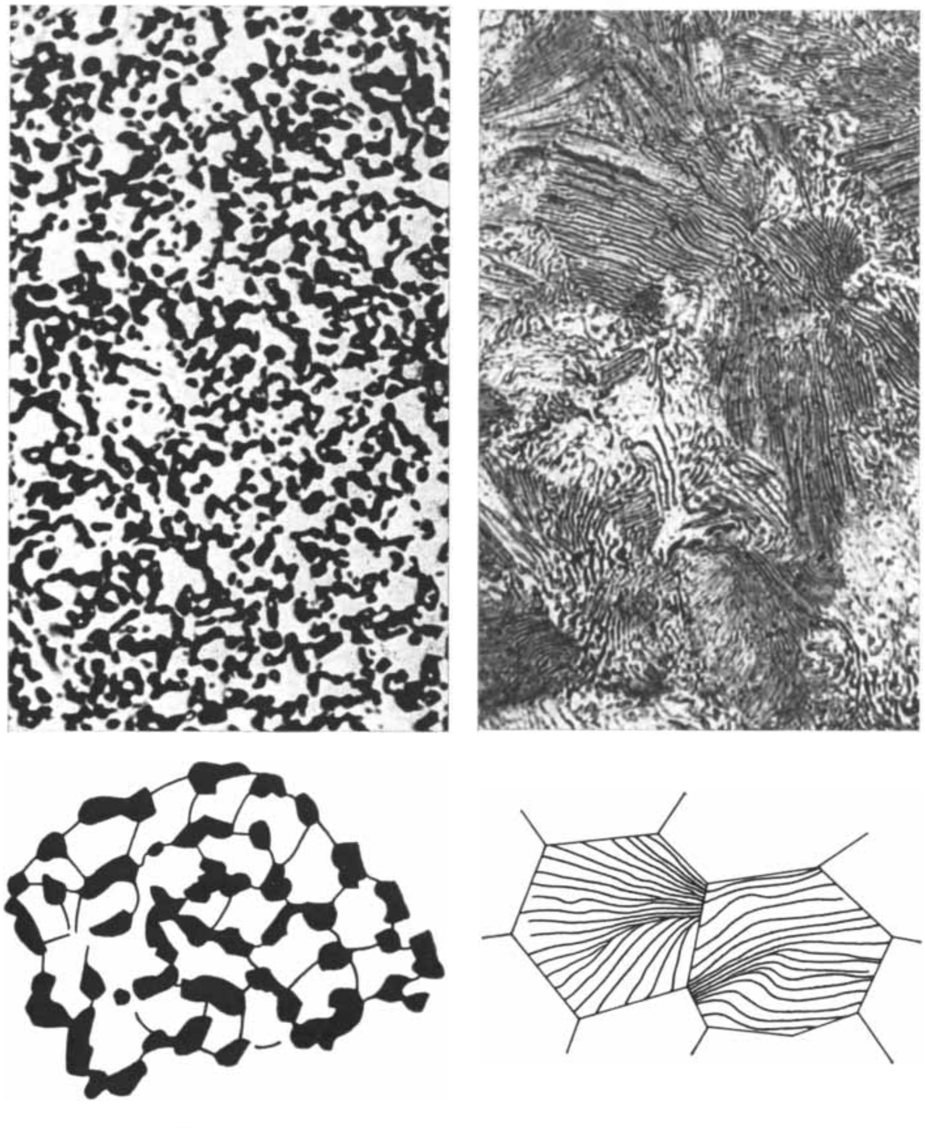
Promising though superplasticity is for new ways of fabricating metals, this promise will not be realized if after fabrication the alloys fail to show such other useful engineering properties as high strength, toughness, resistance to corrosion and quite possibly resistance to deformation at high temperatures. The property of resistance to deformation at high temperatures is precisely the opposite of what one seeks in superplasticity. Hence achieving it would require the ability in a superplastic alloy to "turn off" the superplasticity once it has achieved its purpose.

We have investigated these requirements by experiments with stainless steels that can be shaped superplastically. These experiments have demonstrated a classic relation between structure and properties: strength increases with decreasing grain size. The ultrafine two-phase structure in the superplastic stainless steel results in higher strength, greater toughness and better fatigue resistance at room temperature than are found in conventional grades of stainless steel with the normal microstructural coarseness.

We have also found that recognition of the microstructural conditions that give rise to superplasticity enables us to control its occurrence, that is, to turn it off when it is no longer needed. The technique is to give the material a heat treatment after it has been superplastically shaped. The treatment coarsens the structure, thereby preventing further superplastic deformation. By this means we have developed nickel-base alloys that can be deformed easily and superplastically at an elevated temperature and then put into service at an even higher temperature. With the qualities of versatile shaping and high strength after shaping, superplastic alloys appear to be on the verge of commercial usefulness.



METALLIC BUBBLE was blown by gas pressure in a superplastic zinc-aluminum alloy at the Massachusetts Institute of Technology in work by David L. Holt, Pracheeshwar S. Mathur, Thomas Thomsen and Walter A. Backofen. Initially the alloy was flat. The bubble is virtually impossible to produce by a simple operation in metals lacking superplasticity.



DETAILED VIEW of the internal structure of a superplastic zinc-aluminum alloy (*left*) and an identical alloy that is not superplastic is provided by these photomicrographs and the related drawings. The specimens are shown after they have been subjected to tensile stress. In the superplastic alloy the zinc second phase, which is the dark material, is distributed within the aluminum matrix; in the ordinary metal the zinc phase grows from the grain boundaries. The enlargement of these photomicrographs is about 1,500 diameters.

Phases in Cell Differentiation

By cultivating embryonic tissue in the laboratory one can study the specialization of cells in the mammalian pancreas. There seem to be three regulatory phases, each leading to a new stage of differentiation

by Norman K. Wessells and William J. Rutter

How does a single cell—the fertilized egg—give rise to the many different cell types of a multicellular organism? After fertilization of an egg by a spermatozoon the number of cells in the developing embryo increases dramatically. Soon three classes of cells can be distinguished: ectodermal and endodermal cells respectively make up an outer and an inner embryonic layer, and mesodermal cells constitute an intermediate layer. These cells come to be arranged in groups that develop into recognizable tissues and organs. Ultimately several hundred kinds of cells can be distinguished in an adult mammal. The process of functional and structural specialization of cells is called differentiation, and the mechanisms controlling differentiation remain major mysteries of biology.

Experimental work done by embryologists since 1900 has demonstrated that interactions between cells play an important role in differentiation. The development of some organs, including the pancreas, the liver and the lungs, depends on discrete sets of cells derived from the embryonic gut endoderm and also on adjacent mesodermal cells; the brain, the mammary glands and the limbs involve combinations of certain ectodermal and mesodermal cells. Experiments with intact embryos and with laboratory cultures of combinations of tissues indicate that for normal development the two interacting tissues must be adjacent to each other. A fundamental and still unanswered question is: How does one cell influence the development of another?

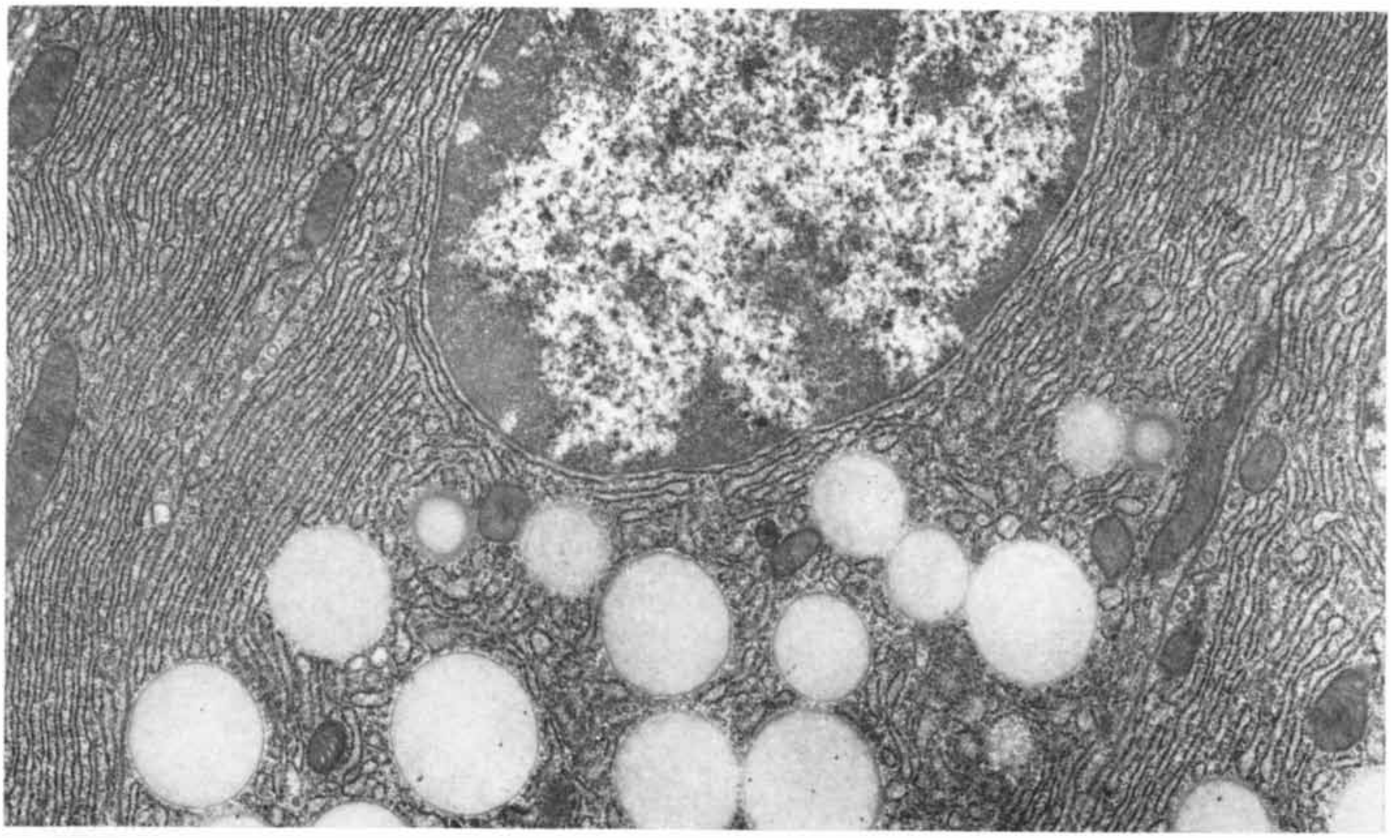
What happens within cells during development to make one cell type different from another? Since nearly all cells have some physiological activities in common, all must contain some of the same enzymes and structures, and mere quantitative differences among cells in this common metabolic machinery would not in themselves confer unique characteristics on a given cell. For that a cell has a group of specific proteins responsible for its specialized functions: muscle cells contain contractile proteins, plasma cells make antibodies, red blood cells form hemoglobin and so on. What controls these qualitative differences? This is a second fundamental and still unanswered question.

In this article we shall approach these two major questions by discussing the differentiation of the mammalian pancreas. First we shall outline the development of the pancreas and then describe a detailed investigation of interactions between cells and the regulation of specific protein synthesis in this organ. The view of differentiation gleaned from this work, carried out in our laboratories at Stanford University and the University of Washington, may be generally applicable to other organ systems.

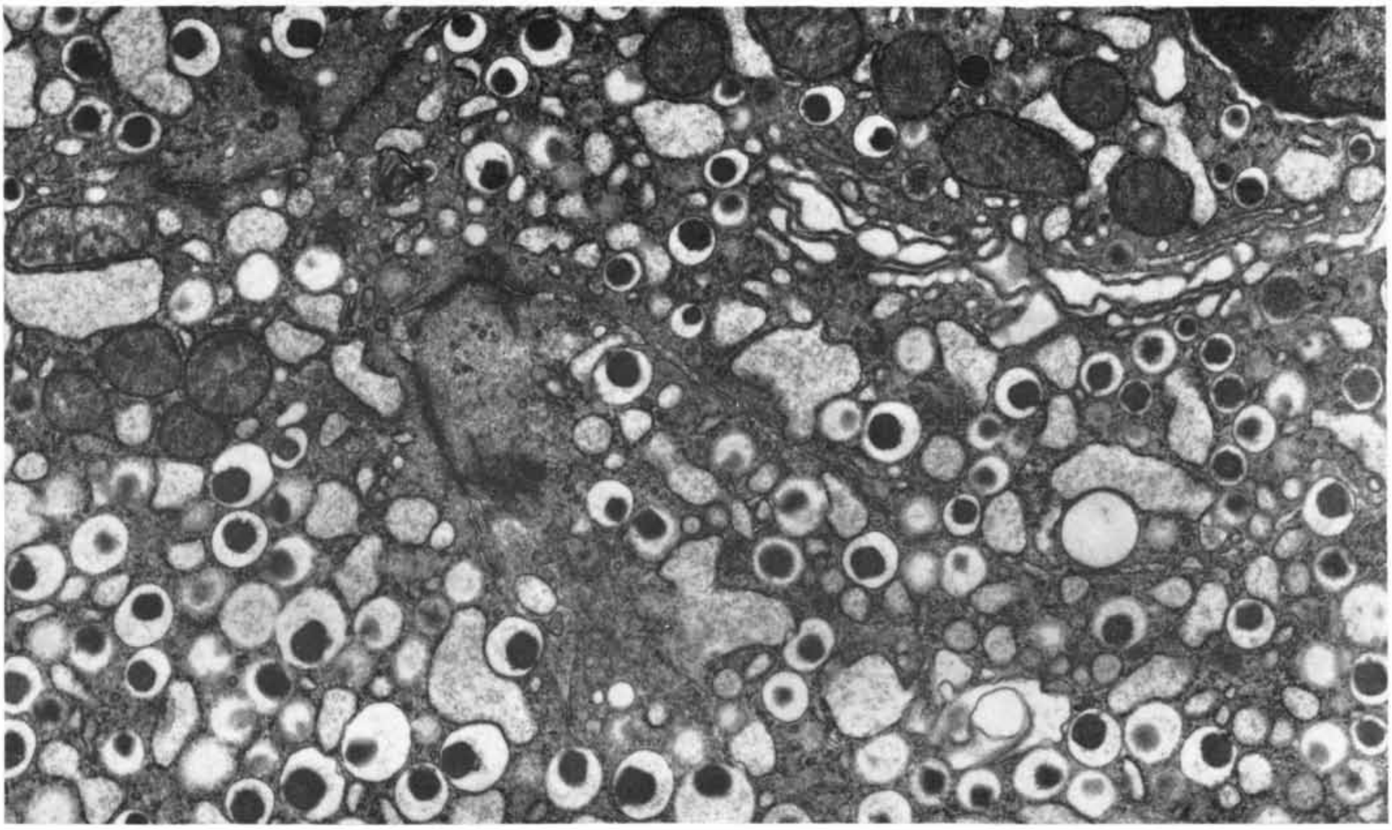
The mammalian pancreas manufactures enzymes that digest foodstuffs and also secretes two hormones that regulate the metabolism of carbohydrates. Different populations of cells are involved in the two functions. The exocrine cells make about a dozen specific proteins that are involved in the breakdown of proteins, carbohydrates, fats and nucleic acids. Some are synthesized as active enzymes and others as inactive precursors called zymogens; all of them are stored within the cells in granules. The exocrine cells are arranged in clusters called acini and secrete the contents of their granules into a duct system that leads into the small intestine. There the zymogens are converted to active enzymes, and digestion takes place.

There are two kinds of endocrine cell. The *B* cells synthesize insulin, the hormone that regulates the uptake of blood glucose and its storage as glycogen in muscle and fat. Insulin is stored in beta granules. Donald F. Steiner of the University of Chicago has demonstrated that the synthesis of insulin (like the synthesis of the protein-digesting enzymes) has two steps: first the formation of a precursor called proinsulin and then its conversion to insulin by an enzyme. The site of this conversion and the specific enzymes involved are not known. The *A* cells produce the hormone glucagon, a protein that controls the degradation of glycogen in the liver. Glucagon is stored in alpha granules. Both *B* cells and *A* cells are in the islets of Langerhans, spherical aggregates of cells that lie near blood capillaries, so that the hormones pass easily into the blood for transport to their sites of action elsewhere in the body.

In the embryo of a mouse or a rat the formation of the pancreas begins about midway through the gestation period. A group of cells in the upper wall of the primitive gut begins to bulge upward on the ninth day of gestation in the mouse, and about 36 hours later in the rat. The base of this evagination constricts, giving rise to the future pancreatic duct; the upper part, the pancreatic epithelium, expands into the surrounding mesodermal tissue. (Note the proximity of endoderm and mesoderm.) In the next three or four days rapid cell division continues and the cells become arranged in typical acini and primitive islets. During the later stages of gestation cell division tends to be restricted to the peripheral regions of the pancreas; near the center one can detect the signs of advanced differentiation: zymogen granules in the acinar cells and alpha



EXOCRINE CELLS of the pancreas secrete zymogens, proteins that are precursors of digestive enzymes. In this electron micrograph made in the laboratory of one of the authors (Wessells) a thin section of exocrine cells from the adult mouse pancreas is enlarged 14,000 diameters. The extensive endoplasmic reticulum, a folded membrane studded with ribosomes (*small black dots*), is the site of protein synthesis. Zymogens are stored in zymogen granules (*spheroidal vesicles at bottom*) until they are secreted into the intestine.



ENDOCRINE CELLS of the pancreas secrete the hormones glucagon and insulin. In this electron micrograph several adult mouse *B* cells, which secrete insulin, are enlarged 17,000 diameters. The light spheroidal vesicles present in large numbers are beta granules and the dense material in many of the granules is thought to be insulin or proinsulin, the precursor protein of the hormone.

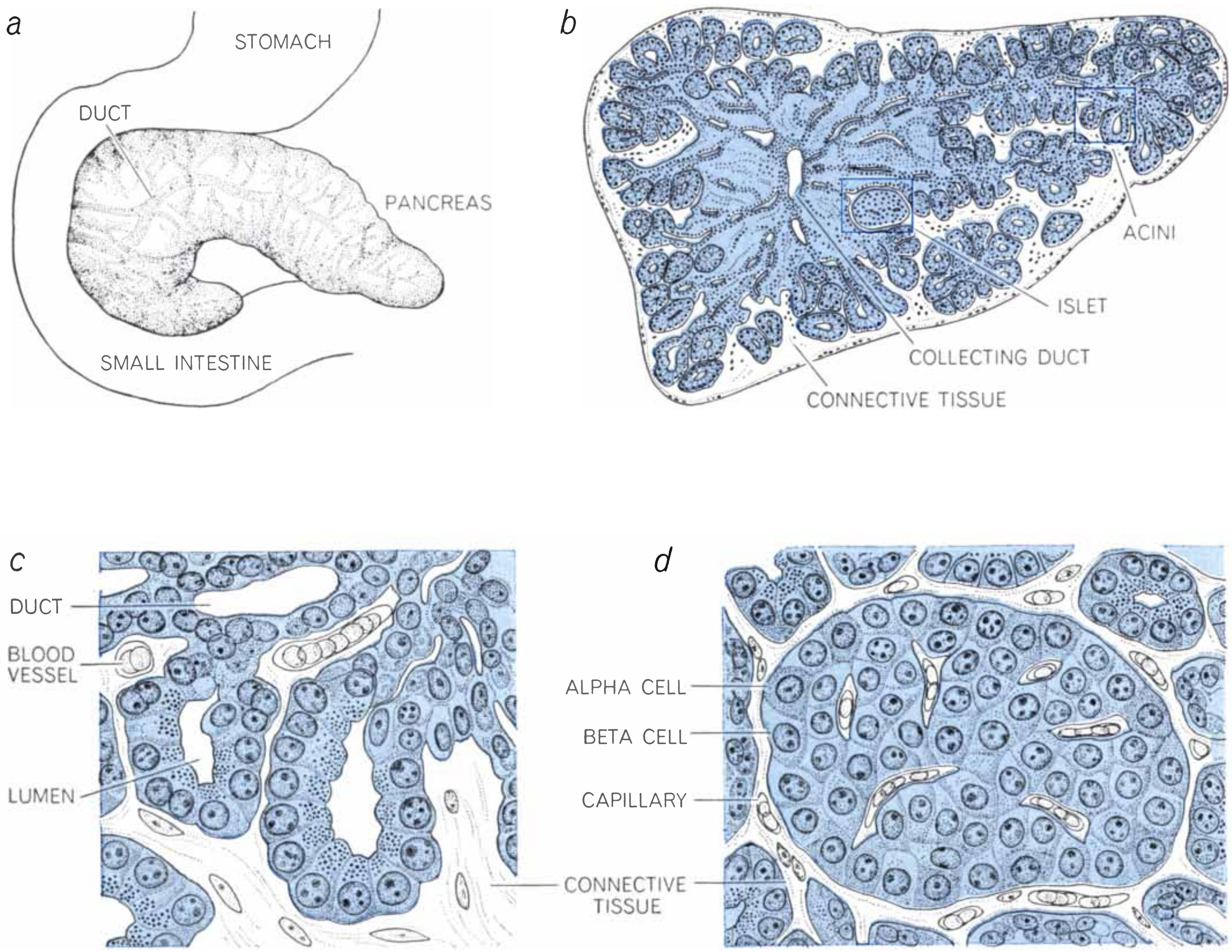
and beta granules in the endocrine cells. As birth approaches, cell division stops in the peripheral acini and they too differentiate.

To establish the course of differentiation at the molecular level one can monitor the specific functional attributes of the various cell types. For the exocrine cells the digestive enzymes produced in the largest quantities are the best indexes of differentiation; for the endocrine cells insulin and glucagon serve the same purpose. The problem is to measure these proteins specifically and sensitively in the mixture of several hundred different proteins that are present in the cells. We developed microassays based on specific catalytic activity (for the individual digestive enzymes), the ability of specific antibodies to recognize single species of proteins (for insulin and glucagon) and such distinctive physical characteristics of the protein molecules as their size and their mobility in an electric field.

With these assay procedures we discovered that both the exocrine enzymes and glucagon and insulin are present at concentrations many thousands of times higher in functional pancreatic cells than in other cells of the adult organism; these proteins are indeed cell-specific. This finding gave us a considerable experimental advantage, since it meant we could attribute the proteins to pancreas cells even when we measured them in extracts of several cell types, such as pancreas plus gut. We set out to determine the pattern of accumulation of the specific products and relate it to the development of acini and islets and the secretory granules.

By measuring the activity of the major exocrine enzymes present in extracts of tissues we established their developmental profiles during gestation [*see top illustration on page 40*]. Two significant and rather unexpected features of their accumulation patterns became apparent. First, the concentration of proteins (expressed as enzyme molecules per cell) does not change in a simple way, rising from zero to finite high levels in one step. Instead there are three discrete stages in the developmental process. In what we call State I the enzymes are present at a relatively low concentration; then the concentration is 1,000 times or more higher (State II), and finally the concentrations of individual enzymes are adjusted a little with respect to one another (State III). Even the low State I level was much higher than the level in other cell types, and the relative proportions of the individual enzymes approximate those found in State II, indicating that State I is a stage in development unique to the pancreas rather than simply a reflection of lack of development.

The second unexpected feature of the developmental profiles is that the concentrations of the exocrine enzymes do not change as a coordinated set during

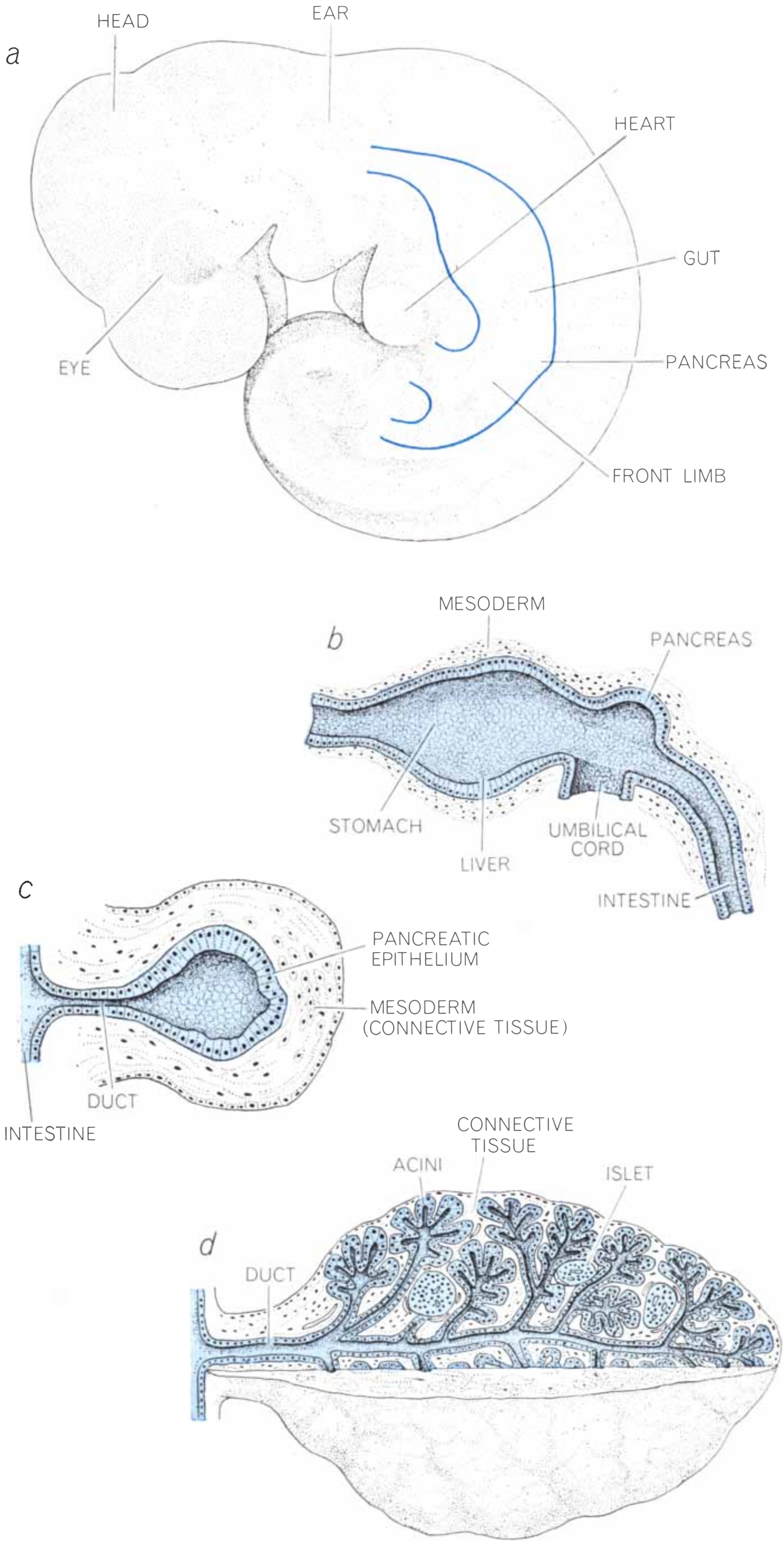


PANCREAS is a secretory organ that lies behind the stomach, as indicated in the schematic drawing of mouse viscera (*a*). A cross section of a small portion of the pancreas (*b*) contains cell clusters called acini, ranged along a duct system, and distinctive cell groups called islets of Langerhans. The acini are composed of exocrine cells with zymogen granules (*c*); islets contain endocrine cells (*d*).

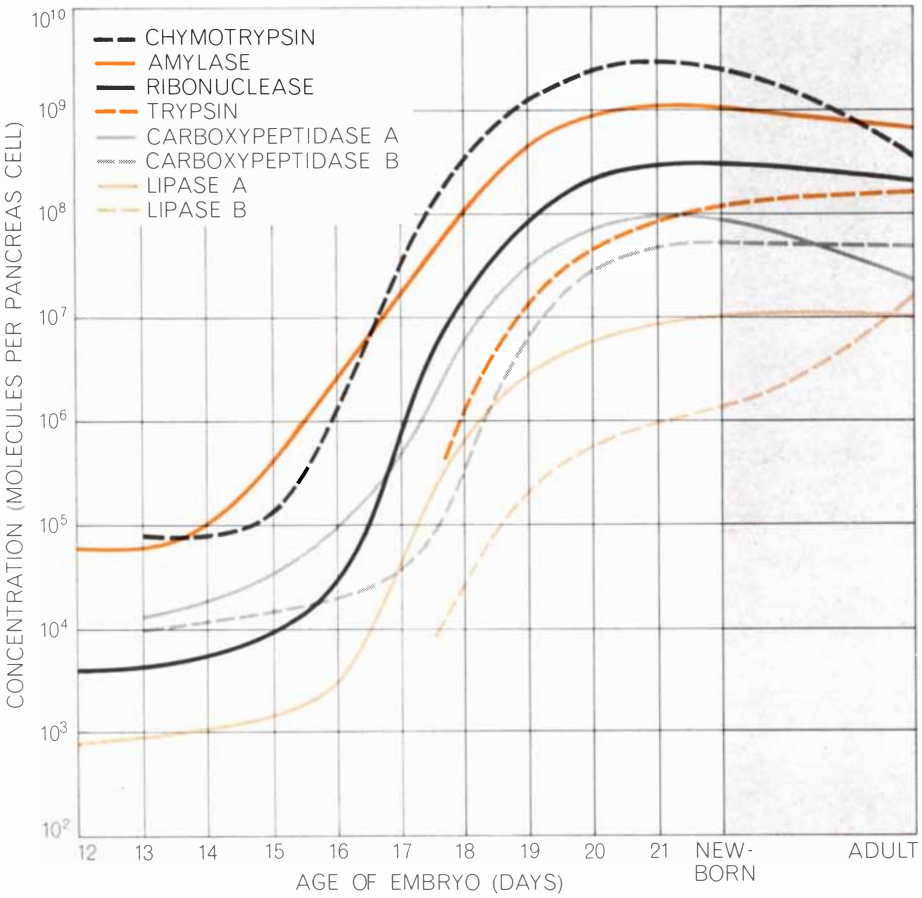
the transition from State I to State II; the pattern of accumulation between 15 and 19 days differs significantly for the various proteins. For example, trypsin and carboxypeptidase *B* lag significantly behind the other proteins; lipase *B* activity is first detected in the rat at 18 days, when the other enzymes have already reached high levels. The exocrine proteins, then, are not all regulated as a unit. On the other hand, the relative concentrations of certain groups of enzymes—for instance lipase *A* and carboxypeptidase *A*, trypsin and carboxypeptidase *B*—appear to change in concert; the curves for the former pair are similar, for example, and their midway points between low and high levels occur at the same time. Perhaps the simplest way to explain the developmental pattern of these proteins is to assume that small sets of them are regulated together and that these sets are synchronized to rise together over a certain period of embryonic development.

With the aid of the electron microscope we correlated these changes in protein content with the appearance of intracellular structures. During the period of great amplification of enzyme content between State I and State II we observed a dramatic formation of the rough endoplasmic reticulum, the site of enzyme synthesis, and saw the first zymogen granules appear. The experiments of George E. Palade and Philip Siekevitz and their collaborators at Rockefeller University had suggested that newly synthesized enzymes first appear within the endoplasmic reticulum and then proceed to the larger cavities of the Golgi apparatus, where they are packaged in zymogen granules. In none of our experiments have we ever observed high levels of specific enzyme formation without the development of these intracellular organelles; the accelerated synthesis of specific proteins appears always to be coupled with the formation of new structures involved in their synthesis, storage and secretion.

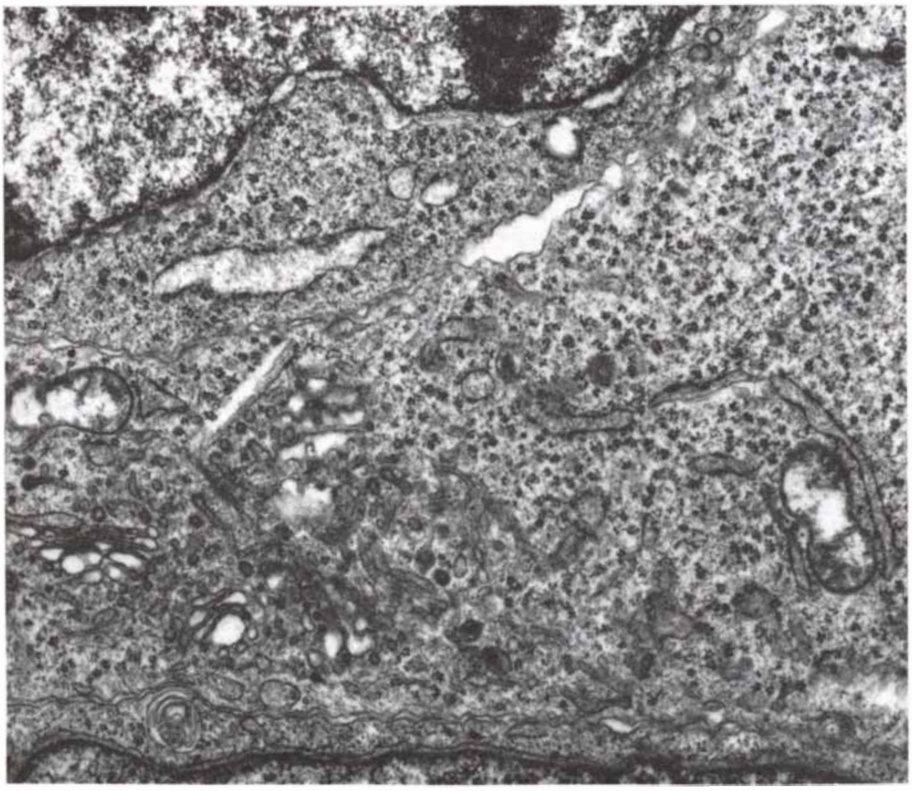
Turning to the endocrine cells, we first established that the insulin content of the adult mouse or rat pancreas is remarkably high: more than a billion molecules per *B* cell. (Very low but significant levels were found in other adult tissues, presumably reflecting the normal dissemination of the hormone from the pancreas.) Our findings in the early embryo were more unexpected: we were unable to detect any insulin before the ninth day, although our assay could have detected one molecule of insulin in 10 cells. Clearly there is little, if any, transfer of



DEVELOPMENT of the mouse pancreas begins when the embryo is in its ninth day of gestation (*a*). The organ forms from the central part of the gut by a process of evagination (*b*): a group of endodermal cells (*color*) bulge into adjacent mesoderm (*black*). By the 11th day the pancreas forms a pouchlike diverticulum (*c*). By the 15th day some pancreas cells have stopped dividing and begun to synthesize large quantities of specific proteins (*d*).



PATTERN OF ACCUMULATION of the digestive enzymes is an index of exocrine-cell development. Extracts of pancreatic rudiments of various ages were prepared and assayed for enzyme content on the basis of specific catalytic activity. The midpoints between low and high levels differ for most of the proteins. Clearly they are not regulated as a single group.



CYTOPLASM of an exocrine cell from a 10-day embryo reflects the cell's differentiation. Some ribosomes are organized into endoplasmic reticulum but there are no zymogen granules yet, perhaps because there is not enough specific protein to require such structures.

insulin from the mother to the embryo and the early embryo itself must synthesize insignificant quantities. Early embryonic development apparently proceeds in the absence of this hormone. There must be remarkably stringent regulation of specific protein synthesis in embryonic cells. Until a given time in development, proteins such as insulin simply are not made.

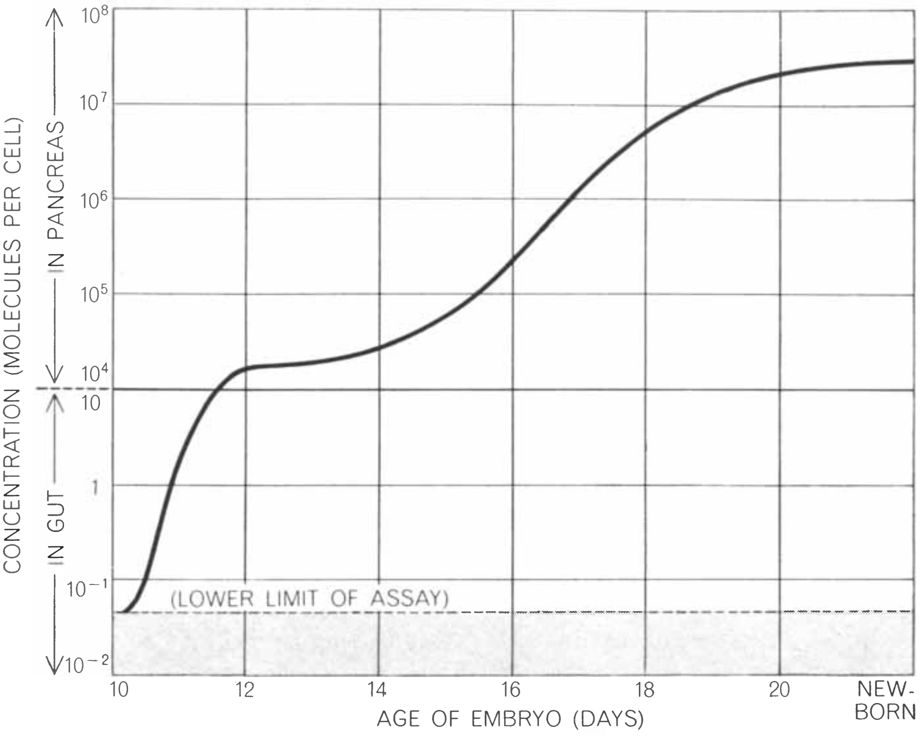
Late on the ninth day insulin is detected in the gut region, where it is presumably present in primitive pancreas cells [*see top illustration on opposite page*]. The concentration then increases until, during the period of State I, there may be as many as a million insulin molecules per *B* cell. Later, at the time the exocrine cells make the transition from State I to State II, the insulin content of the endocrine cells increases another several hundredfold. By applying an antigen-antibody assay for glucagon we have shown that glucagon too is present at the earliest stages of pancreatic development. Electron micrographs show that there are secretory granules in the endocrine cells at this early stage of differentiation. In the exocrine cells, on the other hand, zymogen granules only appear later, during the large rise in specific activity from State I to State II. It seems that the details of the differentiation processes in these two cell types are different.

These correlated studies of pancreatic development suggest a model for the differentiation of pancreatic exocrine and endocrine cells [*see illustration on page 42*]. The model recognizes several distinct stages of development. The protodifferentiated state corresponds to the State I we described earlier. The differentiative state (State II) is found in the late embryo, and the modulated levels found in the adult characterize State III. This implies that there are three regulatory transitions, each introducing a new stage of differentiation.

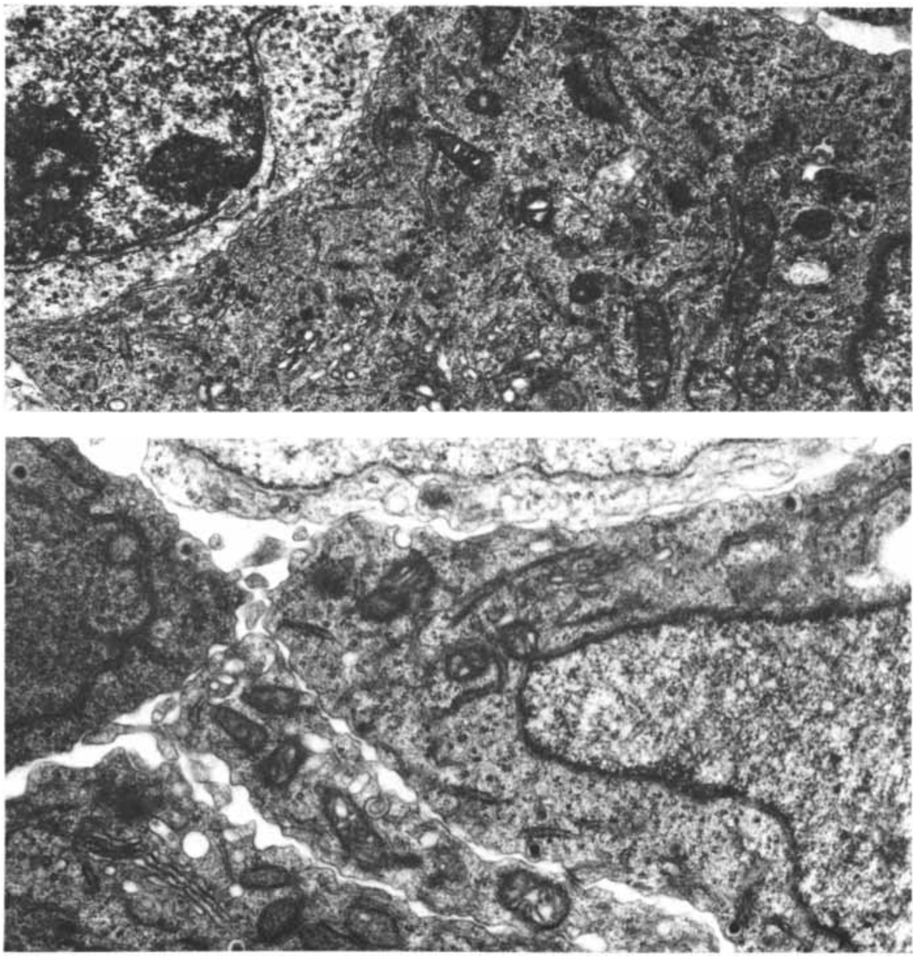
It is difficult, if not impossible, to identify the factors that control these transitions in the intact embryo, because one cannot effectively manipulate the cellular environment or the physical relations among cells. The cultivation of cells or tissues in laboratory vessels seemed to offer a way to obtain further information about the differentiative process, particularly the nature of the regulatory events. We found that a pancreas isolated from a 10-day or 11-day mouse embryo would develop normally in culture, and that both the pattern of increase in enzyme levels and the appearance of zymogen granules closely mimic the developmental sequence in intact embryos. This developmental competence in tissue culture opened up attractive experimental vistas. We could investigate relations among different cell types and also monitor developmental events within cells much more effectively.

First we tested the ability of progressively younger pancreatic tissues and more primitive gut tissues to develop into normal pancreas. We found that even the youngest pancreatic rudiment differentiates normally, provided that both mesodermal and epithelial elements are present. Then we found that gut tissue from the eight-day mouse embryo, excised 18 hours before the pancreas would become visible, develops into normal pancreatic tissue (plus some other tissues) in culture; gut material taken from earlier embryos forms liver-like and stomach-like tissues in culture but no pancreatic cells. It seems, therefore, that some significant developmental event must occur on the eighth embryonic day that confers pancreatic potential on a group of precursor cells. This event presumably requires an input from some part of the embryonic system that is absent in the laboratory culture. After this event, however, the tissue is somehow "determined," so that when it is removed from the embryo, it will develop just as it would in the intact embryo.

The simplest way to test for an interaction between two tissues in an organ is to separate the tissues and see whether each is capable of differentiating normally alone or whether recombination is necessary for normal development. We found we could carry out such an experiment with the pancreatic rudiment in the protodifferentiated state (the 11-day mouse embryo or the 13-day rat embryo). At this point the epithelial cells exist as a bulbar structure encased in mesoderm. We separated the two tissues by first incubating the intact rudiment for a short time in a mixture of protein-digesting enzymes and then stripping off the mesoderm by drawing the rudiment up into a micropipette. We then tested for an interaction between the two components with a technique, perfected by Clifford Grobstein of the University of California at San Diego, in which a thin, porous membrane serves as a platform to support the individual tissues or as a barrier to separate them [*see bottom illustration on page 43*]. We found that the epithelium would not develop normally alone but that it did differentiate if it was cultured directly across the filter from mesoderm. To our surprise,



INSULIN CONCENTRATION in tissue extracts was measured by an antigen-antibody assay and is given here in molecules per gut cell and (when the embryo is large enough for dissection) molecules per pancreas cell. The sensitivity of the assay makes it possible to detect insulin at an earlier developmental stage than one can detect any of the exocrine proteins.



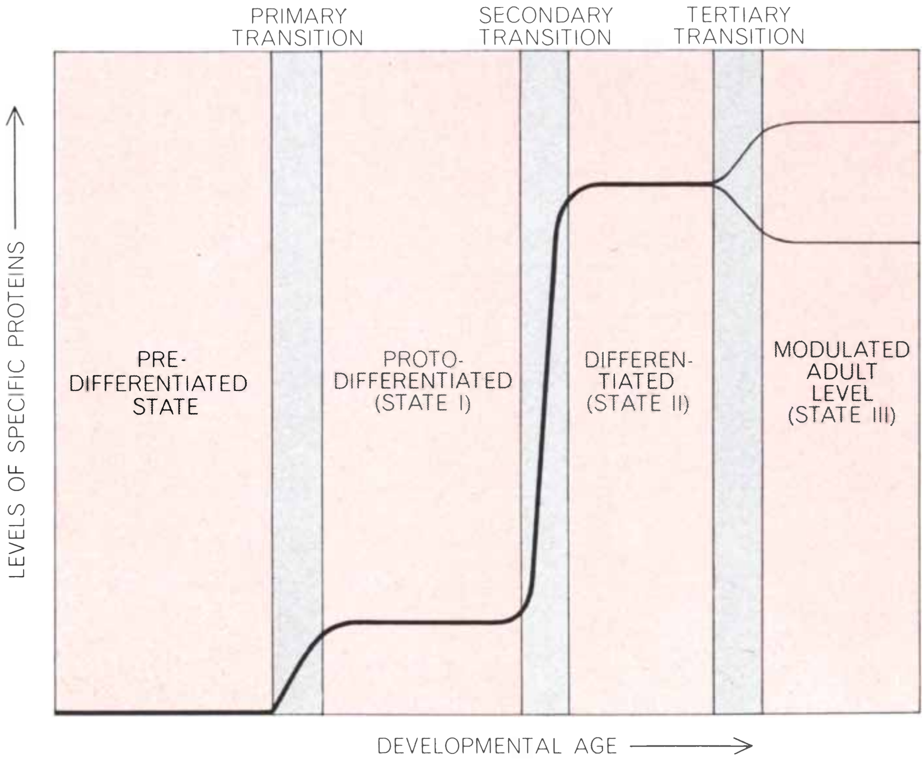
SECRETORY GRANULES appear in endocrine cells as soon as hormones begin to be synthesized. There are no granules in the cytoplasm of future endocrine cells when the pancreas begins to take shape (*top*), but in pancreas cells from an early 10-day embryo, when insulin is first detected (*see illustration at top of page*), a number of granules can be seen (*bottom*).

the mesoderm did not have to be from the pancreas; we learned that mesodermal tissues from salivary gland, kidney, stomach, lung or spleen were just as effective in promoting epithelial development.

Apparently the mesoderm contributed some factor that is required for epithelial differentiation and that is small enough to pass through the filter. To get enough material from which to attempt to isolate the factor, we first prepared a crude homogenate of an entire embryo, which turned out to contain enough of the mesodermal factor to promote pancreatic differentiation. We were able to remove most of the active material from the homogenate by low-speed centrifugation, indicating that the factor was part of (or was bound to) some rather large particulate structure. A number of experiments suggest that it is a protein, or at least that a protein is involved in the activity. The active material has now been solubilized and partially purified.

Is the mesodermal factor required at all times during development of the pancreas or just at a certain time? We have noted an epithelial requirement for the presence of mesoderm as early as the ninth day, when a restricted region of the gut responds to mesodermal tissue by forming pancreatic acini. Nearby epithelium, treated similarly, forms other gut derivatives such as stomach or intestine but does not form pancreatic tissue. Several hours before the pancreas can be seen, then, a discrete group of gut cells has the capacity to form the organ, provided that mesodermal tissue is present. Presumably this ability reflects the determination event that occurred some eight hours earlier. Similar experiments with older pancreatic epithelium indicate that the mesodermal requirement ceases just before the accelerated synthesis of the specific proteins; the mesoderm can be removed from the epithelium late in the protodifferentiated state without any effect on subsequent development. In summary, the mesodermal factor must be present from a time before the pancreas can be seen until about five days later, when cell division ceases and enzymes begin to accumulate at high levels.

How does the mesodermal factor act? It could promote growth, cause the formation of acini or actually cause differentiation in the epithelial cells. Recent results suggest that a major action of the mesodermal factor is to stimulate DNA synthesis and thus the proliferation of epithelial cells. (As we shall see, such DNA synthesis may be required for differentiation to occur.) The eventual availability of pure preparations of the factor may give us clear answers to these questions.

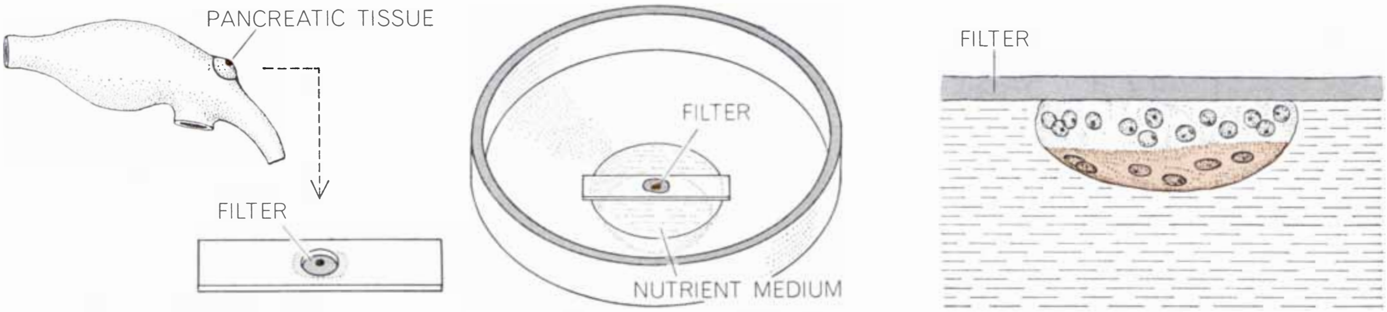


POSSIBLE REGULATORY PHASES in cell differentiation are indicated by this schematic curve. The three stages of differentiation correspond to the numbered states associated with the exocrine enzyme levels. Each of them is introduced by a regulatory transition phase.

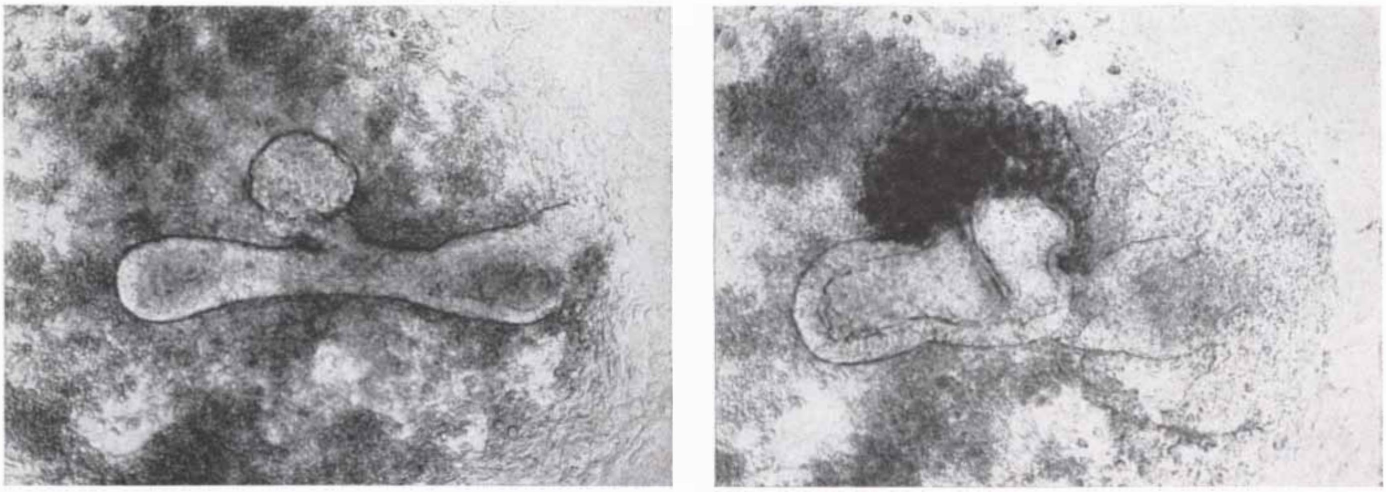
Our efforts to describe and analyze some of the changes that take place at the cellular and molecular level during development have been focused on the secondary transition leading to the differentiated state. First we sought to trace changes in the pattern of protein synthesis during the transition. (What we had measured in intact embryos was enzyme content, not the rate at which enzymes were being synthesized.) We incubated 14-day and 19-day pancreatic rudiments in a medium containing radioactively labeled amino acids—the building blocks of proteins—and then sorted out the resulting radioactively labeled proteins by electrophoresis in polyacrylamide gel. In this technique specific proteins are identified by the rate at which they move through the gel in the presence of an electric field, and the amount of radioactivity at each point shows how much of each protein has been synthesized.

There was a remarkable difference between the two patterns of protein synthesis [*see illustration on page 44*]. Similar experiments with rudiments at various stages of the secondary transition confirmed that there is about a fiftyfold increase in the rate of synthesis of the enzymes during the transition. This increase in rate accounts quantitatively for the increase in enzyme content between State I and State II noted in the earlier assays. There must therefore be little, if any, secretion of these proteins by exocrine cells during this period of development; the proteins that are synthesized simply accumulate within the cells.

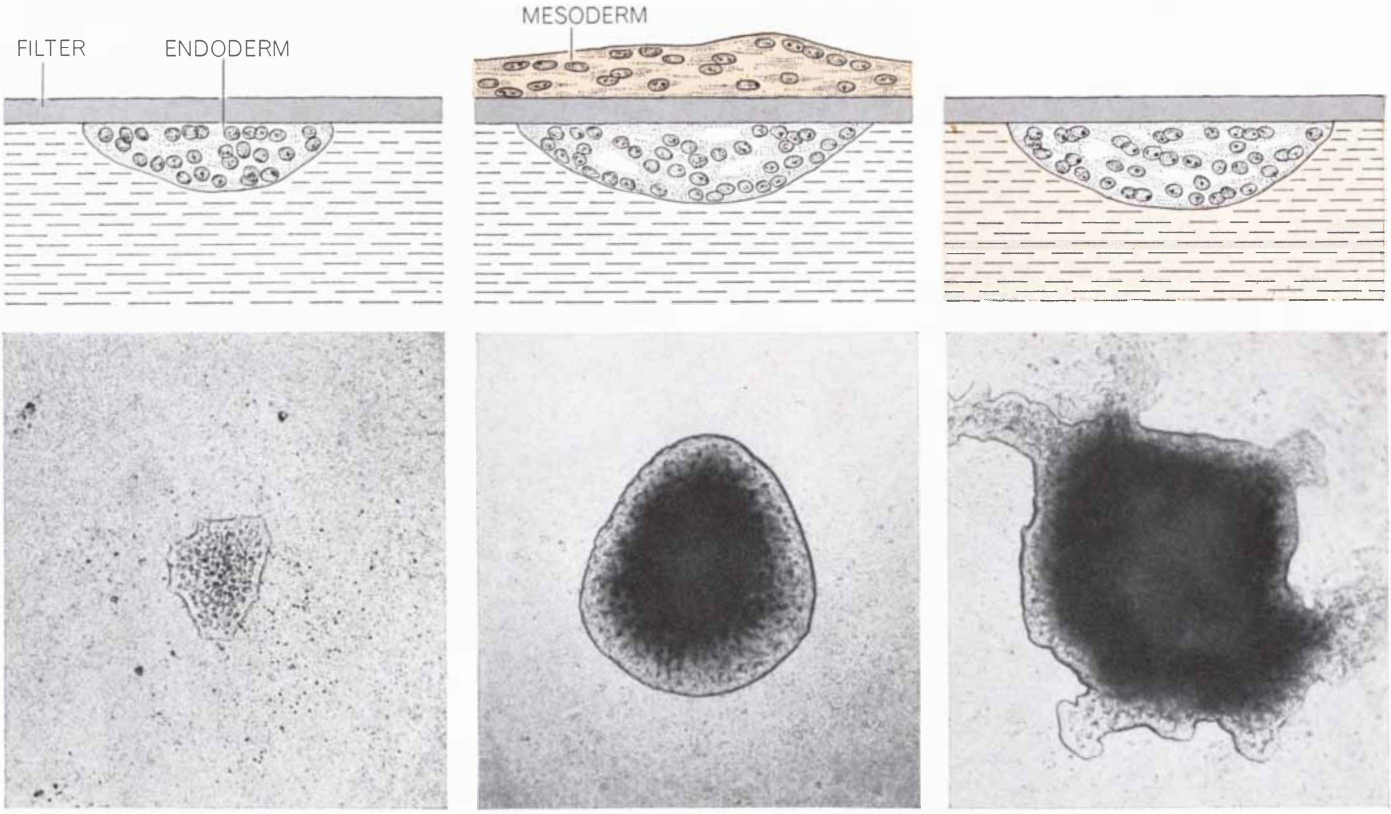
Protein synthesis depends on the activity of a number of species of RNA. We decided to see if the increase in specific enzyme synthesis in pancreas depends on the synthesis of new RNA. For these studies we employed actinomycin *D*, a compound that is bound to the DNA template and thus blocks RNA synthesis. Very low levels of actinomycin applied to dividing cells in the protodifferentiated state inhibited only part (about 70 percent) of the RNA synthesis in those cells but completely prevented the accelerated synthesis of the differentiative proteins. This was a specific effect, since the cells remained healthy and continued to synthesize cellular (as opposed to secretory) proteins at nearly the normal rate. If the tissue was treated with actinomycin at a later time, when the central group of cells in the culture had stopped dividing, those cells differentiated more or less normally and accumulated zymogen granules; cells located on the periphery of the same



PANCREATIC TISSUE is grown in the laboratory. Nine-day gut tissue is excised (*left*) and placed in a culture dish (*middle*). The tissue, mesoderm (*color*) as well as endoderm (*black*), is suspended under a plastic support and incubated in nutrient (*right*).



PANCREAS DEVELOPS in culture. After three days in culture the pancreas bulges upward (*left*). After another four days it has grown, acini have formed and the acinar tissue has been darkened by zymogen granules (*right*). The enlargement is 50 diameters.

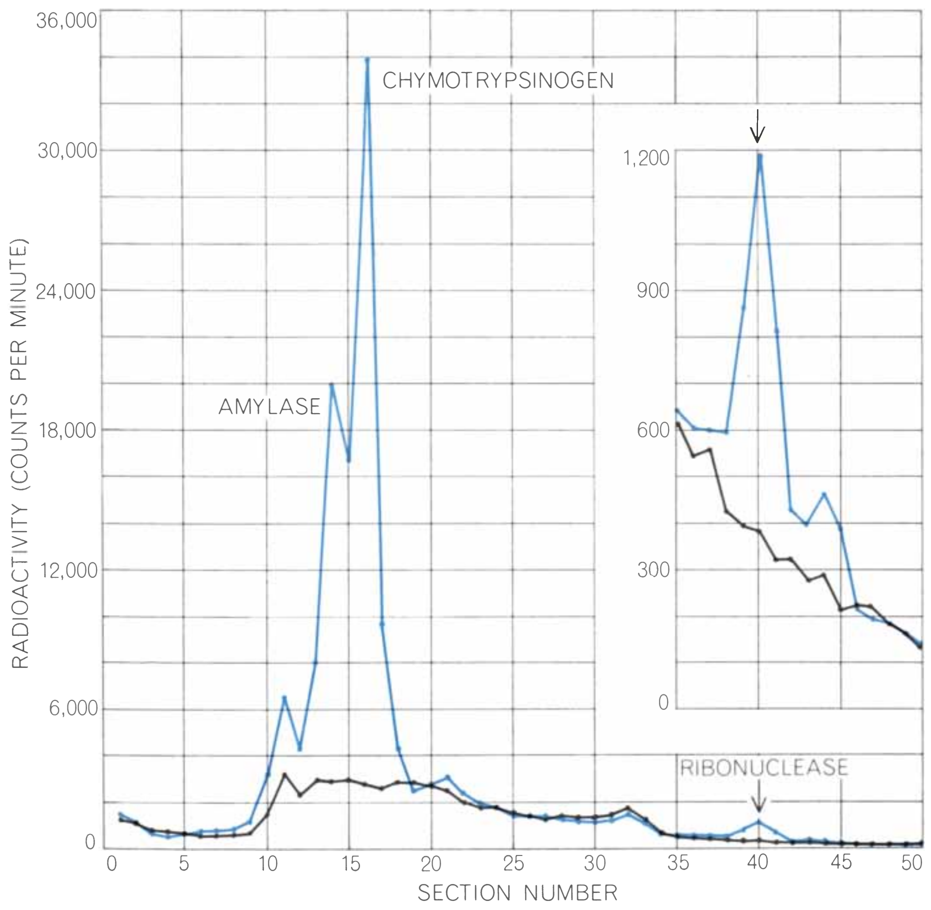


TRANSFILTER CULTURE tests for mesoderm-endoderm interaction. When pancreatic epithelium (endoderm) is cultured alone (*top left*), the cells fail to differentiate in five days (*bottom left*). When epithelium and mesoderm (*color*) are separated by a thin filter and cultured together, there is substantial growth in five days, acini are visible and zymogen granules darken the cells (*center*). Epithelium that is cultured without mesoderm, but with an embryo extract included in the nutrient medium, does equally well (*right*).

rudiments were still dividing and they (like all cells treated in the protodifferentiated state) never made the secondary transition. If cultures were treated still later, when most of the cells had ceased to divide, there was no significant effect of actinomycin on the synthesis or accumulation of the specific proteins. In collaboration with Fred H. Wilt of the University of California at Berkeley we showed that actinomycin has the same inhibitory effect on RNA synthesis in early pancreatic epithelial cells as it does in older cells. The results were therefore not due to a difference in the susceptibility of the two populations of cells to actinomycin but to the fact that the synthesis of new RNA is required for the accelerated synthesis of the pancreatic proteins.

One further experiment with actinomycin is of special significance in relation to our earlier observations about different enzymes being regulated together in groups. When actinomycin *D* is added to cultures early in the secondary transition, the synthesis of the various specific proteins is inhibited to different degrees. This sequential inhibition suggests that the "messenger" RNA's that encode the various proteins are synthesized at slightly different times, supporting our earlier contention that the exocrine proteins are not all regulated as a single set.

Inhibition of DNA synthesis also blocks exocrine cell differentiation. A culture treated with fluorodeoxyuridine (a strong inhibitor of DNA synthesis and hence of cell division) during the protodifferentiated state does not differentiate, although it appears otherwise healthy; cells treated after they have ceased to divide, however, develop normally. Similar results are obtained with bromodeoxyuridine, a structural analogue of thymidine, one of the four nucleosides of DNA. Bromodeoxyuridine is actually incorporated into the DNA during the synthesis period; by forming a faulty DNA that does not allow the synthesis of functional RNA, bromodeoxyuridine blocks the differentiative process. In contrast to the action of actinomycin *D* (which has a sequential inhibitory effect on the synthesis of certain proteins), bromodeoxyuridine exhibits an "all or none" effect, depending on just when it is administered: differentiation is either blocked completely or is unaffected. These experiments suggest that some event that occurs only in dividing cells or is facilitated by DNA synthesis must precede the subsequent changes in RNA synthesis that bring about the new pattern of protein synthesis.



PROTEIN SYNTHESIS in 14-day and 19-day rat pancreatic tissue is compared. The tissues were incubated with a radioactive amino acid, which was incorporated in the proteins being synthesized. When the proteins were subjected to electrophoresis in a gel, they migrated to different points in the gel. The gel was cut into sections and the amount of radioactivity in each section was measured. In the 19-day tissue three new peaks corresponding to specific exocrine proteins are identified. An expanded scale is used (*inset*) to visualize the ribonuclease peak. If 14-day rudiments are cultured for five days, they exhibit a synthetic pattern like that of the 19-day pancreas, unless an inhibiting agent such as actinomycin *D* or bromodeoxyuridine (*see text*) is in the medium during the early part of the culture period.

Our present studies have provided some insight into the basic mechanisms underlying the formation of organs. They demonstrate the importance of cell interaction in the developmental process. Specifically, after the determination event, a mesodermal-epithelial interaction is required to maintain the protodifferentiated state and initiate the secondary transition. It may be that interactions between the exocrine and endocrine cells are also involved in the developmental process.

Perhaps the most striking thing about the molecular events within maturing pancreatic cells is the fact that many complex regulatory events are apparently coordinated at a few distinct times. This is most obvious in the exocrine cells. There some early initiating event in a population of cells starts the synthesis of specific proteins and the morphogenesis of the pancreas. The ensuing protodifferentiated state is primarily a phase of growth and acinus formation. Then there is a transition to the differentiated state, when cells stop dividing and undergo intracellular differentiation, which is marked by the accelerated synthesis and the accumulation of specific proteins in zymogen granules.

The two major developmental phases must involve hundreds of genes. We believe sets of genes are activated and other sets are inactivated simultaneously, rather than each of the hundreds of individual genes' being activated in sequence. One possible mechanism to accomplish such a concerted transition is an alteration of chromosomal structure so that some sets of genes are chemically unmasked, and thus activated, while others are chemically sequestered and thus repressed. These experiments have not yielded a precise description of the molecular mechanisms involved, but they have simplified our view of the problem and suggested a hypothesis that will be the basis for continuing explorations.

In hope of doing each other some good

The never darkened classroom

The "knowledge" industry is reported to have replaced defense as the major growth industry. It demands less complicated engineering. With engineering attracting youth a bit less of late, the resulting engineer shortage may pinch less than had been feared.

In this cheerful mood, it is pleasant to report that audiovisual dealers can now fill orders for KODAK EKTALITE Projection Screens, a significant contribution by our engineers to educational technology. Since the suggested list price of these 40 x 40-inch screens is only $65, they can't be terribly complicated. Nevertheless, their directionality into a space 30 degrees high by 60 degrees wide yields for the same projection six to eight times the brightness of any other screen we know where to buy, or about 16 times the brightness of a matte screen. Ambient light coming from outside the viewing space is reflected outside that space. It cannot dilute the color of the image nor kill its contrast. That's the main point.

So without the fuss and perils of going dark, the teacher moves frequently and easily in and out of the picture mode of communication, with inexpensive, low-powered equipment. Today, of course, kids are totally immersed in an endless stream of miscellaneous images, but the right image at the right moment is still most desirable.

Another reason for calling the KODAK EKTALITE Projection Screen significant is that classrooms which need never be darkened cost less. There needn't be any better reason.

Salt-rejecting membranes

Osmosis isn't as simple as the ninth-grade teacher claimed. No explanation is currently available as to why these unnatural membranes—which are cellulose acetate and quite non-ionic in character—are as good as they are at rejecting salt. No explanation is available, but the membranes themselves are available in three types, along with a fourth that rejects no salt but does reject solutes of molecular weight above 300.

We start with a package deal for $25. The package contains three 8 x 10-inch sheets of each of the four types.

One of the types, when fed 0.5% NaCl solution at 600 psi and 78F, rejects about 97% of the salt in passing about 20 gal/day/ft^2. The second type rejects 89% from 34 gal/day/ft^2. Both of these types come wet in our package, each type separately sealed in polyethylene. We mark the active side—the side bearing the mysterious one-micron-thick layer that must face the feed stock. If you let them dry, give up hope that rewetting will restore them.

NaCl seems to represent the tough case. When working with something easier and you want more flux, the third type comes untempered to let you set the parameter yourself.

Yours in any case is the problem of providing a support for the membrane to withstand operating pressures of 600 to 1500 psi and still permit flow of product water to your collecting system. Don't let pH exceed 8.0.

If your experiments should prove encouraging, base no economic feasibility studies on what you paid for the sample package. Most of the 30 or so companies believed currently busy developing equipment for what is being called "reverse osmosis" seem to be glad we are going to offer the membrane. Now they can concentrate their own efforts on equipment to put it in. Thus will go forward the great endeavor to squeeze good water out of bad water, brackish water, sea water, radioactivity-contaminated water, or water carrying other things in solution. And some will care more about the solutes being concentrated than about the purity of what comes out.

Enter purchase order to Eastman Chemical Products, Inc. (Subsidiary of Eastman Kodak Company), Kingsport, Tenn. 37662 for "Sample Kit of EASTMAN *Membrane." Bibliography will be shipped with order. Good luck!*

Observant, quick, clever, obedient

You can get a new KODAK INSTAMATIC® Reflex Camera body for less than $158.*

Less than $158! Incorporates interesting exposure control system to which you may not have been exposed. Leaves it to you to select aperture. You may wish to do so on depth-of-field basis. Since this is a single-lens reflex, preview on field-lens-brightened ground glass can guide you.

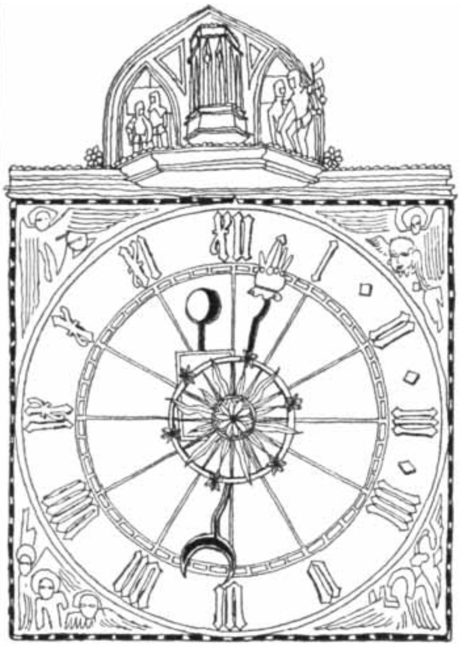
When picture is wanted, press button to open shutter. Will close by itself when enough light has hit film. Bases *its* decision on film speed information conveyed by notch on 126 "drop-in" cartridge. Principle: capacitor fed through photoresistive cell attains limiting charge. This may take 1/217 sec or 9.1 sec, or any other exposure time >.002 sec and <20 sec, *depending on light when picture is actually being taken.* If, however, it appears in advance that exposure will be longer than 1/30 sec, signal in viewfinder warns master to open lens more or rest camera on firm support.

This is a clever camera. Even has button to check battery condition. Camera shop will be pleased to point out other noteworthy features. Do not expect salesman to remind you that photography can also be done with a pinhole and home-made emulsion hand-coated on window glass.

*Less than $200 with 45mm *f*/2.8 Xenar Lens. Needn't get lens if you already have one from a RETINA Reflex Camera. Or if KODAK INSTAMATIC Reflex Microscope Adapter wanted instead of lens. For less closeup than microscopy, other accessories available.

EASTMAN KODAK COMPANY An equal-opportunity employer

Prices subject to change without notice.



SCIENCE AND

Visible Pulsar

After more than a year of fruitless searching at many observatories a pulsar has finally been associated with an optically visible object by astronomers at the University of Arizona. First recognized in the fall of 1967, pulsars emit bursts of radio energy at precisely spaced intervals, ranging from a pulse every 3.7 seconds for the slowest to about 30 pulses a second for the fastest. It is the fastest, NP 0532, that has now been detected by optical telescopes. Like the radio pulse, the optical pulse flashes rhythmically 30 times a second. The optical identification was made by W. John Cocke, Michael J. Disney and Donald J. Taylor, with the aid of a computer linked to the 36-inch reflector at the Steward Observatory.

NP 0532, discovered four months ago, is of particular interest because it lies in the Crab nebula, the remnant of a supernova observed in A.D. 1054. The radio flashes were found to originate near a double star in the center of the nebula. The "south preceding" star of the pair had been identified in 1942 as the probable "cinder" left by the supernova because its spectrum was peculiar: it lacked emission lines. The spectrum of the "north following" star showed lines characteristic of a normal star.

In making their optical identification the Arizona workers could not be sure which of the two stars the light flashes were coming from. Because their telescope was fairly small they had to use a diaphragm whose field was so wide that it included both objects. A subsequent series of observations made with the more powerful two-meter reflector at the Kitt Peak National Observatory by C. Roger Lynds, Stephen P. Maran and Donald E. Trumbo established that the flashes were indeed coming from the south-preceding star. Soon afterward workers at the Lick Observatory snapped NP 0532's picture (*see illustration on page* 49).

The optical flashes closely mimic the radio flashes. In each cycle there is a major pulse followed about 14 milliseconds later by a minor pulse. Moreover, nearly all the light emitted by the pulsar seems to be confined to the flashes; if there is a steady-state component, it does not account for more than 6 percent of the total light emission. This strongly suggests that the source resembles a rotating beacon, perhaps a spinning neutron star, as has been suggested by Thomas Gold of Cornell University. A neutron star would be only five or 10 miles in diameter but would have as much mass as the sun. Under these conditions electrons and protons would be packed together so tightly that they would form neutrons. The cinder left by a supernova explosion might well resemble such an object.

First Synthesis

The first laboratory synthesis of an enzyme has been achieved. The enzyme is ribonuclease, which cleaves ribonucleic acid (RNA). Its synthesis was accomplished independently and by different methods but at about the same time by two groups, one at Rockefeller University and the other at the Merck Sharp & Dohme Research Laboratories. The results have been published in the *Journal of the American Chemical Society*.

The task in synthesizing a protein such as an enzyme is to form one or more linear arrays of amino acid subunits and then cause the chains to be folded and cross-linked into the proper three-dimensional configuration. Fortunately it turns out that if the amino acids constituting a protein chain are put together in the right order and placed in solution under favorable conditions, the chains will fold and cross-link by themselves. The first protein to be synthesized in this way (by groups in the U.S., Germany and China in the early 1960's) was the hormone insulin, with 51 amino acids. The question was then whether more complex proteins, notably the enzymes, could also be made in the laboratory.

Ribonuclease was a good candidate for synthesis because its complete 124-amino-acid sequence and cross-linking pattern had been established in 1960. At Rockefeller University, R. B. Merrifield and Bernd Gutte undertook to synthesize it by the automatic "solid phase" method, which had been developed by Merrifield and by which he had already synthesized insulin (see "The Automatic Synthesis of Proteins," by R. B. Merrifield; SCIENTIFIC AMERICAN, March, 1968). Tiny polystyrene beads served as anchors for the growing ribonuclease chains. The first amino acid was bound to the beads. Then each of the 123 other subunits was in turn protected against unwanted side reactions, bonded to the chains and deprotected in preparation for the next step. The entire operation was achieved automatically in a single vessel; the amino acids and the other reagents and solvents were admitted to the vessel as required and then withdrawn through a filter that retained the beads with their growing chains. Finally the completed chains were removed from the beads, purified and caused to fold into the proper configuration. The 369 reactions, requiring 11,931 steps, were accomplished in three weeks of continuous operation—something that might have taken several years by traditional methods of chemical synthesis.

At Merck Sharp & Dohme a team headed by Robert G. Denkewalter and Ralph F. Hirschmann capitalized on the fact that natural ribonuclease can be cleaved into two segments, the 104-amino-acid S-protein and the 20-amino-acid S-peptide, that can fold together, without actually bonding covalently, to form an enzymatically active protein. The Merck group first prepared a number of S-protein fragments: small chains of from six to 17 amino acids. They solved the difficult protection and deprotection problem by developing the carboxyanhydride method, in which the protecting group is easily removed simply by changing the pH of the solution. Then they combined the fragments to form the complete S-protein, which was allowed to combine with natural S-peptide to produce the enzyme. Both teams' products showed the proper enzymatic activity, cleaving RNA but having no effect on closely related DNA, and passed other tests indicating that the active enzyme has indeed been synthesized.

Pollution by Fertilizer

Man's injection of excess nitrogen into the biosphere not only is seriously polluting rivers and lakes but also has greatly increased the frequency of a rare form of poisoning among both humans and domestic animals. Reporting at a symposium on pollution during the annual meeting of the American Association for the Advancement of Science, Barry Commoner of Washington University noted these results of the rise in nitrogen production in the U.S. during the past 25 years. Nitrogen from sewage, a reflection of the nation's population growth, has increased no more than 70 percent in that time, but the oxides of nitrogen produced by power plant fumes and automobile exhaust have increased by 300 percent and the use of nitrogen fertilizers by 1,400 percent. The last two processes add some 10 million tons of nitrogen compounds to the environment each year, more nitrogen than is cycled annually within the U.S. by all the processes of nature combined.

The excess quickly appears in the form of nitrate in ground water, entering rivers and lakes to nourish algal "blooms" that reduce the oxygen content of the water to the point where processes of self-purification are halted. A measure of the quantities involved is provided in a recent Federal Water Pollution Commission study of Lake Erie, which estimates that municipal sewage annually contributes 90 million pounds of nitrogen and fertilizer runoff from surrounding farmlands 75 million pounds of nitrogen to the lake yearly.

Concentration of nitrate in ground water also affects wells and other potable water supplies. The maximum public health limit on nitrate concentration in drinking water is 10 parts per million. Already the water in a quarter of the shallow wells in Illinois exceeds this limit, as do many wells in southern California. Excess nitrate concentration in the water of one Minnesota town has forced complete replacement of the water system.

The danger to health from nitrate-polluted water arises not from the nitrate, which is relatively innocuous, but from the fact that certain intestinal bacteria convert nitrate (NO_3) into nitrite (NO_2). Nitrite combines with hemoglobin and destroys its oxygen-transporting properties. Commoner notes that polluted water is not the only source of danger. Poisoning from an overabundance of nitrate in foodstuffs has been reported in Europe; children under the age of three months are particularly vulnerable. One French pediatrician has set the maximum permissible concentration of nitrate in foodstuffs at 300 parts per million. A recent study of commercial baby foods, Commoner points out, found 444 parts of nitrate per million in wax beans, 977 in beets and 1,373 in spinach.

Noting that it is at least theoretically possible to halt nitrate pollution from municipal and industrial wastes, Commoner cites the painful dilemma presented by fertilizers: high agricultural yields and thus the whole of the farm economy depend on fertilizers, yet it is physically impossible to prevent the escape of nitrates from field to water table. To impose limitations on the use of fertilizers would be to reduce food supplies at a time of mounting worldwide needs, as well as to engender fierce opposition from farmers and the chemical industry. To fail to do so exposes the environment to steady degradation. "Science can reveal the depths of this crisis," Commoner concludes, "but only social action can resolve it."

One Oceanography

The U.S. Government's many oceanographic programs are now being supervised by a single unit: the Marine Sciences Council. The need for such a coordinator of Federal activities in the fields of oceanography, marine biology and related studies was recognized in 1966, when Congress established the council under the Marine Sciences Act. Until that time six departments and four independent agencies had engaged in oceanographic research with little regard to one another's activities. The council, in a recent report to the White House, presents a unified picture of Government activities budgeted at $471 million in 1968–1969 and projected at $528 million in 1969–1970.

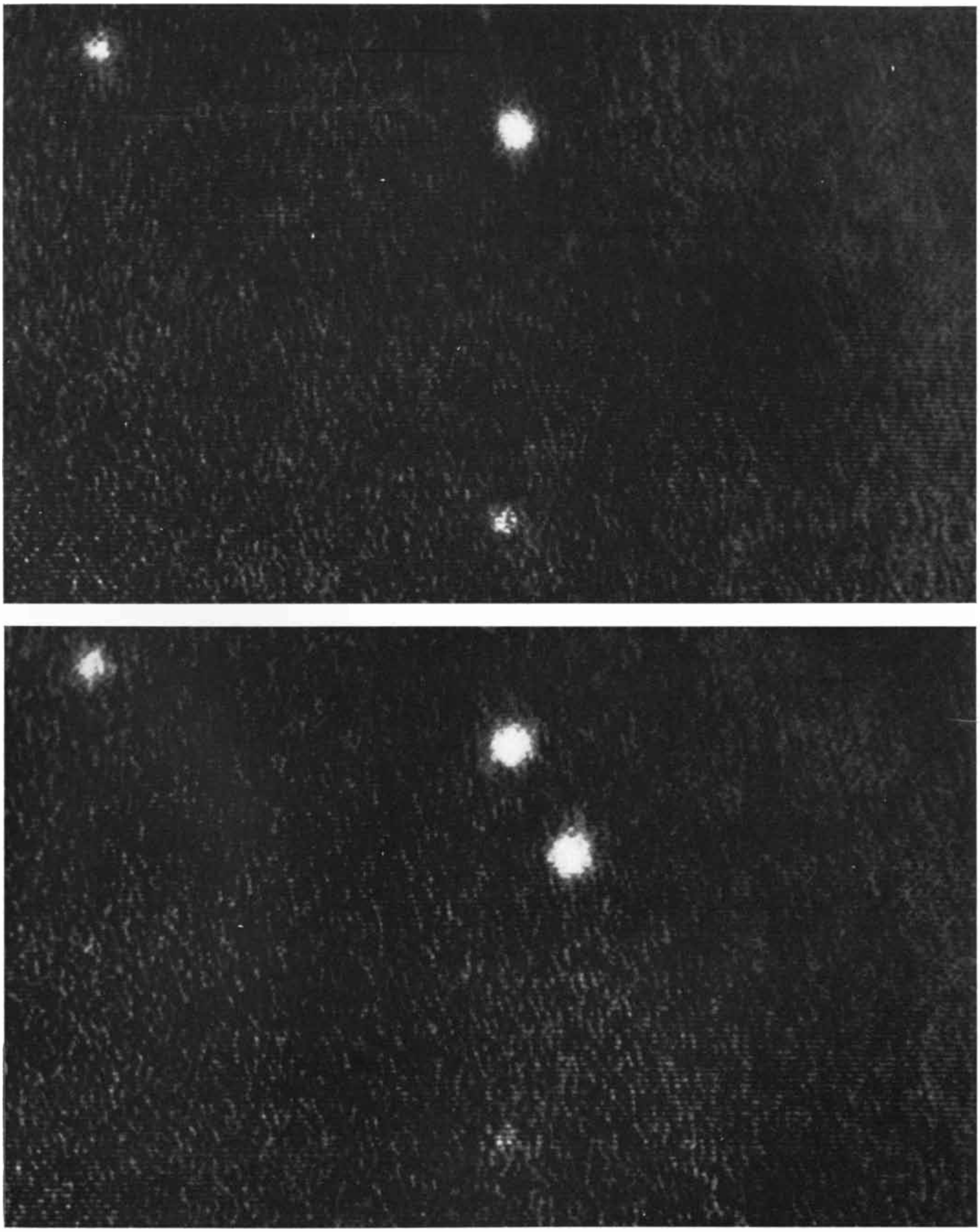
The greatest single projected expenditure, some $300 million, is allocated to the Department of Defense. Roughly half of the amount is for military projects; the remainder is to be spent for research, charting and geodesy. Oceanographic studies by other departments and agencies will cost an additional $150 million; $40 million is earmarked for fisheries research. Some studies have already produced applications. Surveys of the outer continental shelf have been followed by the leasing of mineral rights by the Department of the Interior; three such leases paid the Government $1.7 billion in 1967–1968.

An example of coordinated effort is the progress toward developing a protein-rich food additive from fish meal. The Bureau of Commercial Fisheries, a part of the Department of the Interior, has pioneered production of an odorless, tasteless substance that, in 10-gram units, provides a child's daily requirement of animal protein at a cost of less than a cent. Approval of the fish protein concentrate for human consumption was the work of the Food and Drug Administration, a part of the Department of Health, Education, and Welfare. The purchase and distribution of nearly $1 million worth of the additive for tests of its acceptability in North Africa, Asia and South America is being handled by the independent Agency for International Development.

Two Xis from Brookhaven

From 700,000 photographs of nuclear events produced by the country's largest accelerator (at the Brookhaven National Laboratory) over a three-year period, an international team of 11 physicists winnowed out 150 photographs in which they were able to identify two new subatomic particles and to confirm the existence of a third. The three particles are all members of the xi class, of which three members of lower mass had been known. The existence of the new particles had been predicted by the SU(3) theory of elementary particles—more commonly called the eightfold way—proposed independently in 1961 by Murray Gell-Mann of the California Institute of Technology and Yuval Ne'eman of Israel. In their theory particles that have certain properties in common (spin and parity) can be classified as members of a "supermultiplet," or superfamily (see "Strongly Interacting Particles," by Geoffrey F. Chew, Murray Gell-Mann and Arthur H. Rosenfeld; SCIENTIFIC AMERICAN, February, 1964).

The SU(3) theory applies to the strongly interacting particles, the baryons and the mesons, which respond to the strong, or nuclear, force. More than 100 of these particles are now known. The baryons can all be assigned to six classes: the N, or nucleon, class and the classes designated lambda, sigma, delta, xi and omega. Each class consists of one to four "ground state" members and "resonances" of higher mass. The classes differ in certain quantum properties. According to the SU(3) theory a typical su-



FIRST PICTURES OF A PULSAR were made at Lick Observatory with the aid of a sensitive television system operated like a stroboscope in conjunction with the observatory's 120-inch telescope. In the top picture the pulsar is invisible; in the bottom picture it has flashed on. The pulsar is NP 0532, the "south preceding" star of a pair near the center of the Crab nebula. That one of the pair emitted flashes of visible light about 30 times a second was established at the University of Arizona (see "Visible Pulsar," page 46). At Lick the light from NP 0532 was interrupted by a rotating disk perforated with six evenly spaced holes. When the disk was spun at about five revolutions per second, not quite in synchrony with the pulsar, the light flashes from the object would pass through the holes for several revolutions, then fall out of step and be obscured for several revolutions. When viewed on a television screen, NP 0532 appeared to be pulsing in slow motion. The system was devised by E. Joseph Wampler and Joseph S. Miller.

perfamily of eight members consists of two N particles, one lambda particle, three sigma particles and two xi particles. Other superfamilies have 10 members: four deltas, three sigmas, two xis and one omega. Thus xis are needed to complete both kinds of superfamily. Until recently, however, xis were not being found fast enough to keep up with the rate of discovery of other baryons such as N particles and deltas, which are fairly easy to "manufacture."

Before the present Brookhaven study, which was reported in a recent issue of *Physical Review Letters,* only three members of the xi class had been positively identified: the ground-state member with a mass of 1,320 million electron volts (MeV), one resonance with a mass of 1,530 MeV and a second with a mass of 1,930 MeV. There was also evidence for a third xi resonance with a mass of about 1,830 MeV. The Brookhaven experiment confirmed the existence of the 1,830-MeV xi particle and found two heavier xi resonances: one of 2,030 MeV and another of 2,430 MeV.

Crystalline RNA

An international effort to produce good single crystals of ribonucleic acid (RNA) molecules has met with success almost simultaneously in several laboratories. The way is now open to apply to molecules of nucleic acid the methods of X-ray crystallography that have unraveled the three-dimensional structures of protein molecules. The RNA molecules that have been crystallized belong to the family of "transfer" RNA's (*t*RNA's) that play a central role in the synthesis of protein. The *t*RNA's select specific amino acids from the pool of 20 varieties required to make proteins and deliver them to a specific site on threadlike molecules of "messenger" RNA, which, rather like a punched tape, carry the coded instructions for creating a particular protein molecule. Molecular biologists are curious to know the precise structure of *t*RNA that enables it to bind itself on the one hand to an amino acid and on the other to a particular site on the messenger RNA.

Within the living cell there are 40 to 60 slightly different kinds of *t*RNA, at least one for each of the 20 amino acids. Each *t*RNA consists of some 75 nucleotide units in a chain that seems to be folded in a cloverleaf pattern (see "The Nucleotide Sequence of a Nucleic Acid," by Robert W. Holley; SCIENTIFIC AMERICAN, February, 1966). Success in producing microcrystals of a purified *t*RNA was reported in *Nature* last September by a seven-man group headed by Brian F. C. Clark in the Medical Research Council's Laboratory of Molecular Biology at the University of Cambridge.

Now *t*RNA crystals up to two millimeters long, suitable for X-ray study, are reported by two laboratories in a recent issue of *Science*. One account is by Sung-Hou Kim and Alexander Rich of the Massachusetts Institute of Technology. The other is by a seven-man group at the University of Wisconsin headed by Arnold Hampel. The M.I.T. workers grew their crystals in a two-phase system of chloroform and water, from which chloroform was slowly removed in a desiccator held at four degrees Celsius for a period of two weeks. The resulting crystals contain 87.5 percent water. In the Wisconsin method crystals are produced from an ethanol-water solution held at eight to 10 degrees C. Crystals containing only 58 percent water formed within three days to two weeks. Because of their lower water content crystals produced by the second method are easier to handle and are more suitable for X-ray analysis.

Slow Aging

Experiments with substances that scavenge free radicals in the body have significantly increased the normal lifespan of mice, according to Denham Harman of the University of Nebraska. Writing in the *Journal of Gerontology,* Harman describes work designed to test the hypothesis that free radicals, which are highly reactive molecular fragments, play a role in aging by randomly causing the deterioration of cells. The radicals, which are highly reactive because they have at least one unpaired electron, can be produced by normal metabolic processes or by exposure to ionizing radiation. Whatever the mechanism, a molecule such as water (H_2O) can give rise to a free radical such as OH. The radical reacts readily with other substances, thereby disrupting them or changing their character.

On Harman's hypothesis reactions with free radicals might, for example, change the properties of cell membranes or cause the release of lysosomal enzymes that damage DNA. He believes that if such changes do play a role in aging, the process might be slowed by the addition of nontoxic antioxidants to food. In that way one would raise the concentration in the body of compounds capable of reacting rapidly with free radicals so that fewer of them would be available for harmful reactions. In his experiments with mice Harman has add-

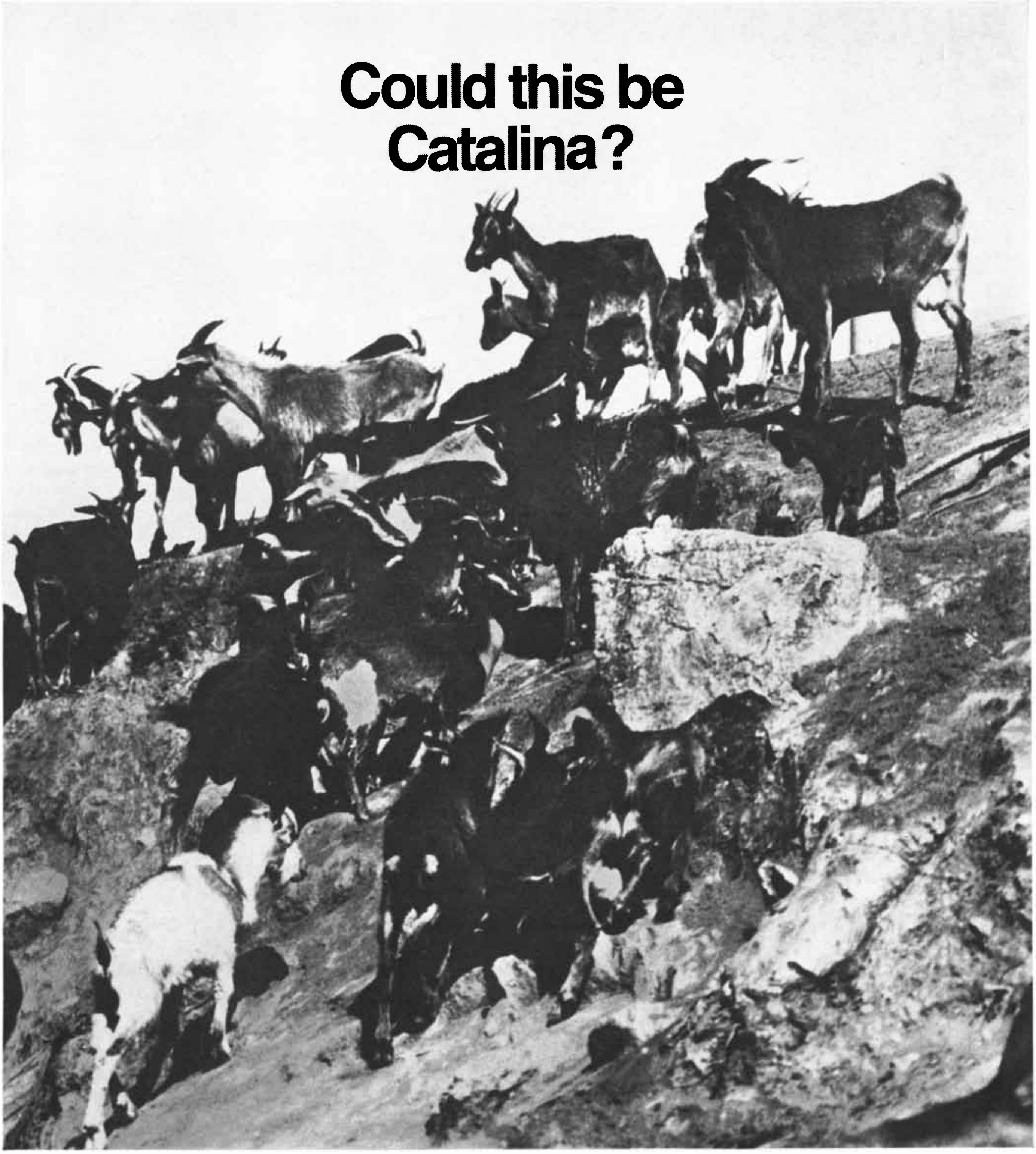
ed small amounts of such free-radical inhibitors as butylated hydroxytoluene (BHT) and 2-mercaptoethylamine to the food given the mice. The results of his most recent experiments, he reports, indicate an increase of 30 to 40 percent in the mean longevity of mice receiving .5 percent BHT in their diet.

Murder at Thank God Bay

The bizarre murder of an early Arctic explorer by members of his expedition has been documented 97 years later by neutron-activation analysis, a sophisticated tool of nuclear physics used in the measurement of trace elements. Analysis of tissue from the frozen grave of Charles Francis Hall on the northwest coast of Greenland has revealed large quantities of arsenic in the hair and fingernails, indicating that his mysterious death was caused by poisoning.

Hall, a somewhat eccentric amateur explorer, was leading an expedition financed by Congress in an effort to have an American party reach the North Pole before any other group. He set sail from Brooklyn in June, 1871, on the former Navy ship *Polaris,* but by the time he had reached the Arctic Ocean it was blocked with ice and he was forced to turn back for the winter. He put in at a place he named Thank God Bay, 500 miles south of the Pole. According to records in the Smithsonian Institution, his decision was challenged by Sydney Budington, the master of the *Polaris,* and Emil Bessels, the head of the scientific staff, who wanted to head south to safer waters. The three men quarreled bitterly. Then Hall became violently sick. He died in two weeks, saying that he had been poisoned. The 30 other members of the expedition buried him and sailed south, and were rescued months later.

Last summer Chauncey Loomis, Jr., of Dartmouth College and Franklin Paddock, a physician from Lenox, Mass., found Hall's grave and removed samples of tissue. At the Toronto Center for Forensic Sciences the tissues were subjected to neutron-activation analysis, in which a sample is bombarded with neutrons in a nuclear reactor. Atoms in the sample become radioactive and each species emits a characteristic spectrum of radiation. The method is sensitive enough to resolve the radiation from as little as a trillionth of a gram of an element. The analysis showed that Hall had ingested massive amounts of arsenic about two weeks before he died. The results, Paddock said, "point to only one conclusion: Hall was poisoned."


Could this be
Catalina?

CONTINENTAL DRIFT AND EVOLUTION

The breakup of ancient supercontinents would have had major effects on the evolution of living organisms. Does it explain the difference in the diversification of reptiles and mammals?

by Björn Kurtén

The history of life on the earth, as it is revealed in the fossil record, is characterized by intervals in which organisms of one type multiplied and diversified with extraordinary exuberance. One such interval is the age of reptiles, which lasted 200 million years and gave rise to some 20 reptilian orders, or major groups of reptiles. The age of reptiles was followed by our own age of mammals, which has lasted for 65 million years and has given rise to some 30 mammalian orders.

The difference between the number of reptilian orders and the number of mammalian ones is intriguing. How is it that the mammals diversified into half again as many orders as the reptiles in a third of the time? The answer may lie in the concept of continental drift, which has recently attracted so much attention from geologists and geophysicists [see "The Confirmation of Continental Drift," by Patrick M. Hurley; SCIENTIFIC AMERICAN, April, 1968]. It now seems that for most of the age of reptiles the continents were collected in two supercontinents, one in the Northern Hemisphere and one in the Southern. Early in the age of mammals the two supercontinents had apparently broken up into the continents of today but the present connections between some of the continents had not yet formed. Clearly such events would have had a profound effect on the evolution of living organisms.

The world of living organisms is a world of specialists. Each animal or plant has its special ecological role. Among the mammals of North America, for instance, there are grass-eating prairie animals such as the pronghorn antelope, browsing woodland animals such as the deer, flesh-eating animals specializing in large game, such as the mountain lion, or in small game, such as the fox, and so on. Each order of mammals comprises a number of species related to one another by common descent, sharing the same broad kind of specialization and having a certain physical resemblance to one another. The order Carnivora, for example, consists of a number of related forms (weasels, bears, dogs, cats, hyenas and so on), most of which are flesh-eaters. There are a few exceptions (the aardwolf is an insect-eating hyena and the giant panda lives on bamboo shoots), but these are recognized as late specializations.

Radiation and Convergence

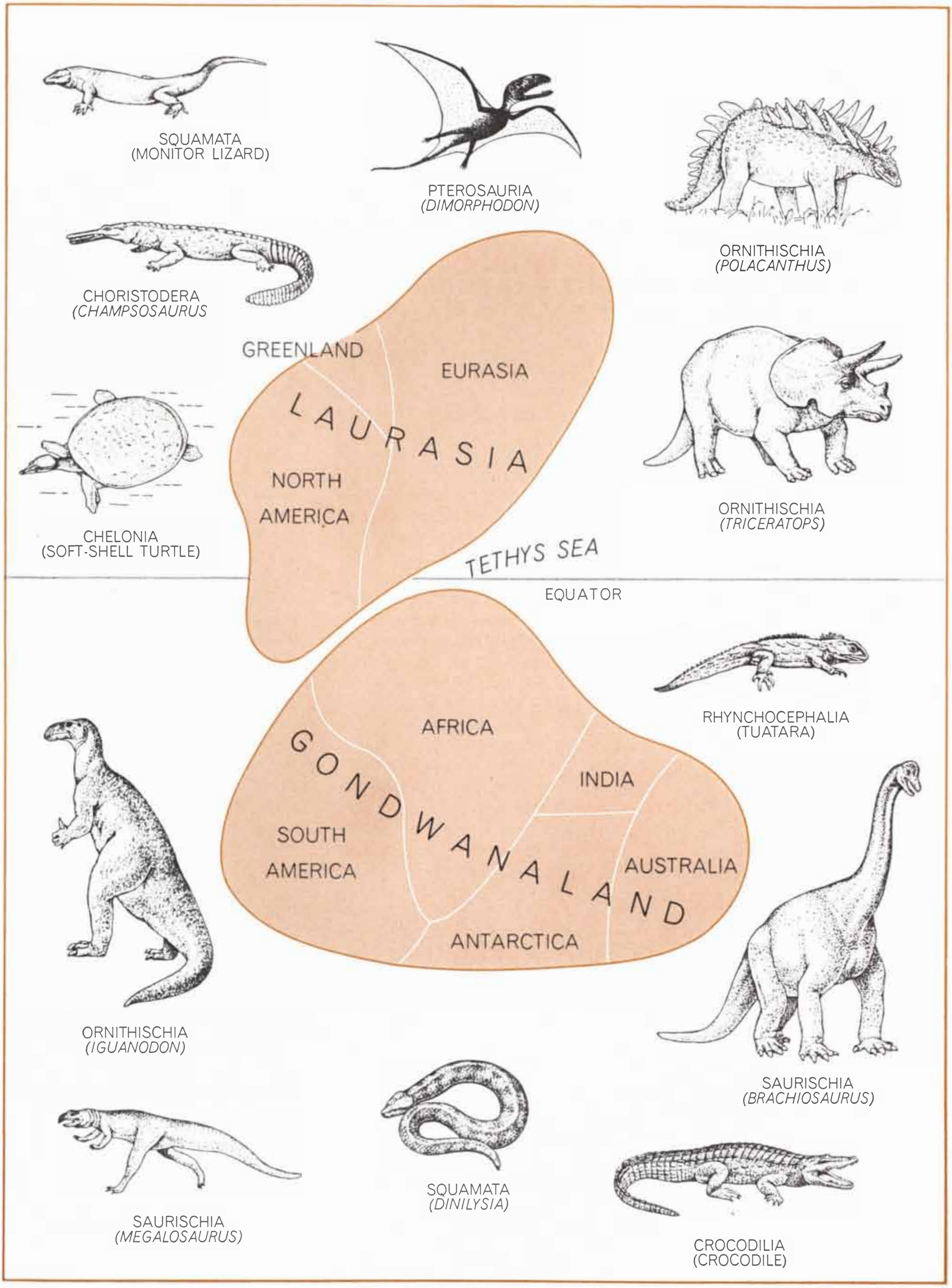
In spite of being highly diverse, all the orders of mammals have a common origin. They arose from a single ancestral species that lived at some unknown time in the Mesozoic era, which is roughly synonymous with the age of reptiles. The American paleontologist Henry Fairfield Osborn named the evolution of such a diversified host from a single ancestral type "adaptive radiation." By adapting to different ways of life—walking, climbing, swimming, flying, plant-eating, flesh-eating and so on—the descendant forms come to diverge more and more from one another. Adaptive radiation is not restricted to mammals; in fact we can trace the process within every major division of the plant and animal kingdoms.

The opposite phenomenon, in which stocks that were originally very different gradually come to resemble one another through adaptation to the same kind of life, is termed convergence. This too seems to be quite common among mammals. There is a tendency to duplication—indeed multiplication—of orders performing the same function. Perhaps the most remarkable instance is found among the mammals that have specialized in large-scale predation on termites and ants in the Tropics. This ecological niche is filled in South America by the ant bear *Myrmecophaga* and its related forms, all belonging to the order Edentata. In Asia and Africa the same role is played by mammals of the order Pholidota: the pangolins, or scaly anteaters. In Africa a third order has established itself in this business: the Tubulidentata, or aardvarks. Finally, in Australia there is the spiny anteater, which is in the order Monotremata. Thus we have members of four different orders living the same kind of life.

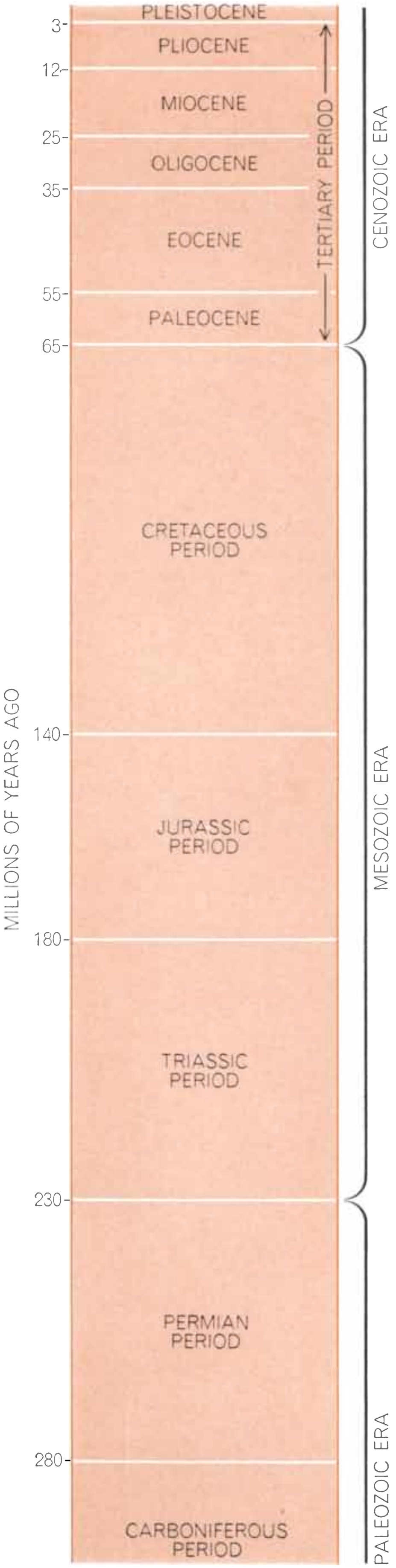
One can cite many other examples. There are, for instance, several living and extinct orders of hoofed herbivores. There are two living orders (the Rodentia, or rodents, and the Lagomorpha, or rabbits and hares) whose chisel-like incisor teeth are specialized for gnawing. Some extinct orders specialized in the same way, and an early primate, an ice-age ungulate and a living marsupial have also intruded into the "rodent niche" [*see top illustration on page 59*]. This kind of duplication, or near-duplication, is an essential ingredient in the richness of the mammalian life that unfolded during the Cenozoic era, or the age of mammals. Of the 30 or so orders of land-dwelling mammals that appeared during this period almost two-thirds are still extant.

The Reptiles of the Cretaceous

The 65 million years of the Cenozoic are divided into two periods: the long Tertiary and the brief Quaternary, which includes the present [*see illustration on page 56*]. The 200-million-year age of reptiles embraces the three periods of the Mesozoic era (the Triassic, the Jurassic and the Cretaceous) and the final period (the Permian) of the preceding era. It is instructive to compare the number



TWO SUPERCONTINENTS of the Mesozoic era were Laurasia in the north and Gondwanaland in the south. The 12 major types of reptiles, represented by typical species, are those whose fossil remains are found in Cretaceous formations. Most of the orders inhabited both supercontinents; migrations were probably by way of a land bridge in the west, where the Tethys Sea was narrowest.



SIX PERIODS of earth history were occupied by the age of reptiles and the age of mammals. The reptiles' rise began 280 million years ago, in the final period of the Paleozoic era. Mammals replaced reptiles as dominant land animals 65 million years ago.

of reptilian orders that flourished during some Mesozoic interval about as long as the Cenozoic era with the number of mammalian orders in the Cenozoic. The Cretaceous period is a good candidate. Some 75 million years in duration, it is only slightly longer than the age of mammals. Moreover, the Cretaceous was the culmination of reptilian life and its fossil record on most continents is good. In the Cretaceous the following orders of land reptiles were extant:

Order Crocodilia: crocodiles, alligators and the like. Their ecological role was amphibious predation; their size, medium to large.

Order Saurischia: saurischian dinosaurs. These were of two basic types: bipedal upland predators (Theropoda) and very large amphibious herbivores (Sauropoda).

Order Ornithischia: ornithischian dinosaurs. Here there were three basic types: bipedal herbivores (Ornithopoda), heavily armored quadrupedal herbivores (Stegosauria and Ankylosauria) and horned herbivores (Ceratopsia).

Order Pterosauria: flying reptiles.

Order Chelonia: turtles and tortoises.

Order Squamata: The two basic types were lizards (Lacertilia) and snakes (Serpentes). Both had the same principal ecological role: small to medium-sized predator.

Order Choristodera (or suborder in the order Eosuchia): champsosaurs. These were amphibious predators.

One or two other reptilian orders may be represented by rare forms. Even if we include them, we get only eight or nine orders of land reptiles in Cretaceous times. One could maintain that an order of reptiles ranks somewhat higher than an order of mammals; some reptilian orders include two or even three basic adaptive types. Even if these types are kept separate, however, the total rises only to 12 or 13. Furthermore, there seems to be only one clear-cut case of ecological duplication: both the crocodilians and the champsosaurs are sizable amphibious predators. (The turtles cannot be considered duplicates of the armored dinosaurs. For one thing, they were very much smaller.) A total of somewhere between seven and 13 orders over a period of 75 million years seems a sluggish record compared with the mammalian achievement of perhaps 30 orders in 65 million years. What light can paleogeography shed on this matter?

The Mesozoic Continents

The two supercontinents of the age of reptiles have been named Laurasia (after Laurentian and Eurasia) and Gondwanaland (after a characteristic geological formation, the Gondwana). Between them lay the Tethys Sea (named for the wife of Oceanus in Greek myth, who was mother of the seas). Laurasia, the northern supercontinent, consisted of what would later be North America, Greenland and Eurasia north of the Alps and the Himalayas. Gondwanaland, the southern one, consisted of the future South America, Africa, India, Australia and Antarctica. The supercontinents may have begun to split up as early as the Triassic period, but the rifts between them did not become effective barriers to the movement of land animals until well into the Cretaceous, when the age of reptiles was nearing its end.

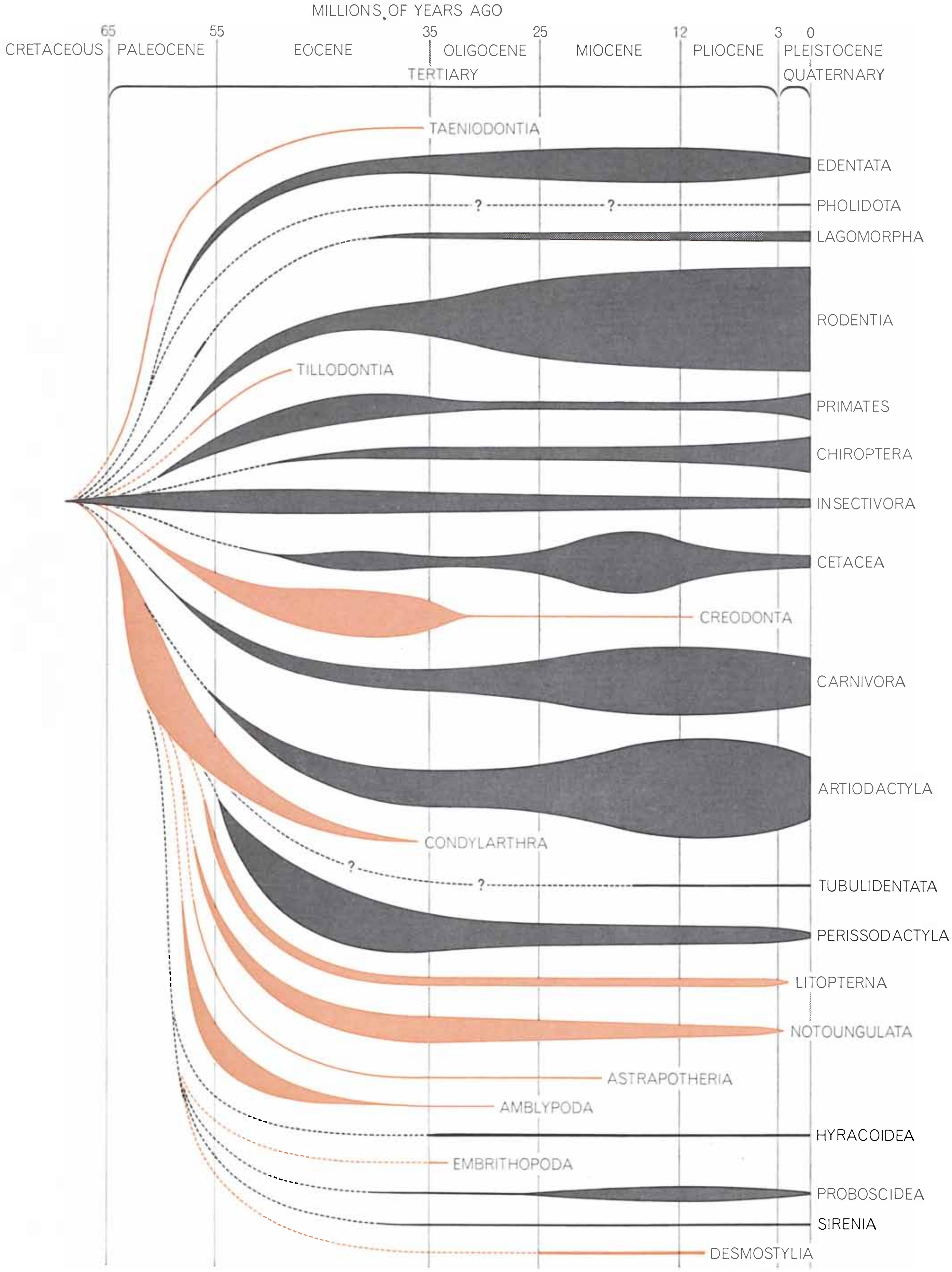
When the mammals began to diversify in the late Cretaceous and early Tertiary, the separation of the continents appears to have been at an extreme. The old ties were sundered and no new ones had formed. The land areas were further fragmented by a high sea level; the waters flooded the continental margins and formed great inland seas, some of which completely partitioned the continents. For example, South America was cut in two by water in the region that later became the Amazon basin, and Eurasia was split by the joining of the Tethys Sea and the Arctic Ocean. In these circumstances each chip of former supercontinent became the nucleus for an adaptive radiation of its own, each fostering a local version of a balanced fauna. There were at least eight such nuclei at the beginning of the age of mammals. Obviously such a situation is quite different from the one in the age of reptiles, when there were only two separate land masses.

Where the Reptiles Originated

The fossil record contains certain clues to some of the reptilian orders' probable areas of origin. The immense distance in time and the utterly different geography, however, make definite inferences hazardous. Let us see what can be said about the orders of Cretaceous reptiles (most of which, of course, arose long before the Cretaceous):

Crocodilia. The earliest fossil crocodilians appear in Middle Triassic formations in a Gondwanaland continent (South America). The first crocodilians in Laurasia are found in Upper Triassic formations. Thus a Gondwanaland origin is suggested.

Saurischia. The first of these dinosaurs appear on both supercontinents in the Middle Triassic, but they are more



ADAPTIVE RADIATION of the mammals has been traced from its starting point late in the Mesozoic era by Alfred S. Romer of Harvard University. Records for 25 extinct and extant orders of placental mammals are shown here. The lines increase and decrease in width in proportion to the abundance of each order. Extinct orders are shown in color; broken lines mean that no fossil record exists during the indicated interval and question marks imply doubt about the suggested ancestral relation between some orders.

varied in the south. A Gondwanaland origin is very tentatively suggested.

Ornithischia. These dinosaurs appear in the Upper Triassic of South Africa (Gondwanaland) and invade Laurasia somewhat later. A Gondwanaland origin is indicated.

Pterosauria. The oldest fossils of flying reptiles come from the early Jurassic of Europe. They represent highly specialized forms, however, and their antecedents are unknown. No conclusion seems possible.

Chelonia. Turtles are found in Triassic formations in Laurasia. None are found in Gondwanaland before Cretaceous times. This suggests a Laurasian origin. On the other hand, a possible forerunner of turtles appears in the Permian of South Africa. If the Permian form was in fact ancestral, a Gondwanaland origin would be indicated. In any case, the order's main center of evolution certainly lay in the northern supercontinent.

Squamata. Early lizards are found in the late Triassic of the north, which may suggest a Laurasian origin. Unfortunately the lizards in question are aberrant gliding animals. They must have had a long history, of which we know nothing at present.

Choristodera. The crocodile-like champsosaurs are found only in North America and Europe, and so presumably originated in Laurasia.

The indications are, then, that three orders of reptiles—the crocodilians and the two orders of dinosaurs—may have originated in Gondwanaland. Three others—the turtles, the lizards and snakes and the champsosaurs—may have originated in Laurasia. The total number of basic adaptive types in the Gondwanaland group is six; the Laurasia group has four. The Gondwanaland radiation may well have been slightly richer than the Laurasian because it seems that the southern supercontinent was somewhat larger and had a slightly more varied climate. Laurasian climates seem to have been tropical to temperate. Southern parts of Gondwanaland were heavily glaciated late in the era preceding the Mesozoic, and its northern shores (facing the Tethys Sea) had a fully tropical climate.

Although some groups of reptiles, such as the champsosaurs, were confined to one or another of the supercontinents, most of the reptilian orders sooner or later spread into both of them. This means that there must have been ways for land animals to cross the Tethys Sea. The Tethys was narrow in the west and wide to the east. Presumably whatever land connection there was—a true land bridge or island stepping-stones—was located in the western part of the sea. In any case, migration along such routes meant that there was little local differentiation among the reptiles of the Mesozoic era. It was over an essentially uniform reptilian world that the sun finally set at the end of the age of reptiles.

Early Mammals of Laurasia

The conditions of mammalian evolution were radically different. In early and middle Cretaceous times the connections between continents were evidently close enough for primitive mammals to spread into all corners of the habitable world. As the continents drifted farther apart, however, populations of these primitive forms were gradually isolated from one another. This was particularly the case, as we shall see, with the mammals that inhabited the daughter continents of Gondwanaland. Among the Laurasian continents North America was drifting away from Europe, but at the beginning of the age of mammals the distance was not great and there is good evidence that some land connection remained well into the early Tertiary. North American and European mammals were practically identical as late as early Eocene times. Furthermore, throughout the Cenozoic era there was a connection between Alaska and Siberia, at least intermittently, across the Bering Strait. On the other hand, the inland sea extending from the Tethys to the Arctic Ocean formed a complete barrier to direct migration between Europe and Asia in the early Tertiary. Migrations could take place only by way of North America.

In this way the three daughter continents of ancient Laurasia formed three semi-isolated nuclear areas. Many orders of mammals arose in these Laurasian nuclei, among them seven orders that are now extinct but that covered a wide spectrum of specialized types, including primitive hoofed herbivores, carnivores, insectivores and gnawers. The orders of mammals that seem to have arisen in the northern daughter continents and that are extant today are:

Insectivora: moles, hedgehogs, shrews and the like. The earliest fossil insectivores are found in the late Cretaceous of North America and Asia.

Chiroptera: bats. The earliest-known bat comes from the early Eocene of North America. At a slightly later date bats were also common in Europe.

Primates: prosimians (for example, tarsiers and lemurs), monkeys, apes, man. Early primates have recently been found in the late Cretaceous of North America. In the early Tertiary they are common in Europe as well.

Carnivora: cats, dogs, bears, weasels and the like. The first true carnivores appear in the Paleocene of North America.

Perissodactyla: horses, tapirs and other odd-toed ungulates. The earliest forms appear at the beginning of the Eocene in the Northern Hemisphere.

Artiodactyla: cattle, deer, pigs and other even-toed ungulates. Like the odd-toed ungulates, they appear in the early Eocene of the Northern Hemisphere.

Rodentia: rats, mice, squirrels, beavers and the like. The first rodents appear in the Paleocene of North America.

Lagomorpha: hares and rabbits. This order makes its first appearance in the Eocene of the Northern Hemisphere.

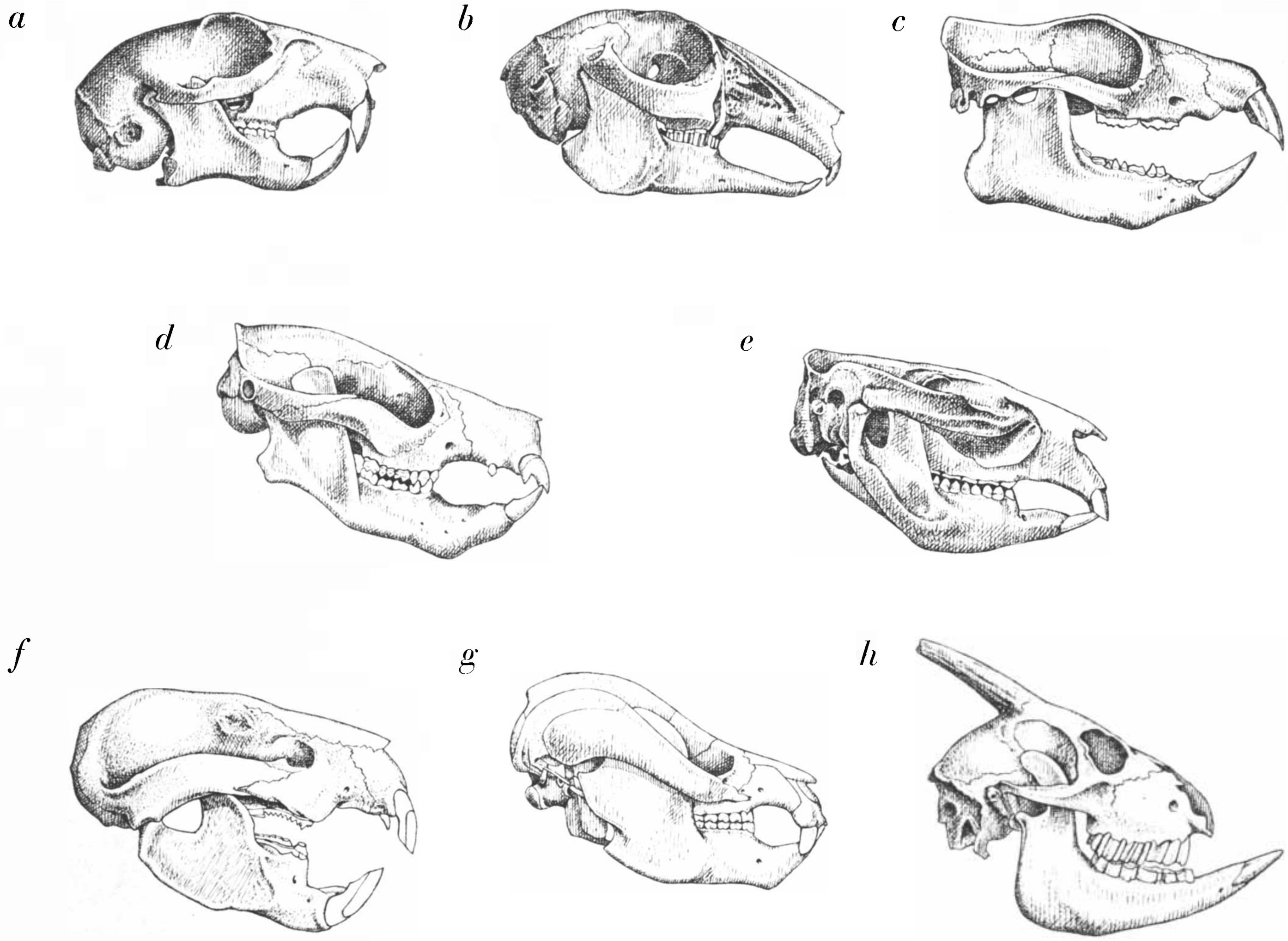
Pholidota: pangolins. The earliest come from Europe in the middle Tertiary.

The fact that a given order of mammals is found in older fossil deposits in North America than in Europe or Asia does not necessarily mean that the order arose in the New World. It may simply reflect the fact that we know much more about the early mammals of North America than we do about those of Eurasia. All we can really say is that a total of 16 extant or extinct orders of mammals probably arose in the Northern Hemisphere.

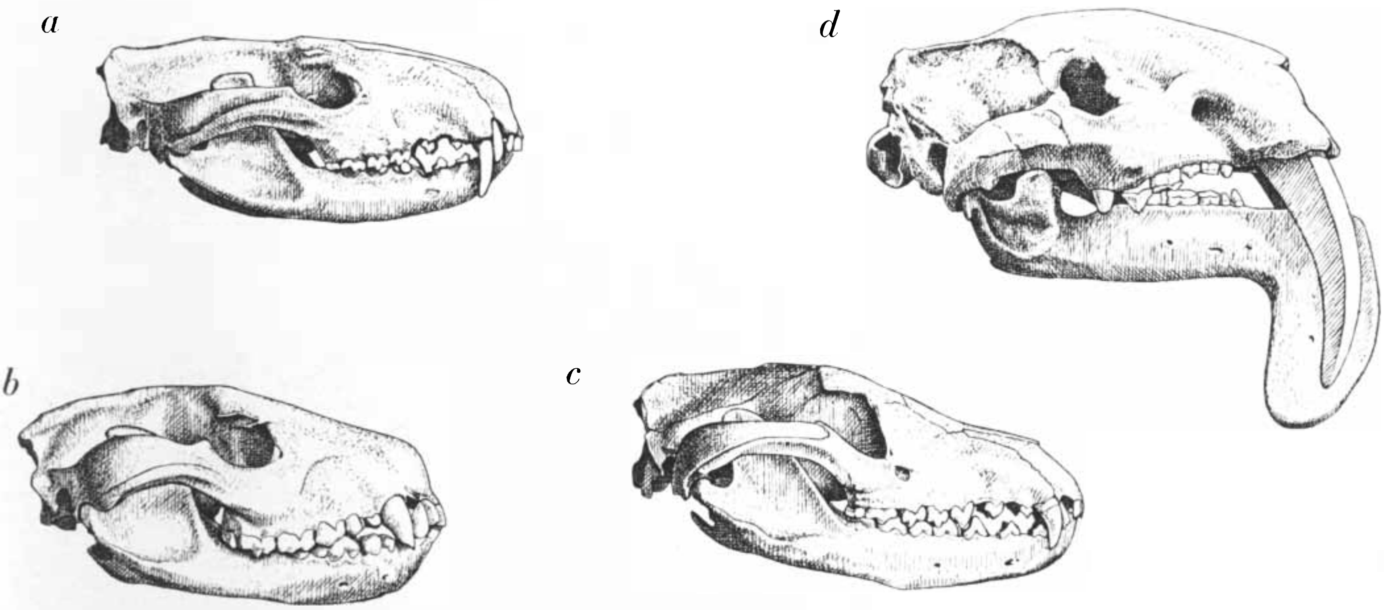
Early Mammals of South America

The fragmentation of Gondwanaland seems to have started earlier than that of Laurasia. The rifting certainly had a much more radical effect. Looking at South America first, we note that at the beginning of the Tertiary this continent was tenuously connected to North America but that for the rest of the period it was completely isolated. The evidence for the tenuous early linkage is the presence in the early Tertiary beds of North America of mammalian fossils representing two predominantly South American orders: the Edentata (the order that includes today's ant bears, sloths and armadillos) and the Notoungulata (an order of extinct hoofed herbivores).

Four other orders of mammals are exclusively South American: the Paucituberculata (opossum rats and other small South American marsupials), the Pyrotheria (extinct elephant-like animals), the Litopterna (extinct hoofed herbivores, including some forms resembling



CHISEL-LIKE INCISORS, specialized for gnawing, appear in animals belonging to several extinct and extant orders in addition to the rodents, represented by a squirrel (*a*), and the lagomorphs, represented by a hare (*b*), which are today's main specialists in this ecological role. Representatives of other orders with chisel-like incisor teeth are an early tillodont, *Trogosus* (*c*), an early primate, *Plesiadapis* (*d*), a living marsupial, the wombat (*e*), one of the extinct multituberculate mammals, *Taeniolabis* (*f*), a mammal-like reptile of the Triassic, *Bienotherium* (*g*), and a Pleistocene cave goat, *Myotragus* (*h*), whose incisor teeth are in the lower jaw only.



CARNIVOROUS MARSUPIALS, living and extinct, fill an ecological niche more commonly occupied by the placental carnivores today. Illustrated are the skulls of two living forms, the Australian "cat," *Dasyurus* (*a*), and the Tasmanian devil, *Sarcophilus* (*b*). The Tasmanian "wolf," *Thylacinus* (*c*), has not been seen for many years and may be extinct. A tiger-sized predator of South America, *Thylacosmilus* (*d*) became extinct in Pliocene times, long before the placental sabertooth of the Pleistocene, *Smilodon*, appeared.

horses and camels) and the Astrapotheria (extinct large hoofed herbivores of very peculiar appearance). Thus a total of six orders, extinct or extant, probably originated in South America. Still another order, perhaps of even more ancient origin, is the Marsupicarnivora. The order is so widely distributed, with species found in South America, North America, Europe and Australia, that its place of origin is quite uncertain. It includes, in addition to the extinct marsupial carnivores of South America, the opossums of the New World and the native "cats" and "wolves" of the Australian area.

The most important barrier isolating South America from North America in the Tertiary period was the Bolívar Trench. This arm of the sea cut across the extreme northwest corner of the continent. In the late Tertiary the bottom of the Bolívar Trench was lifted above sea level and became a mountainous land area. A similar arm of sea, to which I have already referred, extended across the continent in the region that is now the Amazon basin. This further enhanced the isolation of the southern part of South America.

Early Mammals of Africa

Africa's role as a center of adaptive radiation is problematical because practically nothing is known of its native mammals before the end of the Eocene. We do know, however, that much of the continent was flooded by marginal seas, and that in the early Tertiary, Africa was cut up into two or three large islands. Still, there must have been a land route to Eurasia even in the Eocene; some of the African mammals of the following epoch (the Oligocene) are clearly immigrants from the north or northeast. Nonetheless, the majority of African mammals are of local origin. They include the following orders:

Proboscidea: the mastodons and elephants.

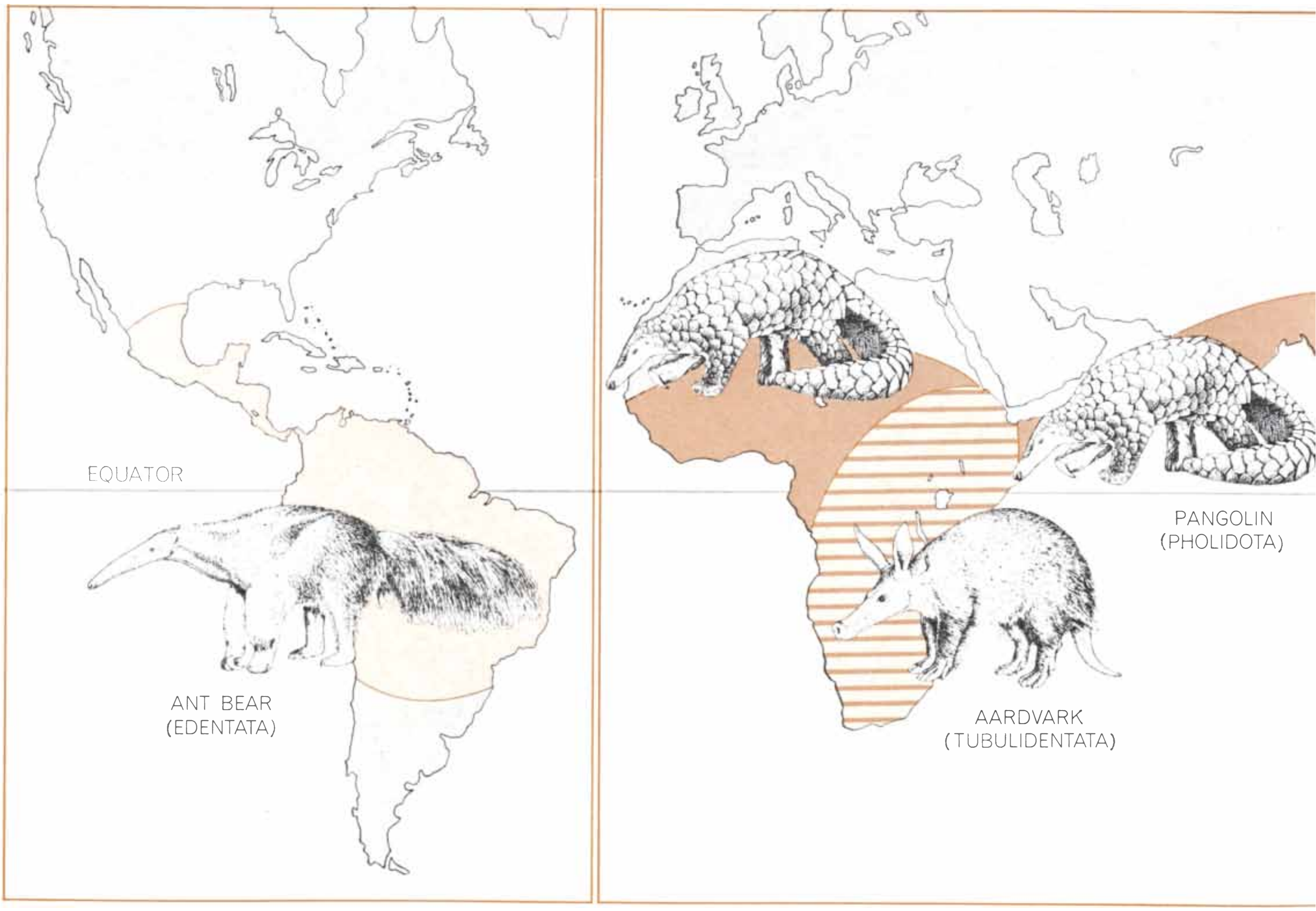
Hyracoidea: the conies and their extinct relatives.

Embrithopoda: an extinct order of very large mammals.

Tubulidentata: the aardvarks.

In addition the order Sirenia, consisting of the aquatic dugongs and manatees, is evidently related to the Proboscidea and hence presumably also originated in Africa. The same may be true of another order of aquatic mammals, the extinct Desmostylia, which also seems to be related to the elephants. The one snag in this interpretation is that desmostylian fossils are found only in the North Pacific, which seems rather a long way from Africa. Nonetheless, once they were waterborne, early desmostylians might have crossed the Atlantic, which was then only a narrow sea, navigated the Bolívar Trench and, rather like Cortes (but stouter), found themselves in the Pacific.

Thus there are certainly four, and possibly six, mammalian orders for which an African origin can be postulated. Here it should be noted that Africa had an impressive array of primates in the Oligocene. This suggests that the order Primates had a comparatively long history in Africa before that time. Even



FOUR ANT-EATING MAMMALS have become adapted to the same kind of life although each is a member of a different mammalian order. Their similar appearance is an example of the process of convergence. The ant bears of the New World Tropics are in the order Edentata. The aardvark of Africa is the only species in the order Tubulidentata. Pangolins, found both in Asia and in Af-

though the order as such does not have its roots in Africa, it is possible that the higher primates—the Old World monkeys, the apes and the ancestors of man—may have originated there. Most of the fossil primates found in the Oligocene formations of Africa are primitive apes or monkeys, but there is at least one form (*Propliopithecus*) whose dentition looks like a miniature blueprint of a set of human teeth.

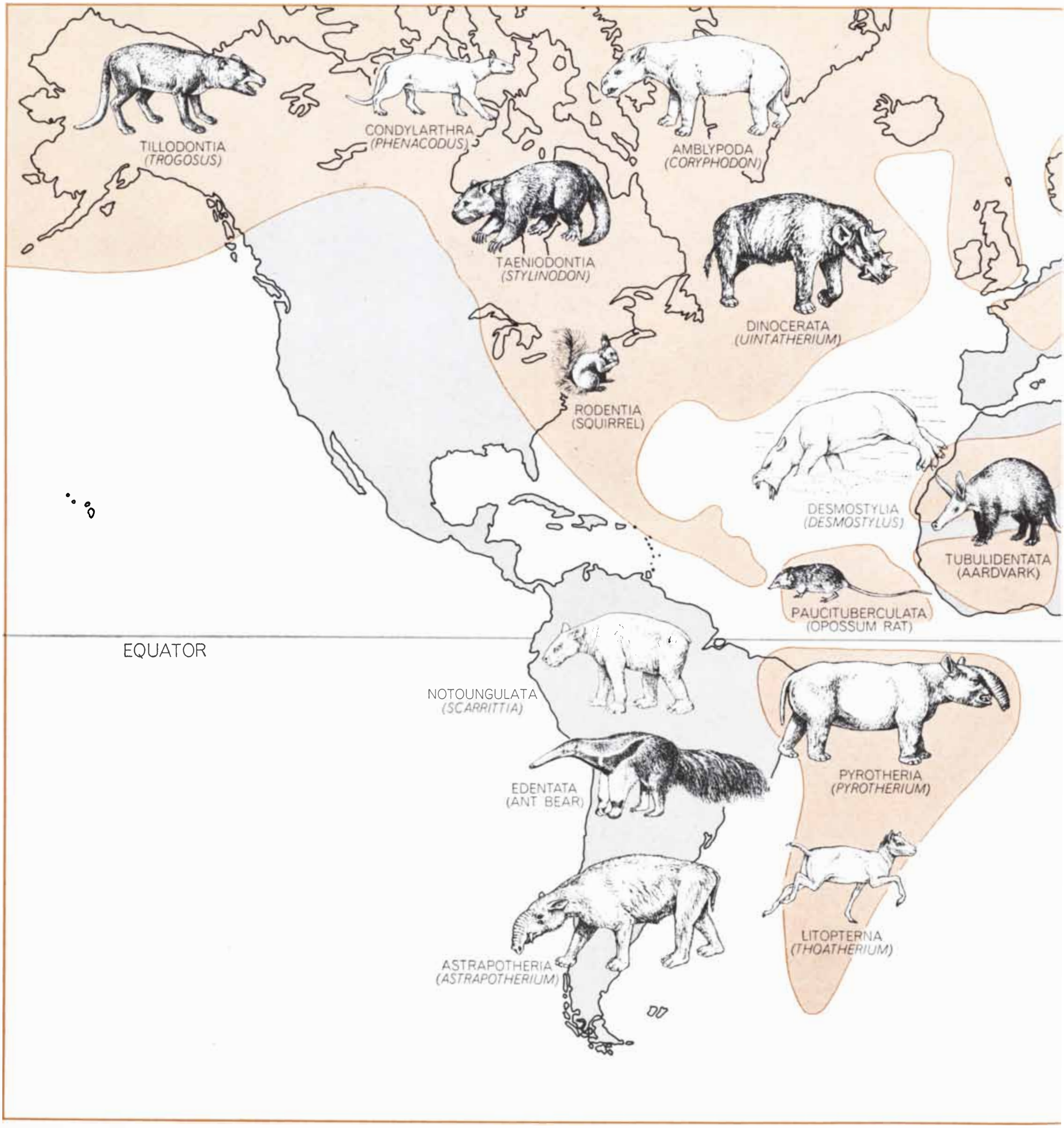
The Rest of Gondwanaland

We know little or nothing of the zoogeographic roles played by India and Antarctica in the early Tertiary. Mammalian fossils from the early Tertiary are also absent in Australia. It may be assumed, however, that the orders of mammals now limited to Australia probably originated there. These include two orders of marsupials: the Peramelina, comprised of several bandicoot genera, and the Diprotodonta, in which are found the kangaroos, wombats, phalangers and a number of extinct forms. In addition

rica, are members of the order Pholidota. The spiny anteater of Australia, a very primitive mammal, is in the order Monotremata.

the order Monotremata, a very primitive group of mammals that includes the spiny anteater and the platypus, is likely to be of Australian origin. This gives us a total of three orders probably founded in Australia.

Summing up, we find that the three Laurasian continents produced a total of 16 orders of mammals, an average of five or six orders per continent. As for Gondwanaland, South America produced six orders, Africa four to six and Australia three. The fact that Australia is a small continent probably accounts for the lower number of orders founded there. Otherwise the distribution—the average of five or six orders per subdivision—is remarkably uniform for both the Laurasian and Gondwanaland supercontinents. The mammalian record should be compared with the data on Cretaceous reptiles, which show that the two supercontinents produced a total of 12 or 13 orders (or adaptively distinct suborders). A regularity is suggested, as if a single nucleus of radiation would tend in a given time to produce and support a



CONTINENTAL DRIFT affected the evolution of the mammals by fragmenting the two supercontinents early in the Cenozoic era. In the north, Europe and Asia, although separated by a sea, remained connected with North America during part of the era. The given amount of basic zoological variation.

As the Tertiary period continued new land connections were gradually formed, replacing those sundered when the old supercontinents broke up. Africa made its landfall with Eurasia in the Oligocene and Miocene epochs. Laurasian orders of mammals spread into Africa and crowded out some of the local forms, but at the same time some African mammals (notably the mastodons and elephants) went forth to conquer almost the entire world. In the Western Hemisphere the draining and uplifting of the Bolívar Trench was followed by intense intermigration and competition among the mammals of the two Americas. In the process much of the typical South American mammal population was exterminated, but a few forms pressed successfully into North America to become part of the continent's spectacular ice-age wildlife.

India, a fragment of Gondwanaland that finally became part of Asia, must have made a contribution to the land

free migration that resulted prevents certainty regarding the place of origin of many orders of mammals that evolved in the north. The far wider rifting of Gondwanaland allowed the evolution of unique groups of mammals in South America, Africa and Australia.

fauna of that continent but just what it was cannot be said at present. Of all the drifting Noah's arks of mammalian evolution only two—Antarctica and Australia—persist in isolation to this day. The unknown mammals of Antarctica have long been extinct, killed by the ice that engulfed their world. Australia is therefore the only island continent that still retains much of its pristine mammalian fauna.

If the fragmentation of the continents at the beginning of the age of mammals promoted variety, the amalgamation in the latter half of the age of mammals has promoted efficiency by means of a large-scale test of the survival of the fittest. There is a concomitant loss of variety; 13 orders of land mammals have become extinct in the course of the Cenozoic. Most of the extinct orders are island-continent productions, which suggests that a system of semi-isolated provinces, such as the daughter continents of Laurasia, tends to produce a more efficient brood than the completely isolated nuclei of the Southern Hemisphere. Not all the Gondwanaland orders were inferior, however; the edentates were moderately successful and the proboscidians spectacularly so.

As far as land mammals are concerned, the world's major zoogeographic provinces are at present four in number: the Holarctic-Indian, which consists of North America and Eurasia and also northern Africa; the Neotropical, made up of Central America and South America; the Ethiopian, consisting of Africa south of the Sahara, and the Australian. This represents a reduction from seven provinces with about 30 orders of mammals to four provinces with about 18 orders. The reduction in variety is proportional to the reduction in the number of provinces.

In conclusion it is interesting to note that we ourselves, as a subgroup within the order Primates, probably owe our origin to a radiation within one of Gondwanaland's island continents. I have noted that an Oligocene primate of Africa may have been close to the line of human evolution. By Miocene times there were definite hominids in Africa, identified by various authorities as members of the genus *Ramapithecus* or the genus *Kenyapithecus*. Apparently these early hominids spread into Asia and Europe toward the end of the Miocene. The cycle of continental fragmentation and amalgamation thus seems to have played an important part in the origin of man as well as of the other land mammals.

RECENT FINDINGS

RESEARCH LABORATORIES

Ford

Heat Torques and Circulation in Superfluid Helium.

Recent experiments at Ford Motor Company have demonstrated that a heated object immersed in flowing superfluid helium will experience a force or torque even in pure potential flow. The existence of such a force or torque is a consequence of the "two-fluid" thermodynamics of helium II, and results from a transfer of superfluid momentum to the object as superfluid is converted to normal fluid at the heated surface.

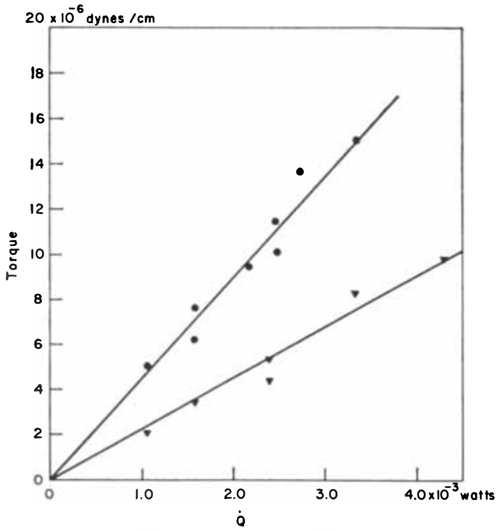
A heated object acts as a heat source to the liquid helium and effectively "creates" normal fluid at the interface. Since there is no mass flow across the solid boundary, the creation of normal fluid is made at the expense of the superfluid component. In the experiments to be described, the heated object is a cylinder and is immersed in liquid helium at temperatures below the lambda point (the point where superfluidity begins). Furthermore, the superfluid has been given a *persistent* net circulation about the object. The conversion of superfluid to normal fluid thus causes a transfer of angular momentum or torque to the heated object. The torque is proportional to the circulation of the superfluid and also to the strength of the heat source, Q.

It is generally expected that the persistent circulation in a superfluid is quantized, i.e. it exists in integral multiples of a basic quantity (h/M), where h is Planck's constant and M is the mass of the helium atom. If the circulation is quantized, then a quantization of the heat torque must also be expected. Although a number of attempts have been made to investigate the circulation in free vortex lines, multiple quantization of circulation in helium II has not yet been clearly demonstrated.

In the Ford experiments, a small heated glass cylinder was suspended by a sensitive quartz torsion fiber in rotating superfluid helium. The torque was measured optically by the deflection of a beam of light. The glass cylinder was painted black and heat was supplied by illuminating it with a calibrated projection lamp. The helium was brought into rotation above the lambda point by stirring it for an hour or more. The bath was then slowly pumped until its temperature fell below the lambda point. The stirring was stopped, and normal-fluid motion allowed to decay for at least an hour before measurements were made. The results show that the torque is indeed proportional to the power of the heat source and that the proportionality constant depends upon the amount of trapped persistent circulation.

Though the basic torque sensitivity of the present system is sufficient to permit the detection of a single quantum of circulation, the average noise level would not permit a clear resolution of the individual quanta.

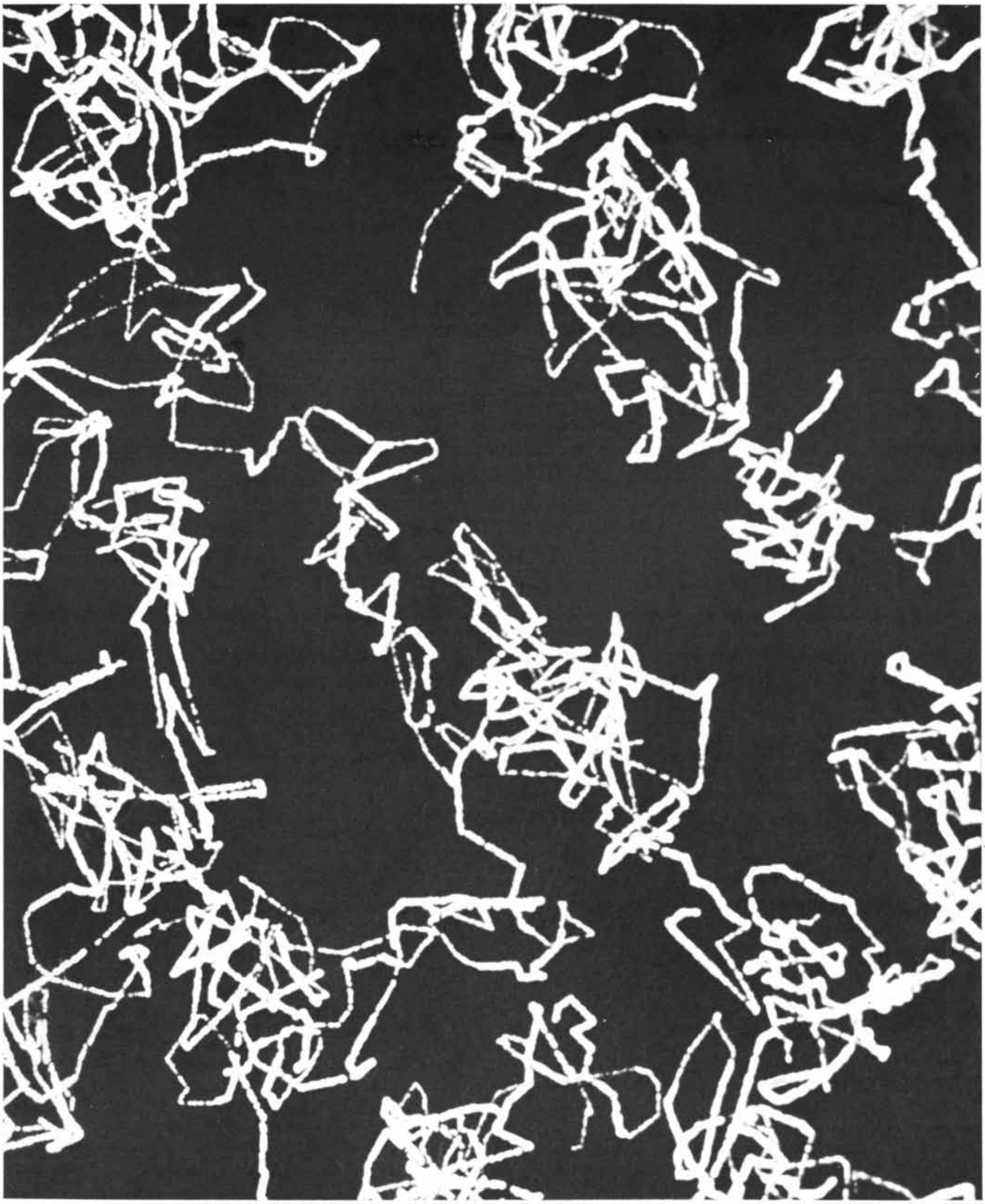
This represents one aspect of Ford Motor Company's continuing study of the fundamental properties of materials to help design better materials and devices.



Examples of the torque on a heated cylinder (2mm diam, 1 cm long) immersed in rotating superfluid helium as a function of the heater power. The agreement with the linear dependence predicted by the analysis is within the limits of experimental error.

PROBING DEEPER FOR BETTER IDEAS

Ford



COMPUTER "SNAPSHOT" of the paths of a number of molecules in a hypothetical fluid symbolizes one of the two main themes of this article: the mathematical analysis of the Brownian motion of a small particle buffeted by a turbulent "sea" of such molecules. In order to make the photograph, each imaginary molecule was represented by a bright dot on the face of a cathode ray tube. By focusing a camera on the screen and leaving its shutter open, it was possible to record the trajectories of the moving dots on photographic film. The molecular-dynamical calculations of the paths were fed into the display system from a computer, in which the location of each molecule was represented by a set of three numbers, specifying the three-dimensional coordinates of its center. The molecules were "shaken" one by one, using a mathematical technique called the Monte Carlo method; according to this technique, a particle chosen at random was displaced by an amount that was determined by picking one of a series of random numbers generated by the computer. Various boundary conditions can then be simulated by making some moves "legal," and others "illegal." The experiment was performed by B. J. Alder and Thomas E. Wainwright at the Lawrence Radiation Laboratory of the University of California at Livermore.

Brownian Motion and Potential Theory

The discovery that these two apparently unrelated branches of physics are in some sense mathematically equivalent has led to a new subject known as probabilistic potential theory

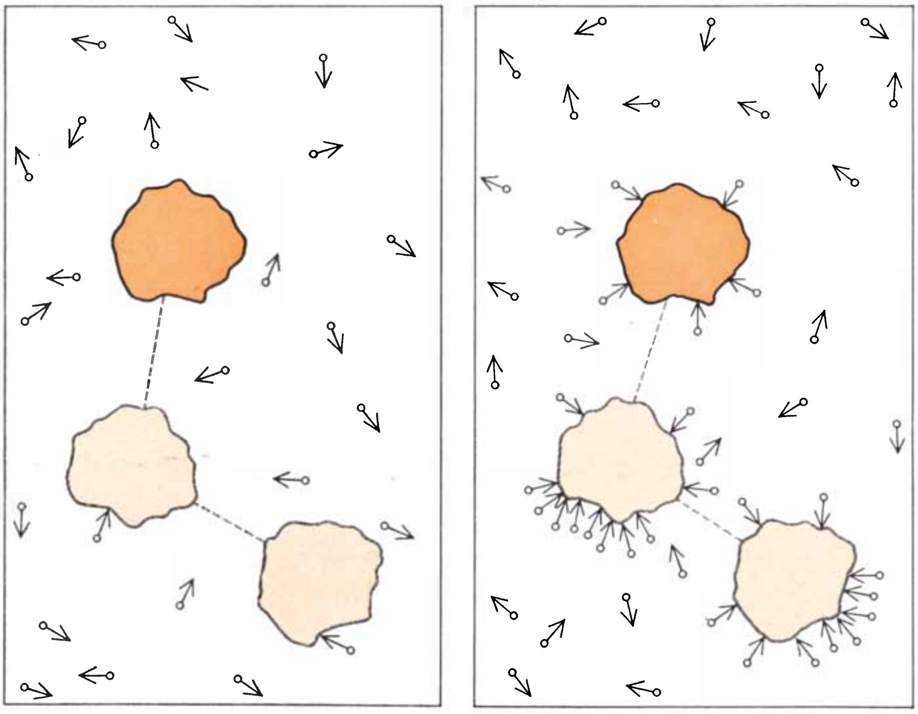
by Reuben Hersh and Richard J. Griego

One of the most exciting events of the past decade or so in the field of mathematical analysis has been the appearance of a new subject called probabilistic potential theory. In essence this subject represents the marriage of two major branches of theoretical physics: the probabilistic theory of random processes, which studies such phenomena as the Brownian motion of a small particle buffeted by a turbulent "sea" of molecules, and potential theory, which studies the equilibrium states of a homogeneous medium, for example the distribution of heat in a solid body at thermal equilibrium. The development of probabilistic potential theory is rooted in the discovery that these two apparently unrelated branches of physics are in some sense mathematically equivalent; in other words, the mathematics of one can be translated meaningfully into the mathematics of the other. From this remarkable circumstance have flowed many unexpected insights into both subjects. Before pointing to the accomplishments of probabilistic potential theory, however, it will be helpful to review briefly the two separate lines of inquiry that the new theory connects, namely the theory of Brownian motion and potential theory.

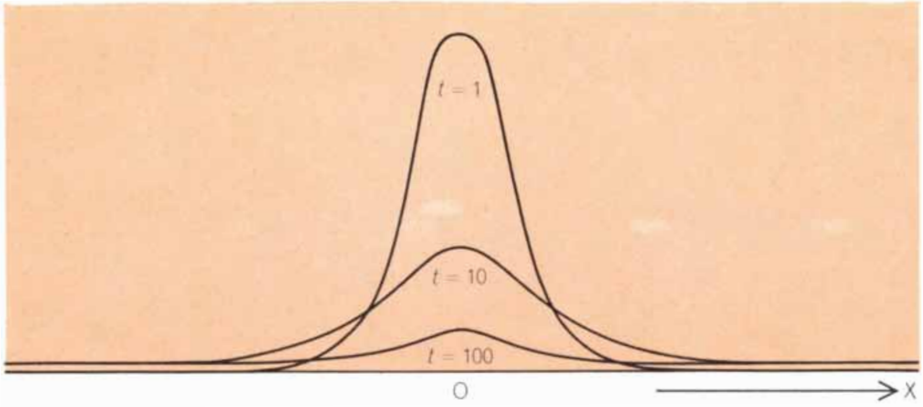
The phenomenon of Brownian motion was described in 1827 by the Scottish botanist Robert Brown in the course of an investigation of the fertilization process in a newly discovered species of flower. Brown observed under the microscope that when the pollen grains from the flower were suspended in water, they performed a "rapid oscillatory motion." At first he believed this motion was peculiar to the male sexual cells of plants. He soon found, however, that particles of other organic substances, bits of petrified wood and even chips of glass or granite exhibited the same motion.

It was not until the 1860's that the problem of the cause of Brownian motion really began to bother theoretical physicists. Early attempts to explain the phenomenon in terms of fluid currents in the host medium had had to be rejected because observation showed that the motions of two neighboring particles seemed quite uncorrelated with each other. Other properties of Brownian motion were equally intriguing; for example, the higher the temperature the faster the motion, the smaller the particle the faster the motion, the more viscous the medium the slower the motion. In addition, the given particle appeared equally likely to go in any direction, and the past motion of a particle seemed to have no bearing on its future motion. Last but not least, the motion never stops.

The relation between temperature and speed seemed to suggest a molecular origin for the Brownian motion, since according to the kinetic theory of heat the temperature of any substance is proportional to the average of the square of the speed of its molecules. Thus a higher



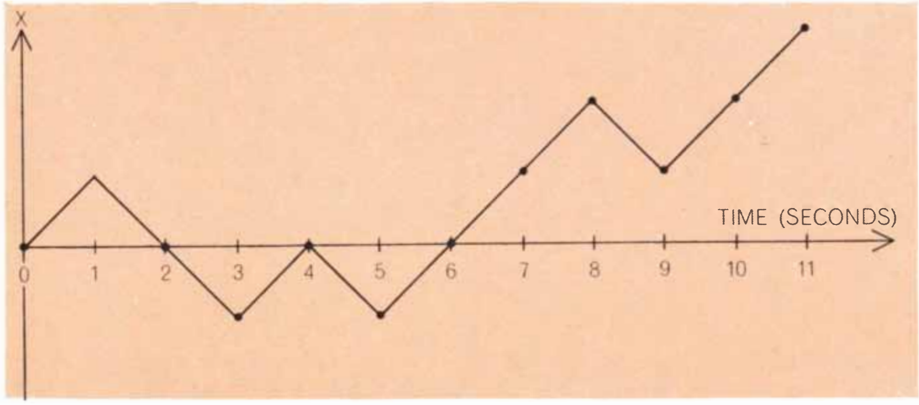
MOLECULAR ORIGIN of Brownian motion was originally suggested by the observation that Brownian particles move faster at higher temperatures; according to the kinetic theory of heat higher temperature simply means more rapid molecular motion. Even the smallest particle observable through a microscope, however, is far too large to show observable motion as a result of a series of kicks from one molecule at a time (*left*). Instead it has been shown that the random motion of a Brownian particle is caused by the random discrepancy between the molecular pressures on different surfaces of the particle (*right*).



BELL-SHAPED CURVES result when one plots the successive positions of a Brownian particle on a graph. The horizontal axis of the graph measures the distance traveled in a given direction, assuming that at a time $t = 0$ the particle is at point $x = 0$. The vertical axis measures the probability that the particle will be at any given point at time $t = 1$, 10 and 100 seconds respectively for the three curves. Such symmetrical curves are called normal or Gaussian densities and represent the quantity $e^{-x^2/4Dt}/2\sqrt{\pi Dt}$, where e is the base (equal to 2.7182818) of a natural system of logarithms and D is the diffusivity of the medium.

temperature means more rapid average molecular motion, and observation had already shown that higher temperature means more rapid Brownian motion. On the other hand, any simpleminded notion that the jerky movement of Brown's particles was due merely to kicks from single molecules was out of the question. Even the smallest particle observable in a microscope is far too large to show observable motion as a result of a series of kicks from one molecule at a time.

A major advance came in the form of a theoretical analysis in statistical mechanics completed by Albert Einstein in 1905, the year he published his first paper on relativity. At the time many leading physical scientists, including Wilhelm Ostwald and Ernst Mach, regarded molecules and atoms not as real entities but as intellectual figments that might be useful for explaining certain natural phenomena. (This view, which is hardly remembered today, is somewhat reminiscent of the way many physicists later regarded the wave-particle duality of quantum mechanics.) Einstein, reasoning on the basis of the kinetic theory of heat, determined that if an observable particle were in the midst of a molecular bombardment, it would then describe a random motion caused by the difference in the number of blows it might receive at any instant on, say, its left and right surfaces. The smaller the particle, the more likely it would be that this difference would be sufficient to cause a detectable push. The less viscous the fluid, the faster and farther the particle would go as a result of each push. Once each tiny step was stopped by the fluid's viscosity, any future motion would depend only on the random discrepancy between molecular pressures on the particle's left and right surfaces, its front and back surfaces or its top and bottom surfaces.



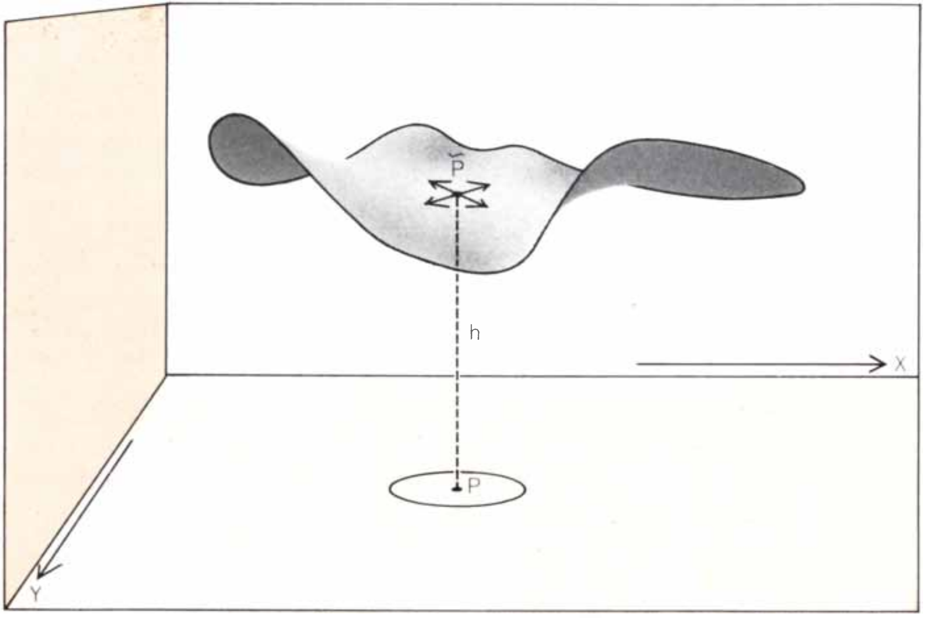
INTUITIVE NOTION of what is involved in Norbert Wiener's mathematical model of Brownian motion can be obtained from this illustration. The black zigzag line is the path traced out in a finite time period (in this case 11 seconds) by a one-dimensional motion that changes direction only at the instants $t = 1$, $t = 2$ and so on. If this process were continued for an hour, say, there would be only a finite number of paths ($2^{3,600}$, to be exact), and one could say that the particle chooses one path at random in the sense that each path has a probability of $1/2^{3,600}$. Such a process is called "random walk" or "drunkard's walk." Wiener solved the problem of going to the limit of infinitely small time increments.

Over a period of time such a hypothetical particle would tend to drift from its original position. Its exact position after a certain number of seconds is of course unpredictable, but it turns out that one can state in a general way where the particle is likely to be. If the experiment is performed many times and the successive positions of the particle are plotted on a graph, one obtains a bell-shaped curve such as the curves shown in the top illustration on this page. The horizontal axis in each case measures the distance traveled in any given direction, say left or right, assuming that at time $t = 0$ the particle is at the point $x = 0$. The vertical axis of the graph measures the probability that the particle will be at any given point at time $t = 1$, 10 and 100 seconds respectively for the three curves.

It is evident from such graphs that the most probable position is always the original position and that the farther away a position is, the less likely the particle is to be there at any given time. Moreover, the graphs are symmetric, since the movement of the particle is unbiased between left and right. As one might expect, the three curves show that the longer the particle drifts, the likelier it is to wander from its starting point. Bell-shaped curves such as these are called normal or Gaussian distributions, and they typically arise in situations where the measured quantity is the sum of a great many independent but essentially identical random variables, in this case the many little pushes that add up to the total motion.

It is remarkable and amusing that only after Einstein had completed his calculations did he learn that the phenomenon he was predicting was already well known! He wrote later: "My major aim in this was to find facts which would guarantee as much as possible the existence of atoms of definite finite size. In the midst of this I discovered that, according to atomistic theory, there would have to be a movement of suspended microscopic particles open to observation, without knowing that observations concerning the Brownian motion were already long familiar."

It was largely this work of Einstein's that finally put out of fashion the view that molecules and atoms might be fictitious. In 1926 Jean Perrin received the Nobel prize for an experimental application in 1909 of Einstein's results. By observing actual Brownian movement Perrin was able to measure the predicted displacements and was thereby able to compute the diffusivity parameter D. According to Einstein's formula, $D = 2RT/Nf$, where R is a universal constant, T is the absolute temperature, N is Avogadro's number (the number of molecules in a gram-molecular weight of a gas) and f is a viscosity coefficient. In this way Perrin succeeded in obtaining a value for Avogadro's number, one of the fundamental constants of nature: he found that N is approximately equal to 6×10^{23}.



ELASTIC MEMBRANE is stretched across a stiff, closed frame that is twisted into some fixed shape in space in this illustration of the role of harmonic functions in potential theory. The configuration of such a membrane is given by the height h of each point $\tilde{P}$ on the surface of the membrane. Directly below each $\tilde{P}$ on the membrane is a point P on the base plane, which has coordinates x, y. Besides being continuous, the function $h(x, y)$ has the following simple property: If P is a point in the x, y plane, and Γ is a small circle with its center at P, then the value of h at P (that is, the height of the membrane above P) equals the average of the values of h for all points on the circle Γ. This is called the mean-value property, and a continuous function h possessing this property is called a harmonic function. In this case the position of $\tilde{P}$ (the point on the membrane above P) is determined by the sum of the tension forces exerted on $\tilde{P}$ by the surrounding portion of the membrane (*arrows*). If the membrane is in equilibrium, these forces must cancel, so that the number of nearby elevations greater than that of $\tilde{P}$ must be matched by corresponding elevations lower than that of $\tilde{P}$, and the average must be just equal to the elevation of $\tilde{P}$, namely, the function h at P.

As far as physicists were concerned, Einstein's and Perrin's investigations left the problem of Brownian motion in reasonably good shape; more recent work has tried to refine and justify Einstein's calculations on the basis of the general Maxwell-Boltzmann equations of statistical mechanics. For mathematicians, however, the story does not really begin until 1920, when Norbert Wiener wrote his first paper on Brownian motion.

Whereas from a physical viewpoint Einstein's calculations and Perrin's experiments had explained Brownian motion quite adequately, from a mathematical viewpoint the subject was still tantalizingly confused. The heart of the difficulty was to make precise mathematical sense out of the notion of a particle moving "at random." Everyone knows what it means to pick between heads and tails at random; it means each alternative has a probability of 1/2 (if we assume that the toss is fair). The Brownian particle follows a path that is in some sense chosen at random from among all possible paths. The set of all possible paths, however, is a very large and complicated one, and it was one of Wiener's major achievements in mathematics to show in what sense one can speak about choosing from this set at random.

We shall not attempt here even to summarize Wiener's argument. Nonetheless, an intuitive notion of what is involved can be obtained by considering the path traced out in a finite time period (say an hour) by a one-dimensional motion, which changes direction only at the instants t = one second, t = two seconds and so on [*see bottom illustration on opposite page*]. In this case there are only a finite number of possible paths ($2^{3,600}$, to be exact), and one could say that the Brownian particle chooses one path at random in the sense that each path has a probability of $1/2^{3,600}$. Such a process, made up of discrete steps, is sometimes called "random walk" or "drunkard's walk." The difficulty is in going to the limit of infinitely small time increments.

Wiener showed how to do this in a mathematically legitimate way, thereby bringing the term Brownian motion into the language of mathematics. In the Wiener process, as one refers to Wiener's model of Brownian motion, the distances traveled are distributed according to a Gaussian curve, just as they are in Einstein's physical model of Brownian motion. Moreover, Wiener proved that almost certainly (with a probability of 1) the path is continuous but nowhere smooth. This also fits very nicely with physical intuition. A particle in Brownian motion surely cannot jump instantaneously from one point to another, so that the path should be continuous; erratic changes in direction seem to be taking place constantly, so that one might expect the path to consist entirely of sharp corners.

Wiener's work has been continued by a long line of successors; in a sense it is the fountainhead of most modern work in random processes. One of the most fruitful outcomes of this work has been its role in the development of probabilistic potential theory. To explain the circumstances that led to this highly successful merger it is necessary to turn briefly to classical, or nonprobabilistic, potential theory.

Potential theory is the mathematics of equilibrium. It studies harmonic functions, which arise whenever one has a homogeneous medium in a state of equilibrium. Consider an elastic membrane stretched across a stiff, closed frame that is twisted into some fixed shape in space [*see illustration above*]. The configuration of such a membrane is given by the height h of each point $\tilde{P}$ on the surface of the membrane. Directly below each point $\tilde{P}$ on the membrane is a point P in the base plane, which has the coordinates x, y. Thus if the coordinates x, y are given, then h is a determined quantity; h is said to be "a function of x and y," or in more concise symbolic terms, $h = h(x, y)$.

It is physically clear, and easy to prove mathematically, that h is continuous, and moreover that it has the following simple property: If P is a point in the x, y plane, and Γ is a small circle with its center at P, then the value of h at P

(that is, the height of the membrane above P) equals the average of the values of h for all points on the circle Γ. This is called the mean-value property, and a continuous function h possessing this property is called a harmonic function. In this case it can be seen that the position of $\tilde{P}$ (the point on the membrane above P) is determined by the sum of tension forces exerted on $\tilde{P}$ by the surrounding portion of the membrane. If the membrane is in equilibrium, these forces must cancel, so that the total of nearby elevations greater than the elevation of $\tilde{P}$ must be matched by corresponding elevations lower than the elevation of $\tilde{P}$, and the average must be just equal to the elevation of $\tilde{P}$, namely h at P.

Another physical problem leading to a harmonic function is the problem of temperature equilibrium. In the theory of heat flow (which long antedates the kinetic theory of heat) it is known that in a homogeneous solid the temperature at any point P tends to fall if the average nearby temperature is lower than that at P; it tends to rise if the average nearby temperature is higher than that at P. If the body is in thermal equilibrium, so that the temperature at any given point does not change with time, then the temperature at that point must equal the average temperature over the surface of a small surrounding sphere. In other words, the temperature T is a harmonic function of the coordinates x, y, z of the point P [*see illustration below*].

The remarkable discovery that all the main problems and features of classical potential theory have a mathematical counterpart in the theory of Brownian motion was foreshadowed in 1928 by the work of Richard Courant, K. O. Friedrichs and H. Lewy in Germany. The mathematical equivalence of the two theories has been fully exploited in the past two decades by a host of mathematicians, including Joseph Doob, Gilbert Hunt and Mark Kac in the U.S., E. B. Dynkin in the U.S.S.R., P. A. Meyer in France and Shizuo Kakutani and K. Ito of Japan.

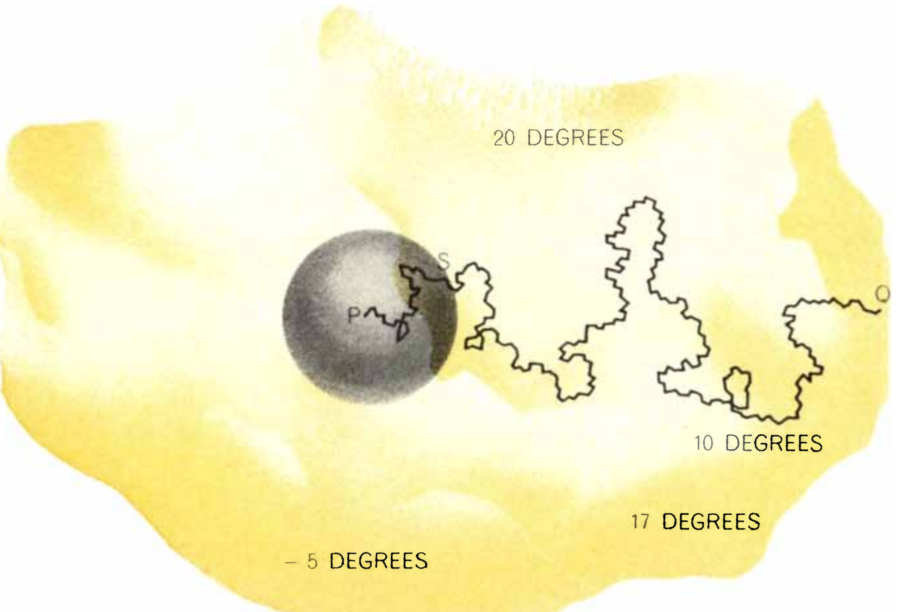
The happy result of all this work is that today any information available in one theory can be translated into a theorem in the other. In particular, it often happens that what is difficult or obscure in one theory is completely transparent in the other. We shall now give several examples to show how light can be shed in either direction by this relation.

The main connection between the theory of Brownian motion and potential theory is made by way of the central problem of potential theory, which is called the Dirichlet problem after the German mathematician P. G. L. Dirichlet. Suppose that in the foregoing example of a body at thermal equilibrium the temperature is measured at all points on the surface of the body. Some points are hot and others are cold, and if the body has been maintained in this state for a while, one can expect that a thermal equilibrium has been attained in the interior. The temperature in the interior varies from point to point, but at each fixed point it does not change with time. From these assumptions can one compute the interior temperature?

In mathematical terms what one is seeking here is a harmonic function, defined in the interior of the body, that takes on certain known values on the surface, or boundary, of the body. (In the membrane example the same mathematical problem has the following physical interpretation: Given the position of the boundary of the membrane, compute the position of the interior points.) The study of Dirichlet's problem has occupied the attention of many leading mathematical analysts of the past century. Wiener made major contributions to this study, yet he never saw how his own theory of Brownian motion could be used to solve the Dirichlet problem; this was first done by Kakutani.

To explain Kakutani's method we resort to the language of Monte Carlo or Las Vegas. Considering once again the solid body of the foregoing example, choose an interior point P at which one wants to know the equilibrium temperature. Having chosen P, one now plays a somewhat unconventional gambling game: Use the point P as the starting position for the Brownian motion of a particle. Watch the particle. It will wander around and eventually (with a probability of 1!) hit the boundary. Say it hits at a certain point Q. Then one wins an amount equal to the known temperature at Q. Of course, it is a matter of chance where the particle first hits the boundary. In short, P is a determined point in the interior, whereas Q is a random point on the boundary.

Obviously one will win at the most the maximum boundary temperature and at the least the minimum boundary temperature. Moreover, as in any gambling game, there is a certain amount one can expect to win in the long run if one plays habitually. This quantity can be found simply by playing a great many times and computing one's average winnings. We shall call this quantity the "expected value" of P, or $E(P)$ for short. The "expectation" of the game is the amount a rational gambler would be willing to pay to the gambling house for the privilege of playing. It is this quantity—the expected winnings if one starts at P—that is precisely Kakutani's solution of the Dirichlet problem.



DISTRIBUTION OF HEAT in a homogeneous solid body at thermal equilibrium is another physical problem that involves a harmonic function. Since the temperature at any given point in such a body does not change with time, the temperature at that point must equal the average temperature over the surface of a small surrounding sphere. In other words, the temperature T is a harmonic function of the coordinates x, y, z of the point P. The problem can be solved by means of the probabilistic theory of the Brownian motion of a hypothetical particle starting at P and hitting the surface of the body at a random point Q.

	5	2	7
	3	−3	−4
	4	0	3
	−1	1	1
SUBTOTAL:	11	0	7
SUBAVERAGE:	11/4	0	7/4

5, 3, 4, −1, 2, −3, 0, 1, 7, −4, 3, 1

TOTAL: 18

AVERAGE: 18/12 = 3/2

AVERAGE OF SUBAVERAGES: 1/3 (11/4 + 0 + 7/4) = 18/12 = 3/2

PRINCIPLE that the average of the subaverages equals the grand average is employed in the solution of the heat-distribution problem by means of probabilistic potential theory.

The idea is so simple that (by neglecting a few mathematical fine points) one can readily verify that $E(P)$ is indeed a solution. First of all, it should be clear that this expectation is a number. As we have defined it, it is a number of dollars, but we can equally well interpret it in degrees of temperature. Moreover, what number it is clearly depends on what starting point one designates for the particle in Brownian motion. If one starts close to a hot part of the boundary, one can expect to win more than if one starts near a cold part of the boundary. Thus a number is actually associated with each interior point P. In order to verify that $E(P)$ is the equilibrium temperature (the solution of Dirichlet's problem) there are only two criteria to check: first, that it matches the known temperature on the boundary, and second, that it is harmonic (that it is continuous and has the mean-value property) in the interior.

With respect to the first criterion, it is clear that if one starts at a point P that is actually on the boundary, then the game is over before it begins, and the payoff $E(P)$ is precisely the known temperature at the starting point; in other words, $P = Q$ with a probability of 1. Moreover, it is plausible (and can be proved rigorously) that if one starts the particle at an interior point P sufficiently close to a particular boundary point Q_0, then it is almost certain that one will first hit the boundary at a point very close to Q_0, so that the expected winnings—the equilibrium temperature assigned to P—is very close to the known temperature at Q_0. Thus Kakutani's solution does have the required boundary behavior. (In this argument it is tacitly assumed that near Q_0 the boundary is smooth and that the boundary temperature is continuous.)

Next one has to show that as a function of P the expected value $E(P)$ in this game is a harmonic function. Again the required continuity is intuitively clear. All this means is that if the starting point P is changed very slightly, then the expected winnings are changed only slightly. The payoff obviously depends on the relative distance of the starting point from the hot and cold parts of the boundary; a slight change in starting position means a slight change in these distances.

What about the mean-value property? This is the only part of the argument that is not perfectly straightforward, in that it requires the introduction of what might be called a gimmick. Draw a small sphere designated Γ around P. Now, in order to show that E is harmonic one has to show that $E(P)$ is equal to the average of all the $E(S)$'s, where S is an arbitrary point on the sphere Γ. $E(P)$ is the expected payoff of the game if the particle starts out at P. Pick a point S on Γ, and suppose that one considers only those plays of the game in which the particle first meets Γ precisely at S. Call the average winnings in these plays $E(P/S)$. Because Brownian motion has no preferred direction, each point S on Γ is equally likely to be the first, and so, by the principle that the average of the subaverages equals the grand average [*see illustration above*], one sees that $E(P)$ equals the average, taken over all points S on Γ, of $E(P/S)$.

The proof would be complete if it were now possible to show that $E(P/S)$ is the same as $E(S)$, $E(S)$ being the expected winnings for a particle starting at S, and $E(P/S)$ being the expected winnings for a particle starting at P and first meeting Γ at S. At any instant, however, the particle in Brownian motion behaves only on the basis of its present position; it is not influenced by its past. (This is called the "Markov property.") Therefore the expected behavior of a path from P through S is no different from the behavior of a path starting at S; consequently $E(S)$ does indeed equal $E(P/S)$. In physical terms this means that one could solve the equilibrium problem for heat flow or for a membrane by observations of Brownian motion, or conversely that one could find the expected outcome of a Brownian-motion experiment simply by observing the equilibrium configuration of a membrane or a heat conductor.

A noteworthy feature of the probabilistic method of solving the Dirichlet problem is that the boundary can be as irregular as one pleases. Other procedures for solving the problem all encounter complications if the surface of the domain is too "spiky" or "hairy." (The surface of a bulldog is all right, but not the surface of a Saint Bernard or a poodle.) In contrast the Brownian-motion solution is meaningful in all cases. For a badly behaved boundary the prescribed boundary values are taken "on the average," but not necessarily at each point.

Having shown how the connection between Brownian motion and potential theory has been exploited to obtain deep insights into classical potential theory by simple probabilistic arguments, we shall now give a few examples, based on the work of Kakutani and Doob, of how, on the other hand, complicated and deep questions in probability are sometimes equivalent to very simple questions in potential theory.

Our first example is the "gambler's ruin" problem. Suppose that one of us (Hersh) plays with an opponent (Griego) at matching quarters. Hersh's fortune at the start is N dollars. Griego has M dollars. Hersh resolves to play until he either "breaks the bank" (by winning M dollars) or is "ruined" (by losing N dollars).

The question is: What is the probability that Hersh will be ruined? It is intuitively clear that the answer depends on the relative sizes of M and N. If N is much less than M, ruin is very likely; if N is much greater than M, Hersh is pretty sure to break the bank. What may not be clear is that it is possible to get an exact answer!

To do this we define a second game. Imagine that you (the reader) are an onlooker at our quarter-matching. You watch our luck, and you bet a dollar that Hersh will be ruined. If Hersh is ruined, you win a dollar; otherwise you win nothing. What are your expected winnings? Clearly they are equal to Hersh's probability of being ruined. Furthermore, if we call $x(t)$ Hersh's net gain or loss at time t, then x changes from one instant to the next in the same way that a Brownian particle changes position. The Brownian particle moves to the left or the right with equal probability; Hersh's fortune increases or decreases with equal probability. The game is over when Hersh's winnings x are either $+M$ dollars or $-N$ dollars. This corresponds to a particle in Brownian motion on the x axis between the points $x = +M$ and $x = -N$. The probability of ruin is the probability that the particle first hits the left boundary ($x = -N$) and not the right boundary ($x = M$). That means your game has a payoff equal to 1 if the particle first hits the left boundary and a payoff equal to 0 if it first hits the right boundary. This corresponds precisely to a Dirichlet problem for a one-dimensional domain (the interval between $x = -N$ and $x = M$) with boundary values 1 at $x = -N$ and 0 at $x = M$.

We could consider the corresponding equilibrium-temperature problem, but it is even simpler in this case to visualize a membrane problem. In fact, since our domain is one-dimensional (a part of the x axis) we must consider a one-dimensional elastic, say a stretched rubber band. Everyone knows that the equilibrium position of a stretched rubber band is a straight line. Since in this case the particle representing Hersh's winnings starts at $x = 0$, we are interested in the height of the rubber band above the point $x = 0$. Simple geometry shows that it is just $M/(M + N)$ [*see illustration below*]. This, then, is Hersh's probability of being ruined.

The method we have employed here is a simple and powerful one. We essentially constructed a special Dirichlet problem, taking care to choose the domain and the boundary values strategically, so that the problem would have an interesting probabilistic interpretation. The solution was available by inspection because the associated equilibrium configuration was so extremely simple.

Our next example will require a little more acquaintance with potential theory, but it will yield a much deeper probabilistic result. Choose a fixed point as an "origin," and consider the domain $\mathcal{D}$ of all points P whose distance from the origin is greater than ε and less than K. Here, as usual, ε is supposed to be a small positive number and K a very large one. In three-dimensional space $\mathcal{D}$ is the region between two concentric spheres, an inner one of radius ε and an outer one of radius K. In two dimensions the same conditions describe a ring between two concentric circles. In one dimension $\mathcal{D}$ is a pair of disconnected intervals, one to the right and one to the left of the origin [*see illustration on page 74*]. In each case we pose a Dirichlet problem by asking for a function u that is harmonic in $\mathcal{D}$ and equal to 1 on the inner surface and 0 on the outer surface. What is the solution?

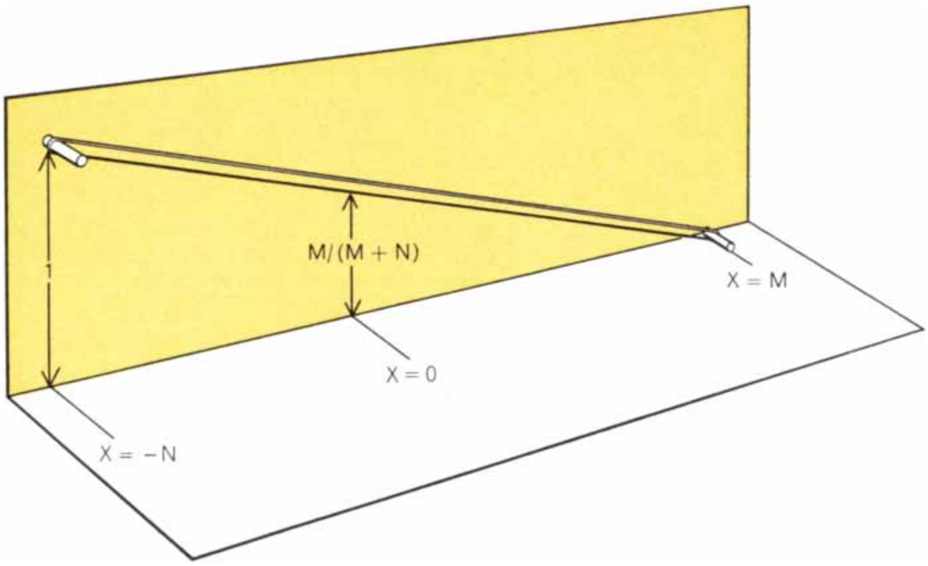
In one dimension, as in the stretched rubber band, the only harmonic functions $u(x)$ are those that have straight lines as their graphs. A comparable construction shows that in the one-dimensional case the solution to the present problem is $u(x) = (K - x)/(K - \varepsilon)$ for x greater than ε, and $u(x) = (K + x)/(K - \varepsilon)$ for x less than $-\varepsilon$.

Just as in the gambler's-ruin problem, $u(x)$ signifies the probability that if a particle starts at position x, it will hit the inner boundary (which is given a payoff equal to 1) before the outer boundary (which is given a payoff equal to 0). The formulas given above show that if K is very large, u is very close to 1. It is possible to take the limit of $u(x)$ as K goes to infinity; then $u = 1$ for all x and all ε. Since the outer boundary has now vanished to infinity, $\mathcal{D}$ is just the set of all points outside the interval from $-\varepsilon$ to $+\varepsilon$, and $u(x)$ is the probability that a particle starting at x will eventually touch that interval. Since $u = 1$, it follows that the particle is almost certain to do so. Because both the origin and the starting position x are arbitrary, the particle will arrive at every point on the line. Having arrived there, the same argument applies once more to the future, so that it will in fact almost certainly return infinitely many times to every point. This property is described by the term "recurrent," and what we have shown is that Brownian motion in one dimension is recurrent.

To consider the same question in two or three dimensions, one needs to know only the solutions of the corresponding two- and three-dimensional Dirichlet problems, with the boundary values 1 assigned on the inner boundary and 0 on the outer boundary. Simple considerations, which would be out of place in this article but which require only elementary calculus to carry out, show that in two dimensions the solution to our Dirichlet problem for a circular ring is $u(r) = (\log K - \log r)/(\log K - \log \varepsilon)$, where log denotes logarithm. In three dimensions, for a spherical shell, the solution is $u(R) = (1/K - 1/R)/(1/K - 1/\varepsilon)$. We use r to denote distance to the origin in the plane, and R to denote distance to the origin in three-space. In each case r and R are numbers between ε and K.

These functions $u(r)$ and $u(R)$ have the same probabilistic meaning as the $u(x)$ we just obtained in the one-dimensional case; they give the probability that a particle, starting at r or R units from the origin, will hit the inner boundary before it hits the outer one.

The interesting question is: What happens as K becomes extremely large? Since log K goes to infinity as K goes to infinity, we see that in two dimensions, as in one, Brownian motion is recurrent, that is, the particle is almost sure to re-



ONE-DIMENSIONAL ELASTIC (in this case a rubber band stretched between two pegs in a wall) is considered in finding a solution to the "gambler's ruin" problem by means of probabilistic potential theory. The game is matching quarters. One player (Hersh) has a fortune of N dollars at the start. The other player (Griego) has M dollars. Hersh resolves to play until he either "breaks the bank" (by winning M dollars) or is "ruined" (by losing N dollars). In the theory of Brownian motion the particle representing Hersh's winnings starts at $x = 0$, and the probability of ruin is the probability that the particle first hits the left boundary ($x = -N$) and not the right boundary ($x = +M$). This means that an onlooker's game has a payoff equal to 1 if the particle first hits the left boundary and a payoff equal to 0 if it first hits the right boundary. Since the equilibrium position of a stretched rubber band is a straight line, Hersh's probability of being ruined is the height of the rubber band above the point $x = 0$. Simple geometry shows that this value is just $M/(M + N)$.

SCIENCE/SCOPE

Apollo 8 voice and television communications depended heavily on Hughes-built equipment:

...the 20-watt microwave traveling wave tube that sent Apollo 8's signals earthward (one of the more than 100 flight-quality TWTs built for NASA since 1962 and used on the Syncom, Early Bird, Intelsat 2, and ATS satellites and the Pioneer, Mariner, Lunar Orbiter, and Surveyor spacecraft);

...the antenna-feed subsystems aboard three special vessels stationed on the high seas around the globe to receive and relay Apollo 8's signals;

...the ATS satellite used in support of the Apollo 8 splashdown; it relayed TV from USS Yorktown to Brewster Flats, Wash., for commercial distribution.

The first European-built equipment for NADGE -- the $300-million air defense system that will guard NATO nations from Norway to Turkey -- is now undergoing integration testing at Hughes in Fullerton, Calif. The data display console built by Selenia S.p.A. of Italy and the video extractor by N.V. Hollandse-Signaalapparaten of The Netherlands are linked with a general-purpose computer and other equipment built by Hughes.

ENGINEERS: Hughes is on the grow! New development programs offer immediate opportunities for Systems, Circuits, Radar, Communications, Electro-optical, Display, Sonar, Electronics Packaging, Test Equipment, and Solid State Microwave engineers. Engineering degree, at least two years of related experience, and U.S. citizenship required. Please send your resume to Mr. J. C. Cox, Hughes Aircraft Co., P.O.Box 90515, Los Angeles 90009. An equal opportunity employer.

An infrared night sight for the Army's Cheyenne helicopter, now being built by Hughes under contract with Lockheed-California, will give the gunner a picture of ground targets nearly as clear as he would see in daylight. The PINE (for Passive Infrared Night Equipment) system enables him to locate targets and fire automatic guns, rockets, grenades, machine guns, or Hughes-built TOW wire-guided anti-tank missiles.

A new method of detecting flaws in metals was presented at the eighth Symposium of Physics and Nondestructive Testing in Chicago recently by a University of Arizona professor and a Hughes engineer. Their method sends ultrasonic Lamb waves throughout a solid material to find defects, much like a submarine sends out sonar waves. Engineers can pinpoint the location and size of flaws by noting the magnitude of the echo signals and the time they take to return.

An orbiting "windowshade" of solar cells, which will capture enough of the sun's energy to produce 1500 watts of power, is being built by Hughes under contract to the Aero Propulsion Laboratories of the U.S. Air Force. Designed to supply future satellites with electricity, it consists of two flat sheets of solar cells (called arrays), each 5½ x 16 feet long, which will unroll into space from a drum. Space testing of the system is scheduled for late 1970.

Creating a new world with electronics

HUGHES

HUGHES AIRCRAFT COMPANY

turn infinitely many times to a small neighborhood of any point.

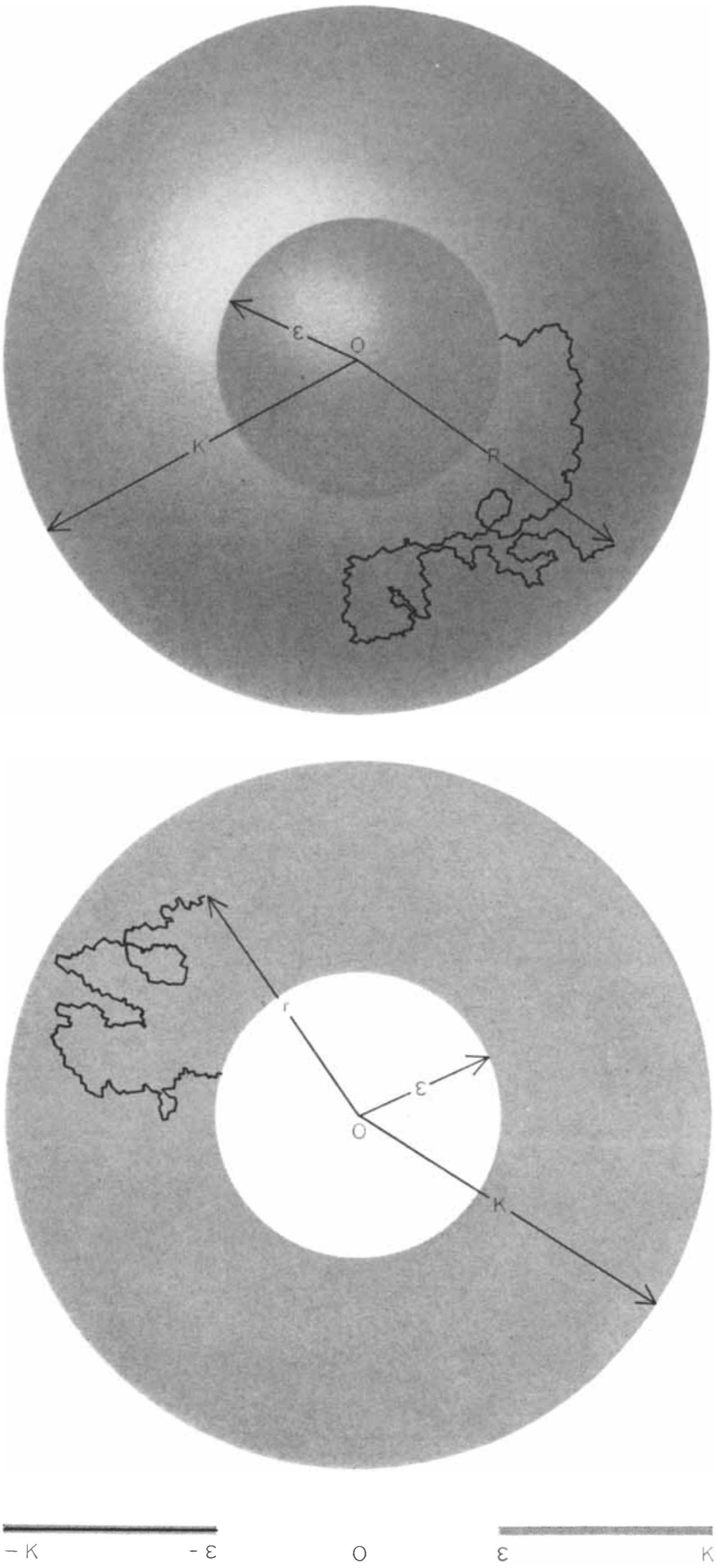
In three-dimensional space, on the other hand, letting K go to infinity yields $u(R) = (-1/R)/(-1/\varepsilon) = \varepsilon/R$. This is the probability that the particle, starting at distance R greater than ε, will ever approach within ε of the origin. Since u is less than 1, there is a positive probability that the particle will wander off and never return. There is, so to speak, more room in three-space to escape. Thus Brownian motion in three dimensions is nonrecurrent. This result, which we have obtained with modest effort, is a deep result in probability theory.

The function $u(R) = \varepsilon/R$, which we have just considered, can be extended by setting it equal to 1 for R less than or equal to ε. The extended function we have defined is known as the capacitory potential of the sphere S_ε, with radius ε and center at 0. The capacitory potential of a set B is an important notion of classical potential theory; it is a function harmonic outside B, equal to 1 inside B, and equal to 0 very far from B (at infinity).

Just as in the special case of S_ε, so in very general cases the capacitory potential of B is simply the probability that a Brownian particle, starting at a given point, will ever hit B. Indeed, almost the same arguments applied to the case of a sphere would show that the probability is equal to 1 for a starting point inside B, is harmonic for a starting point outside B and is small at great distances from B.

Current work in this area has yielded far-reaching generalizations of both Brownian motion and potential theory. The interconnection between classical potential theory and Brownian motion depends heavily on the fact that Brownian motion is a Markov process, that is, its present behavior is not influenced by its past behavior. Recent investigations have shown that in a very real sense every decent Markov process corresponds to some generalized potential theory. For example, the classical theory of Riesz potentials corresponds to what are called the stable processes of probability theory. Moreover, Markov chains (which are discrete Markov processes) have their own potential theories.

Thus the probabilistic viewpoint in potential theory has unified and clarified the underlying principles of potential theory, and conversely concepts borrowed from potential theory and applied to probability theory have demonstrated the deep analytic structure of Markov processes. This has helped to end the isolation in the mathematical realm that probability theory has suffered from to some degree in the past.



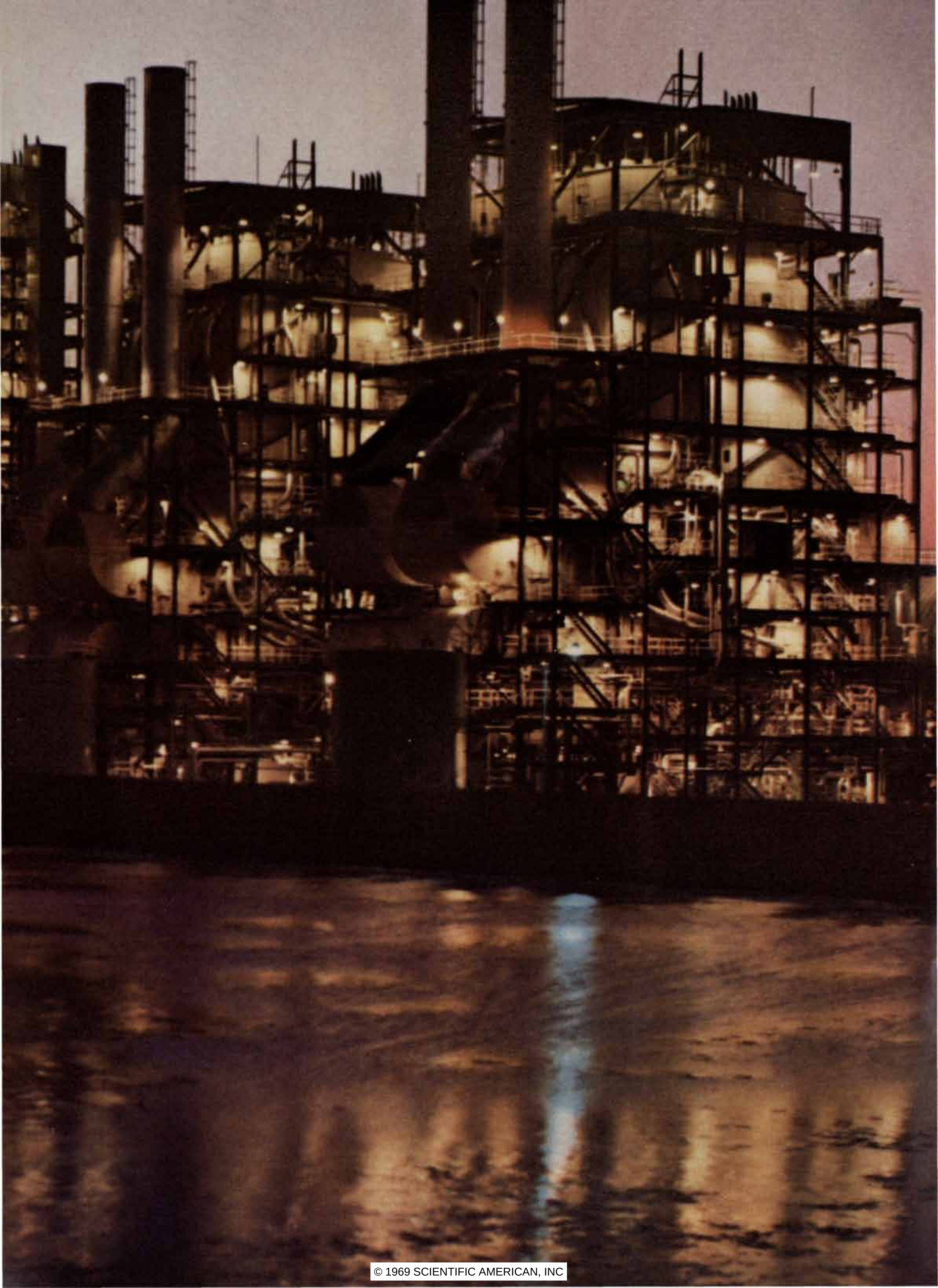
DEEP RESULT in probability theory is obtained in the process of solving the following problem: Choose a fixed point as an "origin," and consider the domain $\mathcal{D}$ of all points P whose distance from the origin is greater than ε and less than K. (ε is supposed to be a very small positive number and K a very large one.) In three-dimensional space (*top*) $\mathcal{D}$ is the region between two concentric spheres, an inner one of radius ε and an outer one of radius K. In two dimensions (*middle*) the same conditions describe a ring between two concentric circles. In one dimension (*bottom*) the domain is a pair of disconnected intervals, one to the right and one to the left of the origin. In each case the problem is to find a function u that is harmonic in $\mathcal{D}$ and equal to 1 on the inner surface and 0 on the outer surface.

We have a fatal attraction for the devastating pine bark beetle

Mountain pine beetles kill over 400,000 mature white pines a year, plus uncounted thousands of ponderosa, lodgepole and sugar pines. Conventional methods of control proved inadequate and too costly. Since 1965, Potlatch foresters have sponsored research by leading biologists and chemists to solve this serious insect problem of federal, state and private timberlands. And now it looks like we have an answer. The female, after she bores into the inner bark, produces an attractant that draws a mass infestation of hungry beetles that kill the tree. This essence has been synthesized, and will now be tested commercially in our forests to lure beetles to selected areas where they can be destroyed. Then everybody will benefit except the beetles. Potlatch Forests, Inc., P.O. Box 3591, San Francisco 94119.

Potlatch

Potlatch, the forests where innovations grow...in wood products, in paperboard and packaging, in business and printing papers



d'Arazien on d'Arazien on Polaroid Land Film

Arthur d'Arazien is a world-famous industrial photographer.

"This photograph is a triple exposure.

And I took 19 more just like it on Polaroid 4x5 Land color film.

Frankly, I didn't think it could be done until I did it. And I've taken a lot of multiple-exposure pictures.

You see, after every exposure I took the film out of the camera and waited until the light had changed to my liking before I put it back and exposed again. It was daylight when I took the first exposure.

What's so unusual about that is the accuracy I got. What with putting in and taking out 20 sheets of film it's remarkable I didn't get a triple image.

If you've ever tried a series of multiple exposures with conventional cut film in a cut film holder and watched that springback snap against the body when you remove the holder, you know how chancy it is.

With Polaroid 4x5 film it's different. Once you've put in the Polaroid 4x5 Land Film Holder, all you're taking out or putting in is the film. And no snapping springs.

To me, getting accurate pictures is important. But that means more than sharpness. The color has to be right too.

The color of Polacolor film is right.

In fact, in some situations it adds a quality to the picture no other film can equal.

On top of all that, as a technically-oriented person, I am just fascinated with the ability to get excellent prints in seconds.

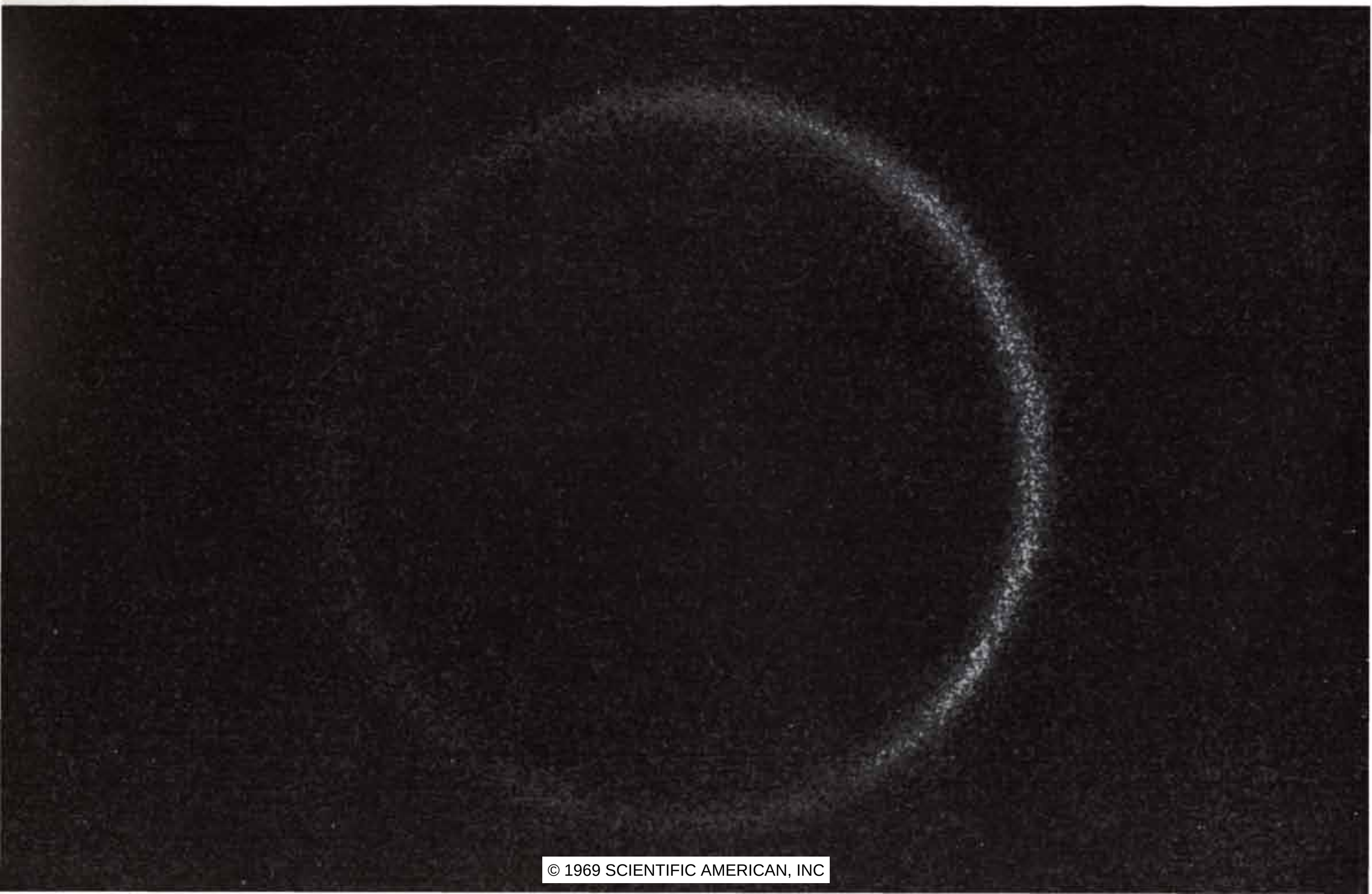
Somehow, it's like having a color lab in your camera. A lab I own and operate. So I can change results at will, by changing development time and temperature. Or by adding or taking away filters.

And the fact that all that is under my control, while I'm shooting, can often lead me into areas I might not have explored otherwise.

It appeals to my sense of adventure."

MARS (*above*) has a thin, cold atmosphere consisting chiefly of carbon dioxide, which in solid form may be the material seen in the planet's permanent polar caps. This color picture is a composite of several 16-millimeter motion-picture frames taken in July, 1956, by Robert B. Leighton of the California Institute of Technology.

VENUS (*below*) has a dense, hot atmosphere, which also consists chiefly of carbon dioxide. This view taken at the Lowell Observatory when Venus was almost directly in line between the sun and the earth shows how the planet's atmosphere scatters sunlight, creating a thin luminous arc opposite the directly lighted crescent.



The Atmospheres of Mars and Venus

Distance from the sun can only partly explain why the atmospheres of these two planets are so different. If life had not appeared on the earth, its atmosphere might resemble that of one or the other

by Von R. Eshleman

The atmosphere of Venus is hot and dense and the atmosphere of Mars is cold and thin. Spacecraft have gathered information about the atmospheres of our nearest planetary neighbors that could not be obtained any other way. The results of these flights, together with new measurements made from the earth, mean that one can begin to speak with considerable confidence about the temperature, density and chemical composition of the atmospheres of Venus and Mars. It also becomes profitable for the first time to consider how the atmospheres of these two planets came to be so different from the atmosphere of the earth. It is now clear that the presence of life has been, and continues to be, a controlling influence on the composition of the earth's atmosphere. Conversely, the apparent absence of life on Venus and Mars may explain much about the nature of their atmospheres. In short, the experiments carried out by American and Russian planetary probes already constitute a preliminary life-detection test for our nearest neighbors beyond the moon.

Not quite four years ago (in July, 1965) the American spacecraft *Mariner IV* passed within 6,118 miles of the surface of Mars. When the spacecraft began its atmospheric measurements, it was early in the afternoon of a winter's day in a "desert" known as Electris in the planet's southern hemisphere; *Mariner IV* indicated that the temperature of the atmosphere at that point was about −113 degrees Celsius, 26 degrees below the coldest air temperature ever recorded on the earth. Nearer Mars's poles during the long winter night some of the atmospheric gas itself may freeze and fall to the surface [*see top illustration on opposite page*].

On successive days in October, 1967, the Russian space probe *Venera 4* and the American probe *Mariner V* found that the temperature of Venus' atmosphere increased nearly 10 degrees C. for each kilometer of decrease in altitude, reaching 100 degrees C.—the boiling point of water at sea level on the earth—at an altitude of about 45 kilometers. It is now apparent that neither spacecraft measured conditions all the way down to the surface, but by extrapolation it seems likely that the temperature there is 425 degrees C., or some 100 degrees above the melting point of lead.

Extrapolations of measurements from *Venera 4* and *Mariner V* also indicate that the atmospheric pressure at the surface of Venus is roughly 100 times that on the earth, comparable to the pressure at a depth of one kilometer in our oceans. The thick air on Venus would seem like a hot, dilute fluid with a tenth the density of water. Such an atmosphere would strongly bend the paths of light rays and create strange optical distortions. In contrast the pressure of the atmosphere at the surface of Mars is only about a hundredth that on the earth.

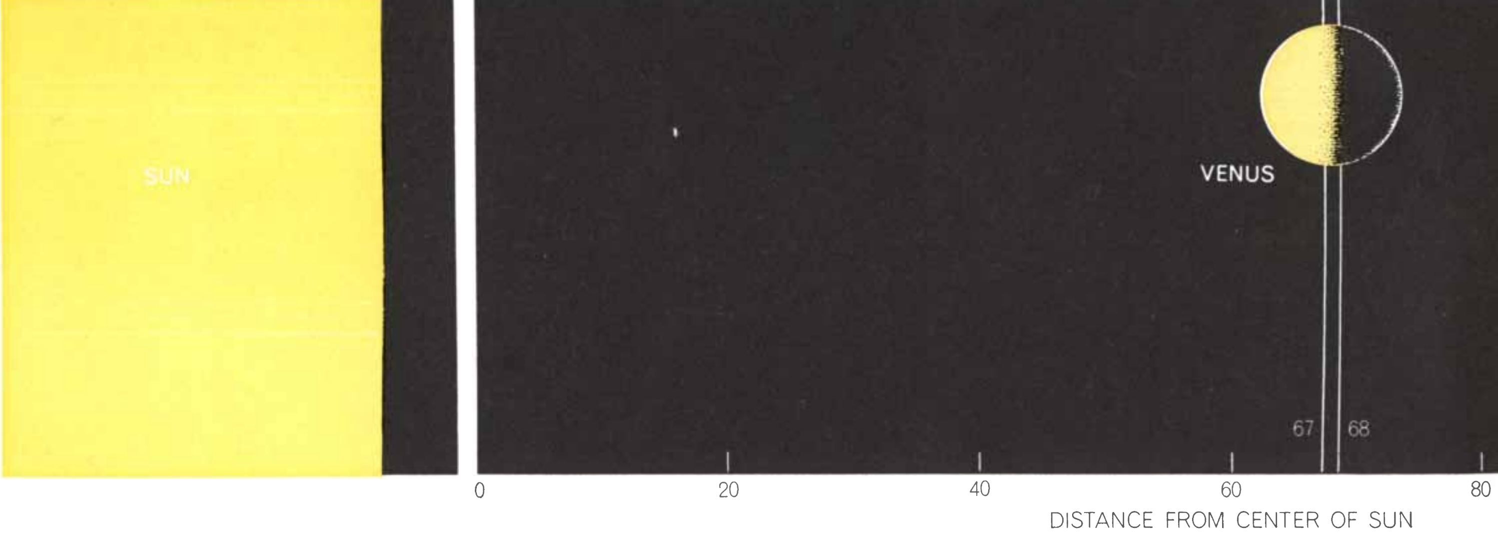
In the atmospheres of both Mars and Venus the main component is carbon dioxide. On the earth carbon dioxide is only .03 percent of the atmosphere; nitrogen accounts for 78 percent and oxygen for 21 percent. On both Mars and Venus free oxygen is certainly rare and may be virtually absent; the proportion of nitrogen in their atmospheres does not exceed 20 percent and could be much smaller.

What about water, which is so plentiful on the earth? If the earth were as hot as Venus, the oceans would evaporate, creating an atmosphere so thick with water vapor that the pressure at the surface would be about 300 times its present value. Apparently Venus has less than .7 percent water vapor in its lower atmosphere, making atmospheric and surface water 400 times less abundant on Venus than it is on the earth and possibly much less so. In Mars's atmosphere water vapor is barely detectable, but large amounts could possibly be frozen below the solid surface of the planet.

I find such comparisons fascinating and also very puzzling. Venus, the earth and Mars are within a factor of two of being the same distance from the sun and the same size [*see top illustration on next two pages*]. All three planets were doubtless formed from the same primordial source of matter at the same time. The atmospheres of all three appear to have been created after the initial supplies of light gases (hydrogen and helium) were largely lost to space. It is generally thought that these secondary atmospheres represent gases "exhaled" from the main body of the planet, with atmospheric evolution proceeding for aeons to produce the markedly different conditions we find on the three planets today.

Observations carried out by the space probes, together with new observations from the earth, have sharpened speculation about the reasons for the present differences in the atmospheres of the three planets. For example, it has recently been suggested that if all life on the earth were destroyed, the atmosphere would slowly begin to change and eventually would resemble the atmosphere of either Venus or Mars. I shall return to such speculations after first describing some of the experiments performed by the American and Russian spacecraft [*see bottom illustrations on next two pages*].

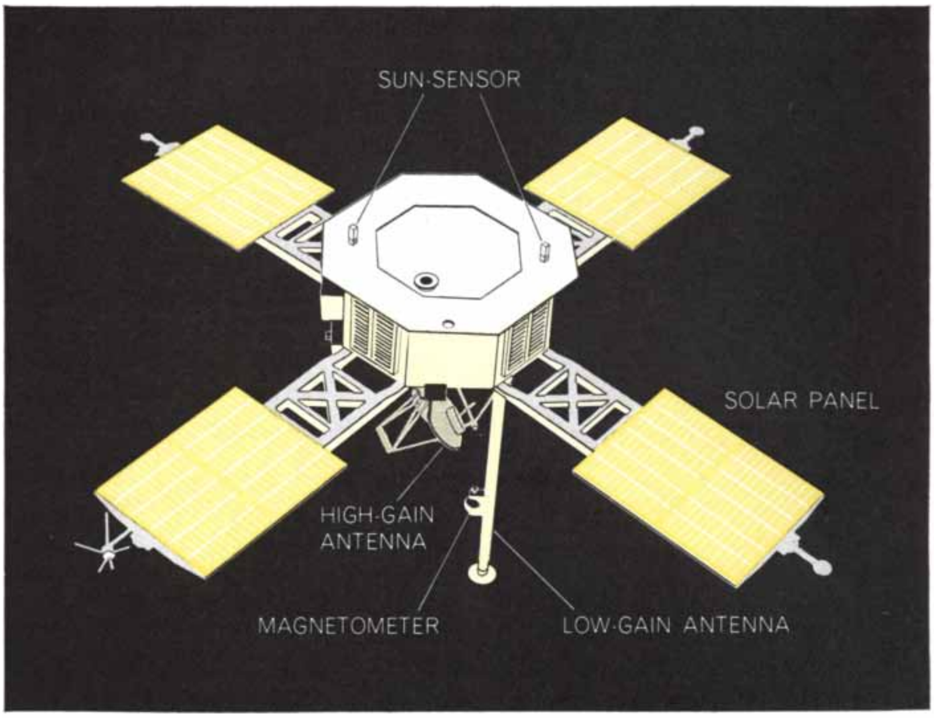
Until a few months before *Mariner IV* was launched in 1964 the plan for the mission did not include any radio measurements of the atmosphere of Mars. Although various methods of mak-



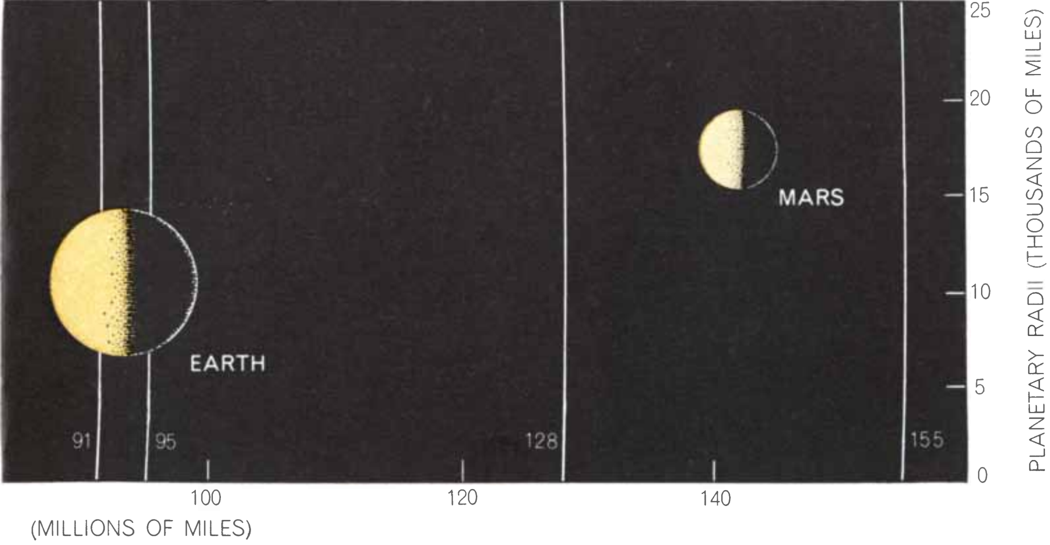
VENUS, EARTH AND MARS resemble one another much more closely than they resemble any of the other planets in the solar system. Their distance from the center of the sun and their size are plotted here on two scales: the size scale (*far right*) is arbitrarily 1,500 times the distance scale (*bottom*). On this size scale the curvature of the limb of the sun (*left*) is barely evident. The minimum

ing such measurements had been discussed since 1960, there was a fundamental problem facing the scientists and engineers concerned with the project. The best experiment would be to have the spacecraft pass behind Mars and be occulted, or eclipsed, so that the radio signals between the earth and the spacecraft would pass through Mars's atmosphere. Because the 1964–1965 flight to Mars was to be only the second interplanetary mission attempted by the U.S., the project managers were understandably reluctant to accept the occultation trajectory. They had plotted the path of the spacecraft specifically to avoid such a temporary but critically timed loss of their long, tenuous thread of radio contact with it [see "The Voyage of *Mariner IV*," by J. N. James; SCIENTIFIC AMERICAN, March, 1966].

A group from Stanford University and the Stanford Research Institute was the first to describe how planetary atmospheres and surfaces could be studied by occultation and related radio measurements. This group had proposed an experiment and instrumentation for this flight that would be particularly sensitive to Mars's ionosphere, the ionized upper part of the planet's atmosphere. Another group involved in the radio tracking of space probes first showed in detail, however, that the radio systems already designed for communications and for tracking the spacecraft could supply such precise data on signal characteristics that valuable atmospheric measurements could be made even without adding any new equipment. During the same period that the two groups were refining their predictions of the scientific potential of radio-occultation measurements, other scientists and engineers were beginning to worry that lack of knowledge about the atmosphere of Mars would make it difficult to plan for missions in which spacecraft would land on the planet. It had been thought for years that the pressure of the atmosphere at Mars's surface was about 8.5 percent of the pressure at the earth's surface, but new measurements from the earth seemed to place the value closer to 2.5 percent. This combination of events led to the 11th-hour change in plan to accept the hazard of the temporary loss of communications in order to make atmospheric measurements with these same radio links in the moments before and after occultation. The men who were assigned the responsibility for this experiment were Arvydas Kliore (the principal investigator), Dan L. Cain, Gerald S. Levy and Frank D. Drake of the Jet Propulsion Laboratory of the



MARINER V **approached to within 2,544 miles of the surface of Venus on October 19, 1967, after a voyage of 127 days. The 540-pound spacecraft was designed by the Jet Propulsion Laboratory of the California Institute of Technology. *Mariner IV,* which passed within 6,118 miles of Mars on July 15, 1965, after a trip of 228 days, was very similar in design.**



and maximum distances of the three planets as they travel in orbit around the sun are indicated by the vertical white lines. The atmospheres of the three planets, now very different, appear to have been formed after the light gases initially present escaped into space.

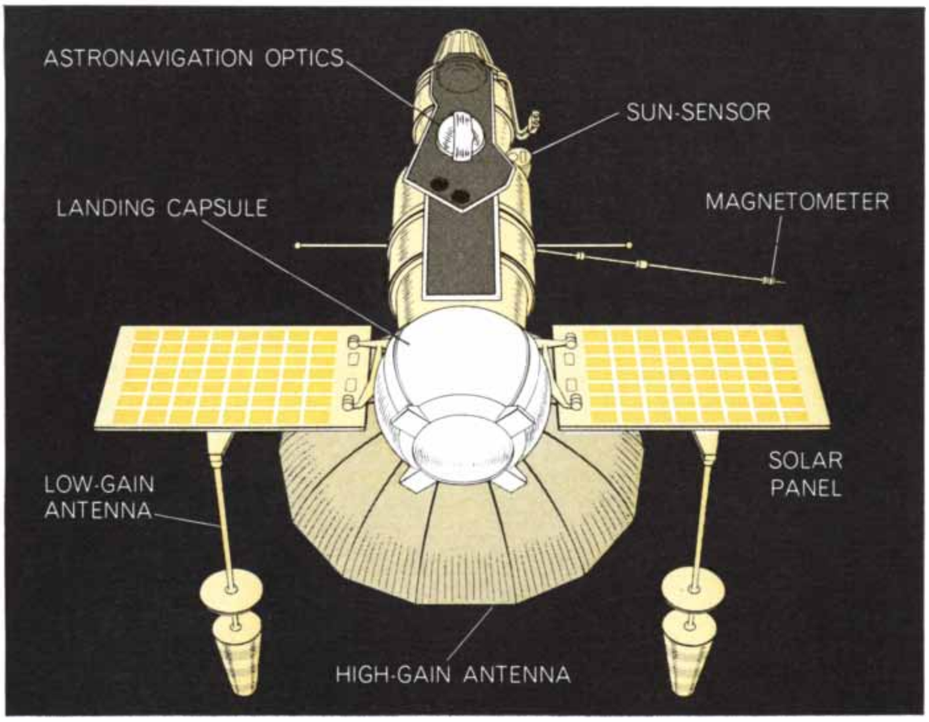
California Institute of Technology, and Gunnar Fjeldbo and myself of Stanford University. (Drake is now at Cornell University.)

The principle involved in the experiment is as follows. The atmosphere of Mars acts as a radio lens, causing the spacecraft to seem to depart from its actual path when it passes behind the atmosphere, just before going into occultation and just after emerging from it. The apparent distance to the spacecraft is determined by sending signals of precisely controlled frequency to the vehicle, which then retransmits the signals to the earth. At the tracking station on the earth the frequency of the returned signal is compared with what was originally transmitted, and the difference (the Doppler frequency) is precisely measured.

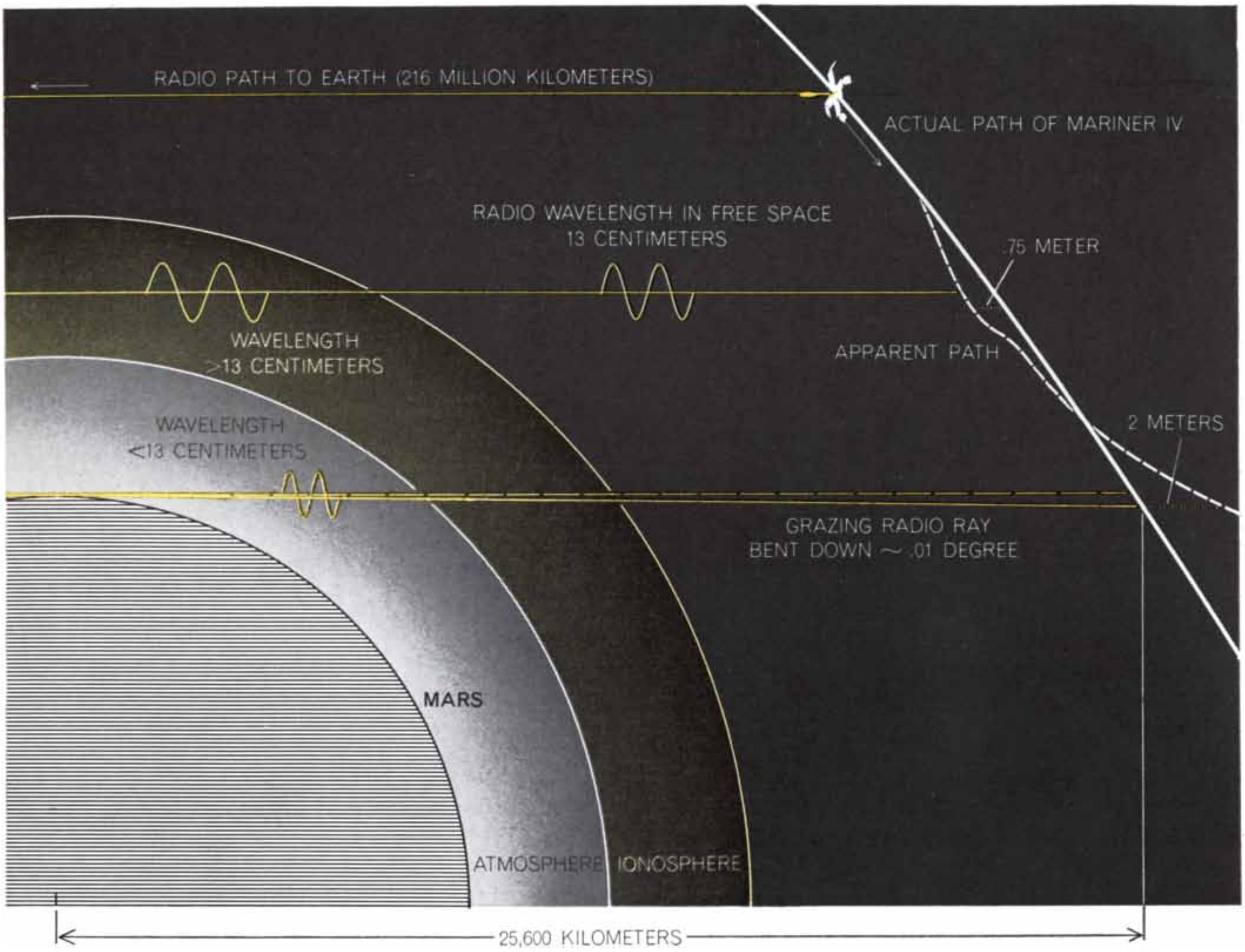
By counting the number of cycles difference in a second, one can determine the apparent distance (in radio wavelengths) the spacecraft has moved in a second, either toward or away from the tracking station. For example, if the frequency of the returning signal is 200,000 cycles per second lower than the outgoing signal and if the radio wavelength of the tracking signals is 13 centimeters, then in one second the spacecraft has apparently moved 13 kilometers away from the tracking station. (The distance is not 26 kilometers, as one might think, because one must divide by two to allow for the two-way path to and from the spacecraft.) If there is nothing but empty space between the spacecraft and the tracking station, the radio wavelength will remain constant at 13 centimeters between the two points, and the apparent motion can be regarded as the true motion of the spacecraft. As the signal passes through the atmosphere of Mars, however, several changes in wavelength occur. In the un-ionized lower part of the atmosphere the phase velocity of the radio waves is less than *c*, the velocity in a vacuum, so that the wavelength becomes less than 13 centimeters. In the ionized upper part of the atmosphere the phase velocity is greater than *c*, and the wavelength becomes greater than 13 centimeters. In addition the path followed by the signals is bent by both the un-ionized and the ionized part of the atmosphere [*see illustration on next page*].

These changes are small, but the accuracy of measurement is great. As *Mariner IV* first moved behind the daytime ionosphere of Mars, the ionosphere made it appear that in one second the spacecraft's motion away from the earth was as much as four centimeters less than the actual increase in distance, which was about 13 kilometers. The sum of such extra effects during the entire period *Mariner IV* was approaching occultation resulted in an apparent displacement of .75 meter toward the earth followed by an apparent displacement of about two meters away from the earth before the signal was completely blocked by the planet. On emergence from occultation the planet's neutral atmosphere made it appear that *Mariner IV* was about 2.4 meters beyond its actual position, but it "came back into focus" without showing a discernible negative shift, which would have been expected if an ionosphere were present on the night side of Mars. In order to appreciate the precision involved in these measurements one should bear in mind that the spacecraft was then about 216 million kilometers from the earth.

It may be hard to believe such small changes could be attributed to the at-



***VENERA 4*, the 2,433-pound Russian probe, reached Venus some 36 hours before *Mariner V* and ejected an instrumented capsule equipped with a parachute. The Russians reported that the capsule measured atmospheric conditions all the way to the planet's surface. There is now good evidence that the last reading was at an altitude of about 25 kilometers.**



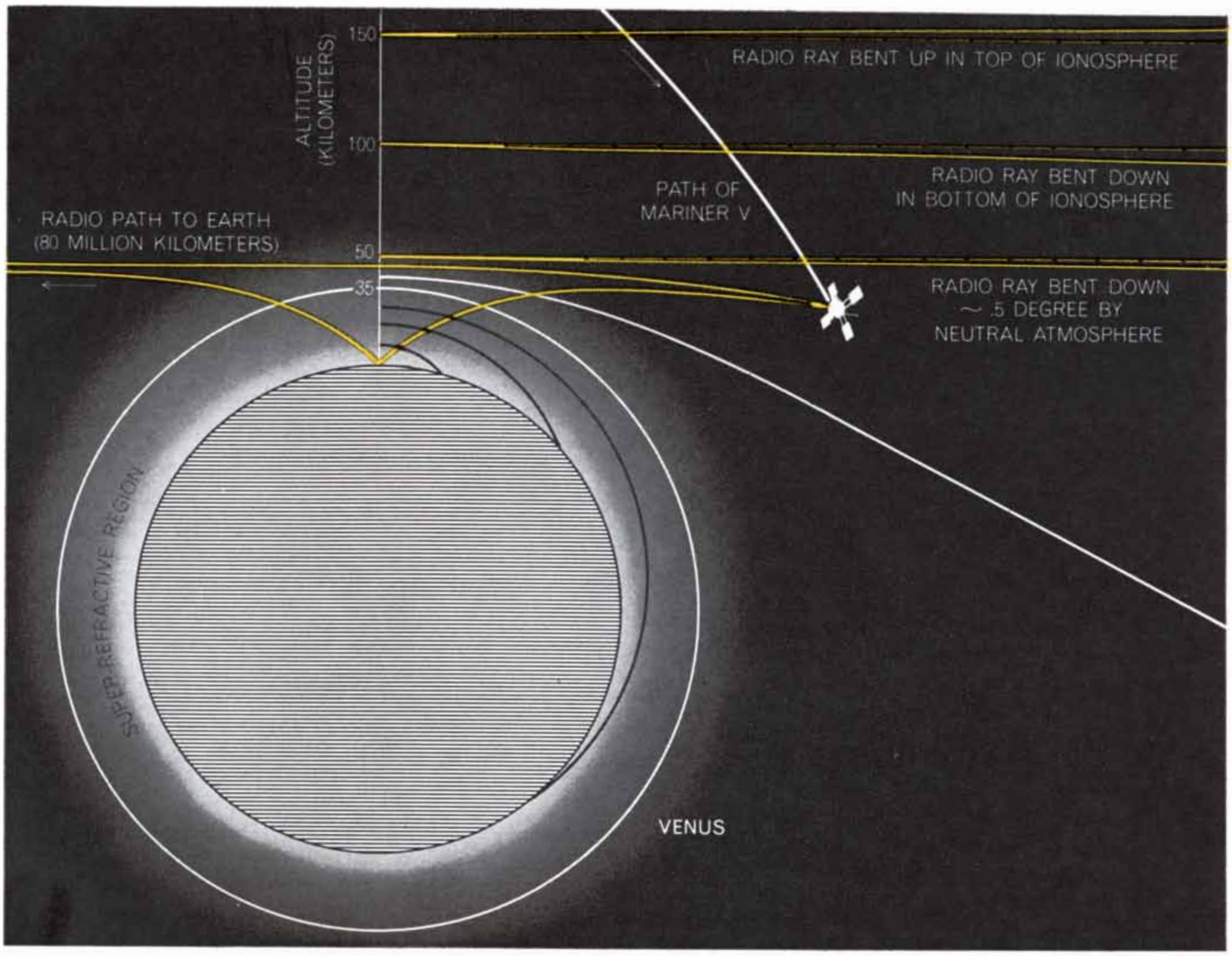
MARS'S ATMOSPHERE was studied by measuring the changes in the frequency of radio signals beamed to *Mariner IV* and retransmitted to the earth in the moments immediately before and after the signals were blocked by the planet. Signals that passed through the ionosphere, the electrically charged part of the upper atmosphere, were increased in wavelength, making it appear that the spacecraft had deviated from its actual path by about .75 meter. Subsequently signals passing through the neutral, lower part of the atmosphere were shortened in wavelength, making it appear that *Mariner IV* had deviated about two meters in the other direction.

mosphere of Mars. After all, the radio signals also had to pass through the much denser ionosphere and lower atmosphere of the earth, and the position and actual path of the spacecraft were not known down to the last meter. In regard to the first point, irregularities in the earth's ionosphere and lower atmosphere did not create a problem because the period of measurement was very short. Most of the effect of the lower atmosphere on both sides of Mars showed itself within 10 seconds, as the spacecraft flashed behind the planet. In such a brief period changes in the earth's atmosphere and in interplanetary space are almost always small enough to ignore, even though they could become important over a longer time. As for the second point, the answer is that, although the exact position and trajectory of the spacecraft are not known on an absolute basis, we do know that the path must be very nearly the path of a point mass in a gravitational field that, at a distance of several tens of thousands of kilometers from the planet, must be highly symmetrical. In the brief period that interests us the law governing motion in such a field can be fitted to the measurements to determine to what extent the law is broken by the extra effects of Mars's atmosphere.

Similar considerations apply to the radio measurements of the atmosphere of Venus by *Mariner V* in 1967, although here the distortion of the radio signals was much greater. Kliore again headed a team that used radio links to study atmospheric effects. In addition a two-frequency occultation experiment was conducted by H. Taylor Howard, Roy A. Long, Fjeldbo, myself and others from Stanford University and the Stanford Research Institute. It measured very tenuous regions in the upper atmosphere of Venus and also provided additional data on characteristics of the lower atmosphere.

One important difference between the results for the two planets arises from the fact that the dense atmosphere of Venus is super-refractive [*see illustration on opposite page*]. It is well known that light and radio rays in the earth's atmosphere are bent by changes of density with altitude; the maximum deflection due to the lower atmosphere is about half a degree when one looks toward the horizon. Thus the disk of the setting sun, which subtends an angle of half a degree, appears to be just above the horizon when in reality it has already gone below it.

Light rays are bent the same amount on Venus at an altitude of about 50 kilometers. At somewhat lower altitudes a horizontal radio ray is bent even more before it escapes into space, and at an

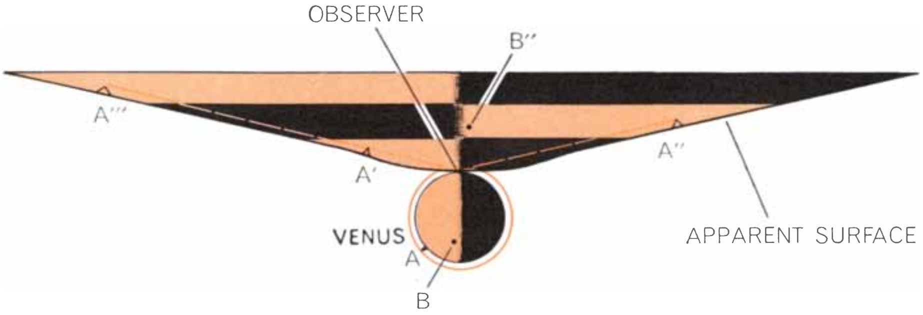


VENUS' ATMOSPHERE, studied with the aid of *Mariner V*, turned out to have a pressure about 100 times that of the earth and some 10,000 times that of Mars. Below an altitude of about 35 kilometers the atmosphere is super-refractive, which means that a ray of light would follow an arc that is sharper than the curvature of the planet. *Mariner V*'s antenna beam was tilted to help compensate for the ray curvature expected in the dense atmosphere. Tape recordings of the signals are still being searched by computer for evidence of the reflected ray that presumably bounced off the planet to demonstrate a possible way to study the super-refractive region.

altitude of about 35 kilometers the ray would curve so much that it would circle the planet and return to its starting point. Below this critical level Venus' atmosphere is super-refractive, which means that the arcs of rays bend more sharply than the arc of the surface of the planet. Hence the radio-occultation measurements of the atmosphere can be made only down to the region of super-refraction. An observer on a planet that had a transparent, super-refractive atmosphere could in principle see the entire surface of the planet spread out around him as if he were on the bottom of a bowl [*see illustration on next page*].

How can these lenslike effects on the radio waves passing through the atmospheres of Mars and Venus be interpreted in terms of density, temperature and pressure? Since the measured effects are the bending of radio rays and apparent changes in distance, the crucial property of the atmospheric gas is its refractive index. This value is inversely proportional to the phase velocity of light through the gas, and it controls ray-bending just as it does in a glass lens. Because the refractive index of a gas is very nearly unity, it is convenient to use the derived quantity called refractivity, which is defined as a million times the difference between the refractive index and unity. Of course the measured refractivity is averaged through almost tangential paths through the atmosphere. Determining the refractivity as a function of height in the atmosphere requires that the sequence of measurements be mathematically inverted. Techniques for solving this complex problem on a digital computer have been worked out by my colleague Fjeldbo. For Mars the maximum atmospheric refractivity turns out to be about four, that is, the refractive index differs from unity by only four parts in a million.

The refractivity of most un-ionized gases is equal to the number of molecules per unit volume (called the molecular number density) times a quantity that is characteristic of that gas. For a gas mixture it is the sum of such products for the individual constituents. Therefore if the constituents and their relative abundances are known, a measure of refractivity directly yields the molecular number density, which in turn is proportional to the ratio of pressure to temperature. Then the refractivity profile, or change of refractivity with altitude, is proportional to the molecular number density profile. In a mixed planetary atmosphere the number density will decrease exponentially with increasing height, falling to 37 percent of the initial value over a height difference called the scale height. The scale height, in turn, is directly pro-



SUPER-REFRACTIVE ATMOSPHERE would persuade an observer on Venus that he was living at the bottom of a gigantic bowl, illustrated here in cross section. If the atmosphere were perfectly clear, he would see the entire surface of his planet endlessly repeated in expanding rings. Because of the bending of light rays an object at position *A* would seem to lie at *A′*, *A″* and *A‴*. Similarly, an object at *B* would appear in the second ring, at *B″*. Its appearances at *B′* and *B‴* cannot be shown because they lie in the missing front half of the bowl.

portional to the gas temperature and inversely proportional to the mean molecular mass of the mixture and the known planetary gravity. If the pressure-temperature ratio is found from the local refractivity and the temperature from the profile, then both pressure and temperature can be determined for various heights in the atmosphere.

The foregoing assumes that the atmospheric composition is known. Alternatively, if other information is available about temperature or pressure at some point in the atmosphere, then some inferences about constituents can be made, but a unique solution is not possible. Fortunately the composition of the atmospheres of Venus and Mars now appears to be known well enough so that pressure and temperature can be derived from occultation measurements with comparatively little uncertainty.

Spectroscopic observations made from the earth before the Mariner flights had indicated that carbon dioxide is a constituent of both atmospheres, but the relative abundance had not been established. For Mars the approximate amount of carbon dioxide and the total refractivity combine to indicate that there is no room for much else—carbon dioxide must be the principal constituent. New observations made from the earth are in agreement. In addition the temperature derived by assuming that Mars's atmosphere consists entirely of carbon dioxide indicates little leeway for lighter constituents; for example, if there were as much as 20 percent nitrogen, the temperature at the southern hemisphere point measured by *Mariner IV* would have been at the freezing point of carbon dioxide. In that case some of the carbon dioxide in the atmosphere should have turned to a snow of "dry ice" and fallen to the surface. Robert B. Leighton and Bruce C. Murray of the California Institute of Technology have presented strong evidence that the polar caps of Mars may consist predominantly of dry ice, with the atmosphere and the caps exchanging carbon dioxide on a seasonal basis. Inasmuch as *Mariner IV* flew over an area well away from the southern polar cap, the atmosphere it measured should have been somewhat above the freezing point of carbon dioxide.

Venera 4 directly measured the carbon dioxide content of Venus' atmosphere and found it to be 90 ± 10 percent, a value that is in general agreement with less direct measurements made with earth-based instruments and by *Mariner V*. *Venera 4* carried two instruments for measuring nitrogen, and both gave negative results; the more sensitive instrument could have detected nitrogen if as little as 2.5 percent were present. The probes to Mars and Venus have overturned the long-held belief that, by analogy with the earth, nitrogen must be the predominant atmospheric constituent on both planets.

The atmospheric-pressure profiles for Mars and Venus show that the surface pressure of Mars is roughly a hundredth of the earth's and that the surface pressure of Venus is 100 times greater than the earth's. The actual measurements at the occultation points on Mars give about .45 and .8 percent of the earth's pressure, but this difference is doubtless due to differences in the height of the particular surface features that cut off the lowest radio path. Radar studies by Gordon H. Pettengill of the Lincoln Laboratory of the Massachusetts Institute of Technology show height variations of at least 11 kilometers in Mars's topography. Hence the atmospheric pressure at the surface must vary at least by a factor of three for different locations on Mars, with the occultation measurements of pressure probably being somewhat below the average.

The pressure curves for Mars, the earth and Venus are remarkably parallel [*see illustration on opposite page*]. The slope of a pressure curve would be steeper with higher temperature, with lower gravity and with a lower mass of the molecules in the gas. The earth and Venus, with nearly the same gravity but different temperatures, still have parallel pressure lines because the average mass of the molecules in the atmosphere of Venus is greater. In comparing the earth and Mars one finds that Mars's lower gravity (only 39 percent that of the earth) is counteracted by the higher molecular mass of its atmosphere and by its lower temperature, keeping the pressure curves approximately parallel.

The temperature of the atmosphere near the surface of Mars is approximately what one would predict on the basis of its distance from the sun [*see illustration on page 86*]. If Mars, Venus and the earth were black spheres without atmospheres, they would reach equilibrium temperatures in which the amount of solar radiation they absorbed would be exactly matched by the amount of heat they radiated into space at infrared wavelengths. Temperature would depend only on distance from the sun. Venus would be 330 degrees Kelvin (degrees C. above absolute zero), the earth would be 280 degrees K. and Mars 225 degrees K. (On the Fahrenheit scale these three temperatures would be 134 degrees, 45 degrees and −54 degrees.) *Mariner IV* provided two values for the lower air temperature on Mars: 160 degrees K. and 235 degrees K. The first figure applies to a winter day at 50 degrees south latitude; the second, to a summer night at a latitude of 60 degrees north.

On Venus the temperature in the lower atmosphere far exceeds the temperature of a simple black body at Venus' distance from the sun. There is still considerable uncertainty about the temperature below an altitude of 25 kilometers. The Russians originally reported that *Venera 4* reached the surface of Venus, where it recorded a pressure of 20 atmospheres (roughly 300 pounds per square inch) and a temperature of 545 degrees K. It now appears that these figures must apply to the region about 25 kilometers above the surface. The figures for the lower atmosphere are determined from extrapolation, for which two different assumptions can be made. One is an "adiabatic" extrapolation, in which it is assumed that the measured rate of change in temperature from 57 kilometers to 25 kilometers continues all the way to the surface of the planet. This is about the maximum temperature gradient possible in the atmosphere of Venus. If the lower region were hotter than this value, the atmosphere would be unstable: the lower, hotter gas would tend to rise through the higher, cooler gas, redistributing the heat energy to re-create an adiabatic temperature gradient at a higher absolute value of temperature.

The other assumption, which I consider the less plausible, is that the rate of change in temperature does not continue all the way to the surface, so that the surface temperature is no higher than about 600 degrees K. In any event, it follows that the temperature near the surface is within 100 degrees of 700 degrees K. Venus, unlike Mars, appears to exhibit little daily or seasonal temperature change and is probably very hot even at the poles. Measurement and theory also indicate that the surface of Venus does not vary greatly in altitude, perhaps less than a few kilometers

It was known soon after the *Venera 4* and *Mariner V* measurements were made that this evidence could not be reconciled with radar determinations of the radius of Venus without a major change in interpretation of at least one of the three sets of measurements. Comparing the atmospheric measurements made by the two spacecraft establishes the connection between the *Mariner V* scale, which is distance from Venus' center of mass, and the *Venera 4* scale, which is purported to be height above the solid surface. From this comparison it follows that the radius of the planet on the *Venera 4* scale would be about 25 kilometers greater than the value determined by earth-based radar.

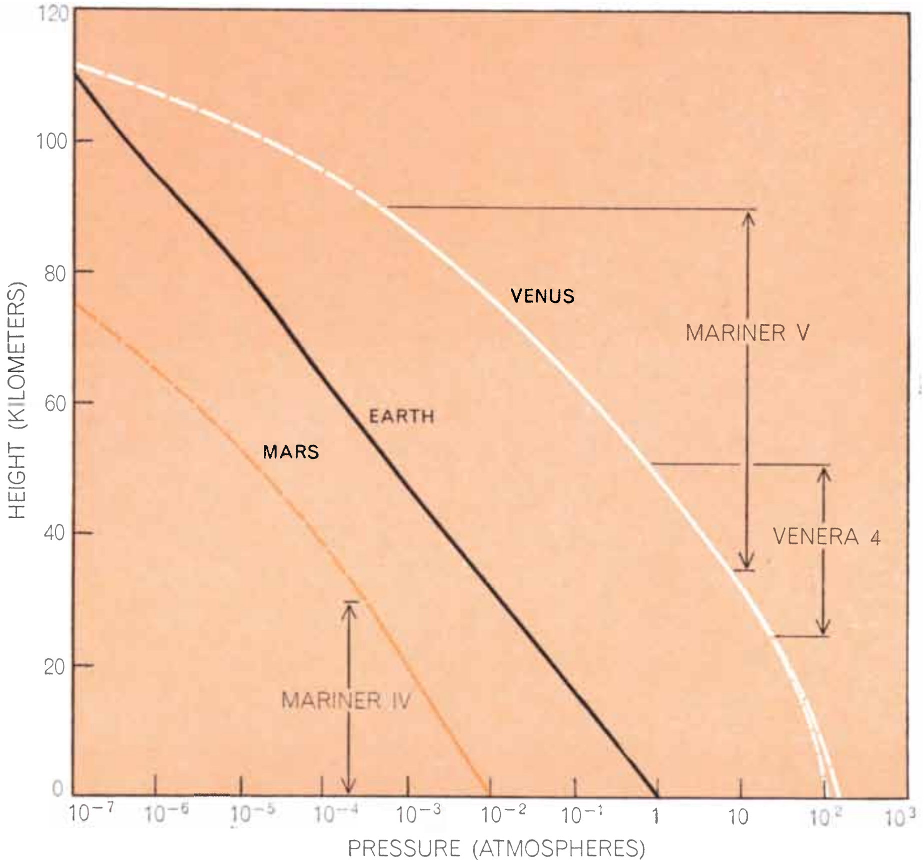
This contradiction was troublesome because it seemed that the radar-determined radius was accurate and that the *Venera 4* data were self-consistent. On the other hand, either could be in error; determining the radius on the basis of radar data alone is quite involved and the *Venera 4* interpretation depends entirely on a single altimeter report. Although a number of other arguments could be brought to bear on the problem, what was really needed was an independent measurement either of the radius of the planet or of the atmospheric properties near the surface.

Such a measurement (for the radius) proved to be feasible by combining the results from the radio tracking of *Mariner V* with measurements of the radar distance to the surface of Venus, which had been made simultaneously [*see illustration on page 87*]. When the spacecraft was near Venus, its trajectory was influenced primarily by the planet's center of mass; hence the radio-tracking data can be used to establish the distance from the earth to the center of Venus. A simultaneous radar echo yields the distance from the earth to the surface of Venus. The difference between these two distances is thus a relatively direct measurement of the radius of Venus. This measurement is largely independent of the complexities of the experiments conducted with radar alone, which involve the simultaneous solution for some 20 parameters associated with the orbits of the earth and Venus, along with perturbations due to other planets. The radar-echo distance to the surface of Venus and the radio-tracking data for calculating the distance to the center of Venus provided a value for the planet's radius that agreed closely with the earlier radar measurements: 6,053 ± 5 kilometers.

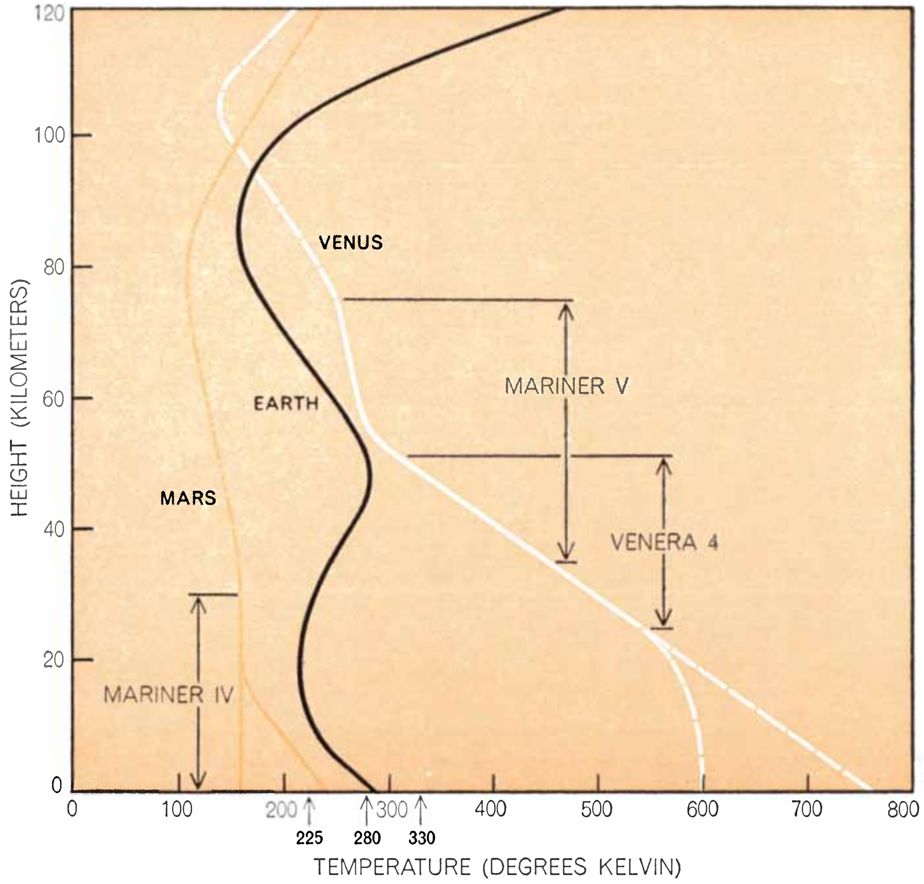
What could have gone wrong with *Venera 4*? Its atmospheric measurements are compatible with the radar and *Mariner V* data only if they actually began 50 kilometers above the surface and ended about 25 kilometers above it. The Russians have stated that the aim of *Venera 4* was not necessarily to sample conditions all the way to the surface but rather to penetrate deep into its atmosphere—to a level where conditions were approximately those of the actual measurements. Nevertheless, they initially concluded that *Venera 4* did reach the surface on the basis of (1) a single altimeter reading, apparently of about 25 kilometers, at the start of the spacecraft's parachute descent and (2) the subsequent elapsed time to the termination of its signal, during which time it must have fallen about 25 kilometers.

Suppose, however, that the initial altitude were 50 kilometers instead of 25. Then all the other data would be in excellent agreement, with *Venera 4* penetrating about as far as it was designed to go and giving its last report at an altitude of 25 kilometers. The logic of this explanation is consistent with an ambiguity in the interpretation of the radio-altimeter reading. It happens that certain simple systems for indicating a particular altitude (say 25 kilometers) are unable to distinguish heights that are exact multiples of the intended altitude. We do not know, however, if this kind of system was used on *Venera 4*.

The temperature of Venus is of partic-



ATMOSPHERIC PRESSURE on the earth is about 100 times that on Mars and the pressure on Venus is about 100 times that on the earth. The values for Venus provided by *Venera 4* and *Mariner V* are in excellent agreement in the region between 35 and 50 kilometers, where the two craft made overlapping measurements. The broken lines are extrapolations.



ATMOSPHERIC TEMPERATURE near the surface of Mars is about two-thirds that on the earth and only about a third that on Venus, as measured on the Kelvin, or absolute, scale. The three values at the bottom of the chart (225, 280 and 330 degrees) are the theoretical temperatures expected at the surface of the planets if they behaved as ideal black bodies in space. The two legs of the actual curve for Mars represent values at different seasons and latitudes as measured by *Mariner IV*. The two legs of the temperature curve for Venus represent different extrapolations for the lowest 25 kilometers of the atmosphere.

ular interest. Why is the planet approximately twice as hot as the black-body equilibrium temperature of 330 degrees K.? Venus rejects (by reflection) about 76 percent of the radiant energy it receives from the sun; this effect alone should make it 100 degrees cooler than the equilibrium value. It would seem that the high temperature of the planet can be explained only in terms of either an extremely strong internal source of heat or an extremely efficient trapping of the infrared radiation that otherwise would carry away the heat energy.

The hypothesis favored at present is that solar radiation is trapped by the "greenhouse" effect. Carl Sagan of Cornell University and James Pollack of the Smithsonian Astrophysical Observatory have studied this problem in detail. Their view, in brief, is that the carbon dioxide and water vapor in the atmosphere of Venus act as a highly efficient glass in the greenhouse, allowing the energy of the sun's visible light to enter and be deposited throughout the lower atmosphere, perhaps down to and including the surface. The energy that is reradiated at infrared wavelengths finds the atmosphere so nearly opaque that the temperature must rise to the point where enough heat is radiated, in spite of this opacity, to balance the incoming energy. The higher the infrared opacity is, the higher the temperature must go before this balance is reached. For the earth and Mars the amount of solar energy lost by reflection from clouds and from the surface of the planet just about balances the weak trapping of infrared in their atmospheres, but for Venus the greenhouse effect greatly predominates and is the determining factor in its temperature.

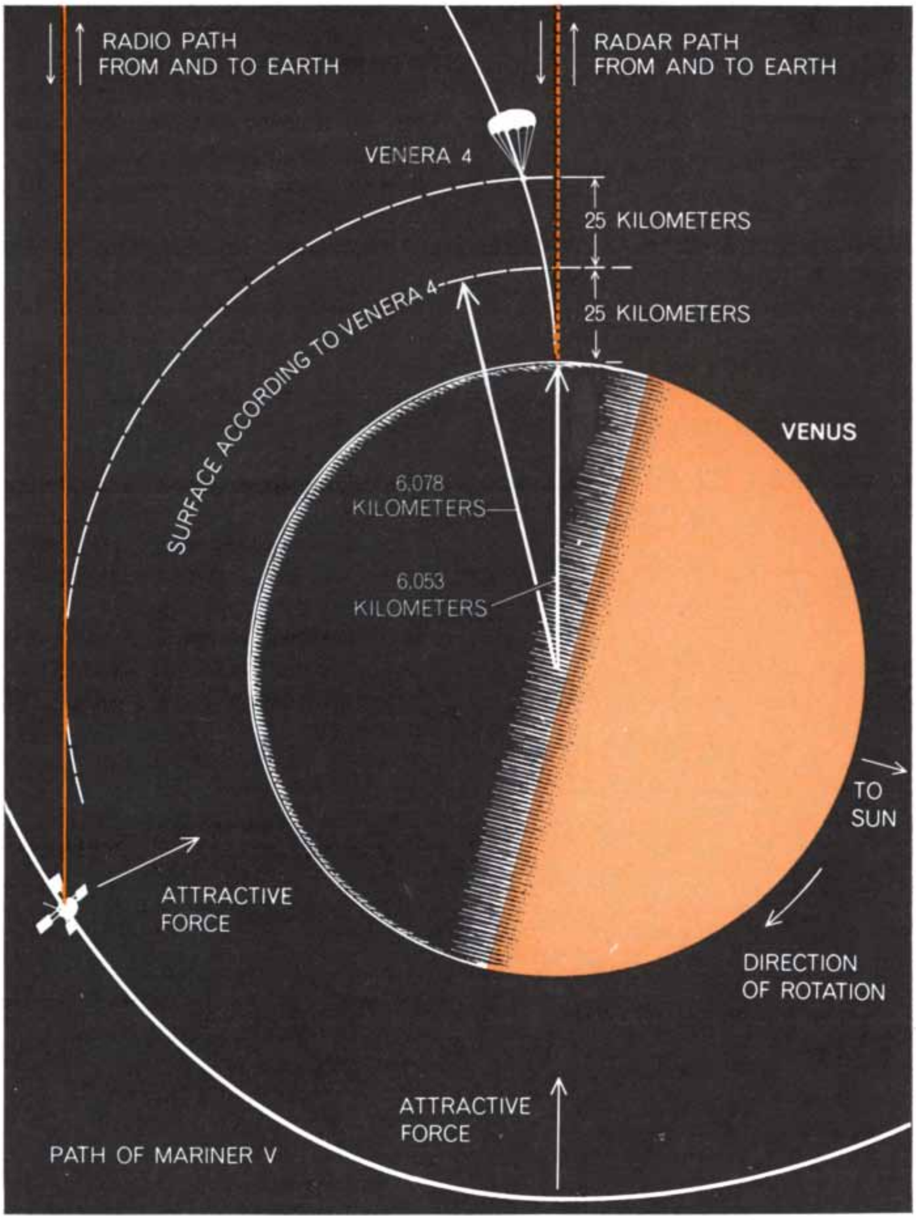
The constitution of the clouds and the amount of atmospheric water are particularly relevant to our understanding the temperature of Venus. *Venera 4* indicated that below the clouds the water content is between .1 and .7 percent, but the reliability of this measurement has been contested. Earth-based spectroscopic observations indicate a very dry atmosphere above the clouds, and there is controversy over whether or not clouds of ice particles are compatible with these observations. Among the alternatives that have been suggested are clouds of dust, carbon suboxide (C_3O_2) and compounds of mercury and chlorine. If the *Venera 4* water measurement is accepted, then it would seem that the clouds are ice crystals, with a lower boundary some 60 kilometers above the surface.

I am struck by the sudden break in the temperature profile near this height as measured by *Mariner V* on both the day and the night side of Venus. The temperature at this break is near the freezing point of liquid water, if one assumes that the atmosphere is almost entirely carbon dioxide. The break suggests to me that the clouds may be a mixture of ice and water, in which case the percentage of water vapor condensing and sublimating at the cloud level should be even higher than the limit established by *Venera 4* at lower altitudes. If raindrops formed in such a cloud, they would not fall to the surface but would vaporize at a height of about 38 kilometers. Alternatively the drops might consist not of pure water but of acid; hydrogen chloride and hydrogen fluoride have been detected in the atmosphere.

Even if one accepts the presence of extensive water clouds on Venus, it is a mystery why the planet should have so much less surface and atmospheric water than the earth. If one assumes that Venus and the earth have exhaled from their interior about the same amount of water since their formation, this means that Venus ultimately lost to space an amount of water comparable to the quantity in the oceans of the earth, while the earth lost essentially none. The upward flow of water vapor in an atmosphere is controlled by a "cold trap," an altitude of minimum temperature; when the vapor reaches this altitude, much of it may condense or freeze, and then fall. The water vapor that reaches higher altitudes is dissociated into hydrogen and oxygen by solar radiation; the free hydrogen may be lost by escape into space and the free oxygen by being bound into compounds, notably in the oxidation of surface materials (such as iron) and of atmospheric constituents (such as carbon monoxide and methane). If atoms in the atmosphere of Venus are ionized by being stripped of electrons by solar radiation, their escape from the top of the atmosphere may be strongly affected by Venus' lack of a magnetic field, which allows the charged particles of the "solar wind" to interact more directly with Venus' atmosphere than they do with the earth's. Clues to the characteristics of such interactions were obtained from three experiments on *Mariner V.* One dealt with solar-wind particles, another with magnetic fields and the third was the two-frequency occultation experiment.

Unfortunately there is disagreement on the magnitudes of the various effects. Some investigators have concluded that Venus could have lost an amount of water equal to the quantity in the earth's oceans over the past 4.5 billion years; others have concluded that it could not have done so. If the second conclusion is correct, it may be that Venus originally condensed from materials poorer in water than those that formed the earth. An interesting alternative to this assumption has recently been outlined by Peter E. Fricker and Ray T. Reynolds of the Ames Research Center of the National Aeronautics and Space Administration. They suggest that Venus still has much of its original water in solution in subsurface molten rock. Because Venus was presumably always hotter than the earth, much of the silica rock at its surface was probably molten over most of geologic time. The early characteristics of the atmosphere would therefore be determined primarily by the relative solubility of water and carbon dioxide in the silica melt. One would expect water to be much the more soluble of the two, with the result that carbon dioxide would have been driven off to form the bulk of the atmosphere. When the surface temperature cooled to the present value of about 700 degrees K., the silica would finally form a thin crust, but one so plastic that it would tend to prevent molten material from escaping to the surface. Most of the small amount of water initially in the atmosphere eventually escaped into space, but the major part of the planet's initial supply of water may remain trapped below the surface. The cooler earth, by comparison, has long had a thick, fractured crust through which molten material was able to escape to the surface. As the molten volcanic effluents cooled, essentially all their water and carbon dioxide entered the atmosphere; eventually most of the water condensed to form the oceans and the carbon dioxide became stored in carbonate rocks near the surface.

Of what use is this intense effort to understand inhospitable environments on our neighboring planets, apart from man's urge to explore the unknown? I think it is clear that we gain a new understanding of our terrestrial environment from these initial but highly informative measurements of other atmospheres. It has often been pointed out that biologists have only one planet as an example of the rules governing the origin and evolution of life, whereas the physical scientist has abundant evidence that his physics and chemistry are applicable in the most distant galaxies. Nevertheless, the origin and evolution of a planetary atmosphere such as our own are so very complex and may be so finely tuned to small changes in temperature, pressure and chemical composition that the physical rules are not sufficient to bestow full understanding. We need examples of other atmospheres to guide our thinking and to throw new light on our attempts to understand puzzling terrestrial phenomena. It has also been suggested that the space techniques developed for studying Mars and Venus could be used to answer important questions about the earth. For example, Bruce B. Lusignan of Stanford and several others have proposed that radio occultation ex-



DISCREPANCY IN *VENERA 4* AND *MARINER V* DATA for atmospheric pressure can be resolved if the Russian craft's readings apply to the region between 50 and 25 kilometers rather than to the region between 25 kilometers and the surface, as the Russians reported. *Venera 4*'s data began with a single altimeter reading indicating an altitude of 25 kilometers. The final pressure reading, presumably at the surface, showed the pressure to be only about 20 times that on the earth, whereas *Mariner V*'s data show the surface pressure to be about 100 times the earth's. If *Venera 4*'s last reading was indeed at the surface, the radius of Venus would have to be some 25 kilometers greater than determined previously by radar measurements from the earth: 6,078 kilometers rather than 6,053. The text explains how simultaneous radio-range data for *Mariner V* and radar-range data for the distance to the surface of Venus confirmed the smaller value for the radius and support the hypothesis that *Venera 4*'s last readings were actually made at an altitude of 25 kilometers. *Venera 4*'s initial reading then agrees with *Mariner V*'s value for the pressure at 50 kilometers.

periments be carried out by satellites in orbit around the earth to provide atmospheric measurements over inaccessible regions so as to serve as a powerful aid in weather prediction.

In a more speculative vein, Venus poses an intriguing challenge for planetary engineering. If the thesis of Fricker and Reynolds is correct, one might be able to create oceans on Venus by finding a way to lower its temperature. Sagan has made an imaginative suggestion for achieving the cooling and at the same time producing atmospheric oxygen. He suggests that microorganisms be introduced into the cool high-level clouds, where they would break down the carbon dioxide in two steps into carbon and oxygen. As oxygen replaced carbon dioxide in the atmosphere the greenhouse effect would diminish, lowering the temperature to the point where Venus might eventually become suitable for human colonization.

The space probes to Mars and Venus have emphasized, I think, that biological processes themselves may play a more central role in the characteristics of planetary atmospheres than we have thought. James E. Lovelock and Charles E. Giffin of the Jet Propulsion Laboratory have advanced the view that measuring atmospheric characteristics of Venus and Mars, including minor constituents, would actually be a powerful experiment for the detection of extraterrestrial life. If these planets support life of any kind, it should develop to the limit set by energy sources and raw materials, with the atmosphere most likely acting as a medium for transferring the products of the life cycle. Lovelock and Giffin hold that this role would so modify the composition of the atmosphere that it would be recognizably different from a composition that had resulted simply from the distribution of available chemical compounds in the absence of life.

A clearer view of the earth's ecosystem—the totality of living organisms and their manifold interactions with their environment—could result from the pursuit of such a thesis. For example, the amount of oxygen, nitrogen, carbon dioxide, methane and nitrous oxide in the earth's atmosphere may be governed by life on the earth. The biological sources and sinks maintain a large and balanced flow of oxygen, carbon dioxide and nitrogen into and out of the atmosphere. The abiological sources and sinks of these atmospheric constituents are weaker than the biological ones by factors of 10 to a million.

In the absence of life on the earth the abiological sources and sinks would be controlling. The abiological sinks for oxygen (oxidation of compounds on the earth's surface and oxidation of atmospheric nitrogen and methane) and the abiological sinks for nitrogen (incorporation into stable nitrate compounds) are apparently stronger than the abiological sources, so that both the oxygen and the nitrogen in the atmosphere should slowly decrease to only trace levels. For carbon dioxide the strengths of abiological sources and sinks appear to be approximately equal, making it difficult to predict the steady-state concentration. If the amount of carbon dioxide in the atmosphere slowly increased, a critical level might be reached where the enhanced greenhouse effect would raise the temperature and drive more water vapor and carbon dioxide into the atmosphere; this would further enhance the greenhouse effect and further raise the temperature. This runaway greenhouse effect would end only when the earth resembled Venus in temperature and atmospheric density. If, on the other hand, the amount of carbon dioxide stabilized at a low level, the lifeless earth would have an atmosphere more like that of Mars. On this basis it can be argued that our first life-detection experiment has already been conducted. The tentative conclusion is that on Mars and Venus life is absent.

It would seem that our relative ignorance of the possible global effects of small changes in the composition and temperature of the earth's atmosphere should give us serious and immediate concern about man's use of the atmosphere as a garbage dump. In the past 100 years the percentage of atmospheric carbon dioxide has increased 15 percent, primarily owing to man's burning of fossil fuels. This same activity now uses 15 percent of the biological production of oxygen and in the future could become the major sink. Smog, jet contrails, deforestation and pollution of the earth's waters may affect the constituents and temperature of the atmosphere by only minute amounts, yet they may nonetheless have far-reaching consequences. Let us hope that continuing studies of the earth and its planetary neighbors will lead rapidly to a fuller understanding of planetary atmospheres in general. At the same time it behooves us to take strong steps to eliminate the known harmful effects of man's activities and to minimize those effects that are not yet fully understood. Whatever else we may wish to know, we hardly wish to test whether a lifeless earth would become like Mars or like Venus.

We'll go that extra mile.

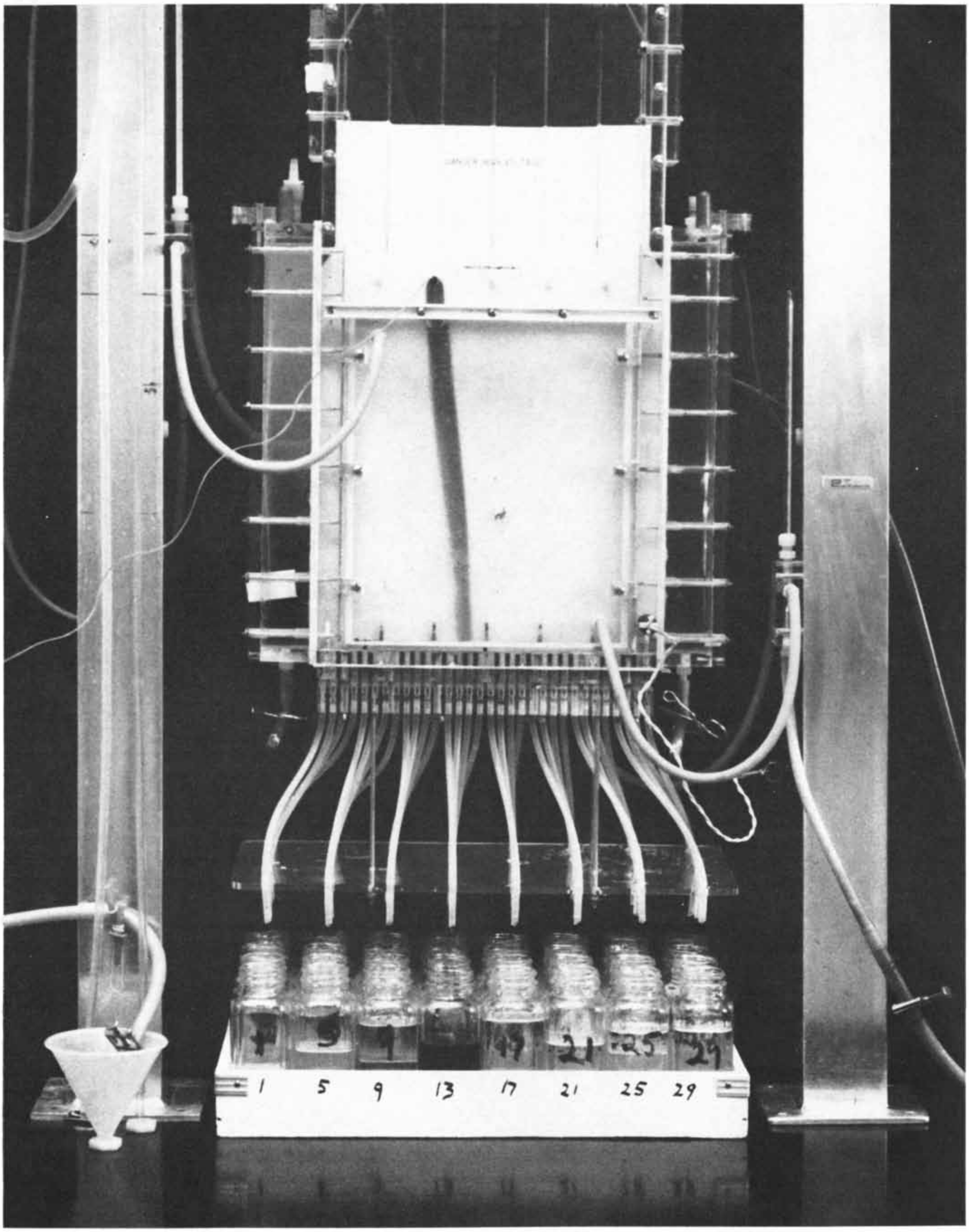
- We feature the cars you prefer—GM.
- We give you S & H Green Stamps.
- We offer you 1600 locations.
- We accept recognized credit cards.

Besides, we're the first in the industry to guarantee reservations through a completely computerized system. And we maintain the best conditioned fleet in the business. No wonder we grew twice as fast as our competitors last year and are now No. 2 in locations. If you're looking for someone who'll take a few extra steps to please and serve, try National.

Heart and sole,
we're in the
customer's shoes.

The customer is always No. 1

NATIONAL CAR RENTAL



PLAGUE TOXIN WAS PURIFIED by continuous-flow electrophoresis, which separates the fractions of a protein mixture. The path of the toxic protein mixture from plague bacilli is shown here by means of a dye marker (*dark stripe*). The mixture, fed into the apparatus through a tube, encounters a medium consisting of tiny glass beads packed between two sheets of plastic (*light rectangular area*). This fractionating plate is connected at each side to electrodes. When a voltage is applied, the mixture spreads across the beads because each protein fraction moves at a characteristic rate toward the anode or the cathode. At the bottom of the apparatus the separated proteins enter individual tubes and are collected in bottles. The fractions are then tested to determine their toxicity.

PLAGUE TOXIN

The bacillus that caused the "Black Death" manufactures a protein that is lethal to some animals and harmless to others. How does this substance exert its effect on the molecular machinery of the cell?

by Solomon Kadis, Thomas C. Montie and Samuel J. Ajl

Plague, the "Black Death" that once (1348–1350) killed a quarter of the people in Europe, is a remarkably virulent infection. In at least one epidemic of modern times (Hong Kong, 1899) 95 percent of those who were infected died. Fortunately major outbreaks of the disease have become increasingly rare. Plague is still, however, far from being a minor medical problem. In 1967, 5,000 cases were reported in South Vietnam, and there is always the possibility that social and ecological conditions will be right for another major outbreak.

Plague is caused by the bacillus *Pasteurella pestis,* which is spread by fleas living on certain rodents. How does the bacillus cause the symptoms of the disease and death? In our work at the Albert Einstein Medical Center in Philadelphia we have approached the question by isolating a toxin produced by the bacillus in infections of mice and rats (murine plague). This has enabled us to study the effects of the toxin at the level of the cell and its molecular machinery.

Current interest in the role of toxin in plague was aroused by observations of medical workers in Madagascar (now the Malagasy Republic), where the disease is endemic. Today plague can be successfully treated with antibiotics, provided that they are administered within 36 hours after the appearance of symptoms. In some cases treated by the Madagascar workers antibiotics could only be given between 36 and 48 hours, and the patients died. When autopsies were performed on these patients, it was found that their blood and their organs were free of the bacillus. In other words, treatment with antibiotics had killed the bacilli, and the patients had apparently died from the effects of substances the bacilli had manufactured.

When we began our work, it was known that the plague bacillus contained several substances capable of stimulating the production of antibodies in human beings and experimental animals. We have focused our attention on one substance: the toxin known to cause death in mice and rats. This murine plague toxin is a protein contained within the bacterial cell. In order to learn how the toxin exerts its effects one must liberate the toxin from the cell and free it from the substances associated with it. If this is not done, of course, one can never be sure that the effects are not partly or wholly due to some other substance.

In our initial attempts at purifying the toxin we dried bacteria with acetone and extracted crude toxin with salt solution. We then subjected the extract to a series of distinct chemical procedures, each of which was designed to remove carbohydrates, nucleic acids, lipids or proteins other than the toxin. The final preparations were highly toxic to mice and lacking in any major class of chemical compounds other than protein. Still, the protein was not homogeneous, that is, one or more nontoxic proteins were associated with the toxin.

We now turned to the technique of continuous-flow paper electrophoresis, a powerful means of resolving the fractions of a protein mixture. The apparatus consists of a filter-paper curtain suspended in an electrically conducting solution and connected to two electrodes. The protein mixture is fed continuously to the filter paper by capillary attraction, and the serrated edges of the paper provide drip points for the collection of individual protein fractions. When an electric potential is applied across the paper, each protein moves toward either the anode or the cathode at a characteristic speed and is separated from the others. This happens because every protein is a unique chain of amino acids, each of which has one or more positive and negative charges (depending on the acidity or alkalinity of the solvent system). A protein with a majority of negative charges will move toward the anode; one with a majority of positive charges, toward the cathode.

The murine plague toxin we obtained in this way is lethal for mice and rats in minute quantities. The LD_{50} (the dosage required to kill half of a randomly selected population) of this toxin administered intravenously to mice is .1 microgram of protein. Sensitive immunological tests revealed that these preparations were free of any components other than the toxin. Thus we had established that the toxin was in a highly purified state. Continuous-flow paper electrophoresis, however, is rather slow; many hours or even days are needed to fractionate large volumes of dilute protein solution. Moreover, some protein is lost by adsorption on the paper. We therefore developed a continuous-flow electrophoresis apparatus in which the paper was replaced by fine glass beads packed between two sheets of plastic [*see illustration on opposite page*]. This apparatus can rapidly fractionate large volumes of material with little loss to adsorption on the glass beads. Since we needed large quantities of toxin for our studies of toxin action, this procedure proved best for our purpose. In this way we secured the highly purified material we used in most of our studies.

Although we had found the toxin protein to be homogeneous by all known standards of protein purity, it was still possible that the toxin was composed of more than one species of protein. This was suggested by the fact that certain enzymes (for example lactic dehydrogenase) are actually systems consisting of two or more molecular forms. These forms differ somewhat in structure but perform the same enzymatic functions.

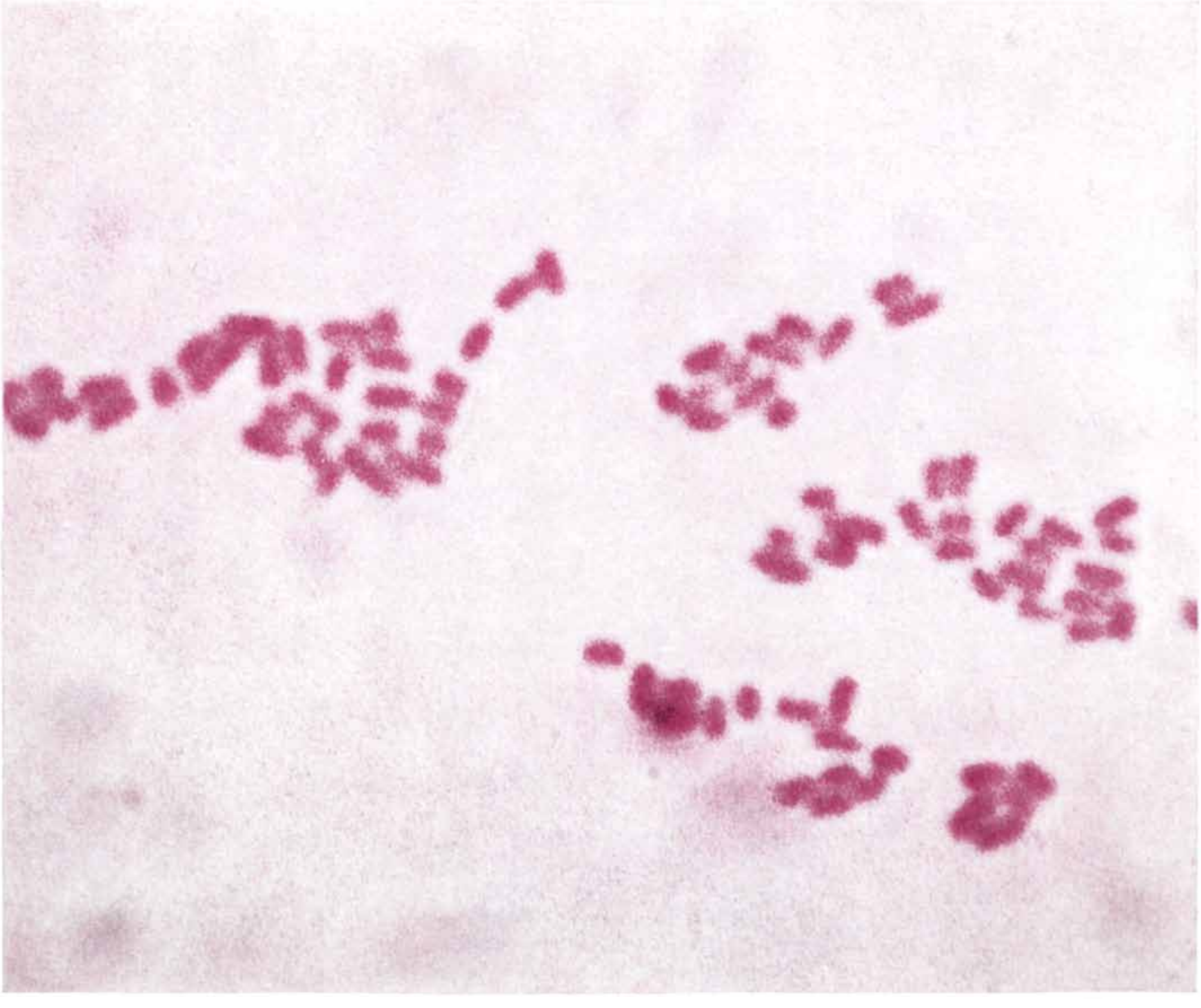
It happens that we were also working with gel electrophoresis, a third electrophoretic method that separates proteins into bands [*see illustration on opposite page*]. While we were testing this method for locating toxin in crude fractions of plague bacilli, we observed that toxin activity was associated with more than one component of the partially purified material. The purest samples exhibited two protein bands, each of which was highly toxic when it was washed out of the gel and injected into mice.

We designated these two proteins Toxin *A* and Toxin *B*. Both toxins may be located in the surface of the bacterial cell; Toxin *A* is the larger molecule of the two. It therefore appears that murine plague toxin consists of two distinct molecular species, each exhibiting identical toxic activity. Further support for this hypothesis is provided by differences in the sensitivity of the two toxins to protein reagents such as deoxycholate, digitonin, urea and various sulfhydryl compounds.

We have estimated the molecular weight of Toxin *A* and Toxin *B* by passing them through a "molecular sieve" of the polymer dextran. The rate at which the two toxins filtered through the sieve was compared with the rate for standard proteins of known molecular weight (a technique that is currently in wide use). On this basis we calculated that the molecular weight of Toxin *A* and Toxin *B* was respectively 240,000 and 120,000. When we spun the two toxins at high speed in an ultracentrifuge, we found that the rates at which they settled in the ultracentrifuge cell were in general agreement with the results we had obtained by filtration.

The two toxins, like other proteins, are composed of amino acids linked in chains by covalent bonds. Many large proteins, however, consist of two or more chains associated by bonds weaker than covalent ones. The fact that Toxin *A* is roughly twice as large as Toxin *B* suggests that Toxin *A* consists of two Toxin *B* units, and that both toxins are made up of smaller chains that closely resemble each other. If one assumes that each of these smaller chains contains one unit of the amino acid cysteine, their minimum molecular weight is about 12,000.

An important difference between the two toxins, apart from their size, is that Toxin *B* contains 30 percent less of the amino acid tryptophan than Toxin *A*. Moreover, if one calculates the number



PLAGUE BACILLI (*Pasteurella pestis*) are enlarged 7,500 diameters in this photomicrograph. The bacilli are stained pink with the Gram stain. The plague organism is transmitted from rodent to rodent (chiefly the black rat) and from rodent to man by the flea.

of tryptophan units in each chain with a molecular weight of 12,000, one finds that there are fewer tryptophans than cysteines. This suggests that certain repeating subunits of Toxin *B* are deficient in tryptophan, and seems to argue against the view that the toxins are made up of the same subunits.

We sought to test the subunit hypothesis by dissociating the toxins with chemical agents that would not break the covalent bonds between the amino acids in a chain but would disrupt the weaker bonds between the chains. Using sodium dodecyl sulfate, we succeeded in dissociating both toxins into subunits of the same size; the molecular weight of these units was between 12,000 and 24,000. When the subunit material was injected into mice, its LD_{50} was 60 percent of the parent toxin. This confirmed the view that the toxins are built up from component chains of the same size. The difference in tryptophan content probably reflects the presence or absence of a tryptophan in certain chains.

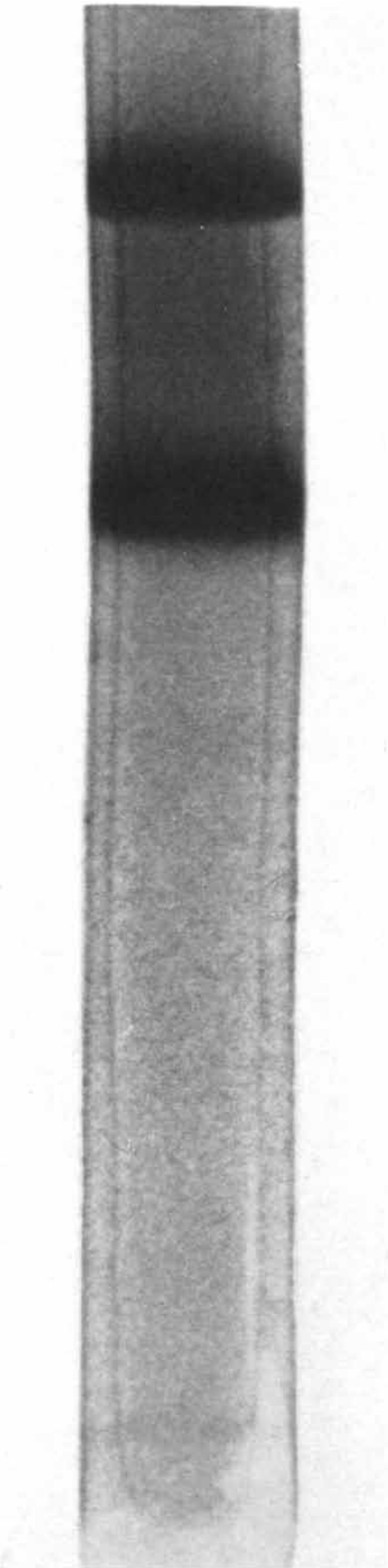
It is difficult to understand why a bacterial cell should manufacture two closely related substances with the same apparent biological effect. It may be that these toxins are pieces of protein that are released when the cell breaks down. Toxicity is a broad phenomenon, and undoubtedly a number of biological and biochemical processes are involved in it.

We now turn to what we have learned about the mechanism of action of murine plague toxin. As we have mentioned, our chief aim was to study this mechanism at the molecular level. We had made progress in experiments conducted outside the living animal, but it was also desirable to examine the mechanism at a higher level—one closer to what might be expected to happen within the animal. Experiments at such a level can still be conducted, however, outside the animal. We undertook to find out how the toxin affects mitochondria, the tiny bodies within the cell that are the active sites of respiration in animal tissues.

It is known that, whereas mice and rats are killed by small amounts of murine plague toxin, other animals (including rabbits, dogs, monkeys and chimpanzees) are not. As we began our work on the action of the toxin, we made an observation that may do much to explain the differences among species in susceptibility to biological poisons. We found an association between the toxin's ability to inhibit the respiration of mitochondria isolated from certain species and the susceptibility of the living animal. The toxin inhibits the respiration of mitochondria from the heart of the toxin-susceptible mouse and rat, and it has little or no effect on similar preparations from the toxin-resistant rabbit, dog, monkey and chimpanzee. Bovine serum albumin, representing another protein, and lipopolysaccharide antigens from *Escherichia coli* do not interfere with mitochondrial respiration. Moreover, inactivating the toxin by heating, with chemical reagents or by altering it in any way that makes it nontoxic to rats and mice also inactivates its ability to inhibit mitochondrial respiration. We therefore put forward the hypothesis that the inhibition of mitochondrial respiration underlies the effect of murine plague toxin in the living animal.

To test this hypothesis we made other efforts to relate the effect on mitochondria with changes in the living animal. Since heart mitochondria from toxin-susceptible animals were affected, the hypothesis suggested that an early sign of toxin action might be detected by electrocardiography. Indeed, we found that there was a change in a specific segment (the S-T segment) of the electrocardiogram of a rat within 60 minutes after the injection of lethal or sublethal doses of the toxin and before the appearance of other symptoms in the animal. In rats surviving sublethal doses of the toxin the electrocardiographic changes were not evident after 24 to 48 hours or after the animal had recovered completely. The specificity of the electrocardiographic changes caused by murine plague toxin is given additional support by the fact that similar changes did not occur in rats dying from a variety of other causes, including hemorrhagic shock, intoxication due to excess glucose in the bloodstream and toxin produced by the bacterium *Escherichia coli*. In addition the rabbit, which is not susceptible to murine plague toxin, shows no electrocardiographic abnormalities when it is injected with large doses of the toxin.

Two questions now arose. The first had to do with the site of inhibition of mitochondrial respiration by the toxin. The second concerned the differences in the structure of mitochondria isolated from different animal species that might explain the susceptibility of some animals to the toxin and the resistance of other animals. It was conceivable that the inability of the toxin to inhibit the respiration of heart mitochondria isolated from the toxin-resistant rabbit was due to the exclusion of toxin by the outer membrane of the mitochondrion. If this



TOXIN WAS SEPARATED into two molecular species with the same biological effect. The two species (*dark horizontal bands*) were identified by a type of electrophoresis that employs a gel as a supporting medium.

were the case, disrupting the membrane of the rabbit mitochondria would make them as susceptible to the toxin as rat mitochondria were. We therefore disrupted heart mitochondria from the rabbit with ultrasound or with detergents, which does not completely stop respiratory activity. When we added toxin, however, such activity decreased further. It thus seems that differences in the mitochondrial membrane may underlie differences in susceptibility to the toxin.

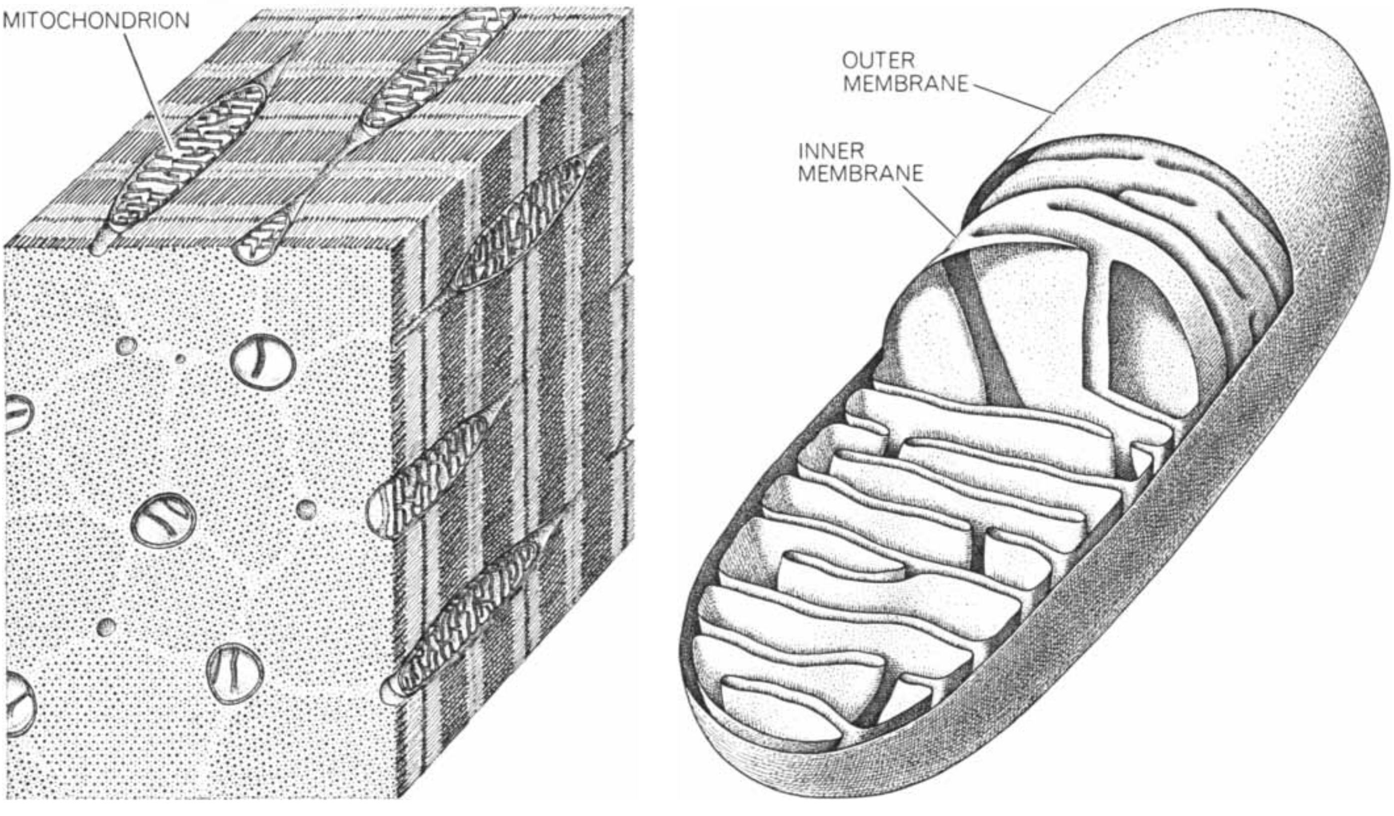
At this stage in our investigations we were ready to seek the precise point in the metabolic machinery where the toxin inhibits mitochondrial respiration. Here it will be useful to briefly review how respiration proceeds in the mitochondrion, that is, how the mitochondrion generates, stores and transforms energy.

The food living organisms eat provides fuel in the form of sugars, fats and proteins; these substances in turn provide the energy required for all the metabolic activities of cells. In the mitochondrion the fuel is oxidized step by step in a series of chemical reactions, each catalyzed by a specific enzyme. The main function of the oxidation reactions is the release of electrons, which serve as agents for the storage and transformation of energy. The electrons are transferred by enzymes through a series of cyclic oxidation-reduction reactions, each set in motion by the one preceding it. The initial electron carrier in this pathway is either succinate or DPNH (the reduced form of the coenzyme diphosphopyridine nucleotide). From this point electrons are passed along by a series of carriers until they are taken from the system by molecules of oxygen, the terminal electron acceptors. The carriers include coenzyme Q, flavoproteins and the iron-containing proteins called cytochromes; the major cytochromes are designated a, a_3, b and c. Associated with the oxidation-reduction cycles of three such carriers (DPNH and cytochromes a and b) are reactions that donate energy in the form of electrons for the synthesis of adenosine triphosphate (ATP). Each molecule of ATP has energy stored within it that is utilized by the cell as it is needed.

Our studies show that murine plague toxin does not interfere with the reactions leading immediately to the synthesis of ATP; rather, the toxin inhibits mitochondrial respiration in the electron transport system per se. Now, cytochromes are among the chief components of the electron-transport system, and they must remain in reduced form in order to transfer electrons from DPNH or succinate to oxygen. Each cytochrome absorbs light at specific wavelengths, and if one records an absorption spectrum of mitochondria in suspension and plots a graph of the light absorbed at various wavelengths, one observes definitive absorption peaks corresponding to each of the cytochrome components. Then if one adds a compound being investigated to such a mixture of enzymes and there is some change in the cytochrome absorption peaks, one gets a general indication of the compound's site of action.

We made suspensions of mitochondria from rat heart and rat liver and incubated them with and without toxin. The absorption spectra showed that the addition of toxin caused the oxidation of cytochromes a, a_3, b and c. This indicated that the toxin exerts its inhibitory effect on the electron-transport system at a point before the reaction catalyzed by cytochrome b, that is, between DPNH or succinate and cytochrome b [*see upper illustration on page 100*].

To locate the site of action of the toxin in the electron-transport system, we undertook to determine the effect of the toxin on the activity of those enzymes known to operate between DPNH or



MITOCHONDRIA, the respiratory organelles of animal tissue, are apparently the site of plague-toxin action. At left is an idealized drawing of a section of heart muscle cells. At right is a detailed drawing showing the pattern of the mitochondrial membranes.

succinate and cytochrome *b*. One such enzyme, the flavoprotein DPNH dehydrogenase, removes hydrogen atoms or their equivalent electrons from DPNH and transfers them to a flavoprotein carrier. We examined the activity of DPNH dehydrogenase in suspensions of mitochondria from rat heart and rat liver, incubated with toxin and without; we found that the activity of the enzyme was identical in both cases. It seemed as though the toxin were acting not on individual enzymes but on complexes of two or more respiratory carriers.

We therefore obtained four such complexes in highly purified form. Their role in the electron-transport system is as follows: When electrons are donated by DPNH, they are transferred by Complex I; when they are donated by succinate, they are transferred by Complex II. In either case the electrons are carried to Complex III by coenzyme Q. This results in the oxidation of DPNH or succinate and in the reduction of coenzyme Q. From Complex III the electrons are passed by cytochrome *c* to Complex IV, so that cytochrome *c* is reduced and coenzyme Q is reoxidized. Complex IV catalyzes the oxidation of reduced cytochrome *c* by molecular oxygen.

When we tested the effect of murine plague toxin on these complexes, we found that the complex predominantly inhibited is Complex I. The term "complex" here denotes a unit consisting of several components, each with one or more functions. Complex I is composed of DPNH dehydrogenase flavoprotein, coenzyme Q and nonheme iron (iron not contained in the four-sided heme ring). We considered the possibility that toxin interferes with the enzymatic activity of Complex I by blocking certain of its individual component functions. We found, however, that the addition of toxin to Complex I had no effect on its DPNH dehydrogenase activity (which confirmed the results of our similar experiments with mitochondrial suspensions). By the same token neither coenzyme Q nor nonheme iron seemed to be affected by the presence of the toxin. It therefore appeared that the toxin was acting not on any one component of Complex I but on the overall architecture of the complex.

It is known that Complex I is a constituent of the convoluted inner membrane of the mitochondrion, and that another constituent of the membrane is a nonenzymatic protein. We wondered if this protein was somehow affected by the toxin. We found that the protein and the toxin do interact; when they are mixed together, the toxin loses much of its toxicity. We are currently trying to find out if the protein and the toxin form a complex that can be isolated. If such a complex could be isolated, one might be able to identify the types of chemical bonds involved in its formation.

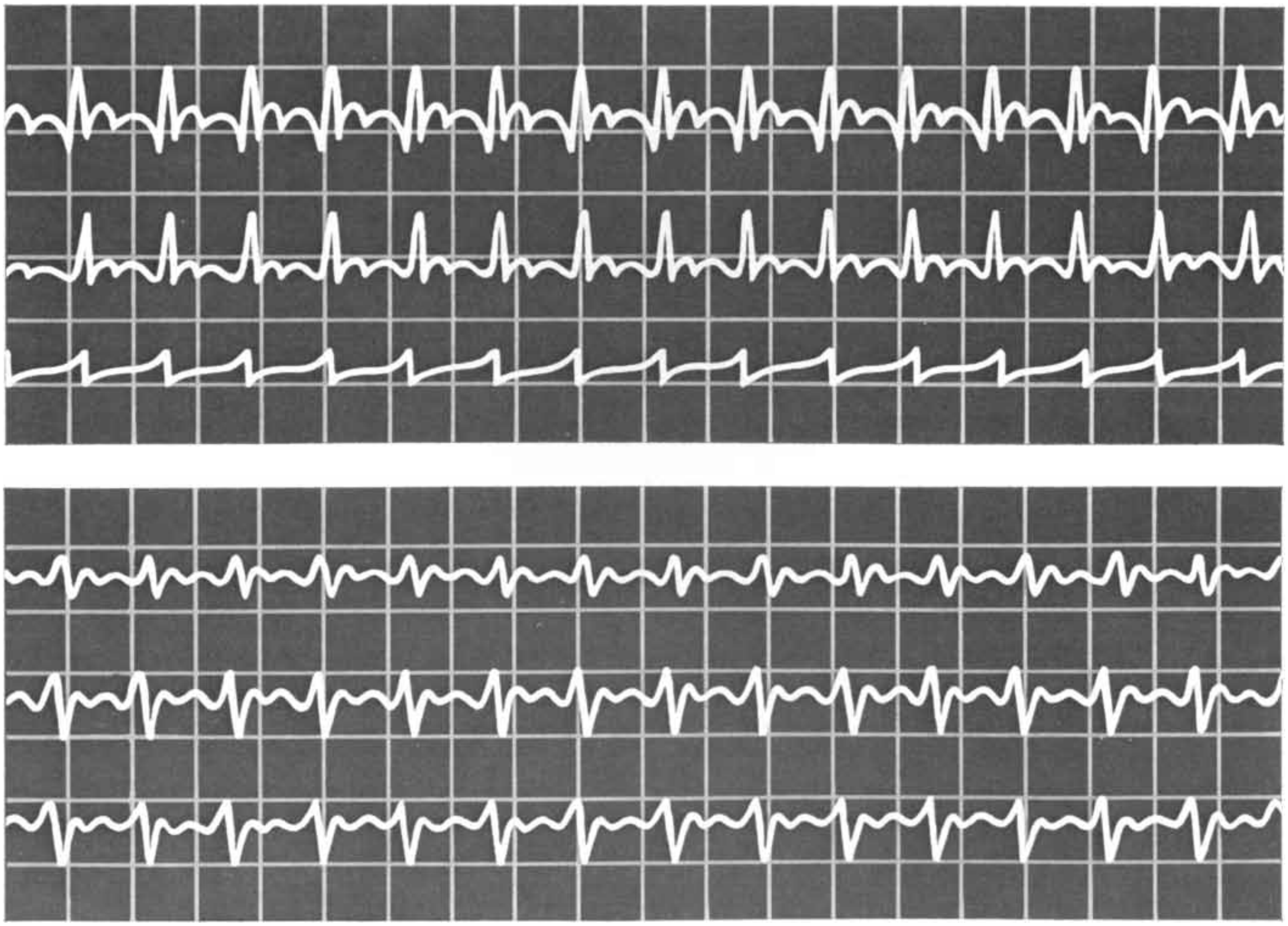
An interesting feature of the mitochondrion is that it can take in water and swell up, and also expel water and contract. These processes require respiratory energy. Many agents can induce mitochondrial swelling; they range from simple metal ions to complex compounds involved in the regulation of biological functions in normal and in pathological states. These substances fall into several categories depending on how they promote swelling. One type of swelling, known as electron-transport-dependent swelling, is believed to result from an increase in the permeability of the mitochondrial membrane.

We have mentioned that murine plague toxin does not inhibit the respiration of heart mitochondria obtained from rabbits. Since it seems likely that the outer membrane of rabbit heart mitochondria can exclude the toxin, we may have here a permeability phenomenon. We now wanted to find out if the toxin could induce mitochondrial swelling and if there was a relation between the ability of the toxin to inhibit mitochondrial respiration and any effect it might have on swelling. When we incubated rat heart mitochondria in an appropriate medium without toxin, we observed a small but consistent amount of spontaneous swelling. We then added toxin, and the amount of swelling increased considerably. The toxin did not, however, cause swelling in rabbit heart mitochondria.

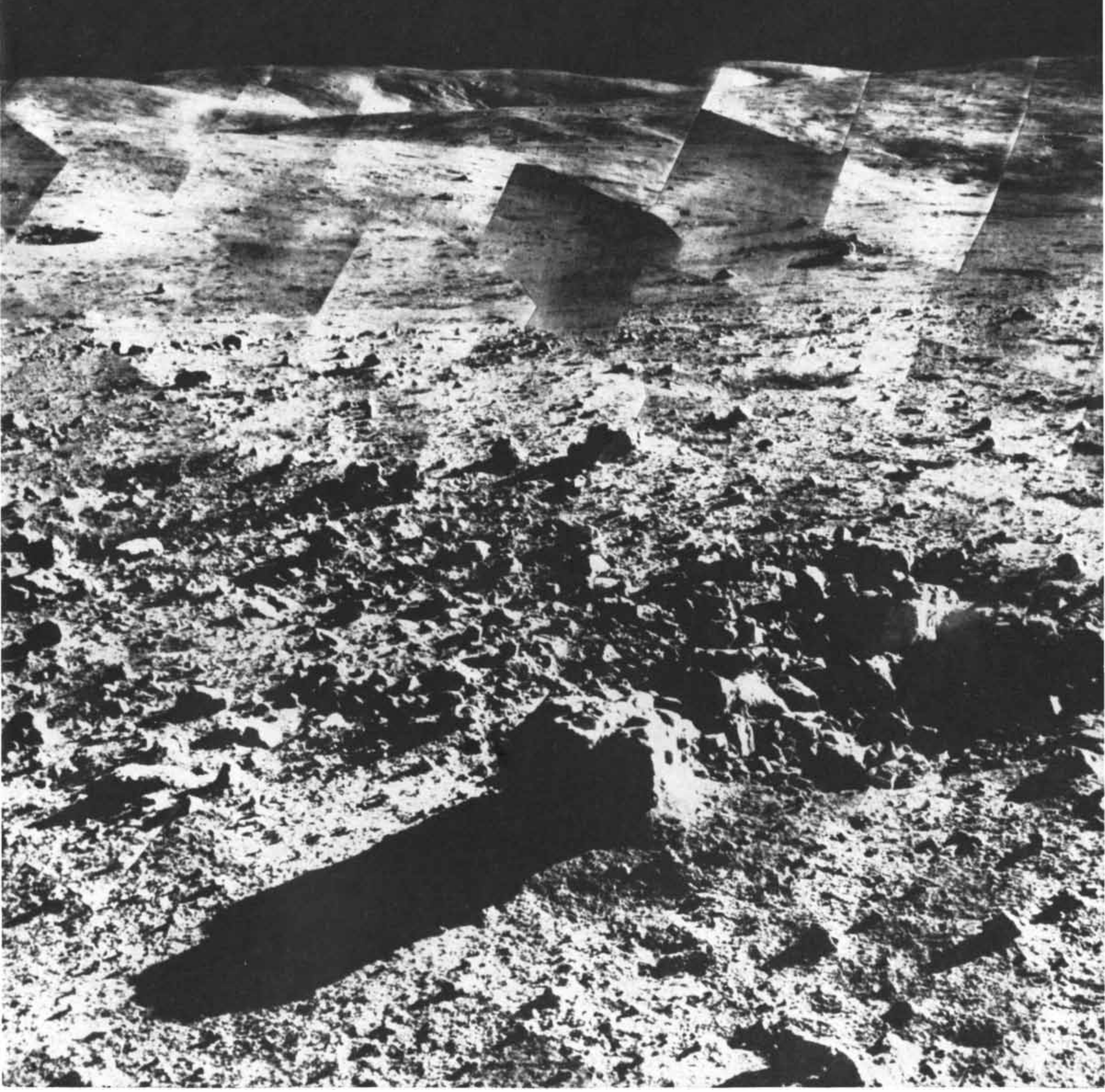
We investigated the mechanism of toxin-induced swelling by adding to the mitochondria various substances that inhibit the respiratory transport of electrons. Each of the compounds we tested (namely cyanide, azide and 2,4-dinitrophenol) prevented the swelling of rat heart mitochondria. This indicated that toxin-induced swelling is related to electron-transport-dependent swelling.

Substances other than water can penetrate mitochondria and accumulate inside them. Such substances, which play a role in the activities of the mitochondrion, include potassium, calcium, magnesium and manganese. Calcium, magnesium and manganese enter the mitochondrion as divalent ions (Ca^{++}, Mg^{++} and Mn^{++}) and are accompanied by inorganic phosphate ions. This process requires energy contributed by a high-energy intermediate similar to the intermediates that lead to the synthesis of ATP.

Other investigators have suggested that changes in the structural integrity of mitochondria may influence their ability to retain accumulated ions. During the swelling process some of the structural components of the mitochondria are altered. We accordingly undertook to determine what effect the toxin had on the accumulation of ions by mitochondria, and to relate this effect to the ability of the toxin to induce mitochondrial swelling. We reasoned that such



ELECTROCARDIOGRAPHIC TRACINGS were recorded from a normal rat (*top*) and a rat that had received an injection of plague toxin (*bottom*). The trace at bottom, indicating an abnormal heart condition, is only exhibited by animals susceptible to the toxin.



Surveyor 7 panoramic view of moon landscape. Photo courtesy of NASA.

Reach for the moon . . . but don't let go of the tail.

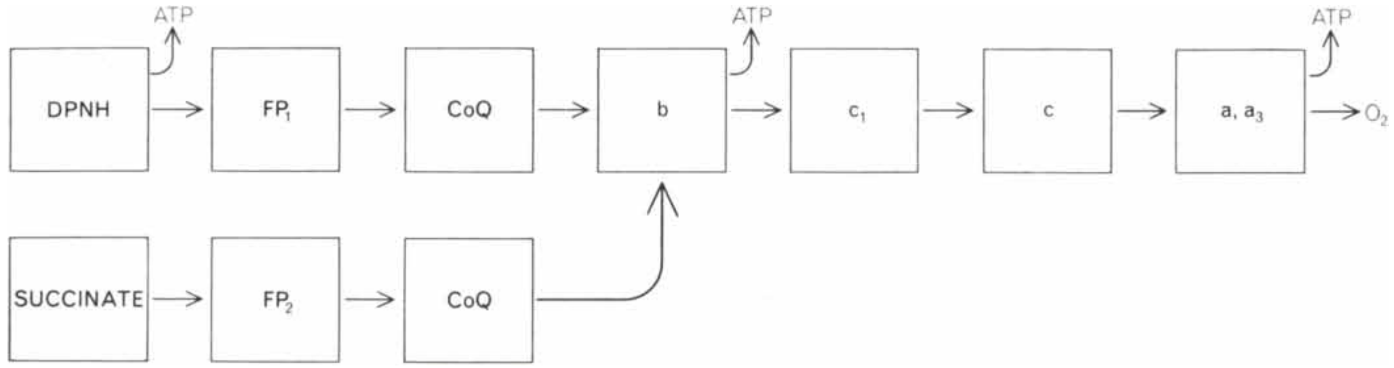
Our country has been described as having the world by the tail but about to let go of it to reach out for the moon. The description had better be wrong . . . for while putting a man on the moon is certainly important, there are matters on earth just as vital to our national security. Like putting more American ships on the sea.

Right now we are relying on foreign ships to carry nearly 95% of our imports and exports. That's a risky degree of dependence, because of 77 strategic materials needed, we must import 66. Worse, only 20% of our present merchant fleet is under 10 years old, while 80% is nearer 25. Meanwhile other nations are building more new ships than we are. Russia is taking delivery of eight ships to our one. She's outbuilding us in tonnage, too. So are thirteen other nations.

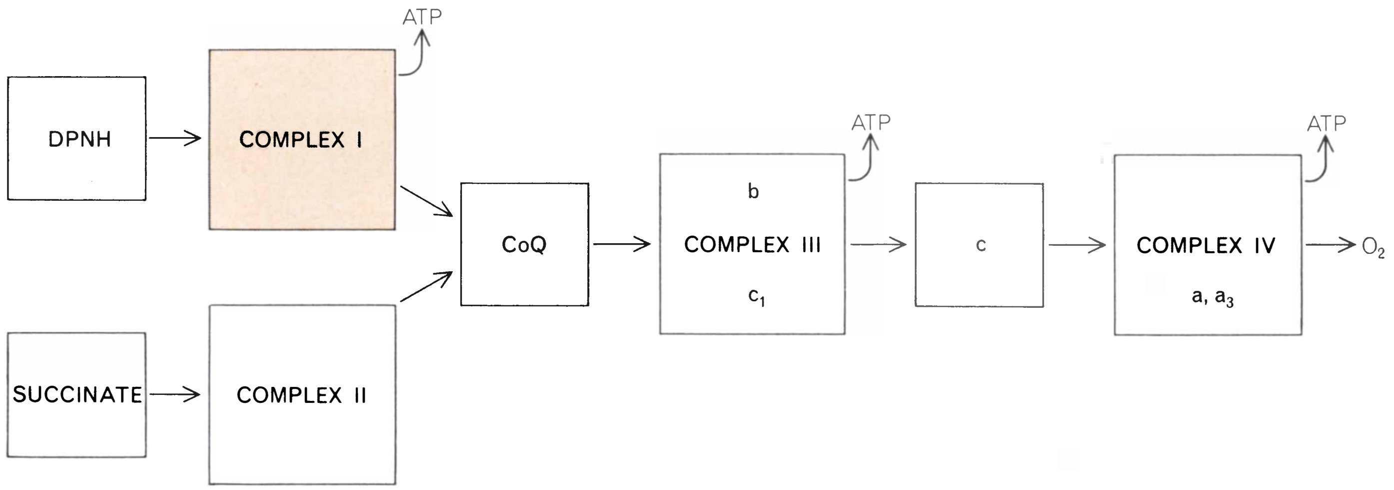
Unless we build more ships in U.S. yards, there's no guarantee we can get either ships or shipping when we might need them even more than now.

One of Todd's contributions to a stronger merchant marine was the creation last year of the Marine Technology Department to advance the technology of ship design and construction and to serve the growing oceanographic field. Executive offices: One Broadway, New York, New York 10004.

TODD
SHIPYARDS CORPORATION
SHIPYARDS: Brooklyn • New Orleans • Galveston
Houston • Los Angeles • San Francisco • Seattle



ELECTRON-TRANSPORT SYSTEM provides energy for reactions that synthesize adenosine triphosphate (ATP), the universal fuel of cellular reactions. From DPNH (the reduced form of diphosphopyridine nucleotide) or from succinate, electrons pass to the flavoproteins (*FP_1 and FP_2*), coenzyme Q (*CoQ*) and the cytochromes (*b, c_1, c, a, a_3*). Oxygen molecules take electrons from the system.



FOUR COMPLEXES within the mitochondrion represent segments of the electron transfer system. Each complex includes electron carriers and other substances with interrelated catalytic activities. Plague toxin primarily inhibits the activity of Complex I.

knowledge would provide deeper understanding of how the toxin influences the physiology of the heart in susceptible animals.

When a suspension of rat heart mitochondria is incubated in an appropriate medium containing calcium and inorganic phosphate ions together with an intermediate of the respiratory cycle to supply energy, a considerable quantity of calcium and inorganic phosphate ions is taken up by the mitochondria. This process was inhibited by the addition of murine plague toxin. The relation between the ability of the toxin to inhibit mitochondrial ion uptake and its ability to induce mitochondrial swelling we examined by incubating toxin-treated rat heart mitochondria with the reagent ethylenediaminetetraacetic acid (EDTA). The EDTA was added at a concentration known to prevent swelling without inhibiting ion uptake, and it prevented the inhibitory effect of the toxin on the accumulation of calcium and inorganic phosphate ions by the mitochondria. This suggested that the EDTA, by preventing the toxin-induced swelling, allowed the mitochondria to retain the accumulated ions. In other words, it appears that the toxin inhibits the uptake of ions by inducing the mitochondrion to swell.

It would clearly be interesting to know how *Pasteurella pestis* makes this potent toxin. We have only begun to investigate this question, but we have some promising leads. In addition to manufacturing toxin the *P. pestis* cell must obviously synthesize many other proteins in order to carry out its various functions. We have employed a number of metabolic inhibitors in the attempt to separate toxin synthesis from total protein synthesis. We have found that analogues of the amino acid tryptophan, which compete with the amino acid in protein synthesis, are most effective for this purpose. When we incubated various tryptophan analogues with *P. pestis* cells, nonspecific proteins continued to be formed but toxin synthesis was essentially stopped. These findings suggest that the concentration of tryptophan in the cell may regulate the rate of toxin synthesis. In other words, the amount of tryptophan in a given protein may determine the extent to which it is synthesized.

To sum up, we and our colleagues have learned the following from our isolation of the murine plague toxin and our experimental work with it. The toxin consists of two proteins, each with the same biological effects. Our current opinion is that the potency of the toxin results from its action on mitochondrial membranes that manifests itself as a disruption of the normal functioning of the respiratory chain. One specific effect of the toxin is the inhibition of ion uptake by the mitochondria of heart cells, which could lead to abnormal heart conditions and death. Much remains to be learned about all these things, but one may hope that such findings, together with knowledge of how *Pasteurella pestis* synthesizes toxin, will lead to improved measures for the control of plague in human populations.

There is a nine-thousand-foot mountain in this picture.

Some 270 miles off the Washington coast, in waters nearly 9000 feet deep, a huge undersea mountain rises to a broad, plateau-like pinnacle just 135 feet beneath the surface. Close enough to be accessible, yet far enough to be beyond the influence of the continental shelf, Cobb Seamount offers unique potential for the scientific exploration of the deep ocean environment.

To exploit this potential, the State of Washington has organized the Sea Use Program. Uniting scientists and technicians from many disciplines, Sea Use is a combined effort of the State and the research components of a number of public institutions and private corporations.

Following preliminary research, Sea Use calls for the erection of a permanent tower. This tower will be instrumented both above and below the water with devices capable of precisely measuring sea currents, salinity, temperature, evaporation rates and wind velocity. Because of the tower's stable base and distance from land mass interference, these basic but very vital readings are expected to be significantly more accurate than any to date.

The next phase of Sea Use is the installation of an underwater habitat. Engineers at the University of Washington are presently studying the feasibility of a two-room dwelling. One room would be kept at surface pressure for technicians and their equipment. The other chamber would be at ambient pressure (a compressed helium and oxygen atmosphere) for the divers.

Operating from this base, the divers could fully investigate the plateau itself, and gradually explore their way down the slopes of the mount to a depth of some three hundred feet. (Scientists at the Virginia Mason Research Center in Seattle are developing a new oxygen-medication mixture that promises to triple the length of time an aquanaut can work under water.)

Beyond the free diving level, the aquanauts will work from diving chambers. Present technology will limit this phase of the exploration at approximately 1100 feet. From there on down, exploration will be conducted from research submarines such as the Deepstar 2000, which is scheduled to make its first visit to Cobb Seamount this summer.

As the name implies, the ultimate aim of the Sea Use Program is economic. But man must first unlock the scientific secrets of the oceans before he can open the door to their utilization. The underwater exploration of this 9000-foot mountain could well provide the key.

The State of Washington is a fertile field for the exciting breakthrough industries of the future. In oceanography, aerospace, nuclear energy, agriculture, fisheries and many other areas, the smart money is betting on the State of Washington.

Daniel J Evans
Governor

STATE OF WASHINGTON

For business location information, write: Daniel B. Ward, Director, Department of Commerce and Economic Development, Olympia, Washington 98501

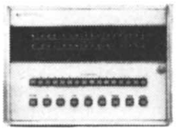
a measure of progress

How to bridge the whole computational gap

Depending on the volume and complexity of your mathematical problems, you might need machine help that runs the gamut from a small desk calculator to a time-shared computer with 16 terminals running simultaneously. Hewlett-Packard feels fortunate that now we can offer you just such a range of computing power.

Our 9100A Computing Calculator, at $4900, is a desk model designed for ease and simplicity of operation, starting with the keyboard. Yet you can perform all the common math, algebra and trig functions on it without learning any special computer language. You simply press the *log* key, the *sin* key, the $\sqrt{}$ key and so on to call these functions out of memory. Using different levels of memory, you can build routines to solve rather sophisticated problems, such as computing the attenuation characteristics of electronic filters with hyperbolic functions. And on one wallet-size memory card, you can store two 196-step routines for future use.

For applications requiring greater power and flexibility, HP offers a family of three computers so you can buy the computer most suited to your needs. All of them can handle three high-level computer languages—FORTRAN, ALGOL and Conversational BASIC. The latter is so close to English you can learn it in three to four hours. The smallest computer, the HP 2114A, costs $9950. You can get it with 8000 words of memory and a teleprinter for $15,950. You also get 16-bit words, 2.0 microsecond memory speed, and 8 channels of input/output capability.



But if you have a number of people who need to use a computer at the same time, then for $89,500 you can get the HP 2000A Time-Shared BASIC System. It can handle up to 16 Teletype terminals at once, with each user thinking he has the computer to himself. In this system we use our largest computer, the HP 2116B, with 16,000 words of core memory and a disc memory with 348,000 words of storage. The same simple Conversational BASIC language is used, but here we have included additional safeguards. The system checks each input statement for format and syntax as it is entered. Additionally, the computer echoes the instruction back to the originator. If there's a transmission error, the originator knows immediately. With these features, 16 persons can have simultaneous error-safe computer power to solve 16 problems at once.

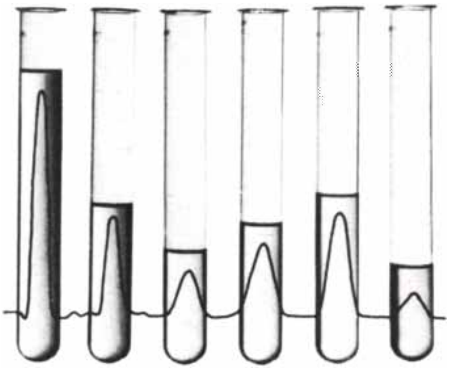
For additional information on any one or all of the Hewlett-Packard solutions to computing problems, write for the 9100A Calculator brochure, a more complete discussion of Hewlett-Packard small computers, or a brochure on HP time-shared computer systems.

Doubling the value of an analytical GC

If you owned a Gas Chromatograph, you'd have one of the fastest, most accurate and economical instruments ever made for chemical analysis. It's a whiz at separating and measuring the chemical components of most any substance. Yet its techniques of vaporizing, separating and detecting can also be used for collecting very pure samples of many chemicals. Because of the highly efficient resolving power of GC, purity of collected fractions can be as high as 99.9%.

Up to now the analytical job and the collecting, or preparative, job were generally done by separate chromatographs, each most efficient at its own task. That is, until Hewlett-Packard developed a reliable, low-cost extension for the analytical GC instrument that makes it a very efficient collector of small samples—without interfering in the least with its analytical role. Since Hewlett-Packard manufactures both analytical and preparative chromatographs, it was just a matter of time until a unit was designed to give both functions to the analytical GC.

The HP 5795A Automatic Preparative Attachment introduces mechanical simplicity and advanced electronic circuits to eliminate any uncertainty or unre-



liability. A programmed Control Unit permits continuous, unattended operation overnight or during weekends. It will direct the collection of any six fractions from up to twenty possible components in the original.

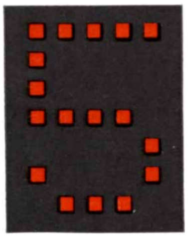
The Injection Unit feeds slugs of the bulk sample into your GC with micrometer accuracy. The Collection Unit distributes the vapor fractions from the GC to the proper trap over and over again, until the required amount has been collected. In one test with rectified turpentine, the instrument cycled automatically 53 times during an 18-hour run. The alpha and beta pinene collected were calculated to be 99.95% and 99.92% pure.

The cost of extending your present GC with the HP 5795A is $2500. For more details, ask for Bulletin 5795.

Solid-state light bulbs

If you want a light that won't blow out or a display that won't fail, you start thinking about devices with no parts to rattle. That immediately eliminates incandescent bulbs with filaments and even glass glow lamps filled with gas. Next, you consider solids or electroluminescent powders that glow when electric current is applied.

One such device Hewlett-Packard has some experience with is the gallium arsenide diode. But it glows in the infrared range, invisible to the eye. Used with a silicon detector it works very well in tape readers or encoders. Still, we wanted a visible light source—so our engineers developed a diode using gallium arsenide phosphide. A small chip 21 mils by 21 mils glows brightly with a soft red light visible for several yards. And it's almost indestructible. Put a diode beside each circuit on a plug-in card, and years from now if the circuit goes bad, the diode will glow to indicate the failure.



Better yet, line some chips up in a matrix five by seven on a side, add a tiny integrated circuit driver for logic control, and you have a device that will flash numbers (shown above, enlarged) from 0 to 9. A somewhat more sophisticated integrated logic circuit could handle letters of the alphabet, as well. The whole assembly mounted on a substrate measures only a half-inch wide, one inch high and less than two tenths of an inch thick. And an entire display needs only five volts to drive it.

This Hewlett-Packard solid-state readout device is ideal for counters, voltmeters or any display that must have high reliability and be insensitive to shock and vibration. But considering its small size, low voltage drive and inherent long life, it may become the economical choice for any small indicator. If you would like more information for your own applications, write for our Solid-State Display pamphlet.

Bedside sentries

Suppose you were in a hospital recovering from major surgery or a heart attack. You wake up—and find yourself surrounded by electronic instruments, with wires of all sorts leading from them to you. It could be frightening. You could feel far worse than you are. And your family might think you've taken a turn for the worse. Not so, and here's why.

An electronic patient monitoring system is measuring, transmitting, displaying and recording heart rate, blood pressure, temperature, and respiration. Second-by-second it forms a continuous communications link between your body's vital functions and the medical staff. Even so, it could still scare you.

But one day a nurse who knows us well suggested another way Hewlett-Packard might help hospitals and their patients. We're old hands at building and installing monitoring systems for hospitals. Why couldn't we publish a booklet telling people how these systems work? A booklet to allay the fears of patient and family alike. We did. If you'd like, write for a copy.

Hewlett-Packard, 1508 Page Mill Rd., Palo Alto, California 94304; Europe: 1217 Meyrin-Geneva, Switzerland.

HEWLETT hp PACKARD

Measurement, Analysis and Computation

00903

THE FIRST ELECTRON TUBE

In 1904 John Ambrose Fleming wrote Guglielmo Marconi: "I have been receiving signals on an aerial with nothing but a mirror galvanometer and my device." The device was the original thermionic vacuum diode

by George Shiers

In October, 1904, John Ambrose Fleming, a British electrical engineer, demonstrated that a high-frequency alternating current could be rectified, or converted to direct current, by what looked like a dimly glowing incandescent lamp in which a flat metal plate was supported between the legs of a carbon filament in the form of a single loop. In his master patent, filed the next month, Fleming disclosed how his primitive vacuum tube could be used as a signal detector in a wireless telegraph receiver circuit. Since his tube acted like a water valve allowing flow in one direction only, Fleming called it an "oscillation valve." In describing it to Guglielmo Marconi, in his capacity as technical adviser to Marconi's Wireless Telegraph Company, Fleming wrote: "This opens up a wide field for work.... I have not mentioned this to anyone yet as it may become very useful." Contrary to some accounts, Fleming did not casually make use of an effect first observed 24 years earlier by Thomas Alva Edison; by virtue of many careful experiments extending over two decades Fleming earned the right to be regarded as the inventor of the original electron tube.

Electric lamps with glowing filaments were developed independently by Joseph W. Swan in Britain and Edison in America at the close of the 1870's. Swan had been on the track of an incandescent-filament lamp for more than 30 years. His precocious American rival was only 32 years old when he demonstrated his version of the incandescent lamp.

Two important defects of the new electric lamp soon became apparent: the filaments broke and dark deposits formed on the inside of the bulb. Edison was particularly troubled by these defects because a reliable lamp with a reasonable lifetime was an essential element in his new electric lighting system. He quickly assumed that the deposit resulted from molecular bombardment. Believing that electrically charged particles of carbon were being ejected from the filament, Edison decided to probe the bulb with a piece of wire to find out if it was possible to prevent "electric carrying," as he called it, and thus preserve the clarity of the glass. The deflection of a galvanometer needle when a galvanometer was placed in a circuit between the wire probe and the positive lead of the glowing filament proved the truth of his notion. This flow of current through a vacuum—the "Edison effect"—contributed nothing, however, toward helping Edison improve his lamp.

This classic experiment was performed during the infancy of the incandescent lamp. It is recorded as Experiment No. 1 and is dated February 13, 1880, in an early Edison notebook [*see illustration on page 108*]. Nothing more was done in the ensuing months, probably because the inventor's time and energy were fully devoted to the creation of his revolutionary power and lighting system.

Edison returned to the lamp problem in the summer of 1882. An entry in one of his notebooks, dated July 5, shows a wire electrode inserted through the top of the bulb. Again the experiment gave way to more urgent work. Early in 1883 Edison, pondering the strange currents within his lamps, went back to the problem and began another series of experiments. A sketch of a new kind of lamp, one with a flat plate between the legs of the filament loop, appears in a notebook entry for March 8. Another entry for March 10 illustrates a similar lamp with a list of metals to be tried as the cold electrode.

Edison found that the bulb current varied according to the temperature of the filament; it was therefore related to the voltage of the filament supply. Ignorant of the true nature of the phenomenon he had uncovered but eager to apply it, he conceived the idea of using the novel lamp as a voltage indicator. A working model was built in the summer of 1883 and was demonstrated in 1884 at an electrical exhibition in Philadelphia (the first such exhibition in America). A patent was granted to Edison for this "electrical indicator" in October of that year. Although the patent was of no commercial value, it is historic: it is the first patent in electronics.

Edison's special lamp and its peculiar behavior naturally attracted attention at the Philadelphia exhibition. The phenomenon was the subject of a note by Edwin J. Houston in the first issue of the *Transactions* of the American Institute of Electrical Engineers, organized in the same year. One may reflect that if a simple experiment had been tried by any one of several able men who attended the exhibition, a practical system of wireless telegraphy utilizing an electron tube could have been demonstrated four years before Heinrich Hertz proved the existence of electromagnetic waves.

The experiment was not done, even though the means and the experimental skills were available. Elihu Thomson had conducted experiments with electromagnetic wave phenomena as early as 1871 (experiments in which Houston played a part) and again in 1875. Edison had toyed with "etheric forces" that year, and his experiments were followed by those of Silvanus P. Thompson in 1876 and of David E. Hughes in 1879. Even then a detecting device was a major need. In most such experiments sparks occurring at some distance from a source were taken as evidence of an etheric disturbance. It remained for Hertz to design a detector consisting of a large loop

FLEMING AND HIS "OSCILLATION VALVE" were photographed in 1923, 19 years after he had discovered that his primitive electronic device could be used to rectify a high-frequency alternating electric current, thus making it suitable as a detector of wireless telegraph signals. His invention of the oscillation valve, better known today as a thermionic vacuum diode, marked the beginning of the age of electronics. At the time this photograph was made Fleming was 74 years old. He died in 1945 at the age of 96.

with a spark gap, and its sparking removed all doubt that electromagnetic radiation had traveled through space. Nonetheless, the rudimentary facts of generating, dispersing and detecting etheric radiations were known to many at the Philadelphia exhibition of 1884. They were not put to practical use until 1896, when Marconi demonstrated his wireless telegraph. The Edison effect remained a laboratory curiosity for another 20 years.

The flow of ideas in electrical science was usually from Europe to the U.S. In the case of the embryonic electron tube, however, the sequence of events was reversed. The tiny currents in the evacuated space of the carbon-filament lamp, ultimately ignored by Edison and his compatriots, became the ward of a British electrical engineer and educator.

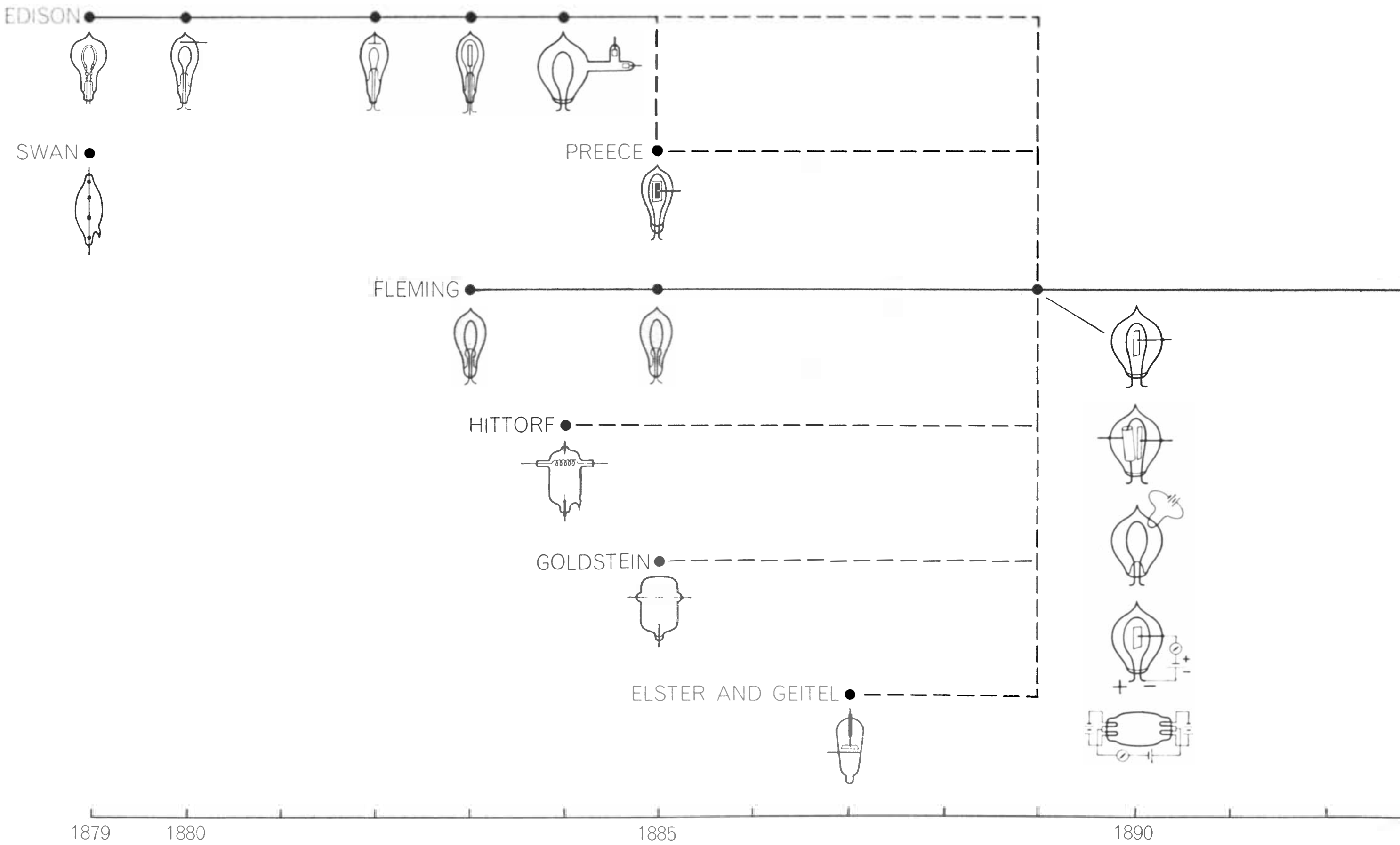
John Ambrose Fleming was born in Lancaster in November, 1849. He began to prepare for a career in electrical engineering when he entered the University College School in London at the age of 14. He continued his studies for several years, occasionally taking time out for commercial work, and started teaching in 1871. After holding several posts as a teacher of science, Fleming entered the University of Cambridge at the age of 28, where he divided his time between his studies and laboratory work under the guidance of James Clerk Maxwell.

Both Edison and Swan formed companies in Britain to develop their electric lamps; eventually they consolidated their interests. Fleming became a consultant to the Edison Electric Light Company in 1882 and was soon involved in problems of the incandescent lamp. His first paper on the subject of "molecular radiation" was read to the Physical Society in May, 1883. He described how volatilization of the carbon filament and of copper from the lead-in clamp produced the dark deposits seen inside the bulb. He identified a narrow line free from deposit as a "shadow" of the loop, and he concluded, as Edison had, that some kind of molecular emission was going on inside the bulb.

Fleming reported similar observations in a second paper, also to the Physical Society, in June, 1885. Since ordinary lamps were used there is no mention of internal currents in these papers, nor is there any reference to the most important clue: the polarity of the leads. This omission seems surprising, but it can be partly explained by the fact that Fleming was examining used lamps and was not working with glowing lamps in a laboratory circuit.

William Henry Preece, then chief electrician of the British Post Office, visited the Philadelphia exhibition of 1884. He obtained some sample lamps from Edison and began to experiment with them soon after returning to Britain. Preece announced his results in a paper read to the Royal Society of London early in 1885. He repeated several of Edison's experiments, including one with a side extension set at right angles to the bulb, and he also tried the shielding effect of a piece of mica over the collector plate. Preece generally confirmed the results obtained by Edison, and it was he who named the basic phenomenon the Edison effect. He also remarked on a "blue glow" that appeared in the lamps at a critical temperature, which was of course directly related to voltage.

Meanwhile independent investigations were being conducted in Germany, in a continuing line of research that had



CHRONOLOGY OF INVENTION from 1879 to 1907 traces the line of work that led from the incandescent lamp of Thomas Alva Edison and Joseph W. Swan to Fleming's oscillation valve. Prior investigations into "thermionic emission" by Edison in the U.S., William Henry Preece and Fleming in Britain, and Wilhelm Hittorf, Eugen Goldstein, Julius Elster and Hans Geitel in Germany were united in Fleming's experiments of 1889. A tube made by Fleming in 1889, which was virtually the same as one devised by Edison in 1883, was

begun in the late 1860's. Wilhelm Hittorf had employed vacuum tubes with cold electrodes in his pioneer studies of gaseous conduction. In 1884 he experimented with platinum electrodes and found that the resistance of the evacuated space was reduced when the negative electrode was heated. Hittorf also observed unilateral conductivity, and he noted that the conductivity could be increased by raising the cathode temperature or by lowering the pressure inside the tube. Eugen Goldstein, another able investigator of electrical conduction in gases, carried out similar experiments the following year and confirmed Hittorf's results.

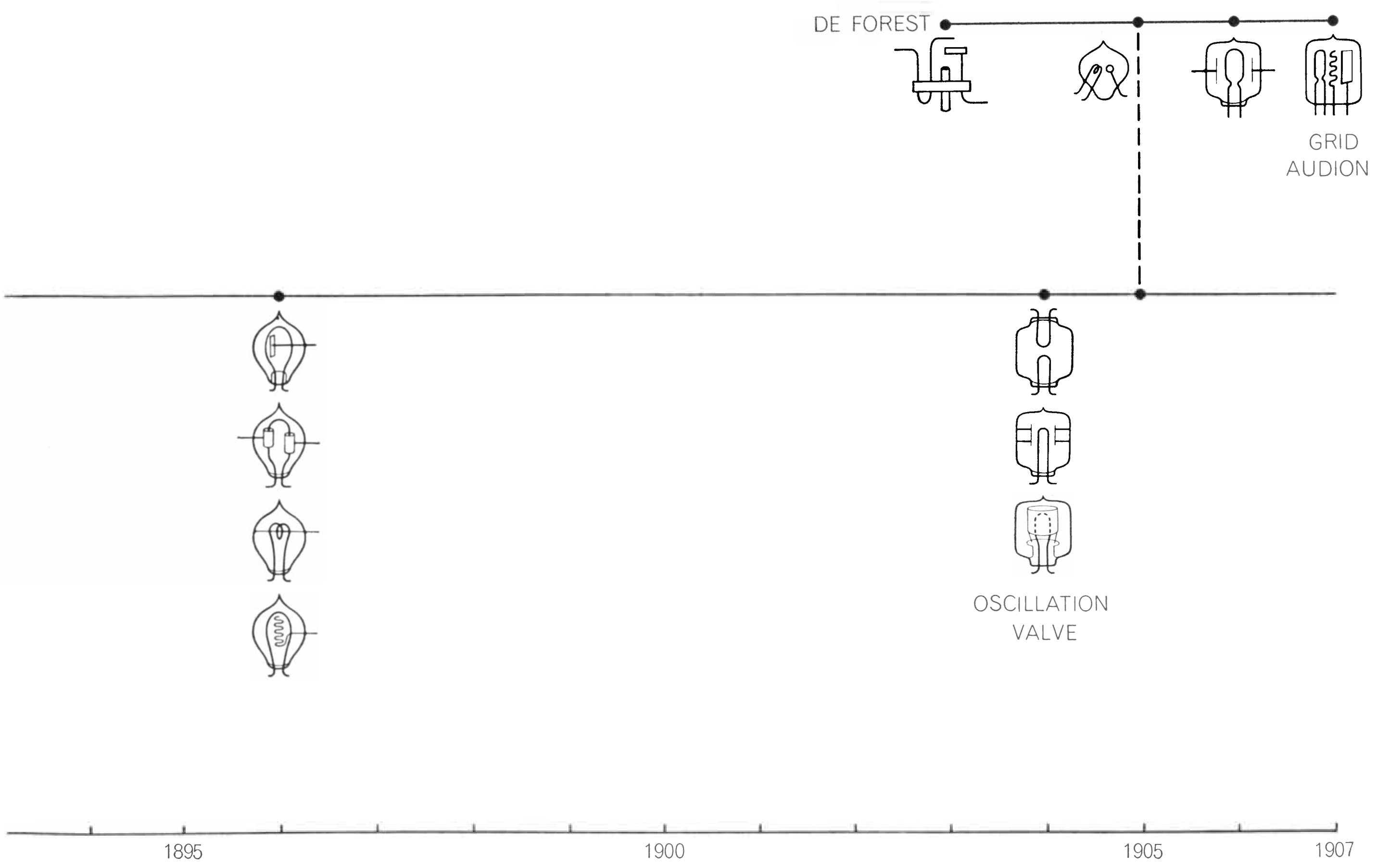
Julius Elster and Hans Geitel, later well known for their studies of photoconductivity, reported a series of experiments on electrical conduction in flames and heated gases from 1882 to 1889. One of their devices had a straight platinum filament and a flat platinum plate in a highly evacuated bulb. After trying a variety of electrode arrangements they concluded that charged particles were being dispersed uniformly from the glowing electrode; they also reported that conductivity was in one direction.

The knowledge gathered by the Germans does not appear to have been utilized by them for further research or for practical ends. Fleming, however, resumed his investigation of the Edison effect in 1889. Building on the American and German discoveries and the work of Preece, he started a line of work that eventually led him to his oscillation valve. Two different fields of endeavor came together in Fleming's hands. The Germans were scientists; their contributions were in the tradition of physical investigation. Edison was an inventor; he made his discovery through curiosity prompted by urgent technical needs. Fleming was neither an inventor nor a scientist but a practicing engineer and a teacher. His interest in the properties of a lamp with an auxiliary electrode arose through his close association with the electric lighting industry and with the training of engineers. Thus Fleming's motivation appears to have been chiefly academic; he had no practical end in view.

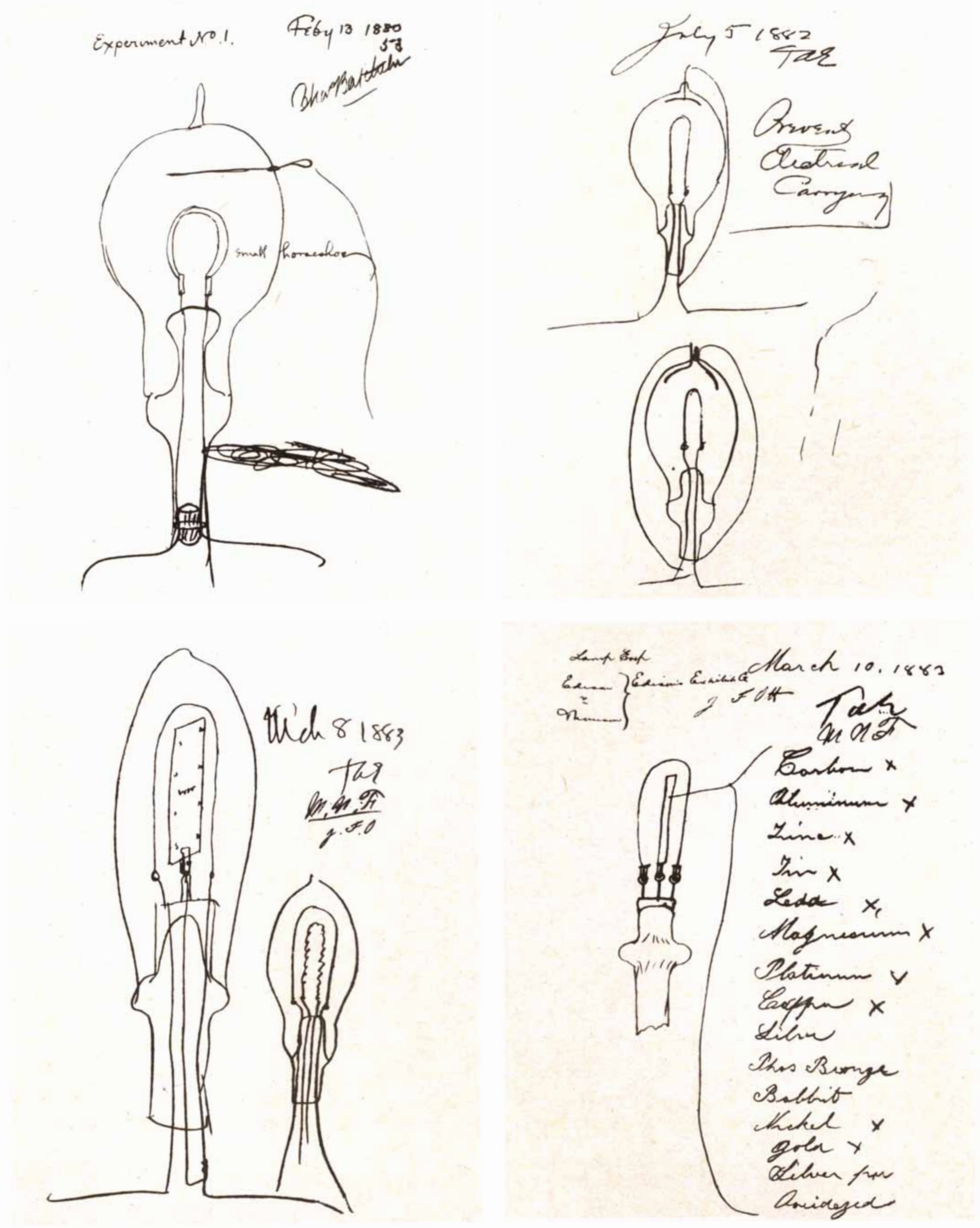
With some special lamps of his own design made at the Edison and Swan Lamp Works, Fleming began a series of thorough experiments to learn more about the curious phenomenon. His results were disclosed in two papers published early in 1890, in which he gave full credit to the earlier work of Edison, Preece and the Germans. Fleming repeated some of the earlier experiments and also devised new ones in an attempt to discover the basic features of the mystifying vacuum currents.

Fleming tried a variety of electrode arrangements [*see top illustration on page 110*], including the use of glass and metal shielding cylinders around the filament legs. He devised lamps with extension tubes, double-ended tubes, heated carbon loops for collectors and center-tapped filaments; with them he used external circuits incorporating batteries and capacitors as well as a galvanometer. Fleming even demonstrated the Edison effect in open air by placing a metal collector plate inside an exposed carbon loop.

Through these and later experiments Fleming became familiar with the phenomenon of current flow through a vacuum. His results supported earlier findings and established unilateral conductivity as a basic property of an evacuated space containing a heated negative elec-



actually the first to achieve high-frequency rectification in 1904. The chart also shows the link between Fleming's 1905 paper and Lee De Forest's early work, which led to the invention of the thermionic triode. Some of De Forest's early devices incorporated open gas flames. With one exception drawings of various tube geometries were adapted from original sources. The drawing of the Goldstein tube, for which no authentic source is available, is based on contemporary tubes and represents a type likely to have been used.



SKETCHES OF EARLY INCANDESCENT LAMPS appear in Edison's notebooks. The entry recorded as Experiment No. 1 and dated February 13, 1880 (*top left*), shows a glow lamp with a "small horseshoe" carbon filament and a wire probe designed to detect the flow of current through the evacuated space of the lamp (the Edison effect); it is probably the first illustration of the bulb that eventually became the electron tube. The sketch is signed by Charles Batchelor, one of Edison's assistants. The entry for July 5, 1882 (*top right*), shows a wire electrode inserted through the top of the bulb. The sketch dated March 8, 1883 (*bottom left*), shows a lamp with a flat plate between the legs of the filament loop. An entry made on March 10, 1883 (*bottom right*), illustrates a similar lamp with a list of metals to be tried as the cold electrode. Edison later abandoned his efforts to find the cause of the Edison effect, and the investigation was taken up by Fleming. The photographs were made at the Edison National Historic Site in West Orange, N.J.

trode and a colder positive plate. Although Fleming thus prepared himself for the invention of the oscillation valve, he did not advance any new theories or suggest any practical applications for the phenomenon or his special glow lamps. He observed in one paper that this field of research was "a region abounding in interesting facts and problems in molecular physics." Nevertheless, he chose to drop his investigations for several years and took up other matters.

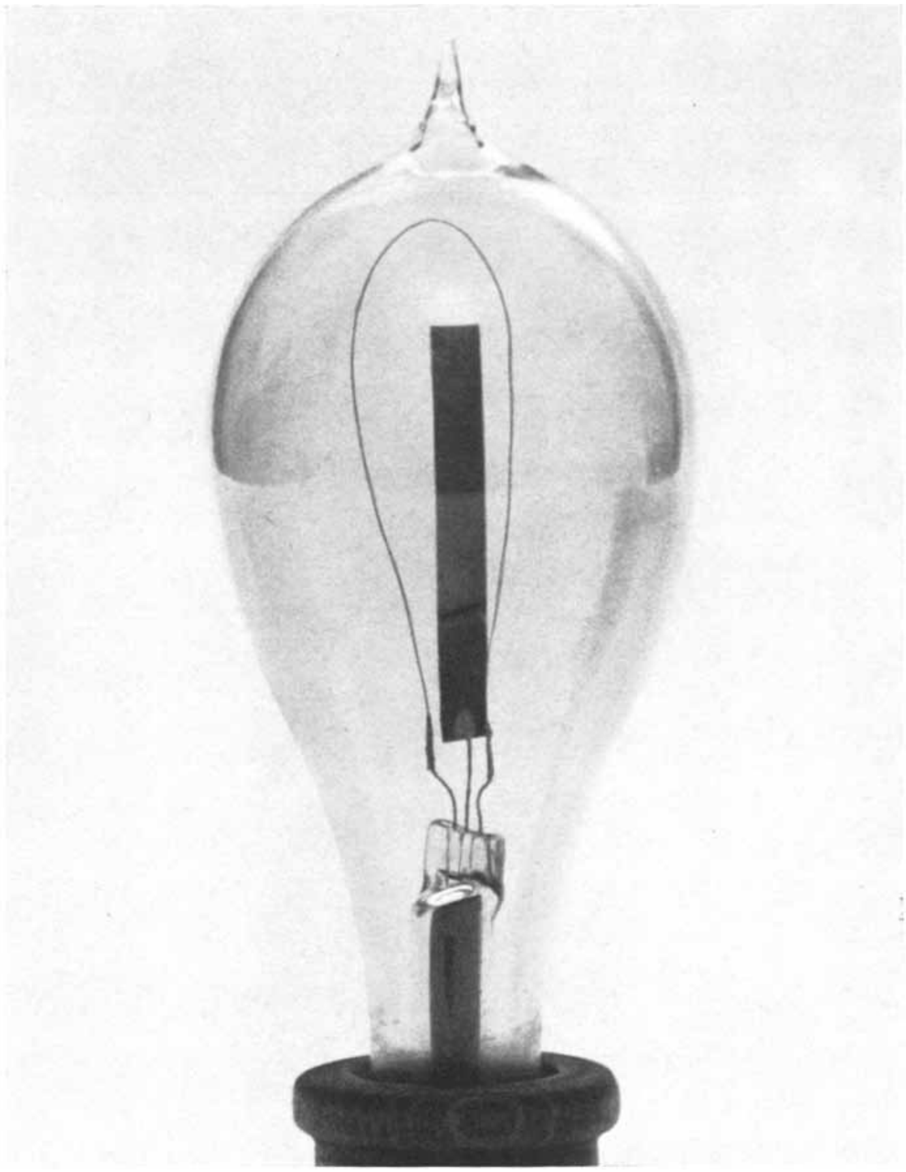
Late in 1895 Fleming once more returned to his current-carrying lamps and initiated an exhaustive series of experiments in which he worked with several new models as well as earlier ones. He presented his new results ("A Further Examination of the Edison Effect in Glow Lamps") at a meeting of the Physical Society on March 27, 1896. In this lengthy paper (profusely supported with diagrams, graphs and tables) Fleming described nearly 30 experiments.

Fleming coined the term "molecular electrovection" to denote the conveyance of electric charge by moving molecules. Molecules dispersed from a glowing electrode were at that time (1896) commonly regarded as being the charge carriers in the Edison effect. It was not until a year later that J. J. Thomson demonstrated the existence of the electron. Fleming clearly described and explained unilateral conductivity, as well as the rectifier action (although he did not name it) of a lamp operated by alternating current. If the word "electron" is substituted for "negatively charged molecules," his paper is a familiar description of the elements of thermionic emission, but of course he was ahead of his time.

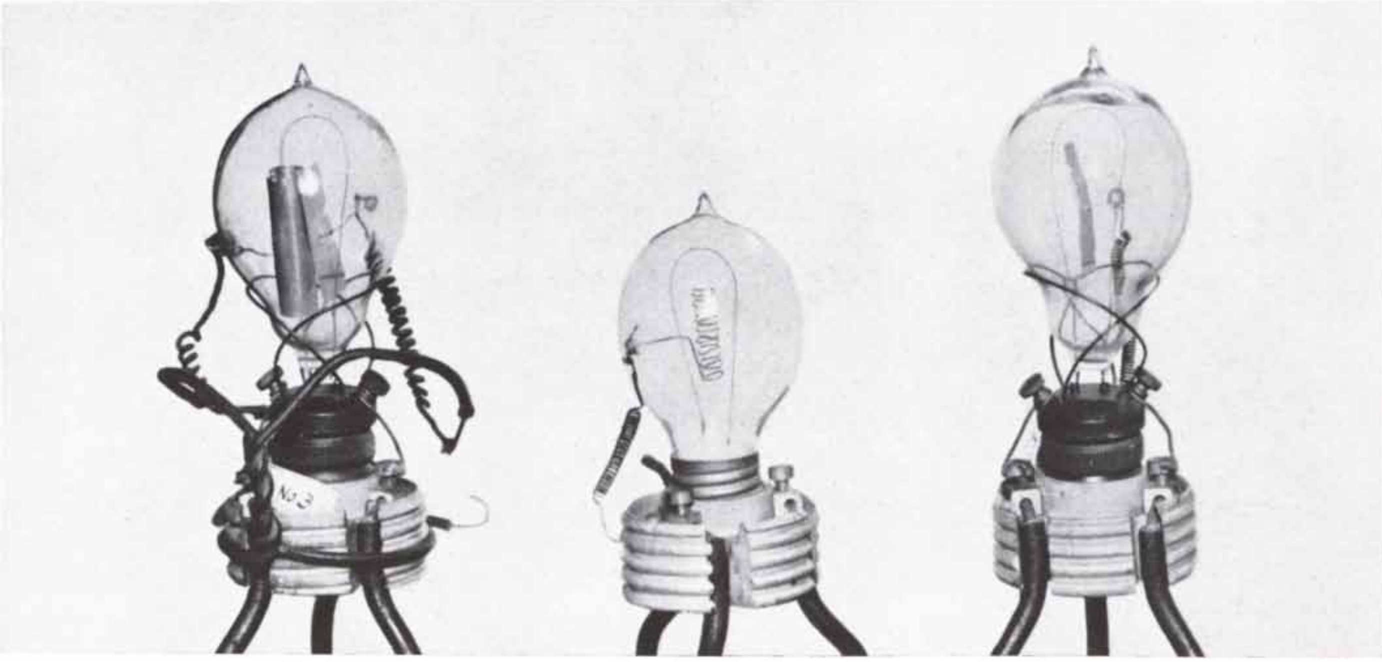
In the last half of the 1890's Fleming's careful experiments were scarcely noted in the press, which was full of news about the X rays that Wilhelm Konrad Roentgen had discovered in 1895 and the wireless telegraph system that Marconi had demonstrated a year later. Fleming's work seemed to be leading nowhere. With Edison's sharp intuition, or the deep perception of Thomson, or perhaps more freedom from pressing duties, Fleming could have leaped into the 20th century simply by combining the elements of his lamps. He could have suggested the possibilities of using his glow lamp as a rectifier for small currents, and since he had built lamps with two anodes, he might have demonstrated full-wave rectification as well as half-wave. Even more important, he could have placed a zigzag wire (which he had put into one of his models) between the other electrodes (just to find out what would happen, as Edison would probably have done) and thereby have anticipated Lee De Forest in inventing the three-element vacuum tube. None of these leaps, however, would have been characteristic of a man as methodical and cautious as Fleming. Although he conducted and documented a series of thorough and well-conceived experiments that apparently covered all aspects of electrical conduction in a vacuum, he did not venture beyond his own experimental limits. Once again, for the sixth time in 16 years, the curious glow lamp was put aside and forgotten.

Over the next three years wireless telegraphy was commercialized. In 1899 Fleming became technical adviser to the newly established Marconi's Wireless Telegraph Company, a post that added to his heavy obligations as professor of electrical technology at University College London. At the end of the century he was busy designing the power system for Marconi's new experimental wireless station at Poldhu in Cornwall. The detector of radio waves in those days was an erratic device known as a coherer, a metal tube filled with iron filings. This was followed by several kinds of electrolytic detector and by Marconi's magnetic detector; both were superior in some respects to the coherer, but they were not satisfactory for regular and dependable service. A new and better detecting device, or signal rectifier, was urgently needed.

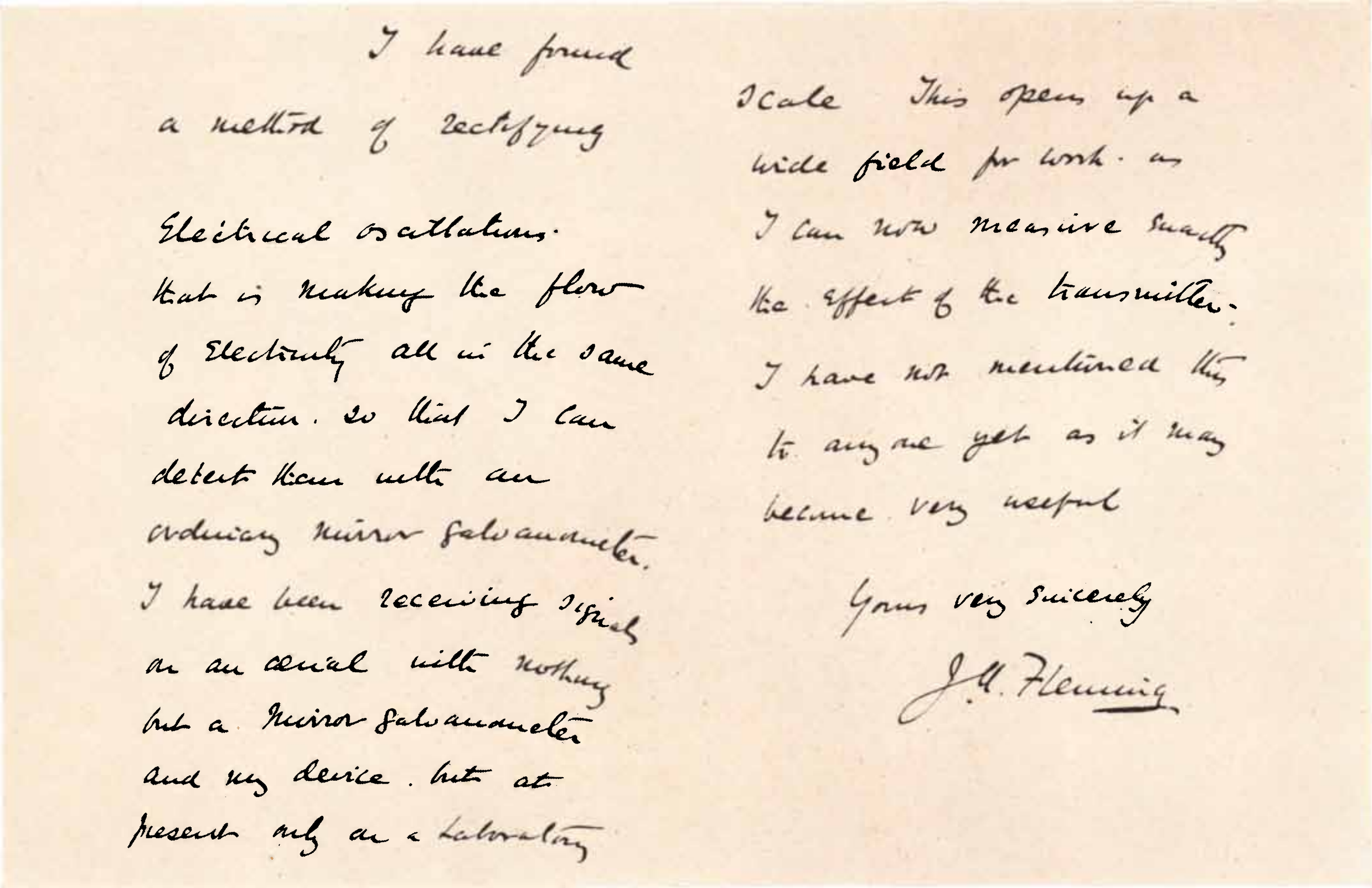
Fleming, a first-rate electrical engi-



REPLICA of Edison's experimental lamp of 1883 shows the simple carbon-filament-plus-metal-plate construction of the original. Conceived of by Edison as a voltage indicator, a working model of this lamp was built in the summer of 1883 and was demonstrated in 1884 at an electrical exhibition in Philadelphia. The patent for this lamp, granted to Edison in October, 1884, was the first patent in electronics. The replica was lent by Vance Phillips.



THREE EXPERIMENTAL LAMPS were made by Fleming in 1889 in the course of his investigations of the Edison effect. The assembly at left contains a metal screening cylinder around one leg of the filament as well as a metal plate inside the loop. The lamp at center has a zigzag of wire inside the loop; the one at right has a metal plate inside the loop. It was this third lamp, inserted in a wireless telegraph receiving circuit by Fleming in his crucial 1904 experiment, that was actually the prototype of his oscillation valve.



I have found
a method of rectifying
Electrical oscillations.
that is making the flow
of Electricity all in the same
direction. so that I can
detect them with an
ordinary mirror galvanometer.
I have been receiving signals
on an aerial with nothing
but a mirror galvanometer
and my device. but at
present only on a laboratory
scale This opens up a
wide field for work. as
I can now measure exactly
the effect of the transmitter.
I have not mentioned this
to anyone yet as it may
become very useful

Yours very sincerely

J. A. Fleming

FLEMING'S LETTER TO MARCONI, written in November, 1904, revealed Fleming's discovery of the rectifying capability of his tube. The key lines in the letter, shown here, read: "I have found a method of rectifying electrical oscillations, that is, making the flow of electricity all in the same direction, so that I can detect them with an ordinary mirror galvanometer. I have been receiving signals on an aerial with nothing but a mirror galvanometer and my device, but at present only on a laboratory scale. This opens up a wide field for work, as I can now measure exactly the effect of the transmitter. I have not mentioned this to anyone yet as it may become very useful. Yours very sincerely, J. A. Fleming." Both photographs were supplied by the Marconi Company Limited.

neer, lecturer and textbook author, rigorously applied classical methods in his work. It was therefore natural for him to ponder the general problem of making quantitative measurements in the new high-frequency apparatus. He sought ways of accurately measuring and evaluating the circuits and devices that were coming into service in wireless systems. Fleming tried the common aluminum-carbon electrolytic cell and found that it, like other detectors, was unsuitable for precision measurements with direct-current meters at radio frequencies.

In considering the problem of converting a feeble oscillatory current into a direct current that could actuate a recording instrument, Fleming later recalled that he had had "a sudden very happy thought": he perceived that his lamp was "the exact implement required to rectify high-frequency oscillations." The device he needed, neglected and all but forgotten, was sitting in one of his laboratory cupboards. Would such a lamp work at Marconi's new high frequencies? Would it be sensitive enough to operate on the extremely low voltages available in a wireless receiving circuit? Finally, could it be used to measure, as well as to detect, feeble currents oscillating at high frequencies?

These questions were soon answered. The inventive act was now nearly complete; all that remained was to unite the strange lamp with a suitable circuit and prove that it would work. Within a few hours a simple spark-coil oscillator and a resonant receiving circuit comprising one of the lamps were set up. The actual lamp that Fleming worked with was one of the 1889 models that had a flat plate supported between the legs of a single-loop carbon filament. It was virtually the same as Edison's lamp of 1883. The experiment was an immediate success; Fleming had discovered a new kind of high-frequency rectifier, one that was eminently suitable for detecting wireless signals.

This memorable event took place one afternoon in October, 1904. Fleming was considering the geometry of the collecting electrode that would be most efficient. He decided on an open metal tube (not unlike the shielding cylinders in some of his 1889 experiments) that would surround the entire filament. The following day he placed an order with the Ediswan Lamp Works for a quantity of special lamps, each with a metal cylinder around the carbon filament. These lamps, which Fleming soon called oscillation valves, were the first thermionic vacuum diodes [*see bottom illustration on next page*].

Early in November, Fleming wrote Marconi: "I have been receiving signals on an aerial with nothing but a mirror galvanometer and my device" [*see bottom illustration on opposite page*]. On November 16 he filed a provisional application for a patent on "Improvements in Instruments for Detecting and Measuring Alternating Electric Currents." This application made direct reference to the use of a galvanometer and a two-element lamp—the first vacuum tube in an electronic circuit—as a receiving instrument in wireless telegraphy. Fleming's master patent was granted on September 21, 1905 [*see top illustration on next page*].

Fleming lost no time in presenting his discovery to his professional colleagues. His paper "On the Conversion of Electric Oscillations into Continuous Currents by Means of a Vacuum Valve" was read at a Royal Society meeting on February 9, 1905, and was published in March. Fleming discussed the rectifying power of the valve and showed characteristic curves relating cathode-anode current and applied voltage (up to 100 volts) at three different filament voltages. He also proposed water cooling for the anode to increase the "rectifying efficiency." Fleming's hopes were only half-realized, however; although he had invented a new wireless detector, it failed him as a device for measuring high-frequency currents.

In his patent Fleming did not describe the mode of electrical conduction in his valve; he referred only to "negative electricity." Nonetheless, in his paper he employed "electron" and "free electrons" in describing how conduction took place. Fleming was the first to use the word "electron" popularly in referring to minute negative charges or particles (also called negative ions or negative corpuscles) in a paper titled "The Electronic Theory of Electricity," first published by the Royal Institution in 1902 and subsequently reprinted in *The Popular Science Monthly*. "Electronic" thus entered the technical literature before the advent of the first practical electron tube.

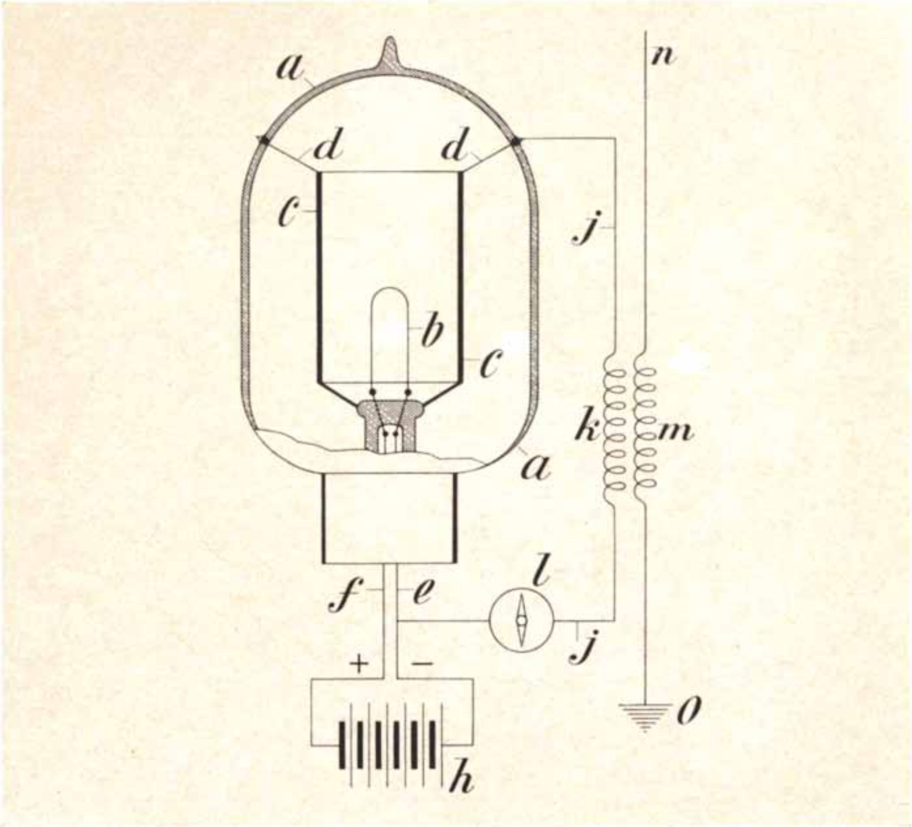
In 1906, at the peak of his career, Fleming published *The Principles of Electric Wave Telegraphy*, an encyclopedic treatise that became the foremost textbook for practitioners at all levels in the new communication field. In this book Fleming described his oscillation valve as one kind of "vacuum tube cymoscope," a generic name that he introduced with his invention of the cymometer, an instrument for measuring the length of electric waves. Curiously, Fleming now omitted all reference to electrons in describing the flow of current from the carbon filament to the metal cylinder in his valve; instead he spoke of a flow of "negative electricity." Nor did he refer to the theory of thermionic emission, published in 1901 by O. W. Richardson, in spite of its relevance to his own work. Fleming made various improvements in his valves early in 1905, and they went into actual service for the first time at Poldhu in the summer. The Marconi organization began manufacturing Fleming valves in 1907 and incorporated them into much of its equipment for several years thereafter.

Few inventions can have brought their inventors so much distress, disappointment and trouble as the oscillation valve gave Fleming. His patent and his interests in it immediately became the property of Marconi's Wireless Telegraph Company under the terms of his contract. Before his "glow lamp detector," as it was sometimes called, had a chance to prove its value it was supplanted by the "cat's whisker" crystal rectifier, which was first employed in 1906. (The crystal rectifier was the first solid-state device in electronic technology.) In later years, when thermionic valves were being produced by the millions annually, Fleming, perhaps understandably, showed a sour side of his nature because he had not received any financial reward for his pioneer labors. Moreover, "his valve," although it was modified and improved a thousandfold by others, was the predecessor and (in his mind) the real source of a booming industry.

The man who reaped some of the fame and fortune that Fleming considered his due was Lee De Forest, one of the most enterprising and determined radio pioneers among the Americans. In 1904, the year Fleming applied for his British patent, De Forest was 31 years old and had decided to establish a wireless system of his own free from patent domination by others, particularly Marconi. De Forest was obsessed with the idea that "a gaseous medium," maintained in a condition of "molecular activity," could serve as an "oscillation responsive device," in other words, as a detector of wireless signals.

In a series of patent applications extending from November, 1904, to the end of 1906, De Forest embodied his ideas in a variety of devices that changed from open gas flames to electrodes enclosed in glass bulbs. About the middle of 1905 Fleming became aware of a De Forest device not unlike one of his own. In response Fleming called De Forest's attention to his Royal Society paper of the previous March. This move by Fleming, obviously a warning that De Forest was poaching on Fleming's preserve, was the beginning of a lifelong feud between the two men.

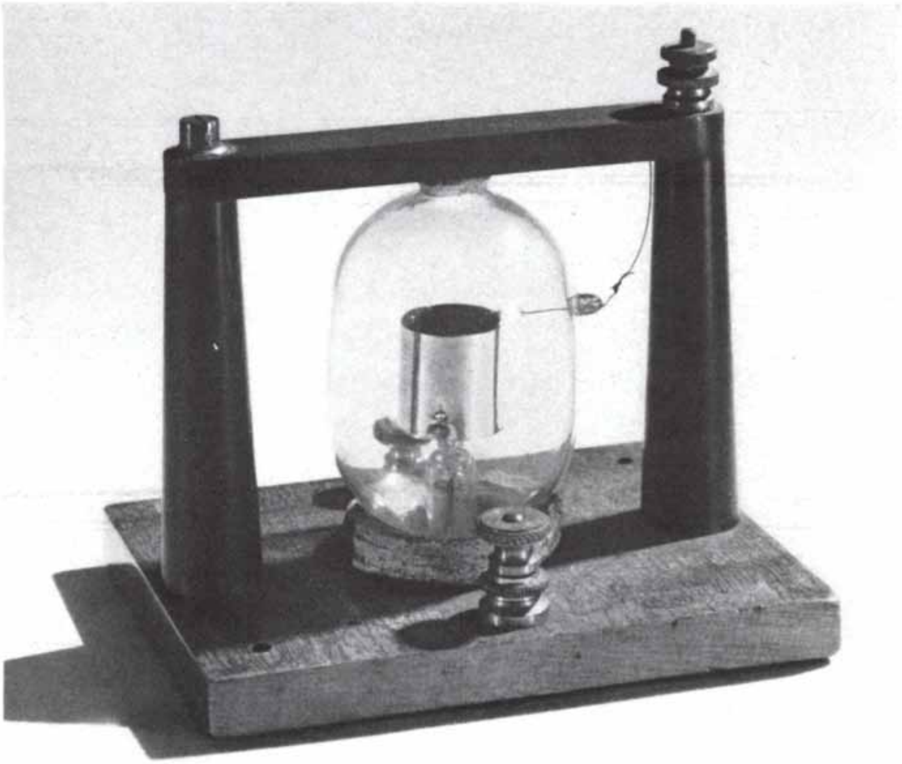
The situation got worse when De Forest described his experiments in a discursive and hazy paper presented to the American Institute of Electrical Engineers on October 26, 1906. This paper ("The Audion") also appeared in *Scientific American Supplement* in November and December, 1907. Unfortunately De Forest was wrong about several historical matters and slighted Fleming by making only a passing and inaccurate reference to his work. Fleming never forgave De Forest for his lack of generosity or for his "exploitation" of the basic thermionic diode. Fleming maintained to the end of his life that De Forest had merely added a piece of bent wire to "his valve."



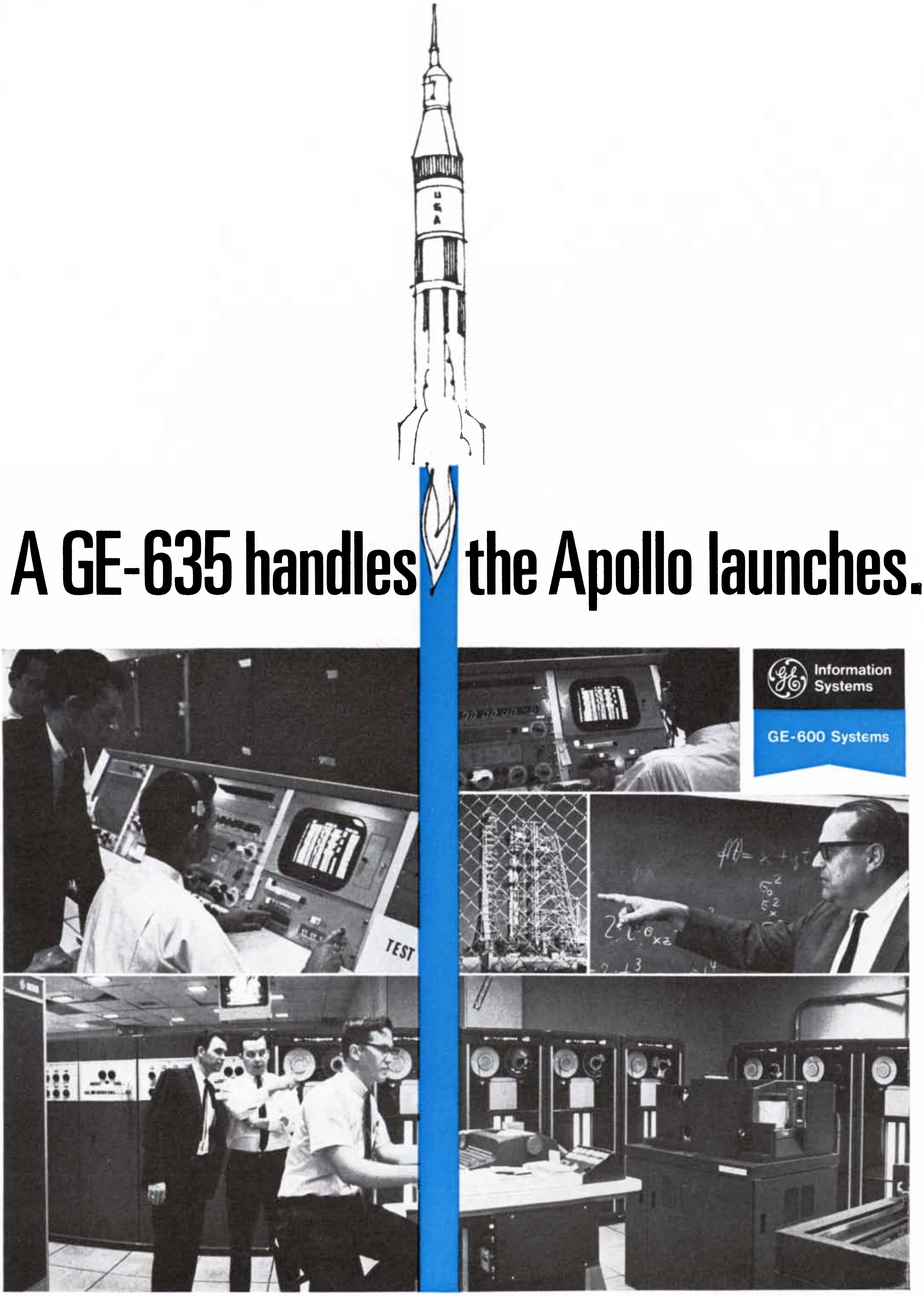
MORE EFFICIENT GEOMETRY was adopted by Fleming in the drawings of the oscillation valve that accompanied his original British patent application in 1904. The design incorporates an open metal cylinder (*c*) that surrounds the entire filament (*b*). Electrons emitted by the filament moved to the cylindrical plate when the cylinder was positively charged by a signal from the antenna (*n*). A galvanometer (*l*) was employed to indicate the arrival of a signal. The patent also included diagrams of a full-wave rectifier with two valves connected to a differential galvanometer, and three valves connected in parallel.

Early in 1907 De Forest applied for an American patent on a wireless telegraph receiver that disclosed the first three-electrode tube with a control element (a zigzag piece of wire) located between the filament and the plate. The tremendous importance of this twin invention—the grid triode and its basic circuit—was not perceived by the inventor or his contemporaries for several years. Suddenly, between 1911 and 1913, a new technology emerged: triode circuits became amplifiers and oscillators, adaptable to both radio and telephone systems.

The new technology grew to maturity as radio broadcasting became a part of everyday life, followed by television, radar, the electronic digital computer and a host of other applications for the ubiquitous electron tube. In the deepest sense these technological advances were launched by Fleming's "very happy thought" and his quick verification that an old idea could serve new ends. There can be no doubt that if he had not devised the thermionic diode, someone else would have done so soon afterward. Still, Fleming was the man who took the immortal first step. In giving the world the first practical electron tube he not only justified his patient guardianship of the Edison effect but also provided an important transitional link between the age of electric power and the age of electronics.



PRODUCTION MODEL of the Fleming oscillation valve is one of the 12 original thermionic diodes manufactured at the Ediswan Lamp Works near London in October, 1904.


A GE-635 handles the Apollo launches.
Information Systems
GE-600 Systems
TEST

But it has other things on its mind.

The same things you worry about every day: printing the payroll, engineering calculations, accounting . . . "The GE-635 is the only information system anywhere that could perform the real-time data processing for the Apollo launch — and handle other necessary processing at the same time," says Dr. R. H. Bruns, chief of the Data Systems Division at Kennedy Space Center. "It is beyond compare."

Credit the people at NASA for a cost-saving decision. Rather than buy individual computers for specialized jobs, they chose a single information system with the ability to handle many programs concurrently — a GE-635 system.

The 3-Dimensional Information System

This same ability could be important to you. Because of it, a single GE-635 can operate in 3 dimensions, all at the same time:

1. **Local processing.** Replace your duplicate computers and their duplicate staffs with a powerful, central GE-635 system. You'll be able to handle dozens of jobs concurrently, so you'll get more work out faster.

2. **Remote processing.** Inexpensive input/output computers link directly with the central GE-635. This gives your satellite operations the full power of the GE-635 from any distance.

3. **Time-sharing.** Easy-to-use time-sharing terminals give your key people instant access to vital information, speed up their calculating ability.

That's why we call the GE-635 the world's only 3-Dimensional Information System. The Apollo launches have proved that it works successfully. Let us prove it to you. Call your GE Information Systems Sales Representative. Or write General Electric, Section 290-27, 1 River Road, Schenectady, N. Y. 12305.

GENERAL ELECTRIC

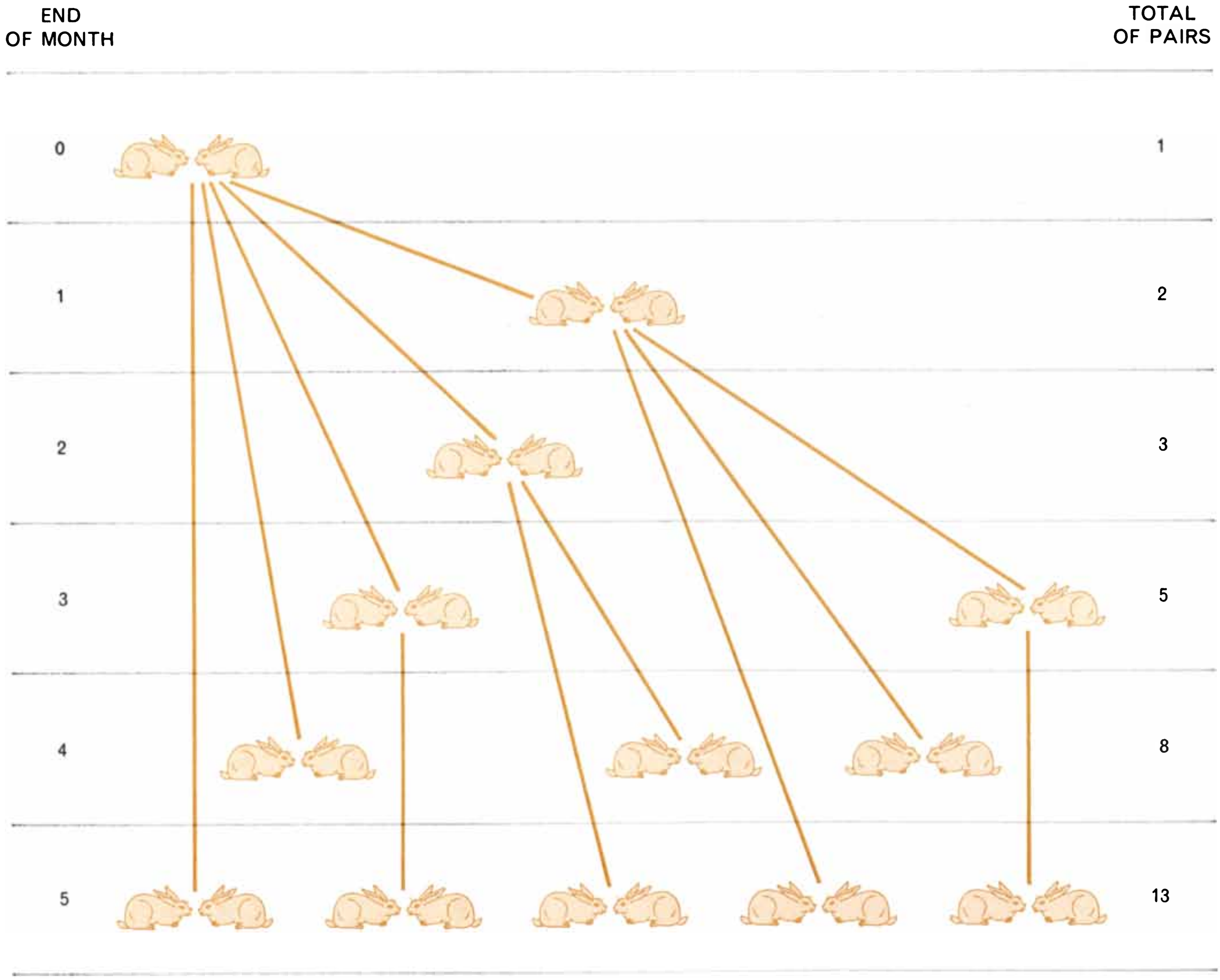
MATHEMATICAL GAMES

The multiple fascinations of the Fibonacci sequence

by Martin Gardner

The greatest European mathematician of the Middle Ages was Leonardo of Pisa, better known as Fibonacci, meaning "son of Bonaccio." Although Leonardo was born in Pisa, his father was an official of an Italian mercantile factory in Bougie in Algeria, and it was there that young Leonardo received his early mathematical training from Moslem tutors. He quickly recognized the enormous superiority of the Hindu-Arabic decimal system, with its positional notation and zero symbol, over the clumsy Roman system still used in his own country. His best-known work, *Liber abaci* (literally "Book of the Abacus" but actually a comprehensive merchant's handbook on arithmetic and algebra), defended the merits of the Hindu-Arabic notation. The arguments made little impression on the Italian merchants of the time but the book eventually became the most influential single work in introducing the Hindu-Arabic system to the West. Although *Liber abaci* was completed in Pisa in 1202, it survives only in a revised 1228 edition dedicated to a famous astrologer of the period. There has never been an English translation.

It is ironic that Leonardo, who made valuable contributions to mathematics, is remembered today mainly because a 19th-century French number theorist, Édouard Lucas (who edited a classic four-volume work on recreational mathematics), attached the name Fibonacci to a number sequence that appears in a trivial problem in *Liber abaci*. Suppose, Leonardo wrote, a male-female pair of adult rabbits is placed inside an enclosure to breed. Assume that rabbits start to bear young two months after



Tree graph for Fibonacci's rabbits

their own birth, producing only a single male-female pair, and that they have one such pair at the end of each subsequent month. If none of the rabbits die, how many pairs of rabbits will there be inside the enclosure at the end of one year?

The tree graph [*see illustration on opposite page*] shows what happens during the first five months. It is easy to see that the numbers of pairs at the close of each month form the sequence 1, 2, 3, 5, 8, . . . , in which each number (as Fibonacci pointed out) is the sum of the two numbers preceding it. At the end of 12 months there will be 377 pairs of rabbits.

Fibonacci did not investigate the series and no serious study of it was undertaken until the beginning of the 19th century, when, as a mathematician once put it, papers on the sequence began to multiply almost as fast as Fibonacci's rabbits. Lucas made a deep study of sequences (now called "generalized Fibonacci sequences") that begin with *any* two positive integers, each number thereafter being the sum of the preceding two. He called the simplest such series, 1, 1, 2, 3, 5, 8, 13, 21, . . . , the Fibonacci sequence. The position of each number in this sequence is traditionally indicated by a subscript, so that $F_1 = 1$, $F_2 = 1$, $F_3 = 2$, and so on. (The first 50 Fibonacci numbers are listed in the illustration on this page). F_n refers to any Fibonacci number. F_{n+1} is the number following F_n; F_{n-1} is the number preceding F_n; F_{2n} is the F-number with a subscript twice that of F_n, and so on.

The Fibonacci sequence has intrigued mathematicians for centuries, partly because it has a way of turning up in unexpected places but mainly because the veriest amateur in number theory, with no knowledge beyond simple arithmetic, can explore the sequence and discover a seemingly endless variety of curious theorems. Recent developments in computer programming have reawakened interest in the series because it turns out to have useful applications in the sorting of data, information retrieval, the generation of random numbers and even in rapid methods of approximating maxima and minima values of complicated functions for which derivatives are not known.

Early results are summarized in Chapter 17 of the first volume of Leonard Eugene Dickson's *History of the Theory of Numbers*. For the most recent discoveries interested readers can consult *The Fibonacci Quarterly*, published since 1963 by the Fibonacci Association. (The annual price is now $6 and subscriptions are handled by the managing editor, Brother Alfred Brousseau, at St. Mary's College in St. Mary's College, Calif.) Edited by Verner E. Hoggatt, Jr., of San Jose State College in San Jose, Calif., the quarterly is concerned primarily with generalized Fibonacci numbers and similar numbers (such as "Tribonacci numbers," which are sums of the preceding *three* numbers), but the journal is also devoted "to the study of integers with special properties."

Surely the most remarkable property of the Fibonacci series (which holds for the generalized series too) is that the ratio between two consecutive numbers is alternately greater or smaller than the golden ratio and that, as the series continues, the differences become less and less; the ratios approach the golden ratio as a limit. The golden ratio is a famous irrational number, 1.61803 . . . , that is obtained by halving the sum of 1 and the square root of 5. There is a considerable literature, some of it dubious, about the appearance of the golden ratio and the closely related Fibonacci sequence in organic growth and about their applications to art, architecture and even poetry. George Eckel Duckworth, professor of classics at Princeton University, maintains in his book *Structural Patterns and Proportions in Vergil's Aeneid* (University of Michigan Press, 1962) that the Fibonacci series was consciously used by Vergil and other Roman poets of the time. I dealt with such matters in an earlier column on the golden ratio, which is reprinted in *The 2nd Scientific American Book of Mathematical Puzzles & Diversions.*

The most striking appearance of Fibonacci numbers in plants is in the spiral arrangement of seeds on the face of certain varieties of sunflower. There are two sets of logarithmic spirals, one set turning clockwise, the other counterclockwise, as indicated by the two colored spirals in the illustration on the next page. The numbers of spirals in the two sets are different and tend to be consecutive Fibonacci numbers. Sunflowers of average size usually have 34 and 55 spirals, but giant sunflowers have been developed that go as high as 89 and 144. In the letters department of *The Scientific Monthly* (November, 1951) Daniel T. O'Connell, a geologist at City College of the City of New York, and his wife reported having found on their Vermont farm one mammoth sunflower with 144 and 233 spirals!

The intimate connection between the Fibonacci series and the golden ratio

n	F_n
1	1
2	1
3	2
4	3
5	5
6	8
7	13
8	21
9	34
10	55
11	89
12	144
13	233
14	377
15	610
16	987
17	1,597
18	2,584
19	4,181
20	6,765
21	10,946
22	17,711
23	28,657
24	46,368
25	75,025
26	121,393
27	196,418
28	317,811
29	514,229
30	832,040
31	1,346,269
32	2,178,309
33	3,524,578
34	5,702,887
35	9,227,465
36	14,930,352
37	24,157,817
38	39,088,169
39	63,245,986
40	102,334,155
41	165,580,141
42	267,914,296
43	433,494,437
44	701,408,733
45	1,134,903,170
46	1,836,311,903
47	2,971,215,073
48	4,807,526,976
49	7,778,742,049
50	12,586,269,025

The first 50 Fibonacci numbers

can be seen in the following formula for the nth Fibonacci number:

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right].$$

This equation gives the nth Fibonacci number exactly but it is cumbersome to use for high F-numbers, although good approximations can be obtained with logarithms. A much simpler formula for the nth F-number is the golden ratio raised to the power of n and then divided by the square root of 5. When this result is rounded off to the nearest integer, it too provides the exact number sought. Both formulas are nonrecursive because they compute the nth F-number directly from n. A "recursive procedure" is a series of steps each of which is dependent on previous steps. If you compute the nth F-number by summing consecutive F-numbers until you reach the nth, you are computing it recursively; a definition of the nth F-number as the sum of the preceding two numbers is a simple example of a recursive formula. (Two highly efficient computer algorithms for computing large F-numbers exactly are given as the answer to exercise No. 26 on page 552 of *Seminumerical Algorithms,* the second volume of *The Art of Computer Programming* by Donald E. Knuth.)

To find the sum of the first n Fibonacci numbers the best procedure is to determine F_{n+2} and then subtract 1. Example: What is the sum of the first 20 F-numbers? Subtract 1 from 17,711, the 22nd F-number, to get the answer: 17,710.

Here are some more well-known properties of the Fibonacci sequence, most of them not difficult to prove:

1. The square of any F-number differs by 1 from the product of the two F-numbers on each side. The difference is alternately plus or minus as the series continues. Like so many properties of the Fibonacci series, this is a special case of a property that applies to the general sequence starting with any two integers. In the general case too the difference is a constant that is alternately plus or minus. For example, the next-simplest series after the Fibonacci, 1, 3, 4, 7, 11, 18, ... (now called the Lucas series after the French mathematician), has a constant difference of 5.

2. The sum of the squares of any two consecutive F-numbers, F^2_n and F^2_{n+1}, is F^2_{2n+1}. Since the last number must have an odd subscript, it follows from this theorem that if you write in sequence the squares of the F-numbers, sums of consecutive squares will produce in sequence the F-numbers with odd subscripts.

3. For any four consecutive F-numbers, A, B, C, D, the following formula holds: $C^2 - B^2 = A \times D$.

4. The sequence of final digits of the Fibonacci sequence repeats in cycles of 60. The last two digits repeat in cycles of 300. The repeating cycle is 1,500 for three final digits, 15,000 for four digits, 150,000 for five and so on for all larger numbers of digits.

5. For every integer m there is an infinite number of F-numbers that are evenly divisible by m, and at least one can be found among the first m^2 numbers of the Fibonacci sequence.

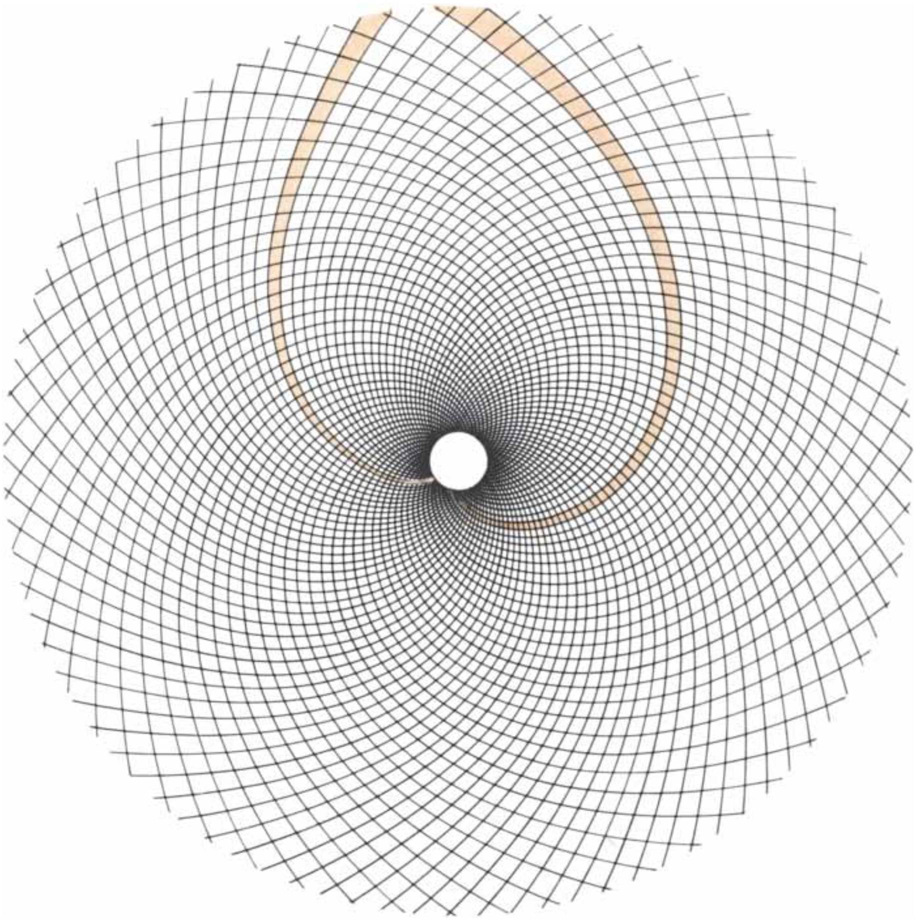
6. Every third F-number is divisible by 2, every fourth number by 3, every fifth number by 5, every sixth number by 8 and so on, the divisors being F-numbers in sequence. Consecutive Fibonacci numbers (as well as consecutive Lucas numbers) cannot have a common divisor other than 1.

7. With the exception of 3, every F-number that is prime has a prime subscript (for example, 233 is prime and its subscript, 13, is also prime). Put another way, if a subscript is composite (not prime), so is the number. Unfortunately the converse is not always true: a prime subscript does not necessarily mean that the number is prime. The first counterexample is F_{19}, 4,181. The subscript is prime but 4,181 is 37 times 113.

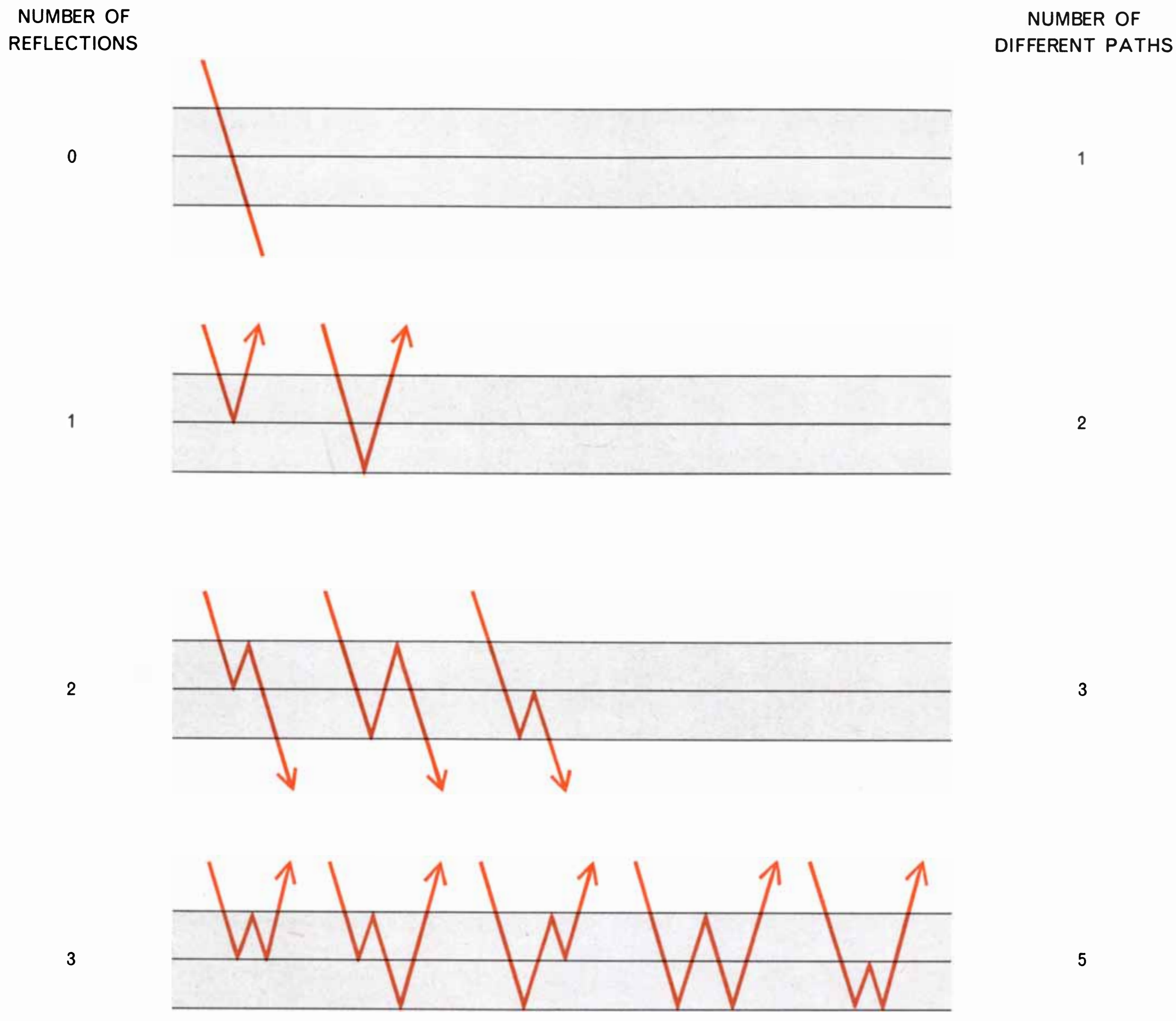
If the converse theorem held in all cases, it would answer the greatest unsolved question about Fibonacci numbers: Is there an infinity of Fibonacci primes? We know that the number of primes is infinite, and therefore if every F-number with a prime subscript were prime, there would be an infinity of prime F-numbers. As it is, no one today knows if there is a largest Fibonacci prime.

8. With the trivial exceptions of 0 and 1 (taking 0 to be F_0), the only square F-number is F_{12}, 144—which, surprisingly, is the square of its subscript. Whether or not there is a square F-number greater than 144 was an open question until the matter was finally settled, as recently as 1963, by John H. E. Cohn of Bedford College in the University of London. He also proved that 1 and 4 are the only squares in the Lucas sequence.

9. The reciprocal of 89, the 11th F-number, can be generated by writing



Sunflower with 55 counterclockwise and 89 clockwise spirals



There are F_{n+2} *paths by which a ray can be reflected* n *times through two panes of glass*

the Fibonacci sequence, starting with 0, and then adding as follows:

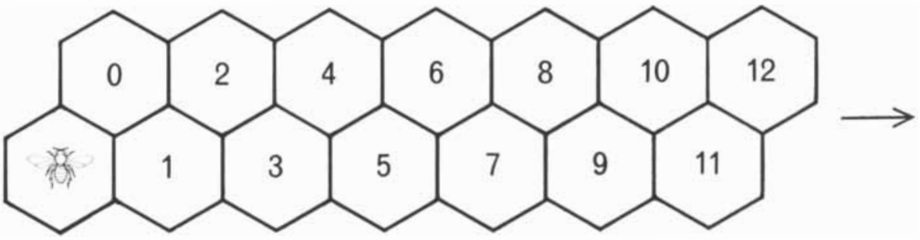
```
.0112358
      13
       21
        34
         55
          89
          144
           233
             . .
.011235955040673... = 1/89 .
```

This list of properties could be extended to fill a book. One could do the same with instances of how the series appears in physical and mathematical situations. (For its appearance in the diagonals of Pascal's triangle see this department for December, 1966.) Leo Moser in 1963 studied the paths of slanting light rays through two face-to-face glass plates. An unreflected ray goes through the plates in only one way [*see illustration above*]. If a ray is reflected once, there are two paths; if it is reflected twice, there are three paths, and if three times, there are five paths. As n, the number of reflections, increases, the numbers of possible paths fall into the Fibonacci sequence. For n reflections the number of paths is F_{n+2}.

The sequence can be applied similarly to the different paths that can be taken by a bee crawling over hexagonal cells [*see top illustration on next page*]. The cells extend as far as desired to the right. Assume that the bee always moves to an adjacent cell and always moves toward the right. It is not hard to prove there is one path to cell 0, two paths to cell 1, three to cell 2, five to cell 3 and so on. As before, the number of paths is F_{n+2}, where n is the number of cells involved.

Consider Fibonacci nim, a counter-removal game invented a few years ago by Robert E. Gaskell. The game begins with a pile of n counters. Players take turns removing counters. The first player may not take the entire pile, but thereafter either player may remove all the remaining counters if these rules permit: at least one counter must be taken on each play, but a player may never take more than twice the number of counters his opponent took on his last play. Thus if one player takes three counters, the next player may not take more than six. The person who takes the last counter wins.

It turns out that if n is a Fibonacci number, the second player can always win; otherwise the first player can win. If a game begins with 20 counters (not an F-number), how many must the first



There are F_{n+2} *paths by which the bee can crawl to cell* n

player take to be sure of winning? The answer will be given next month along with a simple strategy employing *F*-numbers.

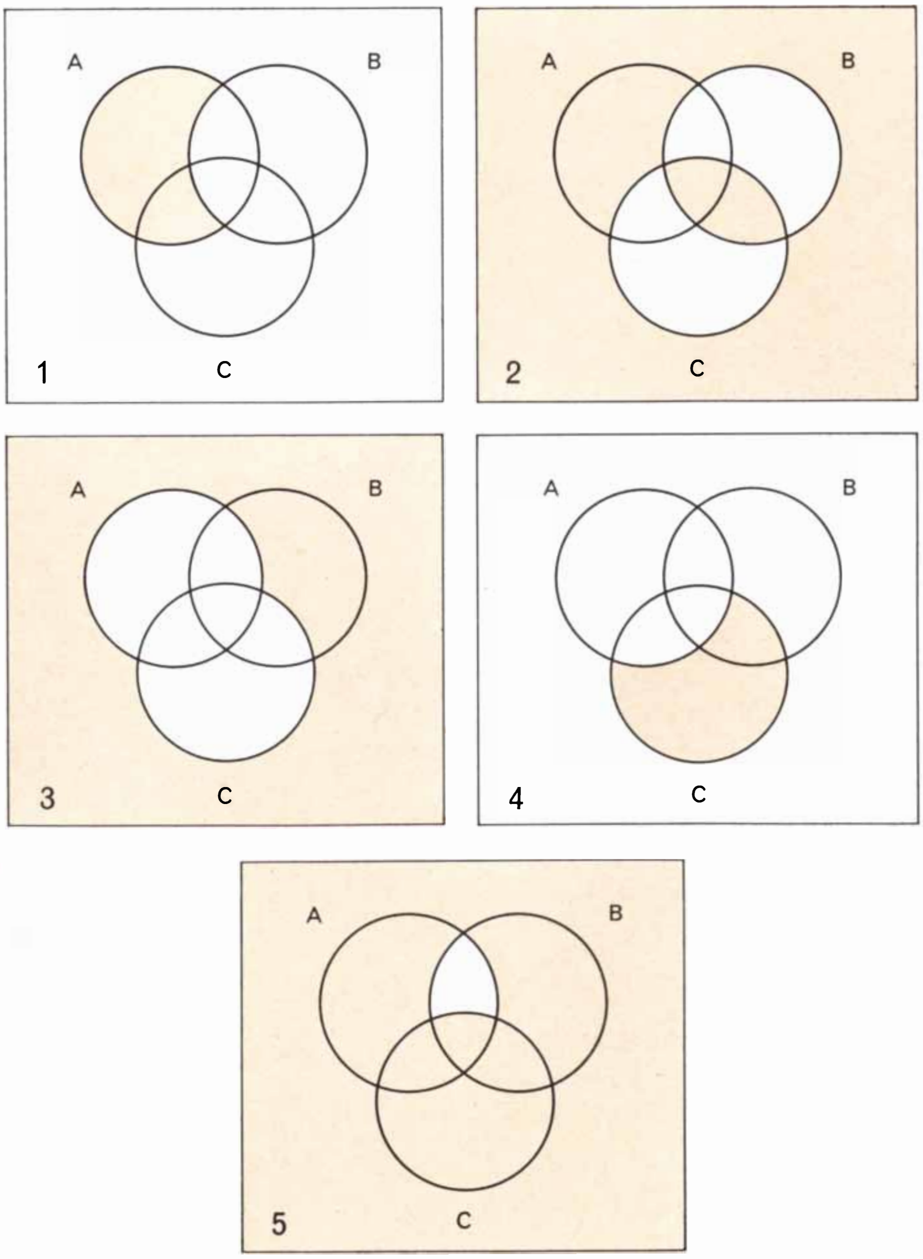
A second problem, also to be answered next month, concerns a little-known lightning calculation trick. Turn your back and ask someone to write down any two positive integers (one below the other), add those two numbers to get a third, put the third number below the second, add the last two numbers to get a fourth, and continue in this way until he has a column of 10 numbers. In other words, he writes 10 numbers of a generalized Fibonacci sequence, each the sum of the preceding two numbers except for the first two, which are picked at random. You turn around, draw a line below the last number and immediately write the sum of all 10 numbers.

The secret is to multiply the seventh number by 11. This can easily be done in your head. Suppose the seventh number is 928. Put down the last digit, 8, as the last digit of the sum. Add 8 and 2 to get 10. Put 0 to the left of 8 in the sum, carrying the 1. The sum of the next pair of digits, 9 and 2, is 11. Add the carried 1 to get 12. Put 2 to the left of 0 in the sum, again carrying 1. Add the carried 1 to 9 and put down 10 to the left of 2 in the sum. The completed sum is 10,208. In brief, you sum the digits in pairs, moving to the left, carrying 1 when necessary, and ending with the last digit on the left.

Can you prove, before the simple answer is given next month, that the sum of the first 10 numbers in a generalized Fibonacci sequence is always 11 times the seventh number?

Three Venn circles are shaded as in the illustration at the left to solve last month's problem about the three men who lunch together. Each of the first four diagrams is shaded to represent one of the four premises of the problem. Superimposing the four to form the last diagram shows that if the four premises are true, the only possible combination of truth values is *a*, *b*, $\sim c$, or true *a*, true *b* and false *c*. Since we are identifying truth with ordering a martini, this means that Abner and Bill always order martinis, whereas Charley never does.

The method of generating 2^n integers to form Boolean algebras, as explained last month, was given by Francis D. Parker in *The American Mathematical Monthly* for March, 1960, page 268. Consider a set of any number of distinct primes, say 2, 3, 5. Write down the multiples of all the subsets of these three primes, which include 0 (the null set) and the original set of three primes. Change 0 to 1. This produces the set 1, 2, 3, 5, 6, 10, 15, 30, the first of the examples given last month. In a similar way the four primes 2, 3, 5, 7 will generate the second example, the $2^4 = 16$ factors of 210. A proof that all such sets provide Boolean algebras, when the algebra is interpreted as explained last month, can be found in *Boolean Algebra*, by R. L. Goodstein (Pergamon Press, 1963), page 126, as the answer to problem No. 10.



Venn-diagram solution to last month's martini problem

Racing–the ultimate proof

■ No proving ground can duplicate the elements which make competition the final test of a car's performance. The rivalry of premier drivers, the unexpected moments, the constant stress on the entire machine, and the incentive to win are present only in racing.

Research, not publicity, has been the prime objective of Porsche's competition program since the firm's founding. Win or lose, Porsche races to prove out engineering and design concepts under the toughest of all possible conditions.

Take one example. The Sportomatic semiautomatic transmission was installed in a Porsche 911 and raced in the Marathon de la Route, 84 hours over the demanding Nurburgring course. It met the test. The car won.

Porsche prototype racers, last year, won the Daytona 24-hour; Sebring 12-hour; Targo Florio; Nurburgring 1,000 kilometer and other major races. The earlier developments perfected in these unique cars brought victory to virtually stock Porsche sedans in the Trans-American championship and to hundreds of amateur owner-drivers who race their own Porsches.

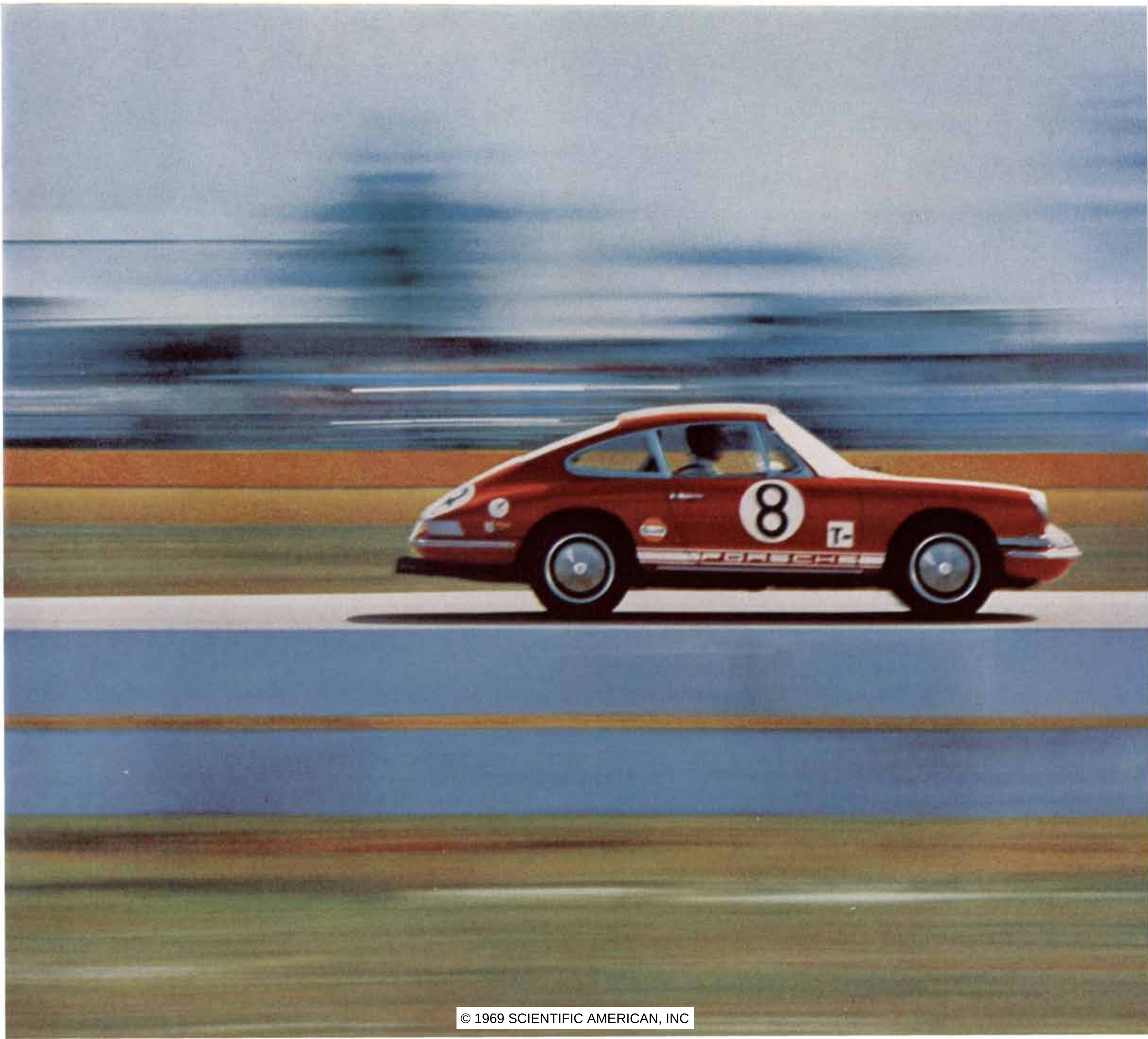
Not all Porsches are raced, of course. But the Porsche you drive is raceworthy; able to take the punishment of high speed racing. Engine, brakes, suspension, electrical system—the total design—are based on race-bred research and built to racing standards.

Can a car be built too good for everyday use? Porsche doesn't think so.

If you're serious about your driving, you can have a lot of fun driving a Porsche, the car that's good enough to race.

Prices start at about $5,100, East Coast P.O.E. See your Porsche dealer or write to the Porsche of America Corporation, 100 Galway Place, Teaneck, N.J. 07666.

PORSCHE



Now, the muscle in the mushroom can turn deserts into wheatfields, light the world, feed it, and fuel it!

(The steels are ready whenever you are)

Man's most horrendous creation is being transformed from slaughterer into servant.

Already, nuclear power is supplementing traditional means of generating electricity for electric-utility companies. The huge quantities of heat released can be put to work to desalt sea water. Resulting fresh water could make the arid regions of the earth bloom into fertile fields to help feed a hunger-haunted world.

Now comes Project Gasbuggy, the Atomic Energy Commission's first experimental project in underground engineering. Gasbuggy used nuclear explosives to free natural gas from tight rock formations that would otherwise have held billions of cubic feet of fuel captive forever.

In other activities, nuclear engineers are working to release oil from shale, and to gouge out giant excavations for canals and harbors with nuclear explosions.

Clearly, nuclear energy is on its way to becoming a major industrial tool. Scores of American companies are intensively engaged in improving and broadening the use of nuclear power for the benefit of mankind.

With this comes a new axiom: Where nuclear energy goes, there goes steel — housing it, channeling it, controlling it. The tough, rugged steels needed in the future are ready now, because Republic Steel has anticipated future needs for the new and improved steels that will serve mankind in harness with nuclear power.

Republic, for example, has a new high strength steel originally developed to withstand the high external pressures found deep in the ocean. By test, this new steel has been found to be equally dependable in containing the terrific internal pressures built up in nuclear reactor vessels.

At this moment, the long reach of steel from Republic is probing into areas wherever man's imagination needs it — from beneath the land to beyond the moon, from the heartbeat of man to the drumbeat of defense. Republic Steel Corporation, Cleveland, Ohio 44101.

You Can Take the Pulse of Progress at

REPUBLIC STEEL

CLEVELAND, OHIO 44101

Steel

THE AMATEUR SCIENTIST

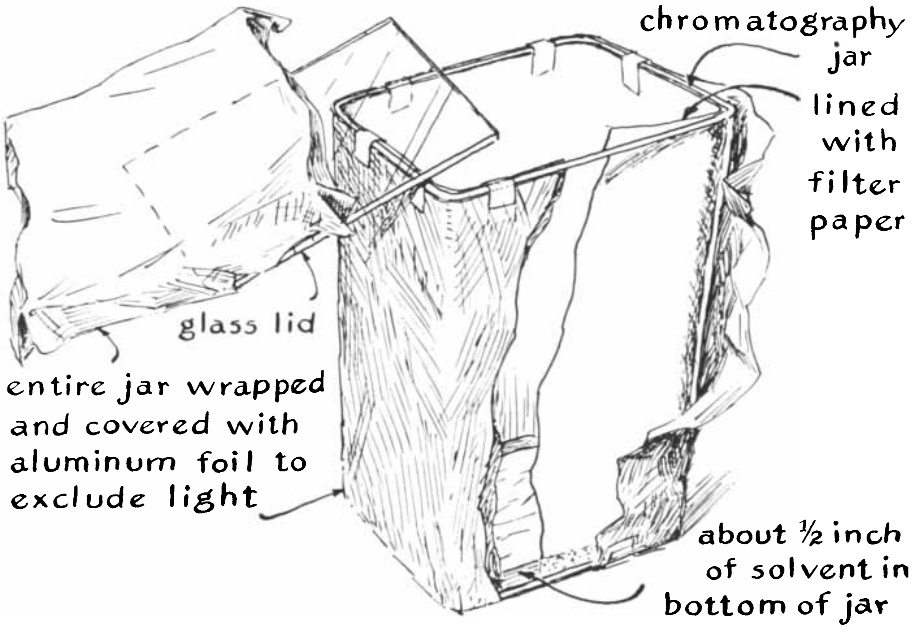
Various kinds of chromatography, especially the thin-layer method

Conducted by C. L. Stong

A colorful experiment involves the separation of pigments in a complex mixture. From each of several kinds of soft white paper cut a strip about an inch wide and five inches long. The materials might include paper toweling, a piece of soft writing paper and the margin of a newspaper. Across each strip, an inch from one end, deposit three evenly spaced drops of food coloring. Use three different colors on each strip. Let the coloring dry.

Select a tall drinking glass for each strip and put an ounce of tap water in each glass. After the drops of coloring have dried suspend a strip in each glass so that the edge at the colored end just touches the water. The paper can be suspended by a toothpick pushed through the upper end of the strip and laid across the mouth of the glass. Water drawn into the strip by capillary attraction will migrate to the top of the paper, carrying most of the color partway. If your drops of food coloring are red, blue and green, the red dye will probably spread completely across the strip and migrate as a band of color that diminishes in intensity toward the bottom. The blue will move similarly, but it will probably separate into two ragged bands, one of bright blue and one of deep red, showing that it is a mixture of colors. The green will separate into adjoining patches of bright blue and lemon yellow.

The order in which the colors separate depends partly on the nature of the paper. On the strip cut from newspaper the blue patch will most likely lead the parade, but blue will probably trail red on the strip cut from the paper towel. The colors may fail to separate on white writing paper. If so, prepare an identical strip and dip its lower edge in a solution made by mixing one part of household ammonia with three parts of rubbing alcohol. The colors will then doubtless separate cleanly, with blue in the lead. The behavior of the mixture depends on the fluid as well as on the paper.



Materials for paper chromatography

Better separations can be achieved with strips of white blotting paper. The colors migrate on this material in the form of oval or comet-shaped patches, whereas the colors on newsprint and paper toweling tend to form in irregular bands with ragged edges. The clearest separations can be accomplished with Whatman's No. 1 filter paper, which is available from dealers in chemical supplies. On 10-inch strips of this material the colors separate in distinct, widely spaced patches. You can cut patches from the strips and recover the pure dyes by soaking each patch in water, removing the paper and evaporating the solution.

Inks can be analyzed by the same technique. Most inks, particularly the brown and blue-black varieties, are mixtures of dyes. Inks can usually be separated on Whatman filter paper with a wash fluid that consists of six parts of *n*-butyl alcohol, two parts of rubbing alcohol and two parts of ammonia water that contains 34 grams of ammonia per liter of water. Bank checks that have been altered by the addition of words or numbers can often be detected by soaking off the suspected ink, separating its dyes and comparing them with the dyes of the other ink on the check.

This method of analyzing mixtures is known as chromatography and is based on a technique invented in 1906 by Michael Tswett, a Russian botanist working at the University of Warsaw. Tswett packed a glass tube with calcium carbonate and washed through it, using petroleum ether as a wash fluid, pigments extracted from green leaves. The pigments separated into distinctly colored zones as they migrated down the column of calcium carbonate. Impressed by the colors, Tswett named the array a chromatogram and the procedure chromatography.

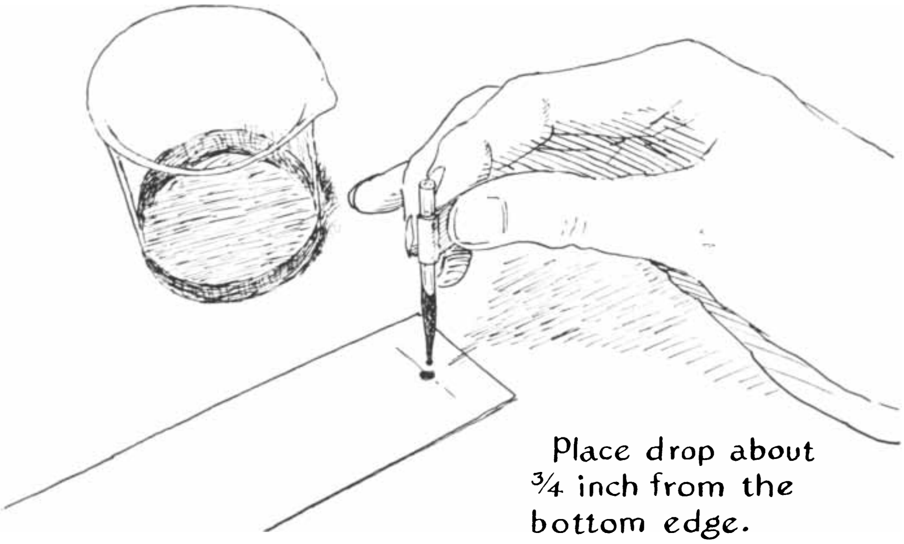
Why do substances that start as a mixture at one point on a sheet of paper, or at the top of a Tswett column, migrate at differing rates in the direction of the wash fluid's movement and ultimately concentrate in distinct zones? Specialists explain that two forces act on the substances. The moving fluid exerts a force that tends to drive the substances in the direction of the flow. The second force is essentially electrical; in the case of paper strips it exists between the moving substances and the moist fibers of cellulose. This binding force—adhesion—tends to oppose the force exerted by the fluid and hence to arrest the movement of the substances.

Periodically, however, adhering molecules of the substances are thrown into the moving stream by the ceaseless vibration that characterizes matter at all temperatures above absolute zero. Some substances are more strongly attracted to cellulose than others; when they are thrown into the stream of wash fluid, they quickly find a new resting place on the cellulose downstream. Substances less strongly attracted by the cellulose come to rest at greater distances downstream. As a result particles that make a series of long hops soon outdistance their more sluggish neighbors.

The chromatographic column invented by Tswett can be packed with cellulose as well as with calcium carbonate. Indeed, it can be packed with almost any powdered substance: sugar, alumina, silica and so on. Experiments have demonstrated that almost any mixture of substances can be separated by packing a Tswett column with a suitable adsorption material and washing the column with a fluid. The column must be uniformly packed or the components of the mixture will collect in poorly separated zones of irregular shape.

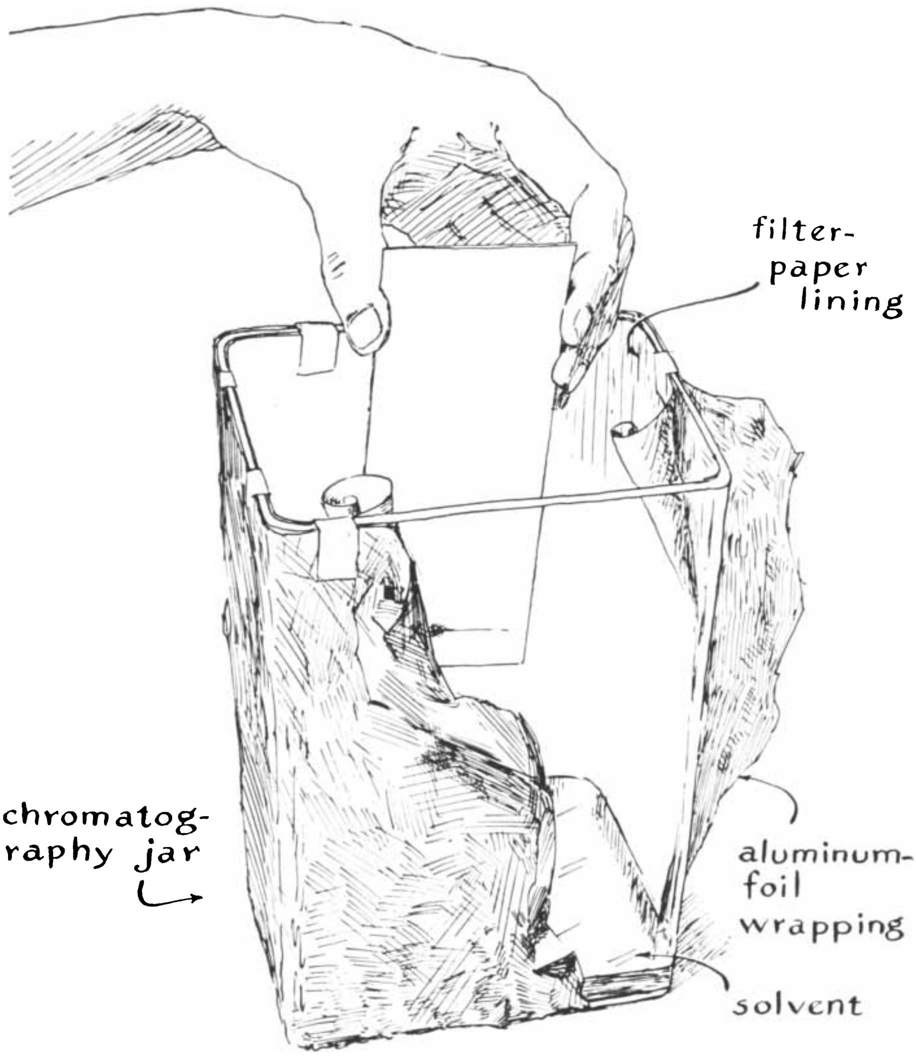
Packing a column, however, is something of an art and is both difficult and time-consuming. For this reason chromatography using paper became universally popular following its discovery during World War II by the British investigators Raphael Consden, A. H. Gordon and A. J. P. Martin. A number of applications were found for the technique even though many mixtures will not separate on cellulose.

Ultimately a third chromatographic technique was devised. It combines the versatility of the chromatographic column with the convenience of the paper strip. Although this technique, known as thin-layer chromatography, was first described in 1938 by the Russian chemists N. A. Izmailov and M. S. Shraiber, its development was delayed until the mid-1950's. Izmailov and Shraiber coated a glass plate with a slurry of alumina and water two millimeters thick and used it for the separation of medicinal preparations of vegetable origin.

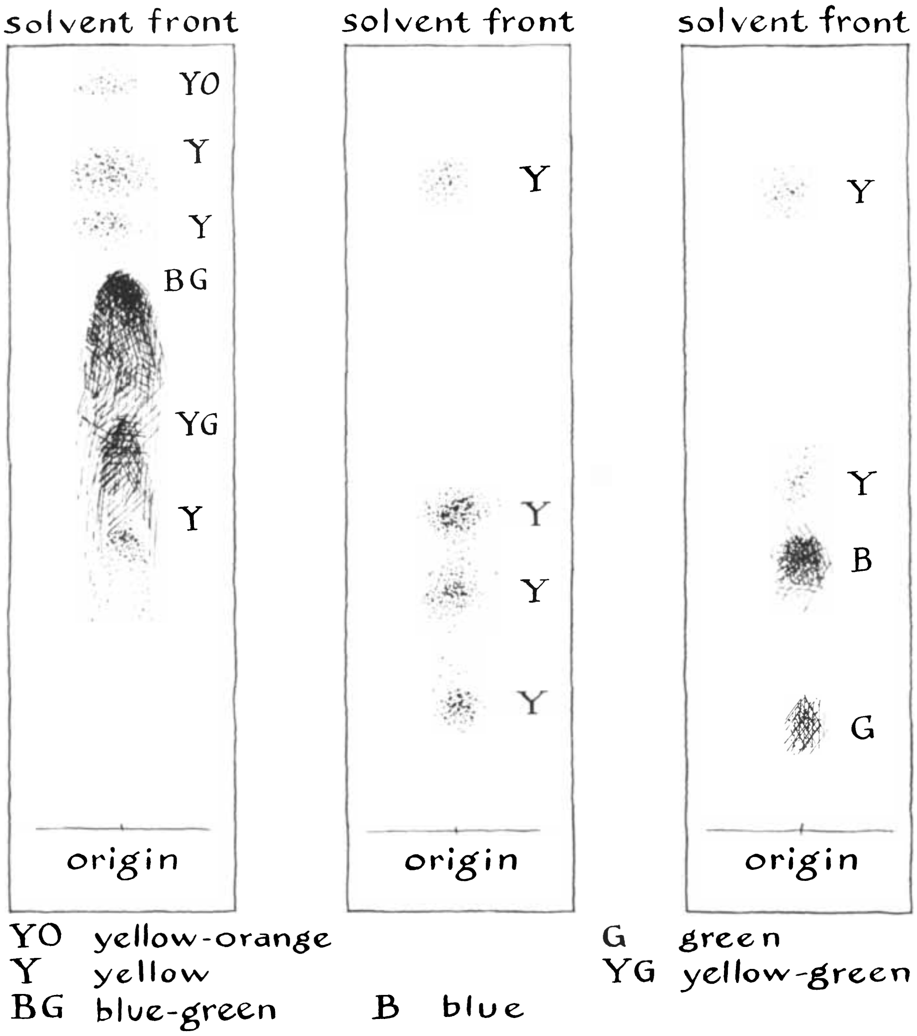


Applying a specimen to a paper strip

The technique suffered from various limitations. Difficulty was experienced in producing uniform coatings, and the plate had to be used in a horizontal position. As a result thin-layer chromatography was all but forgotten until an improved method was discovered in 1951 by Justus G. Kirchner of the U.S. Department of Agriculture. Five years later Egon Stahl of the University of the Saarland developed methods for applying to glass plates a uniform layer of silica gel, an effective adsorbent, that was held in place by plaster of paris. The final step



Putting a strip in the jar



Separated pigments of spinach leaves

of transforming thin-layer chromatography into a routine procedure was taken six years ago when the Eastman Kodak Company developed methods for precoating flexible sheets of transparent plastic with a variety of adsorbents, including silica gel, alumina and cellulose. A number of firms now manufacture such films, which can be bought from dealers in scientific supplies.

An introductory experiment in the use of thin-layer chromatography has been specially developed for amateurs by Gunter Zweig of the Syracuse University Research Corporation and Joseph A. Sherma of Lafayette College. Essentially they update Tswett's classical analysis of the chloroplast pigments in green leaves. Zweig and Sherma describe the experiment as follows:

"Plant pigments are destroyed easily by heat and light. The solvents used for extracting them should be ice cold. When you make extractions, work as quickly as possible, avoid direct sunlight and store the final solution in the dark. Use distilled water in all experiments.

"Obtain some fresh spinach. Cut away the stem and midribs with scissors, chop the rest into small pieces and weigh out about two grams of the small pieces. Place the fragments in an electric blender, add 40 milliliters of acetone and let the blender operate at its highest speed for two minutes. (Caution: The motors of most blenders generate sparks that may ignite spilled acetone. Blenders equipped with screw caps are the safest.) Centrifuge the pulp for a few minutes at 2,000 revolutions per minute to separate the solids from the liquid. Pour the green fluid into a 250-milliliter separatory funnel. Do not use grease on the stopcock or the stopper of the funnel. Add to the extract 40 milliliters of petroleum ether and 100 milliliters of brine that contains 10 percent by weight of table salt. For extracting plant pigments use petroleum ether that boils between 20 and 40 degrees Celsius. Shake the funnel and allow the layers to separate. The top layer (dark green) contains most of the chloroplast pigments; the bottom layer is colorless. The addition of salt reduces the solubility of the pigments in the water-acetone solution. If an emulsion forms, try to dissipate it by using more brine.

"Discard the bottom layer. Wash the remaining extract down the sides of the funnel with two successive 100-milliliter portions of water. Swirl rather than shake the contents in the funnel. Pour the washed solution through the top of the funnel into a round-bottomed flask. The solution must be evaporated to dryness under vacuum. Close the flask with a perforated rubber stopper. Insert a glass tube in the perforation and connect it by flexible tubing to a water aspirator or some other type of air pump. While the fluid is evaporating swirl the flask gently in a bath of tepid water at a temperature below 40 degrees C. Just before making the chromatographic analysis dissolve the dry residue in one milliliter of petroleum ether that boils at a temperature of 60 to 110 degrees C. Label the container Extract I. This final solution contains all the chloroplast pigments: the green chlorophylls as well as the yellow xanthophylls and carotenes.

"As an aid to identifying the pigments, prepare a second test solution from which the chlorophylls are eliminated. Again extract the pigments from two grams of spinach leaves with acetone and, after centrifuging, transfer the fluid to a 500-milliliter separatory funnel. Add 25 milliliters of methyl alcohol that contains 2½ grams of potassium hydroxide. The solution will turn brown and then green. Let the mixture stand for 20 minutes but swirl the funnel occasionally. Add 40 milliliters of a solution made of equal parts of diethyl ether and petroleum ether. Add 200 milliliters of a solution that contains 10 percent by weight of table salt. Shake the funnel to mix the contents. Let the mixture settle. The solution will separate into a golden yellow layer that floats on a dark green one. Discard the green layer. Add 200 milliliters of fresh water, swirl the flask, let it settle and discard the water. Repeat the washing. Transfer the yellow solution, thus washed, to a round-bottomed flask and evaporate it to dryness under vacuum. The extract now contains yellow pigments only. Just before making the analysis dissolve the concentrated pigments in one milliliter of a solution that contains equal parts by volume of diethyl ether and petroleum ether (boiling temperature 60 to 110 degrees C.). Label the container Extract II.

"The chromatographic analysis of Extract I is made with Eastman Chromagram sheets that are coated with cellulose. They are available from Eastman Organic Chemicals, Eastman Kodak

Company, Rochester, N.Y. 14650. The sheets are supplied in eight-inch squares that can be cut into one-inch strips eight inches long.

"By means of a micropipette apply five microliters of Extract I to the center of a strip about 3/4 inch from one end. Let the spot dry in the dark. In the meantime make a wash solution (or solvent) of petroleum ether, benzene, chloroform, acetone and isopropyl alcohol in the proportions of 50 : 35 : 10 : 5 : .17 parts by volume respectively. Before making up the solution mix the chloroform with an equal quantity of water in a separatory funnel, swirl the mixture, let it settle and separate and then discard the upper water layer. Wash the chloroform twice by this procedure to remove traces of ethanol, then dry the chloroform over anhydrous calcium sulfate ('Drierite').

"Line a glass container, preferably a rectangular jar 10 inches deep, with filter paper. Cover the outside of the jar with aluminum foil to exclude light. Pour wash solution into the jar to a depth of about 10 millimeters and allow the filter paper to become saturated. Stand the Chromagram sheet in the jar, with the spotted end down. The top of the solvent must be well below the level of the origin line on the sheet. Close the jar with a snug-fitting lid. In about 45 minutes the solvent will have migrated almost to the top of the sheet. Remove the sheet and let it dry. Do not expose it to strong light. The order of pigment migration, beginning from the bottom, will be neoxanthin plus chlorophyll-b (yellow and yellow-green), chlorophyll-a (blue-green), violaxanthin (yellow), lutein (yellow) and carotenes (yellow-orange).

"Extract II is analyzed with Eastman Chromagram sheets that are coated with silica gel. Follow the procedure used in the preceding analysis but substitute for the wash solution a mixture consisting of three parts of isooctane and one part each of acetone and diethyl ether, by volume. The pigments will separate into four yellow zones that are, beginning at the origin, neoxanthin, violaxanthin, lutein and carotene.

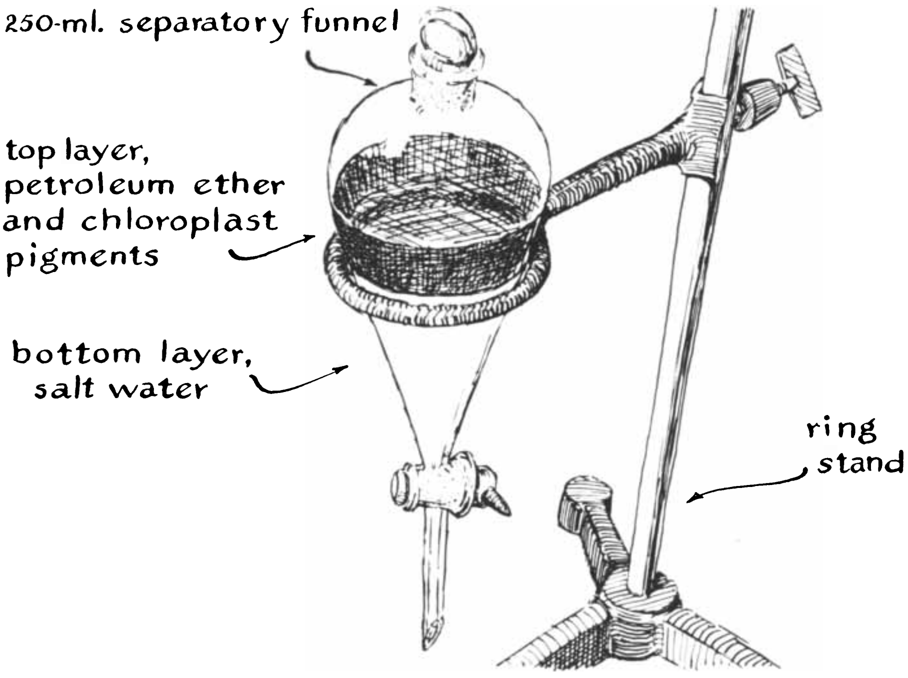
"The identification of these pigments can be confirmed by exposing the developed sheet to the fumes of concentrated hydrochloric acid. Pour a little acid into a shallow rectangular container and cover it with the Chromagram sheet, with the coated side facing the acid. The yellow patch of neoxanthin will turn a bluish green, the violaxanthin will turn blue and the remaining pigments will stay yellow. An additional analysis can be made by scraping the colored patches from the sheet and mixing the scrapings with ethyl alcohol. The pigments will enter solution. A spectrophotometric analysis can then be made of the stained alcohol by means of a spectrophotometer of the kind described in this department last year [see "The Amateur Scientist," May, 1968].

"Other plant groups such as diatoms, dinoflagellates, red algae and so on can be investigated in the same way, although the extraction procedures and the wash fluids may have to be modified in some cases. The pattern of pigments obtained can aid in the taxonomic classification of the plants. One cannot predict the best procedure for each species of plant. Chromatography is largely an empirical technique."

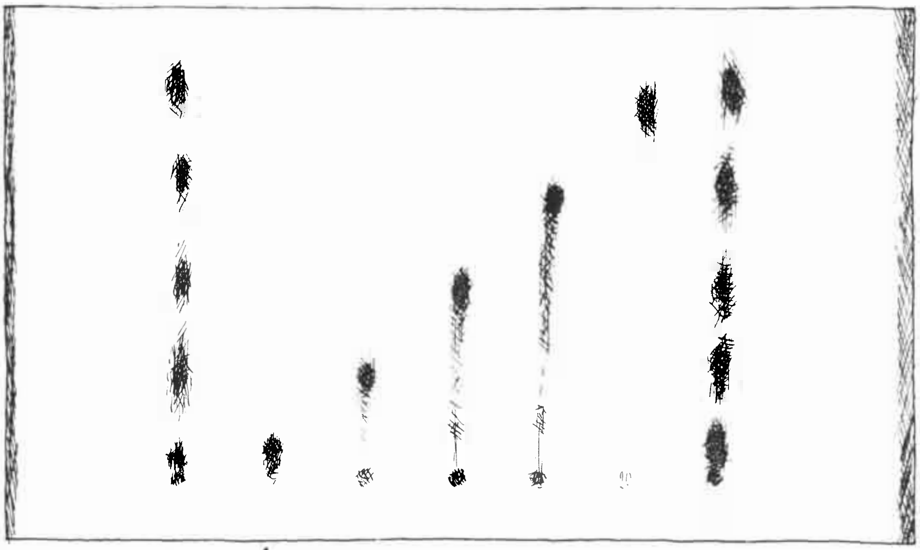
Chromatography is used routinely for separating colorless substances of numerous kinds that occur as natural mixtures, including the amino acids. Albert Baitsholts and Richard Ardell of Eastman Organic Chemicals have found that the thin-layer chromatographic technique separates amino acids in less time and more effectively than the older techniques do. Baitsholts and Ardell have developed a two-part experiment that amateurs can perform. The first part separates free amino acids on thin cellulose layers, and the second part separates derivatives of the acids on silica gel. The derivatives are prepared by letting the acids react with 1-dimethylaminonaphthalene-5-sulfonyl chloride. The products are called "dansyl" derivatives and are of particular interest because they are highly fluorescent. Dry amino acids, as well as solutions of their dansyl derivatives, are available from the Mann Research Laboratories, Inc., 136 Liberty Street, New York, N.Y. 10006. Baitsholts and Ardell write:

"Sample mixtures of amino acids are prepared for analysis by placing one millimole of each of the acids in a container and dissolving the mixture in 100 cubic centimeters of water to which .364 gram of hydrochloric acid has been added (.1 normal solution). We make two batches. Mixture *A* contains lysine, aspartic acid, glycine, threonine, proline, valine, tryptophan, phenylalanine and leucine. Mixture *B* contains cystine, histidine, arginine, serine, glutamic acid, alanine, tyrosine, methionine and isoleucine. These mixtures are representative of the combinations encountered routinely in samples taken from living organisms.

"One-microliter samples of the test mixtures are spotted on a Chromagram sheet that is coated with cellulose. The wash fluid is prepared by shaking together a mixture of butanol, acetic acid and water. The mixture separates after standing for a few minutes. The top layer is used as the wash fluid. Place the wash fluid in a rectangular container to a depth of a few millimeters, as in the preceding experiments, and saturate the filter paper that lines the jar to produce a saturated atmosphere. Insert the Chromagram sheet in the jar and let it develop for about two hours. Remove the



Details of the separatory funnel



Separated dansyl derivatives

sheet and let it dry for about 15 minutes. To the eye the sheet will appear blank. The acids must be made visible by chemical treatment.

"Spray the coated side of the dried strip with a .2 percent solution of ninhydrin in acetone. Colors will form at the sites of resolved acids when the sprayed sheet is heated in an oven at 100 degrees C. for a period of from two to five minutes. The color can be enhanced by adding collidine (2,4,6-trimethylpyridine) to the ninhydrin-acetone solution just prior to the spraying operation. Use one drop of collidine per 100 cubic centimeters of the ninhydrin-acetone solution.

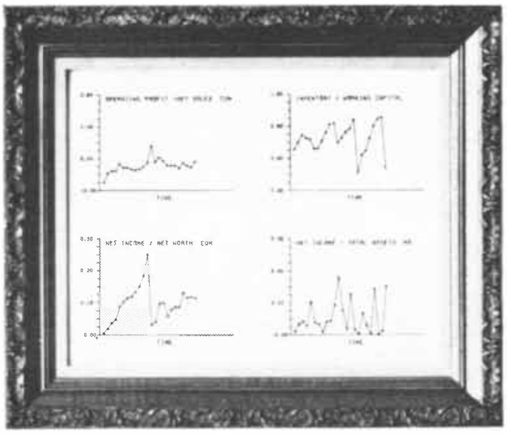
"In the second experiment we made a chromatogram that displayed five dansyl derivatives both as individual substances and as the separated components of a mixture. This scheme enabled us to identify the separated components of the mixture by comparing the distance each fraction migrates with the migration distance of each known substance. The analysis is made on an uncut Eastman Chromagram sheet (an eight-inch square) that is coated with silica gel. Just before use the sheet must be heated for 15 minutes in an oven at 100 degrees C. to activate the coating. The mixed derivatives are applied to adjacent corners of the uncut sheet, say an inch from the bottom and an inch from the sides. Make the mixture by successively applying to each location one microliter of each of the five substances. Allow each specimen to dry before applying the next one. Make a record of the sequence, such as 'Left to right: dns-cystine, dns-serine, dns-glycine, dns-alanine and dns-valine.' Reference to this record will enable you to identify the separated components of the mixture.

"Prepare the wash solution by adding to 80 parts of benzene 20 parts of pyridine and five parts of acetic acid, by volume. Line the rectangular jar with filter paper that has been saturated with wash solution, immerse the lower edge of the sheet in solution to a depth of approximately 10 millimeters and cover the jar. The solution will migrate a sufficient distance within 30 minutes. Remove it and let it dry in the air. The coating will show no visible trace of the derivatives.

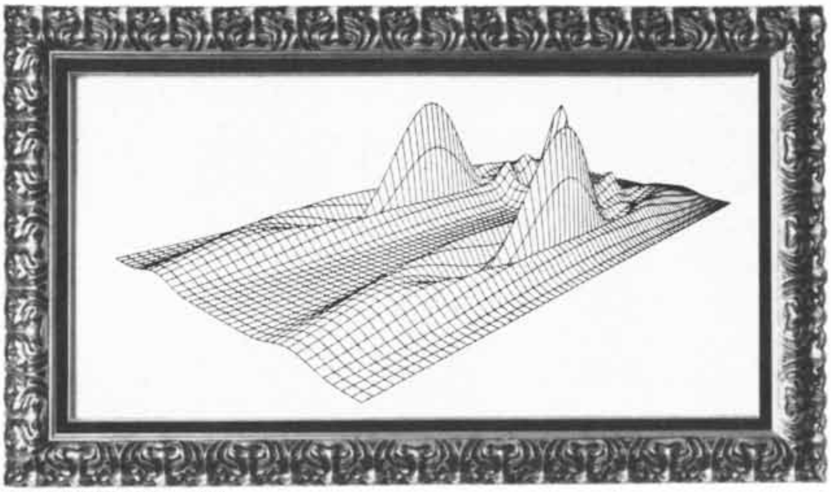
"To see the resolved specimens, examine the sheet under an ultraviolet lamp in a dimly lighted room. Any kind of ultraviolet lamp can be used because the specimens fluoresce when exposed to either long or short ultraviolet rays (254 or 360 millimicrons). Observe that the serine and glycine derivatives are clearly resolved, although they do not separate as free amino acids. It can also be shown that other amino acids, such as proline and phenylalanine, separate well as free acids but migrate close together as dansyl derivatives. Differences of this kind suggest how thin-layer chromatography, combined with the procedures of classical chemistry, has become a powerful tool for unscrambling mixtures of complex substances."

Most of the fluids and reagents used in these experiments are toxic and some are highly flammable. Avoid inhaling the fumes. Work in a well-ventilated room and keep all materials away from sparks and open fires.

Teach your computer to draw



FINANCIAL TREND CHARTS

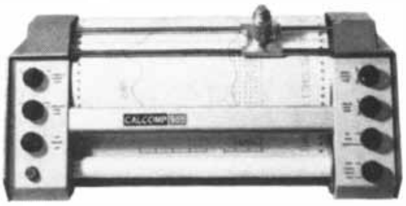


PERSPECTIVE DRAWINGS

ARCHITECTURAL DRAWINGS

CONTOUR MAPS

with a CalComp Plotter



With a CalComp Plotter and CalComp Software, your computer can draw.

Anything.

Accurately, efficiently, economically, automatically.

CalComp offers 13 digital plotters. Drum, Flatbed and Microfilm. Plus a complete range of on-line/off-line plotter controllers. Also the functional and application software required.

For more information call or write Dept. N3.

CALCOMP
LEADER IN COMPUTER GRAPHICS

California Computer Products, Inc., 305 Muller Street, Anaheim, California 92803 Phone (714) 774-9141

BOOKS

On teaching biology in a biological revolution

by Salvador E. Luria

BIOLOGY, by Helena Curtis. Worth Publishers, Inc. ($11.50).

The publication of this new biology textbook is a welcome occasion to take a fresh look at some aspects of teaching introductory biology in an American college. I have two reasons for wishing to formulate my ideas on the subject. First, my qualifications for doing so are both minimal and optimal: I have been involved in teaching introductory biology for the first time in my life after teaching other courses in biology for the past 25 years. After being for so many years a client of the introductory biology courses, which prepared the ground for the more specialized crops I tried to sow, I am now myself a tiller of the (almost) virgin soil.

Second, the work I shall review here is a somewhat unusual textbook of biology. It represents an experiment in collaboration, having been written by an experienced science journalist working under the sponsorship of a young publishing company and with the expert advice of a distinguished roster of consultants who are active investigators in different areas of biology. The result is in my opinion the forerunner of a new crop of biology textbooks in which several practitioners of experimental biology cooperate in producing a single work (not a series of pamphlets) that addresses the beginning biology student with a voice from the laboratory rather than from the classroom.

These two reasons, my involvement in teaching introductory biology and the trend toward a new kind of collaborative textbook writing, are not unrelated. The process that creates the demand for such textbooks is the same one that brings the likes of me—microbiologists, biochemists, biophysicists, geneticists and physiologists—to the beginning-biology classroom. It is the outcome of developments that have taken place in biology in the past quarter-century (more particularly since 1953, when molecular biology came of age).

Why are these developments critical in the teaching of biology? They are critical in the same way, I believe, that developments in physics and physical chemistry between 1925 and 1950 were critical in reshaping the teaching of chemistry all the way to the beginning college course level.

The developments in biology do not represent a "revolution" but are rather a fulfillment of the biological revolution that began in the 19th century with Darwin and Mendel. It is worth reflecting on the nature of these events if we are to understand the demands they make in the classroom on student, teacher and textbook.

What has happened is that the structure of biology as a science has not only become unified but also been made explicit. What do I mean by this? Biology as a science was unified by the theory of evolution. Since Darwin's time we have known that biology is the science of the single historical process—evolution—that has created the array of organisms that exist today, flowing from past arrays of organisms and representing the unique source for all future forms of life. Biology is history in the same sense that the record of human affairs is history—the history of successful or abortive lines of descent, of new inventions (such as photosynthesis or the spinal cord), of their trial in the marketplace of natural selection, of the conquest of new ecological niches and of the effects of such conquests. It is the story of how the present state of the living continuum is related to its past states and of how the present acts as the gateway through which the future is formed.

The biologist has one advantage over the historian. He knows the motive force of his process—natural selection—and he understands its workings and the material on which it works much better than the historian knows or understands the forces that shape the course of human events. Natural selection works on organisms and determines the frequencies with which an individual organism existing at time t_1 will be represented by its heredity in the array of organisms that will exist at time t_2. The evolutionary record, however incompletely it is known, is unified by our certainty of its historical continuity and by our awareness of the directing process.

Yet the theory of evolution by natural selection can exist, and in fact it did exist for a long time, in the absence of any explicit statement about the material substrate of the underlying continuity. We can state that flagellates were probably ancestors of algae, or that a planarian worm may have been the common ancestor of most of today's higher animals, without postulating anything about the molecular continuity between the members of these lines of descent. To put it another way, evolution as the unifying principle of biology can be studied and verified without knowing any genetics; the probability that organism X existing at time t_1 will be represented by n descendants at time t_2 can be defined without knowing how X conveys its hereditary specifications to its heirs.

The emergence of genetics was the next step in biology. The substrate of the continuity of life processes now became the Mendelian factors, or genes, whose persistence and mutation were revealed (indirectly!) by the observation of detectable biological traits and by statistical inferences on their combinatorial behavior. This "Mendelian view of the world," to use James D. Watson's phrase, was still basically formal. It first acquired material content with the identification of chromosomes as the carriers of genes in specific sequences. Next came the conversion of the Mendelian factors into the Muller-Delbrück genes: large molecules with a precise individual structure, whose mutations were changes in that structure. In time these molecular genes became identified with segments of nucleic acid fibers characterized by specific sequences of nitrogenous bases. With the clarification of the molecular structure of the gene came the understanding of gene function and of its thrilling evolutionary aspect: the universality, or at least near-universality, of the genetic code. This proved that evolution had perfected and preserved very early (and had since then jealously preserved) not only a unique structural basis of specificity but also a unique way of using it.

This is what I mean when I say that the structure of biology has become unified and explicit. It is the science of a historical continuum, whose continuity is embodied in the specific mutable structure of the genes, and whose evolution occurs by natural selection acting on individual organisms that are the integrated expression of specific sets of genes. What biologists do is to explore the nature and function of the genes, their expression in the growth and function of organisms, the production of new generations of organisms from the preceding ones, the array of organisms that actually exist, the ancestral relations of existing organisms and the mechanisms of natural selection that bring about a specific state of the living world.

If this picture of the content and practice of biology is accepted, it becomes the task of any biology course, whatever the preparation and goals of the students may be, to enable them to understand the structure of biology and the role the various specialized activities of biologists play in the exploration of the living world. This formulation may seem obvious, particularly to those who are well acquainted with the molecular, genetic and evolutionary aspects of biology. It nonetheless calls for a number of changes in the tradition of certain branches of biology and in the practice of biology teaching. For example, under the impact of molecular genetics biochemistry is cautiously reorienting itself around the central questions of the structure, synthesis and function of specific large molecules and assemblies of large molecules: nucleic acids, proteins and membranes. Biochemistry has just begun to view its traditional concerns with energy metabolism, the interconversion of metabolites and the storage of reserve substances as being important but not central aspects of the chemical drama of life. It was not long ago that one heard statements such as "Life is the story of the flow of energy in the organic world." Today this statement can be regarded as a grave distortion, focusing as it does on the means instead of the goal.

Another example is what is happening to the cell theory, one of the main generalizations of biology. With the growth of molecular biology the cell theory is being converted from a morphological concept to a molecular one. The cell is now viewed operationally as a chemical domain within which specific membrane structures segregate, orient and position the functional components of the chemical machinery.

A third significant change is in the role that certain organisms—bacteria and bacterial viruses—are now playing in biology and must play in any course that wishes to present biology as it really is. Most of our knowledge of the synthesis of proteins and nucleic acids and of the structure, function and regulation of genes comes from work on these organisms. It is clearly important for the student to understand their physiology if he is to grasp the role they play in molecular genetics.

My thesis is that biology has a well-understood structure and that it is possible to organize a course presenting that structure in a straightforward fashion. Such a course, whether it is directed to nonscience students or to future scientists, should convey the feeling of why biology is a pursuit that interests people other than "born" naturalists. Today biology is the frontier science, the science where progress is being made at the fastest pace and whose call attracts some of the most brilliant minds of the younger generation.

In outlining one possible version of such courses I should emphasize my debt to others who have given much thought to these problems, notably Cyrus Levinthal, who designed the first modern course in introductory biology offered at the Massachusetts Institute of Technology, and James Watson, who with his book *Molecular Biology of the Gene* has provided a model for the presentation of molecular biology to students.

I believe it is important to introduce at the outset of a course the fundamental concepts: organism, gene and evolution by natural selection; the organism as the individual protagonist of the evolutionary process and the gene as the material element of continuity in evolution. Then one introduces the cell as the domain of specific chemical processes. The fact that bacteria are single-cell organisms makes it possible to begin our story with the life cycle of a bacterium such as *Escherichia coli,* whose growth processes can be quantified and interpreted in terms of the chemical conversion of nonspecific substrates to specific large molecules. With *E. coli* as a model system it is quite easy to illustrate concepts such as nutritional requirements, the transport of substrates, the role of intracellular pools of metabolic intermediates, the utilization of energy to drive synthetic reactions and the interplay of energy metabolism with biosynthetic processes. The important thing is that by dealing with bacteria one can present these topics in terms of specific experimental measurements and operations even in a course for beginners. The reason is that in bacterial physiology the experimental relations among nutrition, growth and biosynthesis are more immediately obvious than they are in work with higher organisms. Moreover, a comparative discussion of the metabolism of a few bacterial species and of other cell types introduces topics such as photosynthesis and the evolution of metabolic pathways.

Having focused from the beginning on specific large molecules as the goal of the entire metabolic process, one can then discuss the structure of proteins and nucleic acids as the level at which the significant specificities appear. It is confusing in a general biology course to deal with these substances on the same basis as fats or complex sugars, as though molecular size rather than specificity were the criterion. The concepts of protein structure and conformation and the roles of various types of chemical bond in determining them are best illustrated in terms of enzyme function. By discussing the molecular regulation of enzyme catalysis (such as feedback activation and inhibition) one gives the student a feeling for the precision with which natural selection influences specificity at the molecular level. The discussion of protein and nucleic acid structure leads directly to molecular genetics: the replication of DNA, the synthesis and function of RNA, protein synthesis, the genetic code and the regulation of gene action—topics that for the time being are known almost exclusively from experiments on bacteria and bacterial viruses.

The transition from the molecular genetics of bacteria and viruses to the formal genetics of higher organisms is then made by explaining the alternatives of haploid and diploid (one set of chromosomes v. two), the meaning of meiosis (the reduction of two sets of chromosomes to one), the evolutionary implications of heterozygosity (different genes in a pair instead of identical ones) and other concepts of population genetics. A critical discussion of the alternation of haplophases and diplophases in the life

cycle of different organisms clarifies the regularities behind the seemingly bewildering array of life histories. It also introduces in a natural way the dichotomy between plants and animals and provides a rational transition to the topics of development, differentiation and the genesis of organs.

Embryology can now be presented as the answer to the problem of delegating different functions to different organs as the size of the organism increases. Much of animal and plant physiology can be discussed in terms of the differentiation of specialized organs for food processing, movement, sensation, coordination and reproduction. The description of the main plant and animal phyla and their subdivisions, to whatever depth it is carried, then becomes the story of how the problems of differentiation have been solved. Within this framework one can discuss the main areas of evolutionary biology: taxonomy (the rational assessment of relatedness among organisms), paleontology (the study of the record of past solutions) and ecology (the analysis of the elements through which natural selection chooses some favorite sets of solutions). As a complement to such a course some discussion of other historical sciences may be valuable: cosmology and the problem of the origin of life on the one hand and the relation between biological and cultural evolution on the other.

This outline is obviously not meant as a dogmatic prescription. It is just one example of how the contents of a modern biology course can be formulated once we accept certain basic premises of what the structure of biology is. The essential feature is that the organismic and the molecular-genetic outlooks are not placed side by side in uneasy cohabitation but are fully integrated by the persistent use of genetic concepts, by the use of bacteria as model one-cell organisms and by the emphasis on specific large molecules as both the master plan and the production goal of life processes. In my opinion a textbook should concentrate on the kind of dual unity I have tried to emphasize: the evolutionary unity and the molecular one.

How does the Curtis book fare in this respect? In spite of its earnest intentions in the right direction, which are quite unusual in a text for a first course in college science, it is not wholly successful. The unifying role of evolution emerges, but the essential nature of what is unified and how is not made explicit enough to provide the backbone for a fully coherent course. The attention drifts too soon from the simple to the complex; one has gone from molecules to cells to organisms before one knows what the relations among them are. The presentation of biochemical processes, although it is fairly detailed for a text at this level, does not come into focus because it does not keep in view the main goal of these processes: the synthesis of specific large molecules within the chemically controlled environment of the cell. The section on molecular genetics is well organized but it comes too late; it is too isolated from its biochemical background and too weak in its foundation of bacterial physiology to be convincing. It seems more an introduction to an exciting world of work in progress than an account of a main foundation of biology.

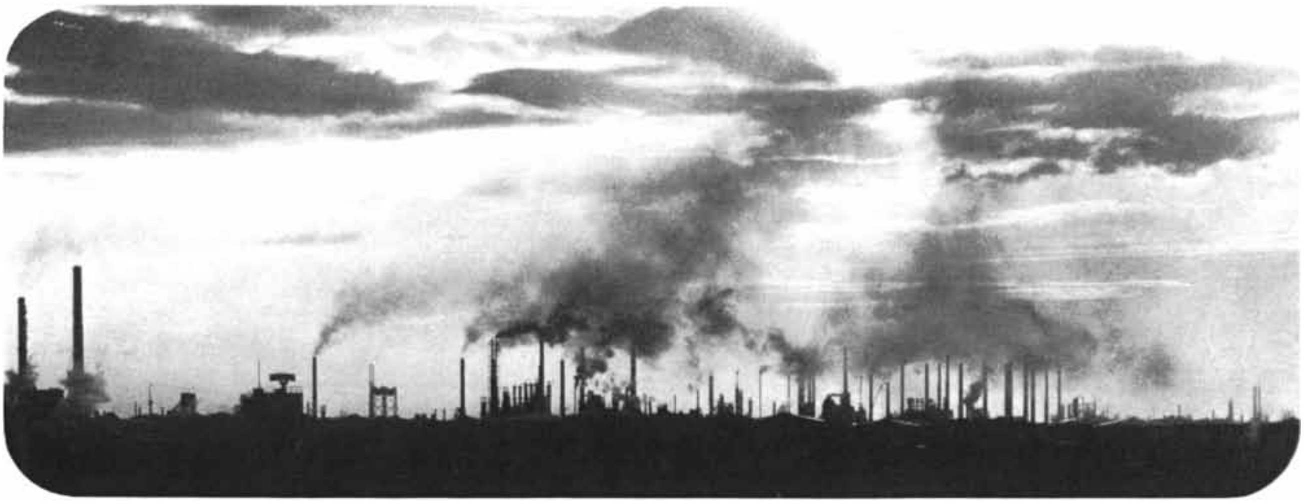
Some features of the Curtis book are excellent. The style is charming, and the writing is uniformly clear and engaging. The passages on development are clearer than they are in most texts, those on the various types of organism are readable and entertaining, and the section on behavior is a welcome experiment in coordinating the psychological aspects of biology with the organismic ones (even though the author is too generous in her comments on some biochemical studies of memory). There are very good explanatory diagrams, particularly those accompanying the electron micrographs. I should like to see an even greater use of diagrams to explain certain key experimental procedures and findings step by step.

Another good feature of Mrs. Curtis' book are the occasional short essays on special topics, which make it possible to interject comments and digressions without breaking the continuity of the text. They should also be useful in teaching students to organize their thoughts clearly and succinctly in scientific writing. All in all, Mrs. Curtis and her publishers are to be complimented on this book, which will certainly see many editions. Possibly some of my comments will be helpful in future revisions. I only wish that all textbooks were written with the same verve, enthusiasm and charm.

Shorter Reviews

by Philip Morrison

CHEMICAL AND BIOLOGICAL WARFARE: AMERICA'S HIDDEN ARSENAL, by Seymour M. Hersh. The Bobbs-Merrill Company ($7.50). THE SILENT WEAPONS, by Robin Clarke. David McKay Company, Inc. ($4.95). UNLESS PEACE COMES: A SCIENTIFIC FORECAST OF NEW WEAPONS, edited by Nigel Calder. The Viking Press ($5.75). All of us, like sheep



A Keener 'Nose' for SO_2

The air we have to live on is becoming more and more polluted by the by-products of industry and transport. Monitoring the level of pollution could be a first step towards its control. Sulphur dioxide has long been recognized as a reliable indicator of the state of air pollution. Information about the SO_2 level in the air, as a function of place and time, could indicate, and even predict, dangerous situations. However, SO_2 detectors have hitherto required too much supervision to be suitable for a large operating system.

When Mr. H. J. Brouwer at Philips Research Laboratories, in consultation with Mr. H. Zeedijk from the Eindhoven Technical University, started the design of an improved SO_2 monitor, he realized that a built-in, reliable chemical calibration source would greatly reduce problems of long term accuracy. Furthermore, the measuring cell proper should require a minimum of maintenance.

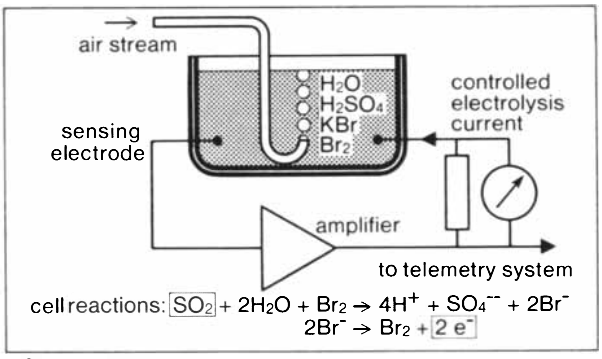
Unlike most of its fore-runners, the new monitor functions on the principle of continuous coulometric titration. When SO_2 enters the cell, it reacts with bromine, thereby lowering the Br_2 concentration and thus the redox potential. This is sensed by the measuring electrode, which, via an amplifier, controls an electric current into the solution. This current replenishes Br_2 by electrolysis of KBr and is a direct measure of momentary SO_2 contents of the air.

The SO_2 calibration source consists of a vessel, with a small hole, filled with SO_2. The hole is covered with a plastic membrane into which SO_2 may dissolve. SO_2 diffuses through the membrane and evaporates at the other side. The product of solubility and diffusion constant, which determines the 'leakage' has a temperature coefficient of only one per cent per degree C. A simple thermostat suffices to keep the required accuracy.

Dust, O_3 and H_2S are filtered from the incoming airstream, because the latter two would react with the cell contents in a way similar to SO_2. An air flow switching system provides for three states- 'zero', 'calibrate', and 'measure'. In the 'zero' position the air first passes an active coal filter, which removes SO_2. The loss of bromine then indicated by the current is due to evaporation. In the 'calibrate' position the cleaned air passes the SO_2-source and is thus polluted to a known extent. This fixes a second point on the linear scale. In the 'measure' position the air stream, with unknown SO_2 contents, is led through the cell.

The detection limit, as well as the zero drift per day, obtained with the new monitor, is down to below 0.01 parts per million, a level typical for a nature park - in a non-volcanic region, of course.

We now have a monitor, able to provide stable, unattended operation for periods up to three months, as required in a national system for automatically monitoring the air above the Dutch polders.



SO_2 measuring cell

In the Research Laboratories of the Philips group of companies, scientists work together in many fields of science. Among these are: Acoustics, Cryogenics, Information Processing, Mechanics, Nuclear Physics, Perception, Solid State, Telecommunications and Television.

This work was carried out at the Philips Research Laboratories, Eindhoven, the Netherlands.

PHILIPS

This bird sanctuary

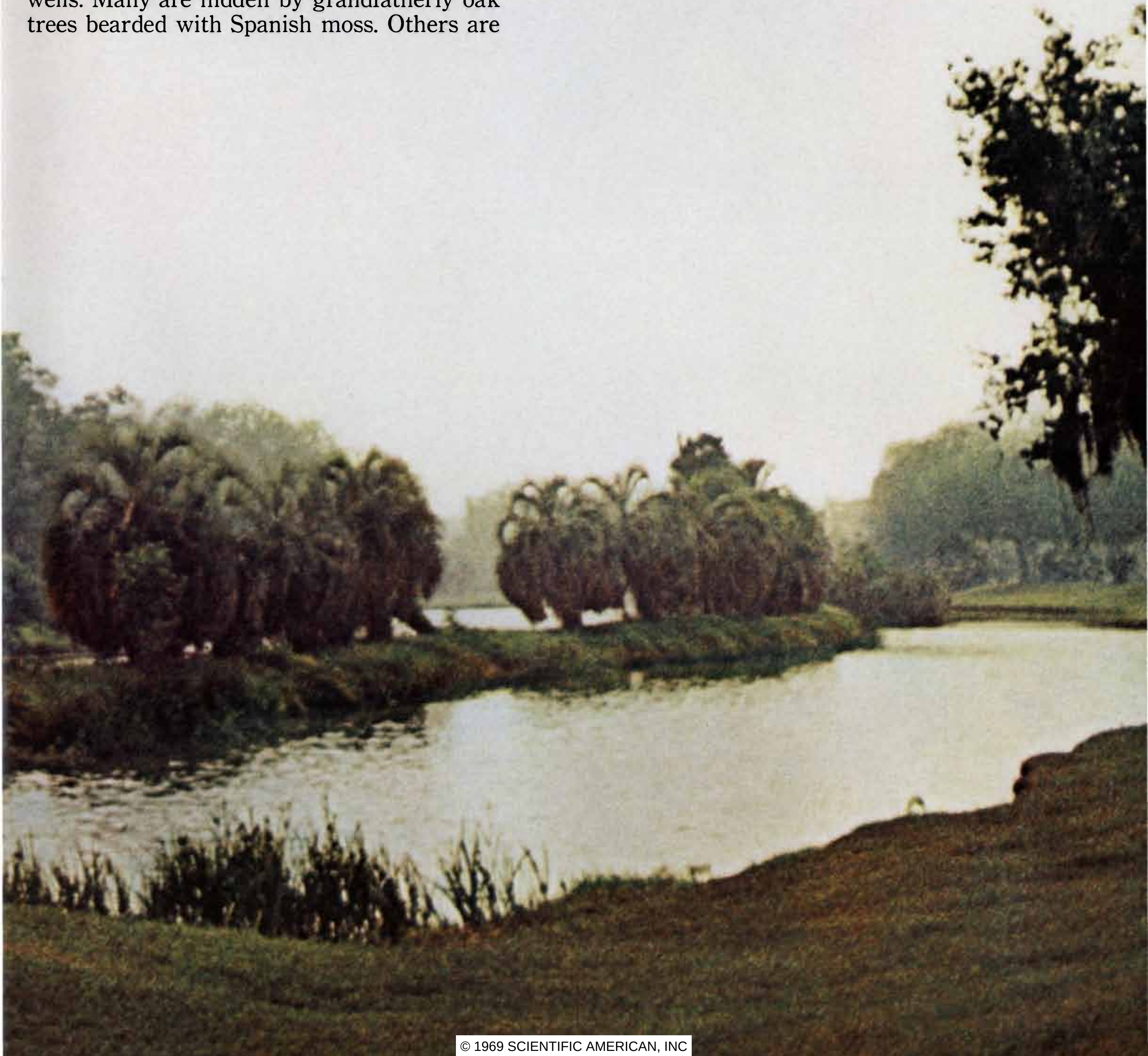
Imagine a tiny green hump of an island in a Louisiana swamp. Its total area is less than five square miles.

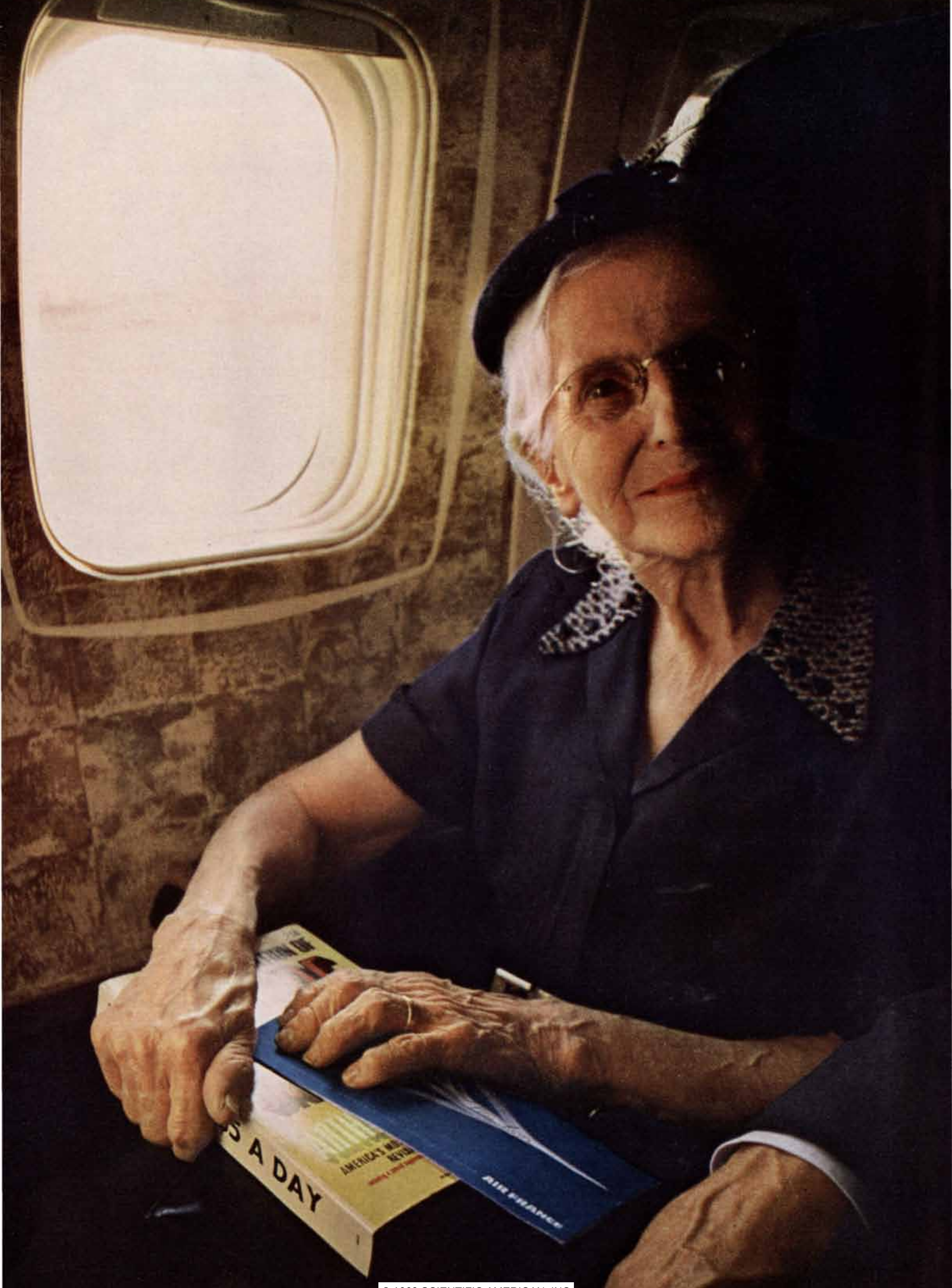
Put two hundred houses on it and seven hundred people. Add one of America's largest rock salt mines, the Tabasco® sauce factory and over a hundred oil wells. And what have you got? Overcrowding?

Quite the opposite. Avery Island seems almost undiscovered. A place for the painter and the poet.

Its bird sanctuary sits in a 200-acre garden. Here you find irises from Siberia. Grapefruits from Cochin. Evergreens from Tibet. Bamboo from China. Lotuses from the Nile. Soap trees from India. Daisies from Africa's Mountains of the Moon. And the world's most complete collection of camellias.

The sanctuary itself is a sight for any sore-eyed conservationist. It was established twenty-six years ago by Mr. Edward A. McIlhenny, a member of the family that has owned the island




A DAY
AIR FRANCE

and garden beetles, are wired with neurons. At the splices a special messenger substance is secreted to bridge the gaps. That substance is removed by a particular enzyme in the normal process of turning each nerve signal on and then off; a very small number of enzyme molecules suffice for the task. Organic phosphates of one kind or another can quite specifically bind up the enzyme, inactivate it and so promote unending firing of the nerves until sheep, beetle or man dies, twitching, secreting, erectile and self-strangled. It takes five or 10 grams ingested or absorbed to kill perhaps 6,000 sheep, which seems to be the story of the 1968 incident at Dugway Proving Grounds. Atropine—the active principle of belladonna—is a special antidote: if a good dose is given promptly, one survives a few milligrams of the poison. These weapons—they are German developments from World War II, improved in persistence and contact effectiveness in Britain and the U.S. in the 1950's—are operational. They are made here from Muscle Shoals phosphate, after R. and D. at Edgewood Arsenal, synthesized into organic molecules at plants near Denver and Terre Haute, loaded into a wide variety of shells and grenades and canisters, then stored at those depots or shipped closer to the front.

The biologicals are subtler and chancier. Gruinard, an offshore island near Scotland, is still closed to visitors because of tests with anthrax spores carried out there 25 years ago. The experts think it may remain infected for a century. Fort Detrick at Frederick, Md., develops such weapons with great resources and considerable success; they have been produced in quantity, loaded for battle in the Army's plant at Pine Bluff, Ark., and stored; none have been sent, they say, to Vietnam. The identity of specific operational biological weapons is not public knowledge, but there are hints and signs. A Fort Detrick researcher survived an infection of pneumonic plague in 1959; the disease is normally 90 percent fatal. The commonest accidental infection reported among the Fort Detrick workers is tularemia; half a dozen other infections are seen. The Army gave its highest honor for civilians a few years back to a Fort Detrick scientist "for her contribution to the development of the fungus" of rice blast, a common and costly disease of paddy rice, the world's principal food crop. Psychochemicals such as LSD are still too uncertain in their effect, but given 15 years of study they might become practicable. A secret substance called BZ is now operational in the U.S. as an incapacitating agent. It may be an anesthetic or a psychochemical. Its effects include disorientation, drowsiness, hallucination "and sometimes maniacal behavior." The effects are temporary.

The chemicals require doses two orders of magnitude larger in mass than the infectious agents, but drying, sunlight and oxygen soon inactivate most of the biologicals. An aerosol—fine dust or tiny droplets—is the chief method of delivery for both. Continent-wide attacks still seem beyond man's skills; the chemicals would be too heavy for anything but a huge, vulnerable air fleet, and most biologicals cannot be spread so widely and remain active. Locally, against this province or that or against enemies without air defense and public health and stores of vaccines and atropine and optical aerosol detectors—who knows? One should also bear in mind the possibility of covert use; these weapons are silent.

Are they humane? Of course not; they preferentially kill the frail, the old and the poor. They are far more likely to kill the civilian than the soldier. The line between "nonlethal" antiriot and defoliant chemicals (substances heavily used in Vietnam) and the intended lethal agents is fuzzy and shifting. There is strong evidence that our military organized a sham punishment for an officer who first used tear gas in Vietnam as a trial balloon for the later escalation. The military results have proved minor; the precedent is dangerous; the morality remains unspeakable.

The Egyptians appear to have used organic phosphates made in the U.S.S.R. against the Yemeni royalists; the Chinese manufacture the same stuff; the Japanese tested biological agents on prisoners in Manchuria in 1945, according to a plausible Moscow trial of 1949. Thus the technique unites East and West.

Nearly 100 American sailors died in Bari harbor in 1943 when 100 tons of mustard gas spread over the water from an American cargo ship broken open by a German air attack. Modern CBW—chemical and biological warfare—has yet to be tested on the scale of its first use 50 years ago at Ypres. The U.S. was indicted during the Korean war for the use of a strange set of insect vectors and other curious means by a study group on the other side in a report described by Clarke as "a curious mixture of blatant propaganda and what appears to be accurate observation."

The three books reviewed here undertake distinct tasks. The Hersh book is the most completely documented and

fully detailed; it cites officials, describes visits and interviews, and is particularly useful and specific about the U.S. and its CBW efforts. The Clarke book (he is a British editor and science writer) is much more discursive; Clarke, basing his account on a careful statement of the facts, mainly seeks to put clearly the arguments concerning the morals, law, effectiveness and consequences of the manufacture and use of such weapons. It is a book of informed and humane reflection and historical analogy. The third book, a series of short essays by an international set of experts, devotes only a sixth of its space to CBW; the rest of the essays treat of the other known and foreseeable kinds of warfare that contemporary technology and nationalism imply. They do not make a very convincing case for new horrors different from those in use or already under study; it does not seem that the ozone layer can be removed, except quite transiently, or that robot tanks can march slowly and invincibly across borders to the capitals with their cargo of bombs. The new tetrodotoxin from the blowfish is worse only in size of dose, and not in kind, than the old botulinus toxin. And so on.

Underneath these cool accounts there runs a deep current of revulsion. All science and technology share some responsibility for this strange and terrible growth. Two chemists from the French national laboratory at Orsay, Marcel Fetizon and Michel Magat, close their essay in the Calder collection with the disgusted outcry: "The question may arise: Is all science damned? We must either eliminate science or eliminate war. We cannot have both."

MECHANICAL DESIGN OF SMALL PRECISION MECHANISMS, by Howard Jarmy. Chilton Book Company ($5.95). Try though they may, the electronic engineers with their transistors cannot replace all moving parts by electron pulses. The world of automation and "space-age gimmickry" has plenty of mechanical systems: film transport, positioners for valves and lenses, components of analogue computers and much more. Usually these small, intricate machines are not swift; neither are they strong. What they are is precise. This book is really a knowing personal essay on the art of precision design. It is directed to the designer familiar with the usual analysis of machine design, to call to his attention the special problems in the world of precision.

It begins with structure. The key elements are usually steel, which is the stiffer, or aluminum, which is the lighter. Deflection must be prevented by the choice of shape and section. Thermal expansion is often highly significant; sometimes the designer "must assume that heaters and coolers will be used faithfully, the enclosures kept closed . . . the temperature settings . . . not changed, and [that] the power will be turned on. . . . The designer . . . clearly demonstrates an outstanding and refreshing faith. On occasion . . . he has no other choice." Flat surfaces, parallel plates and parallel bored holes from shafts and bearings are urgent. A photograph of two expensive bookends shows what use can be made of spoiled mounting plates.

There is a checklist for cabinet design; it is not to be dismissed lightly. A single low, heavy mounting surface is a good one; the needs of the people who use and maintain the equipment must always be considered with forethought.

Gears and couplings, reducers and special drive linkages and belts are usually the product of specialized venders. Spur gears are the mainstay of the precision designer; backlash is his problem. Accurate ball bearings are the recommended antifriction elements; gas bearings are too special for any treatment here. Preloading by offsetting the ball races and changing the ring sizes can reduce play below the tenths of a mil the manufacturers allow. The highest grade of instrument bearings will meet tolerances of a tenth of a mil in bore.

Solid straight pins help to align mating parts; solid tapered ones hold gears neatly in place. Vibration, rotational balance, inertial effects are all found here as they are in the usual machine design, but with care their proper control is not difficult. Lubrication too is relatively easy: these precision machines need operate only "under limited adversities" and not on dusty or rocky roads. Sealed or oil-impregnated bearings may solve the problem; if care is needed, the lubrication period and lubricant ought to be legibly marked right on the access door as "golden information." Then there is the game of tolerances. For small-lot production the designer can add errors to design for the worst case; for larger runs he can use simple statistics. He is urged to study this complex topic of describing parts with room for shop errors from a lengthy Department of Defense document (MIL-STD 8: *Dimensioning and Tolerancing*).

At length the expensive and delicate device will work. It will have various plays around a tenth of a mil. Errors in angle readings and setting will come out to 10 arc-seconds or so, and the design will be a success, based on the remarkably wide skills and experienced intuition of the designer, who might even learn that "there are more important things in the world than small precision mechanisms."

TRADITIONAL MEDICINE IN MODERN CHINA: SCIENCE, NATIONALISM, AND THE TENSIONS OF CULTURAL CHANGE, by Ralph C. Croizier. Harvard University Press ($6). Medicine runs deeper than science and engineering; its intimacy of concern and its antiquity as an intellectual endeavor give it a special place. Chinese medicine is about as old as Galen; modern medicine, whatever it owes to Greece and Islam, is the product of Harvey, Pasteur, Banting and Burnet, and it is still growing. This penetrating and intriguing little book tells the tale of a century of conflict, still unresolved, between the two medicines in China.

Chinese medicine has two roots: first the empirical folk medicine—mainly a pharmacology—that has produced a famous stock of herbals, many of genuine therapeutic value, and second the organic philosophy that by the Han time was codified into a medical theory flowing both from the five elements with their elaborate isomorphisms in the body and from the sought-for balance of yin and yang. The two chief arts of this branch are the pulse-reading diagnosis (a subtle art of checking the pulse beat at three spots along each wrist artery, both deeply and by light touch) and acupuncture, the insertion of fine needles several millimeters into the skin at specified points along 12 "meridians" running over the body. There is no doubt that some psychotherapeutic and neurological effects are to be expected from the needle therapy, and the pulse diagnosis at least gives the physician a good long look at the cues his patient's bearing and behavior bring him, unconsciously or otherwise.

The issues of thought and feeling are too complicated to be briefly summarized, but there are a few events of significance. The pneumonic plague hit Manchuria in 1910, at the dusk of the Chinese empire. A specially chosen team of 80 traditional physicians went bravely off to stem the epidemic. All of them died. There was then no Western cure for the disease either, but modern public health measures kept the outbreak away from China proper.

When the present government of China took power, there were perhaps 10,000 physicians trained in Western methods (for more than 600 million people) and 50 times as many practicing in the old manner. The intellectual policy has been to preserve and advance the Chinese ways: they are national in origin, they speak for folk practice and against the educated and outward-looking expert, they are familiar to the peasant. Chinese medicine has had the full support of the regime in word and deed. There were interesting syncretisms: electric acupuncture, a mass test of swallowing tadpoles as a means of feminine contraception (half of the users became pregnant) and bones set by X ray and splinted with the old flexible willow twigs. Of course, the empirical values of the old medicine need not be denied; the issue is joined with the thought rather than the effects. Yet only a tenth of the national medical budget is spent on traditional medicine, and now perhaps half of the doctors are followers of the modern way, although they are tolerant by training and common consent of the pulsologist, the herbalist and the needler. The direction is clear.

The contrast with Taiwan and with India, where the "pure Ayurvedic" medicine still wins intellectual support on quasi-political grounds, is strong. The recent Chinese synthesis of insulin, like Chinese work on the surgical restoration of lost limbs, has made its own contribution to modern medicine, which is therefore no longer purely Western. The book is a fascinating essay in intellectual and social history. The editors have not given us a precise date for the work; the copyright is 1968, but such a coarse figure is useless for books about volatile China! No references seem to be newer than the spring of 1966; it would appear that the book represents the state of affairs before the turmoil of the great cultural revolution really took hold.

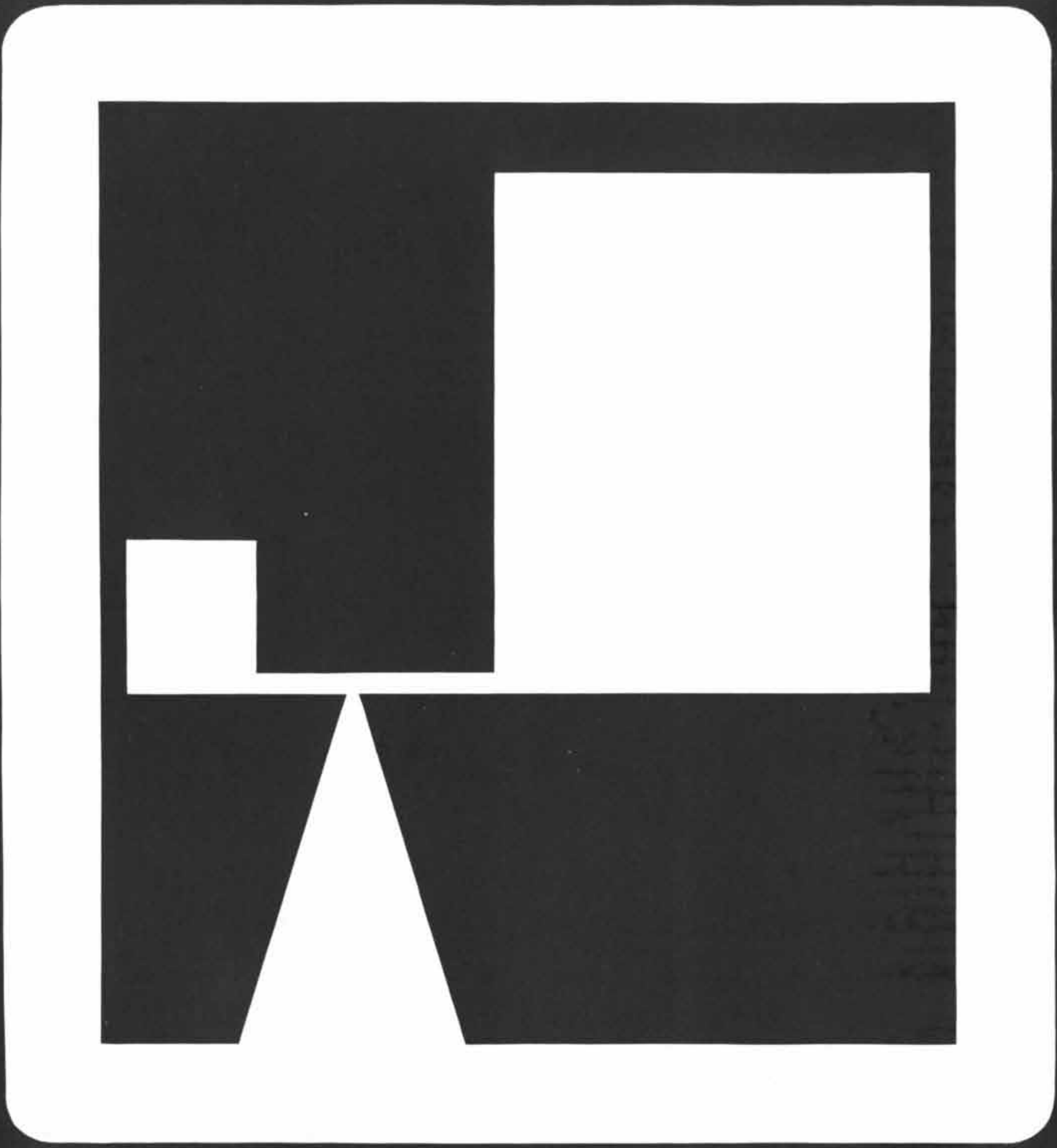
An Introduction to Tree-Ring Dating, by Marvin A. Stokes and Terah L. Smiley. The University of Chicago Press ($5.85). The Navajo tribe has used the precise dating of structural wood from old hogans to establish before the U.S. Indian Claims Commission the time of their occupancy of the land. These authors, from the Tucson laboratory that has developed these methods, describe the work in clear and brief but adequate detail. The photographs, mainly by Clifford Gedekoh, make the account a genuine guide to the method.

A sharp-edged coring tool can remove a thin core sample without permanent injury from a living tree trunk up to several feet in diameter. So the count starts, with care to avoid coring near branches or in directions of irregular growth. Not all trees will do; the species must add a single ring each year. A hand lens or a low-power microscope helps in the counting. A ring is formed by the contrast in cell size between thick-walled, small-diameter longitudinal cells, which are the last to form during the growing season, and thin-walled large ones, which mark the start of the next season. This change—the layer of cells grows as a tall sharp cone added over the previously grown tree stem—is seen as a ring in cross section.

The core is mounted in a grooved wooden holder and sanded smooth in several stages. The outside of the tree is labeled by bark; the ring under it was formed in the current year. Samples of dead wood from old houses are now sought out; long continuous sequences can be fitted together step by step, 100 or 200 at a time for the piñon tree, the species most used in the Navajo country. The dead-wood samples may be sections, not cores; the trick with a time-eroded sample is to learn which is the final ring; bark, bark-beetle borings or a final complete ring will give a good clue. It is unlikely that any natural erosion would leave a single complete outer ring unless it is the original terminal ring itself.

There are traps. In any tree a ring is often locally missing, since growth is not fully symmetrical and equally vigorous at every height. Missing rings are established by repeated sequences of consistent cross dating. "False," or doubled, rings appear to be genetic in origin; they are usually established by close examination of the unusual transition edge of the ring. Experience enables a man to be able to find results that cross-check to within one year with a wholly independent tree-ring chronology. Naturally the scheme does not cohere reliably over distances that are large enough to bring out major climate differences.

The rings can vary only if some variable factor limits growth. In Arizona that factor is rainfall; in Alaska it is temperature. The tree must not be a complacent one, a tree that has prospered year after year on flat land drinking ground water and produced only a blandly uniform alternation of rings. Hardy, sensitive trees on slopes fighting steep drainage for water in the dry months are the best; the ring-width variation is strong in the Southwest because the rains are highly variable. Trees in clumps compete and show modified patterns. The trend of growth varies with age, of course; a curve that expresses each ring width as a fraction of the slow averaged trend is perhaps the best means for building a long sequence. Cross dating works well for a homogeneous locality such as a



small valley; it becomes unreliable over a 600- or 800-mile spread in location.

There is a drawing of a century of matched rings, but one actual long chronology ought to have been included in this careful and interesting survey of a marvelous dating method. Nor does any statistical discussion appear, although such analysis is used in the successful dating laboratory.

EARTH, MOON, AND PLANETS, by Fred L. Whipple. Harvard University Press ($7.25). MYSTERIES OF THE SOLAR SYSTEM, by R. A. Lyttleton. Oxford University Press ($7). This third edition of Professor Whipple's likable introductory account of the bigger members of the sun's family is excellent. Always easy for beginners to read, the text now contains a wealth of fresh pictures, mostly the products of NASA (the moon from low altitude and high, the surface of Mars) or of the big radar sets (for example the face of Venus, with detail). The main facts are all here, with a clear account of how we infer them in an optimistic view of their plausibility and coherence. Mathematics and even close argument are not much used; the text is frankly and winningly descriptive. There are tables for finding the planets in the sky, and tables of the important numbers. The book is up to date: it tells the strange and still unfinished story of the rotation speeds of Venus and of Mercury, each locked by tidal forces, Mercury to the 3/2 harmonic of its own year, Venus to the 3/2 harmonic of the *earth's* year (and retrograde at that).

The second book is a keen complement to the first. It has few pictures and not much more algebra; it is a set of essays by an iconoclast. These pieces are closely argued; they present mathematical results clearly and deeply and replace much manipulative algebra by physical reasoning. The problems discussed are few but wide: the origin of the solar system, the interior of the terrestrial planets, the nature and origin of comets and of tektites (perhaps the result of a comet striking the earth) and the discovery of Neptune. The origin of the planets remains uncertain; the smaller planets may all be the same within, except that the earth and Venus were large enough to melt and collect their slag on top and their metal below. The comets are not dirty ice left over from the time of planet-building but are accreted interstellar dust. John Couch Adams and Urbain Leverrier were mainly lucky—their famous Neptune prediction could have been made by their methods only at a favorable epoch. All of this is told in a personal, dry and cheerful tone. Some of it is surely right, and no student can fail to benefit by seeing how Professor Lyttleton sifts the relevant from the clamorous mass of purported data. He has probably not said his last word, and some of his sallies at the academy have now been contradicted, but his edged thought cuts deep into the fashions that always influence the assumptions required by genetic studies in astronomy.

MY LIFE AND MY VIEWS, by Max Born. Introduction by I. Bernard Cohen. Charles Scribner's Sons ($4.95). Professor Born has earned our attention by his 60-odd years of work in physics, work that includes not only such epochal discoveries as the statistical interpretation of the wave function but also the best textbook on optics, several other outstanding textbooks and first-class popular writings. These are his latest essays. A third of the book, his autobiography, has already appeared in this country in periodical form; the rest, on the political, historical and philosophical implications of science, has been published only in German. He was head of the Göttingen theoretical physics department during the Weimar years when quantum theory grew up; his first assistants there were Wolfgang Pauli and Werner Heisenberg! He writes that "I never learned nuclear physics properly," because the lecturer on radioactivity, Johannes Stark, "did not satisfy my mathematical mind." Nevertheless, "both... Oppenheimer and Teller, as well as Fermi and... some of the Russian physicists, were once my collaborators.... It is satisfying to have had such clever and efficient pupils, but I wish they had shown less cleverness and more wisdom. I feel that I am to blame if all they learned from me were methods of research and nothing else."

He sees fear and hope in the state of the world, and expresses them both with simple clarity in an attractively modest and tentative style. Here his insight into philosophical issues is given popular form, with emphasis on the objectivity not of individual sense impressions but rather of the relations among them, in which he sees the proper subject matter of science. He discusses the world's dilemmas with what may be a professor's lack of emphasis on the genuine conflicts of material and national interest, but he ends in hope, a reasoned hope for the change of human nature, away from war and "against official lies and encroachments... suppression... *gloire*, greed... and infallibility."

INDEX OF ADVERTISERS

MARCH 1969

ABINGDON PRESS ... 143

ALITALIA AIRLINES ... 64
Agency: Gardner Advertising Company, Inc.

AMERICAN BOOK CLUB ... 133
Agency: The Kaplan Agency, Inc.

AMERICAN PLYWOOD ASSOCIATION ... 152
Agency: Cole & Weber, Inc.

BARNES, A. S., & CO., INC. ... 148
Agency: Kameny Associates, Inc.

BASIC BOOKS, INC. ... 146
Agency: Wenk/Schwartz & Associates, Inc.

BELL TELEPHONE LABORATORIES ... 9
Agency: N. W. Ayer & Son, Inc.

BERLITZ SCHOOL OF LANGUAGES ... 10
Agency: Independent Media Services, Inc.

BOLAND ENTERPRISES ... 88
Agency: Graphic Design

BRUSH BERYLLIUM COMPANY, THE ... 145
Agency: Meldrum & Fewsmith, Inc.

CABOT CORPORATION ... 7
Agency: J. G. Kasten & Company

CALIFORNIA COMPUTER PRODUCTS, INC. ... 129
Agency: Carson/Roberts/Inc.

CARRIER AIR CONDITIONING COMPANY ... 2
Agency: N. W. Ayer & Son, Inc.

CELESTRON PACIFIC, INC. ... 88
Agency: Baytron, Inc.

COLUMBIA UNIVERSITY PRESS ... 142
Agency: Franklin Spier Incorporated

CROWN PUBLISHERS ... 134
Agency: Sussman & Sugar, Inc.

CURTA COMPANY ... 134
Agency: Eastman Advertising Agency

DIGITAL EQUIPMENT CORPORATION ... 51
Agency: Kalb & Schneider, Inc.

EASTMAN KODAK COMPANY ... 45
Agency: Rumrill-Hoyt, Inc.

EDMUND SCIENTIFIC CO. ... 150
Agency: Walter S. Chittick Company, Inc.

FORD MOTOR COMPANY ... 65
Agency: Grey Advertising, Inc.

FREEMAN, W. H., AND COMPANY ... 136, 137

GARRARD, BRITISH INDUSTRIES CO., A DIV. OF AVNET, INC. ... 4
Agency: Cole Fischer Rogow, Inc.

GENERAL ELECTRIC CO., INFORMATION SYSTEMS ... 114, 115
Agency: Robert S. Cragin, Inc.

HARVARD UNIVERSITY PRESS ... 128
Agency: Franklin Spier Incorporated

HAVERHILL'S, INC. ... 143
Agency: Kadeemah Associates

HERTZ SYSTEM, INC. ... 1
Agency: Carl Ally Inc., Advertising

HEWLETT-PACKARD COMPANY ... 102, 103
Agency: Lennen & Newell, Inc.

HUGHES AIRCRAFT COMPANY ... 73
Agency: Foote, Cone & Belding

HUGHES AIRCRAFT COMPANY, AEROSPACE DIVISIONS ... 151
Agency: Foote, Cone & Belding

LTV ELECTROSYSTEMS, INC., A SUBSIDIARY OF LING-TEMCO-VOUGHT, INC. ... Inside Back Cover
Agency: Lennen & Newell, Inc., The Wyatt and Williams Division

LIBERTY MIRROR, A DIVISION OF LIBBEY-OWENS-FORD COMPANY ... 97
Agency: Fuller & Smith & Ross Inc.

LIBRARY OF SCIENCE, THE ... 130
Agency: Henderson & Roll, Inc.

LOCKHEED AIRCRAFT CORPORATION ... 16, 17
Agency: McCann-Erickson, Inc.

LOCKHEED-CALIFORNIA COMPANY, A DIVISION OF LOCKHEED AIRCRAFT CORPORATION ... 150
Agency: McCann-Erickson, Inc.

M.I.T. PRESS, THE ... 144
Agency: Franklin Spier Incorporated

MONSANTO COMPANY ... 53
Agency: Foote, Cone & Belding

NATIONAL CAR RENTAL SYSTEM, INC. ... 89
Agency: Campbell-Mithun, Inc.

NAVAL LABORATORIES OF THE EAST COAST ... 149
Agency: Allerton, Berman & Dean

NIKON INCORPORATED, SUBSIDIARY OF EHRENREICH PHOTO-OPTICAL INDUSTRIES, INC. ... 5
Agency: Gilbert and Felix Inc.

OXFORD UNIVERSITY PRESS ... 6
Agency: Denhard & Stewart, Inc.

PHELPS DODGE INDUSTRIES, INC. ... Inside Front Cover
Agency: Compton Advertising, Inc.

PHILIPS RESEARCH LABORATORIES ... 135
Agency: T.A.G. deLaMar—Intermarco

PLANNING RESEARCH CORPORATION ... 113
Agency: MacManus, John & Adams, Inc.

POLAROID CORPORATION, THE ... 76, 77
Agency: Doyle Dane Bernbach Inc.

PORSCHE OF AMERICA CORPORATION ... 121
Agency: Henry J. Kaufman & Associates

POTLATCH FORESTS, INC. ... 75
Agency: R. G. Williams Company

QUESTAR CORPORATION ... 52

RAYTHEON COMPANY ... 138, 139
Agency: Hoag & Provandie, Inc.

RENAULT, INC. ... 47
Agency: Gilbert Advertising Agency

REPUBLIC STEEL CORPORATION ... 122, 123
Agency: Meldrum and Fewsmith, Inc.

STANDARD OIL COMPANY (New Jersey) ... 136, 137
Agency: LaRoche, McCaffrey and McCall, Inc.

TODD SHIPYARDS CORPORATION ... 99
Agency: O. S. Tyson and Company, Inc.

TUBE METHODS INC. ... 46
Agency: John Miller Advertising Agency

UNION CARBIDE CORPORATION ... Back Cover
Agency: Young & Rubicam, Inc.

UNITED AIRCRAFT CORPORATION ... 12, 13
Agency: Cunningham & Walsh Inc.

UNITED AUDIO PRODUCTS, INC. ... 8
Agency: Gordon and Weiss, Inc.

UNIVAC DIVISION, SPERRY RAND CORPORATION ... 15
Agency: Daniel and Charles, Inc.

UNIVERSITY LABORATORIES ... 61
Agency: West Associates

UNIVERSITY OF CHICAGO PRESS ... 140
Agency: Franklin Spier Incorporated

WASHINGTON, STATE OF, DEPARTMENT OF COMMERCE AND ECONOMIC DEVELOPMENT ... 101
Agency: Kraft, Smith & Lowe, Inc.

WESTERN ELECTRIC COMPANY ... 141
Agency: Cunningham & Walsh, Inc.

WFF'N PROOF ... 148
Agency: Ad-Com Agency

WRIGHT LINE, A DIVISION OF BARRY WRIGHT CORPORATION ... 64
Agency: John F. Norris & Company, Inc.

XEROX CORPORATION ... 90, 91
Agency: Needham, Harper & Steers, Inc.

BIBLIOGRAPHY

Readers interested in further reading on the subjects covered by articles in this issue may find the lists below helpful.

THERMAL POLLUTION AND AQUATIC LIFE

THE PHYSIOLOGY OF FISHES. Edited by Margaret E. Brown. Academic Press, Inc., 1957.

THE PHYSIOLOGY OF CRUSTACEA, VOL. I: METABOLISM AND GROWTH. Edited by Talbot H. Waterman. Academic Press, Inc., 1960.

FISH AND RIVER POLLUTION. J. R. Erichsen Jones. Butterworths, 1964.

A FIELD AND LABORATORY INVESTIGATION OF THE EFFECT OF HEATED EFFLUENTS ON FISH. J. S. Alabaster in *Fishery Investigations, Ministry of Agriculture, Fisheries, and Food,* Series I, Vol. 6, No. 4; 1966.

THERMAL POLLUTION—1968. Subcommittee on Air and Water Pollution of the Committee on Public Works. U.S. Government Printing Office, 1968.

SUPERPLASTIC METALS

SUPERPLASTICITY IN AN AL-ZN ALLOY. W. A. Backofen, I. R. Turner and D. H. Avery in *American Society for Metals Transactions Quarterly,* Vol. 57, No. 4, pages 980–990; December, 1964.

THE ORIGIN OF SUPERPLASTICITY IN THE SN-5% BI ALLOY. T. H. Alden in *Acta Metallurgica,* Vol. 15, No. 3, pages 469–480; March, 1967.

SUPERPLASTICITY IN THE NI-FE-CR SYSTEM. H. W. Hayden, R. C. Gibson, H. F. Merrick and J. H. Brophy in *American Society for Metals Transactions Quarterly,* Vol. 60, No. 1, pages 3–14; March, 1967.

THE INTERRELATION OF GRAIN SIZE AND SUPERPLASTIC DEFORMATION IN NI-CR-FE ALLOYS. H. W. Hayden and J. H. Brophy in *American Society for Metals Transactions Quarterly,* Vol. 61, No. 3, pages 542–549; September, 1968.

PHASES IN CELL DIFFERENTIATION

EARLY PANCREAS ORGANOGENESIS: MORPHOGENESIS, TISSUE INTERACTIONS, AND MASS EFFECTS. Norman K. Wessells and Julia H. Cohen in *Developmental Biology,* Vol. 15, No. 3, pages 237–270; March, 1967.

MULTIPHASIC REGULATION IN CYTODIFFERENTIATION. William J. Rutter, William R. Clark, John D. Kemp, William S. Bradshaw, Thomas G. Sanders and William D. Ball in *Epithelial-Mesenchymal Interactions,* edited by Raul Fleischmajer and Rupert E. Billingham. The Williams & Wilkins Company, 1968.

ULTRASTRUCTURAL STUDIES OF EARLY MORPHOGENESIS AND CYTODIFFERENTIATION IN THE EMBRYONIC MAMMALIAN PANCREAS. Norman K. Wessells and Jean Evans in *Developmental Biology,* Vol. 17, No. 4, pages 413–446; April, 1968.

REGULATION OF SPECIFIC PROTEIN SYNTHESIS IN CYTODIFFERENTIATION. W. J. Rutter, J. D. Kemp, W. S. Bradshaw, W. R. Clark, R. A. Ronzio and T. G. Sanders in *Journal of Cellular Physiology,* Vol. 72, No. 2, Part II, pages 1–18; October, 1968.

CONTINENTAL DRIFT AND EVOLUTION

VERTEBRATE PALEONTOLOGY. Alfred Sherwood Romer. The University of Chicago Press, 1966.

THE AGE OF THE DINOSAURS. Björn Kurtén. World University Library, 1968.

BROWNIAN MOTION AND POTENTIAL THEORY

CONNECTION BETWEEN BROWNIAN MOTION AND POTENTIAL THEORY. Anthony W. Knapp in *Journal of Mathematical Analysis and Applications,* Vol. 12, No. 2, pages 328–349; October, 1965.

WIENER AND INTEGRATION IN FUNCTION SPACES. M. Kac in *Bulletin of the American Mathematical Society,* Vol. 72, No. 1, Part II, pages 52–68; January, 1966.

DYNAMICAL THEORIES OF BROWNIAN MOTION. Edward Nelson. Princeton University Press, 1967.

THE ATMOSPHERES OF MARS AND VENUS

MARINER IV MEASUREMENTS NEAR MARS: INITIAL RESULTS. *Science,* Vol. 149, No. 3689, pages 1226–1248; September 10, 1965.

LIFE DETECTION BY ATMOSPHERIC ANALYSIS. Dian R. Hitchcock and James E. Lovelock in *Icarus,* Vol. 7, No. 2, pages 149–159; March, 1967.

MARINER V. *Science,* Vol. 158, No. 3809, pages 1665–1690; December 29, 1967.

THE ATMOSPHERES OF VENUS AND MARS.

Edited by John C. Brandt and Michael B. McElroy. Gordon and Breach, 1968.

THE ATMOSPHERE OF VENUS: PAPERS FROM THE SECOND ARIZONA CONFERENCE ON PLANETARY ATMOSPHERES. *Journal of the Atmospheric Sciences,* Vol. 25, No. 4; July, 1968.

VENUS: LOWER ATMOSPHERE NOT MEASURED. Von R. Eshleman, Gunnar Fjeldbo, John D. Anderson, Arvydas Kliore and Rolf B. Dyce in *Science,* Vol. 162, No. 3854, pages 661–665; November 8, 1968.

PLAGUE TOXIN

PLAGUE. R. Pollitzer. World Health Organization, 1954.

THE MURINE TOXIN OF PASTEURELLA PESTIS: A STUDY IN ITS DEVELOPMENT. Solomon Kadis, Thomas C. Montie and Samuel J. Ajl in *Bacteriological Reviews,* Vol. 30, No. 1, pages 177–191; March, 1966.

ISOLATION OF TOXIC SUBUNITS FROM TWO MURINE-TOXIC PROTEINS FROM PASTEURELLA PESTIS. Thomas C. Montie, Diane B. Montie, Shalom A. Leon, Carole A. Kennedy and Samuel J. Ajl in *Biochemical and Biophysical Research Communications,* Vol. 33, No. 3, pages 423–429; November 8, 1968.

THE FIRST ELECTRON TUBE

FIFTY YEARS OF ELECTRICITY: THE MEMORIES OF AN ELECTRICAL ENGINEER. J. A. Fleming. The Wireless Press, Ltd, 1921.

MEMORIES OF A SCIENTIFIC LIFE. Sir Ambrose Fleming. Marshall, Morgan & Scott Ltd., 1934.

THERMIONIC VALVES 1904–1954: THE FIRST FIFTY YEARS. The Institution of Electrical Engineers, 1955.

MATHEMATICAL GAMES

FIBONACCI NUMBERS. N. N. Vorob'ev in *Popular Lectures in Mathematics Series: Vol. II,* edited by I. N. Sneddon and M. Stark. Blaisdell Publishing Company, 1961.

AN INTRODUCTION TO FIBONACCI DISCOVERY. Brother U. Alfred. The Fibonacci Association, 1965.

FIBONACCI AND LUCAS NUMBERS. Verner E. Hoggatt, Jr. Houghton Mifflin Company, 1969.

THE AMATEUR SCIENTIST

THIN-LAYER CHROMATOGRAPHY. Kurt Randerath. Academic Press, 1966.

Would your family like it if you took a job at Hughes in California?

It's easy to find out.

Let's exchange resumes.

Ours includes a 126-page book on Southern California. Do you want to know about living costs? Weather? Taxes? Cultural advantages?

Want to know how to pick a hilltop homesite? How to catch a shark? What to expect on your driver's license test? How much it costs to build a pool? How reading is taught in the public schools?

It's all here, in a good book about a good life. (You'll also find California isn't mere tinsel and glitter. Actually one-fourth of America's physicists live here. So do one-fourth of the members of the National Academy of Sciences and almost half of America's Nobel Prize winners.)

And of course Hughes has a story too. You can imagine the benefits of working with the men who designed Surveyor, Syncom and dozens of other highly advanced systems. We've gathered lots of material for you—facts on everything from programs to fringe benefits.

If you're an E.E. from an accredited university with 2 or more years of professional-level experience applicable to aerospace research and development technology, please send us your resume.

We'll send you ours by return mail.

Mr. Robert A. Martin
Head of Employment
Hughes Aerospace Divisions
11940 W. Jefferson Boulevard
Culver City, Calif. 90230

U.S. citizenship is required
An equal opportunity employer—M&F.

HUGHES

HUGHES AIRCRAFT COMPANY
AEROSPACE DIVISIONS