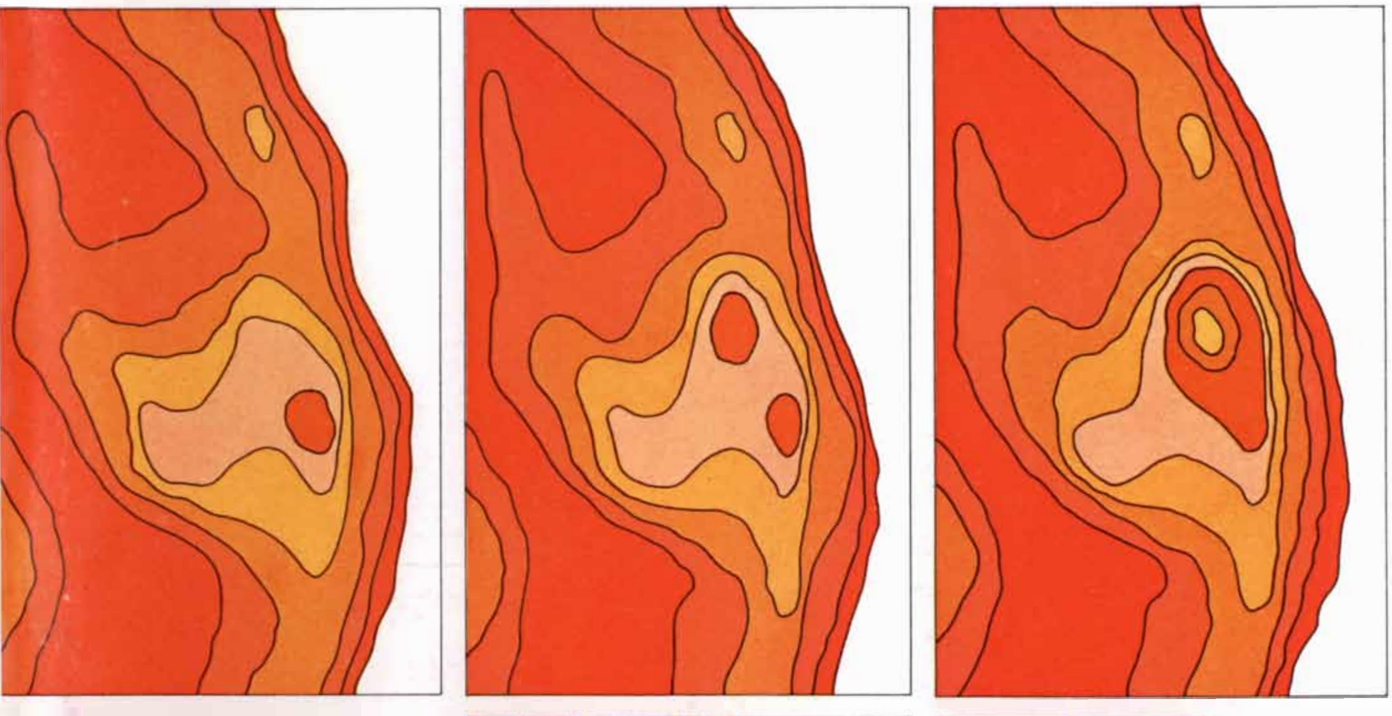
SCIENTIFIC
AMERICAN

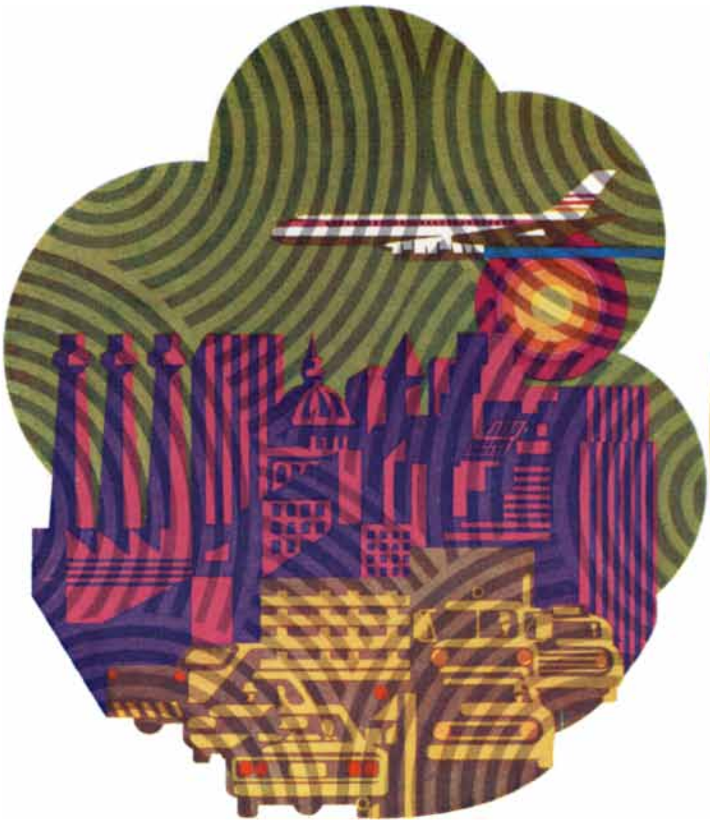


ULTRAVIOLET SUN

SEVENTY-FIVE CENTS

June 1969

from fumes



to fragrances

UOP has a piece of the action in your life.

We've all got problems, right?

Los Angeles, Chicago, your own home town, your own industry want to quit smoking and smogging. Pollution problems.

Your wife ponders which of a thousand scents will accentuate her personality. That's her problem.

UOP wants to help you understand that many of your problems are our problems.

UOP is a company with a nose for the new. A sense of now. A driving scientific curiosity.

Against air pollution UOP Air Correction Division offers the widest variety of equipment to control solid and gaseous pollutants. Entire systems for air control, industrial or urban, have been developed. Consultation, source-testing, start-up assistance are all part of UOP Air Correction's service.

On the sunnier side, you might be surprised to know that UOP makes fragrance compounds for use in sexy perfumes, toiletries, cosmetics. Very likely the famous name perfume your wife uses gets its magic from a compound developed by UOP Fragrances.

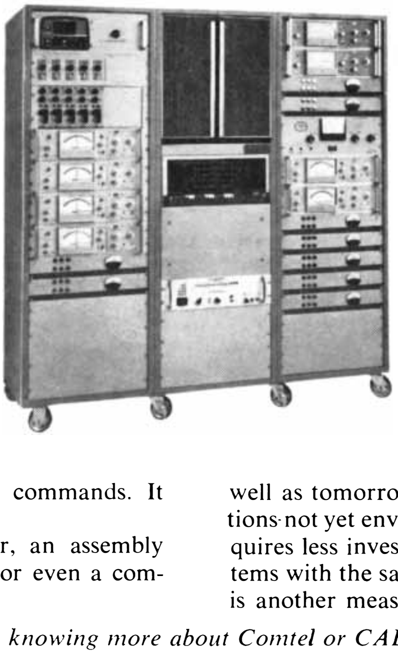
From fumes to fragrances is just one band in UOP's creative spectrum. Remember Platforming®? UOP Process Division did it: designed, engineered and perfected this and other revolutionary processes for refining high-energy fuels.

And Universal Oil Products Company is also poking its nose for the new into such areas as air and ground transportation equipment, fabricated metal products, plastic laminates, advanced petroleum processes.

No end in sight. Whoever you are, UOP has a piece of the action in your life.

© 1969 SCIENTIFIC AMERICAN, INC

Comtel has developed a production line control that thinks for itself



Here is a control concept that takes a giant step ahead of conventional digital, relay-actuated equipment. Our Technical Sciences Division calls it CABOT,* Control and Basic Operational Techniques.

CABOT will react to standard logic inputs, but it also has innate analog flexibility, being able to accept analog data and create analog commands. It can be programmed to interpret and react proportionately to a multiplicity of infinitely variable inputs, such as temperature, pressure, torque, etc. CABOT's response sequence, degree, or mix can be quickly altered by teletype commands. It remembers these, too.

CABOT can control a machining center, an assembly machine, an automatic inspection process—or even a complete production line.

For a tire manufacturer, CABOT controls an inspection and correction station that can heft and spin a newly made tire, pinpoint out-of-round force characteristics, and prescribe the exact grinding needed to achieve perfect balance.

For a brake manufacturer, CABOT provides on-line testing, dynamic balancing, and inspection, as well as being able to duplicate any combination of in-service conditions.

The unique blending of analog control concepts and computer technologies makes CABOT a tool for today's as well as tomorrow's problems. It can adapt to control functions not yet envisioned! And, in many instances, CABOT requires less investment than relay or logic panels or test systems with the same capacity but only digital capability. This is another measurement of success for the men at Comtel.

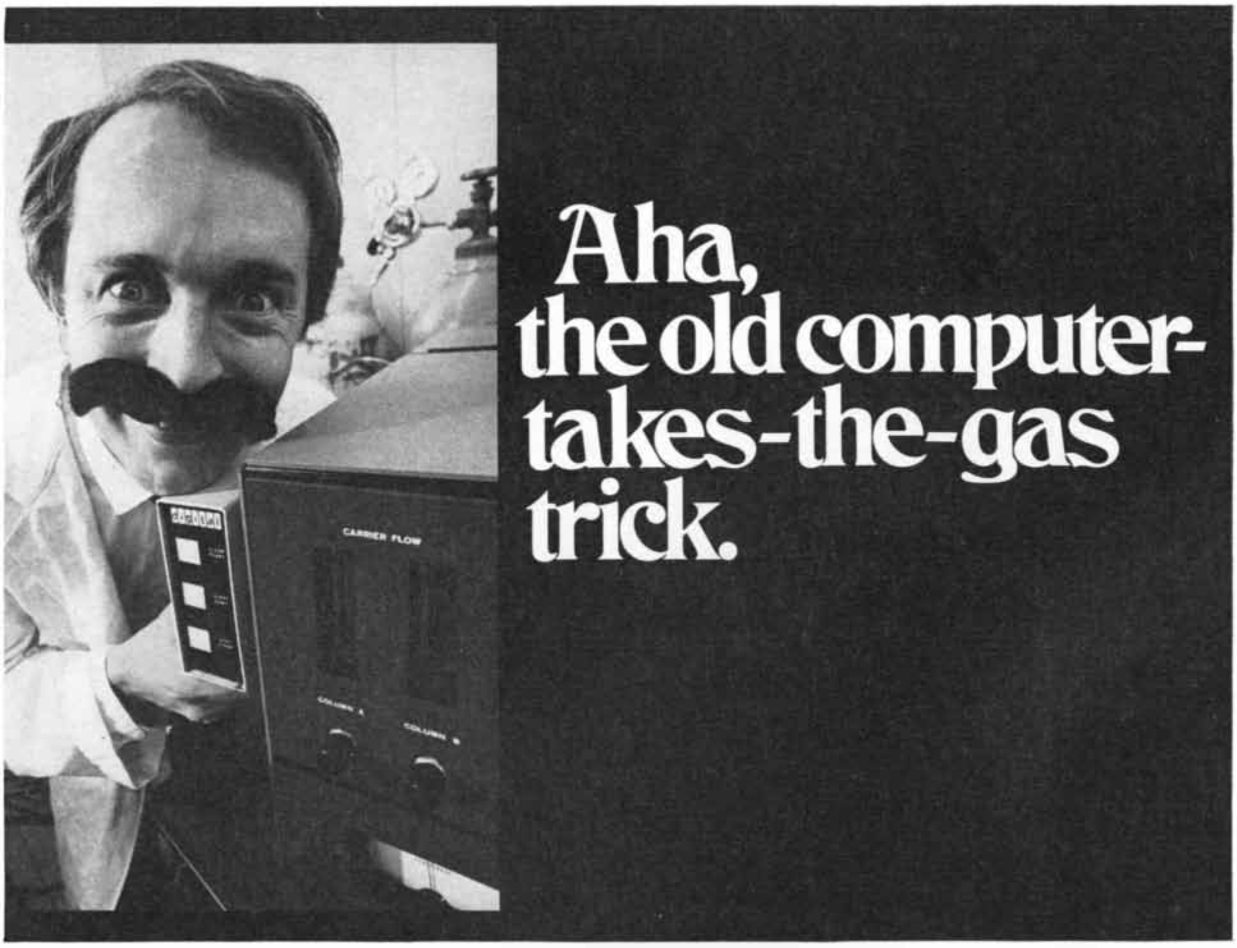
Interested in knowing more about Comtel or CABOT? Write.

COMTEL CORPORATION
21250 W. 10½ MILE ROAD, SOUTHFIELD, MICHIGAN 48075, 313-352-4650

Total Systems in Electronic Sciences

Detroit Testing Laboratory, *Material and Product Evaluation Center* • **S. Sterling Company,** *Electronic Instrumentation* • **Southfield Electronics,** *NationalInstrument Marketing* • **Technical Sciences Division,** *Specialized Testing Systems* • **Comtel Components,** *Precision Component Sales* • **Compar Midwest Corporation,** *Component Distribution* • **Oceanography & Pollution Research Division,** *Undersea and Atmosphere Studies* • **Comtel Engineering, Inc.,** *Television Communications Systems* • **Metrolonics Associates, Inc.,** *Linear Measurements* • **National Astro Laboratories, Inc.,** *West Coast Standards Calibration and Certification, Traceable to NBS* • **SSCO Standards Laboratory, Inc.,** *Midwest Standards Calibration and Certification, Traceable to NBS* • **Comtel Leasing, Inc.,** *Electronic Equipment Leasing*

*TM

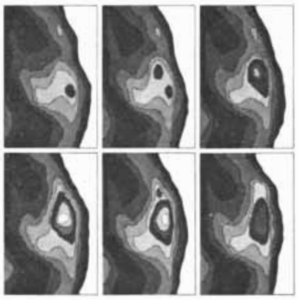

Aha,
the old computer-takes-the-gas trick.
CARRIER FLOW
COLUMN A
COLUMN B

Established 1845 SCIENTIFIC AMERICAN June 1969 Volume 220 Number 6

ARTICLES

DEPARTMENTS

PUBLISHED MONTHLY BY SCIENTIFIC AMERICAN, INC., 415 MADISON AVENUE, NEW YORK, N.Y. 10017. SECOND-CLASS POSTAGE PAID AT NEW YORK, N.Y., AND AT ADDITIONAL MAILING OFFICES. AUTHORIZED AS SECOND-CLASS MAIL BY THE POST OFFICE DEPARTMENT, OTTAWA, CANADA, AND FOR PAYMENT OF POSTAGE IN CASH. SUBSCRIPTION RATE: $8 PER YEAR.



THE COVER

The colorful patterns on the cover represent the development of a bright flare on the sun as observed in the ultraviolet part of the spectrum by a spectroheliograph aboard Orbiting Solar Observatory IV, 480 miles above the earth's surface (see "Ultraviolet Astronomy," page 92). The sequence, from left to right and top to bottom, shows events on the west limb, or edge, of the sun recorded at five-minute intervals on November 29, 1967. The measurements were made at a wavelength of about 1,035 angstroms, deep in the ultraviolet region that is blocked by the earth's atmosphere. Intense radiation of this wavelength originates some 2,000 kilometers above the sun's visible surface. Each contour interval represents an increase of 1.6 in brightness, progressing in nine equal steps from the coolest region (*largest red areas*) to the hottest region in the center of the flare (*yellow*).

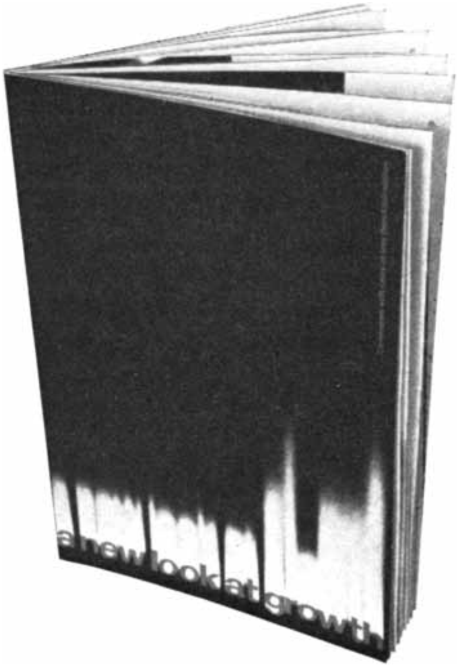
THE ILLUSTRATIONS

Cover drawing by Jerome Kühl

Page	**Source**
20–21	Sovfoto
22	Enid Kotschnig
23	*Soviet Life*
24	Sovfoto (*top*), *Soviet Life* (*bottom*)
25	Sovfoto
26–27	Enid Kotschnig
28–29	Sovfoto
30–31	R. L. Fleischer, General Electric Company
32–33	Jim Egleson
34	R. L. Fleischer, General Electric Company (*top*); Jim Egleson (*bottom*)
35	Jim Egleson
36	R. L. Fleischer, General Electric Company
37–38	Jim Egleson
39	R. L. Fleischer, General Electric Company
41	James D. Huber, University of Washington
42	Bunji Tagawa
43	Russell Ross, University of Washington
44	Russell Ross, University of Washington (*top*); Bunji Tagawa (*bottom*)
45–47	Russell Ross, University of Washington
48	Bunji Tagawa
49	Russell Ross, University of Washington (*top*); Bunji Tagawa (*bottom*)
50	Russell Ross, University of Washington (*top*); Bunji Tagawa (*bottom*)
79	John Tyler Bonner
80–81	Dan Todd
82–83	Patricia Collins
84	Dan Todd
88	Dan Todd (*top*), Patricia Collins (*bottom*)
91	Patricia Collins and Dan Todd
93	Air Force Sacramento Peak Observatory (*top left*), Allen Beechel
94–101	Allen Beechel
102	Project Stratoscope, Princeton University
104–111	Eric Mose
113	Jon Brenneis
114	Thomas Prentiss
115	L. E. Maley, Anacon, Inc.
116–118	Thomas Prentiss
120	Jet Propulsion Laboratory, California Institute of Technology
122–124	Alan D. Iselin
130	Ben Rose
131	Roger Hayward
132	Roger Hayward (*top*), Ben Rose (*bottom*)
133–134	Ben Rose

Just printed.

"A New Look at Growth"



This 48-page booklet brings you key facts about 64 smallish companies that look to us like up-and-coming growth stocks. We think many could double earnings in the next 5 years—and some could triple. If you'd like a free copy, send us the coupon below.

We describe each company in about 300 words—enough to give you a taste of its record, products, competitive position, and prospects. There's also a chart showing the price of the stock for the last 10 years and the first quarter of 1969.

The back of the booklet has a handy statistical section that covers all 64 companies. It shows earnings per share in '67 and '68, annual dividend rate, recent price, number of shares outstanding, and the P/E ratio based on '68 earnings.

Why we took a new look at growth

Things have changed.

It used to be that practically all you needed to sniff out a growth stock was a good nose for technology. Xerox smelled good to some people. Polaroid smelled good to others. And sure enough, both rose higher than a soufflé at Pavillon.

But technology's not enough anymore. The time has come, as the Walrus said, to speak of many things. Like marketing. And management. And even of specialization.

It may seem odd in this age of conglomerates, but some companies got into our booklet mostly because they specialize so profitably. Certain software companies, for example.

Why we left out the giants

We wanted companies that could show dramatic gains in earnings. That's a lot easier for Davids than for Goliaths.

During the last five years, for example, about half the companies in our booklet increased their earnings by 10 to 30 percent a year.

Some have boosted earnings by over 70 percent a year. That's no guarantee of future performance, but it sure sounds encouraging, doesn't it?

Although we focused on growth, we didn't forget about quality. Among others, we left out companies with spotty earnings records, limited marketability for their shares, and relatively weak financial positions.

We put a lot into this booklet. We hope you get a lot out of it.

841EA

MERRILL LYNCH,
PIERCE, FENNER & SMITH INC

To: Merrill Lynch, P.O. Box 333
Wall Street Station, New York, N.Y. 10005

Please send me "A New Look at Growth—Companies with Sales of Less Than $100 Million."

Mr.
Mrs
Miss ______________________
(Please print)
Address ______________________
City ____________ Phone ____________
State ____________ Zip ____________

NL

LETTERS

Sirs:

In the April *Scientific American* George W. Rathjens presents an interesting balanced view on the psychology underlying the competitive development of nuclear armaments, the vehicles for their delivery as well as countermeasures related thereto.

One problem, which is unfortunately quite real but is nonetheless completely disregarded in connection with all these calculations and estimates, is that of the chronic delayed effects that would result from an exchange involving nuclear weapons. A Chinese source indicated that even if a man-for-man fatality rate resulted in the U.S.S.R. and the U.S. in a nuclear war with China, they, the Chinese, with their large numbers, would be the obvious survivors and beneficiaries.

On reading Rathjens' paper as well as other articles in this area I have the uncomfortable feeling that others concerned with defense may be laboring under similar misconceptions. Even if a defensive system were perfect and no land target was hit, the fallout and radiation from a massive utilization of nuclear weapons, either in the stratosphere or in the lower strata of the atmosphere, would so hopelessly contaminate land and foodstuffs that survivors would not have access to clean, radiation-uncontaminated food and water for a long time. Decontamination on such a large scale seems a rather futile task, particularly if some of the industrial capability was rendered inoperative....

J. HENRY WILCOX

Bethesda, Md.

Sirs:

Karl H. Pribram's article "The Neurophysiology of Remembering" [SCIENTIFIC AMERICAN, January] attempts to relate holography to memory. Pribram believes that supporting evidence for this hypothesis comes from work by J. Stone and myself on retinal neurons in the cat. We found the assumption of linearity to have good predictive value in describing the responses of these cells to low-contrast stimuli. The relationship is not exactly linear but the assumption of linearity serves as a good approximation. Pribram notes that "convolution integrals" appear in the calculated response curves of retinal neurons and adds: "Convolution integrals and Fourier transformations provide the mathematical basis on which holography was founded. Thus at least a first step has been taken to show that interference effects may operate in the central nervous system."

This analogy does not survive scrutiny. The common factor between holography and calculated retinal responses is the assumption of linearity. Convolution integrals occur *whenever* linearity is assumed. But a great many physical systems besides holograms are commonly described by linear analysis including many problems in electronics, mechanics, heat conduction and hydraulics. One could, I believe, make as strong a case for heat conduction as for holography as a model of remembering.

There is no doubt that memory is somehow distributed in the brain. But what does the hologram hypothesis add? The hypothesis is so vaguely formed that there does not appear to be any possible experiment that could disprove it. It is not so much a hypothesis as the ghost of one.

R. W. RODIECK

University of Sydney
Sydney, Australia

Sirs:

Two interesting letters concerning "The Neurophysiology of Remembering" have come to *Scientific American*. One from Stephen Grossberg, published in your March issue, claims "the existence of a mathematical theory of learning...and recall...that has already given a rigorous form to Dr. Pribram's suggestion"; the other, printed above, that "the hypothesis is so vaguely formed that there does not appear to be any possible experiment that could disprove it."

Grossberg has developed a long and detailed mathematical description of the organizational properties of "embedding fields" that is meant to simulate occurrences in the brain. He does not, however, describe simple experiments to test the assumptions or consequences of the model that can be accomplished by neurophysiologists with today's techniques. Rodieck admits that memory is somehow distributed in the brain but fails to find the holographic hypothesis useful in generating testable ideas as to how memory might become distributed.

These criticisms are not new to me and afford a good opportunity to dissociate Dr. Nico Spinelli, whose work I quote so liberally in my article, from the holographic hypothesis. In my laboratory he has for years played devil's advocate by taking both the Grossberg and the Rodieck position on many occasions. He wants a sharper engineering model that simulates the brain's rather than an optical information-processing system; he wants hypotheses framed so that they can be experimentally tested. I am thus well accustomed to the arguments made. If there were no other reason, the fact that the holographic hypothesis can generate interest, excitement and criticism would make it well worth holding.

But of course there is more. One of the major questions concerning coding in the central nervous system is whether or not the transformations involved are more or less linear. There is good behavioral evidence that both linear and nonlinear processes must be taking place. For example, many of the facts of perception suggest linearity (see Floyd Ratliff, *Mach Bands*, Holden-Day, 1965), whereas language is obviously nonlinear. The question arises therefore as to the particular processes served by the linear and nonlinear mechanisms. My article attempts to present the evidence that the distribution of information in memory is accomplished by virtue of some linear mechanism analogous to that by which holograms are produced. Experimental results such as those that show the imperviousness of the discriminative process to my aluminum hydroxide implantations of striate cortex found no ready explanation until it was shown that holographic reconstructions of images were similarly resistive to punctate damage. This does not say that all transformations accomplished by the brain are holographic in nature, but that some are likely to be.

As to the precision of the hypothesis, I mean to attribute to the neural hologram (1) linearity of transformation, (2) distribution of information and (3) the fact that spatial phase relationships as well as intensity are coded. As Rodieck points out, linear analysis applies to many types of problem. The virtue of the holographic model is that it restricts the range of linear transformations needing to be considered and that it emphasizes the spatial interactions among neural elements. In doing this the holographic hypothesis provides a useful alternative to detector theory in accounting for some (though not all) of the facts of perceptual imagery and recognition.

Finally, there is the "ghost" brought up by Rodieck's letter. One of the most appealing arguments for the neural hologram is a philosophical one. Gilbert

Ryle, in his famous term "the ghost in the machine," encapsulated the problem psychology faces. Brain processes can be studied, but where are the "images" we see, the "feelings" we hold, the wishes and wants, etc.? They are, of course, no-*where,* although some boundary conditions can be specified. One of the exciting things to do with a holographic image is to try to grab hold of it. One can't. The interference patterns are recorded, but the image becomes constituted in front of or behind the record, shifts around according to one's position in space and disappears unless conditions are just right.

The issue is whether nonbiological mechanisms such as computers and holograms serve as useful models for biological (and specifically brain) processes. I believe they do, but I agree with Rodieck that hypotheses derived from studies of the models must be tested against the results of actual experiments on organisms. But then this is what my laboratory is all about.

KARL H. PRIBRAM

Stanford University School
of Medicine
Stanford, Calif.

Scientific American, June, 1969; Vol. 220, No. 6. 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.

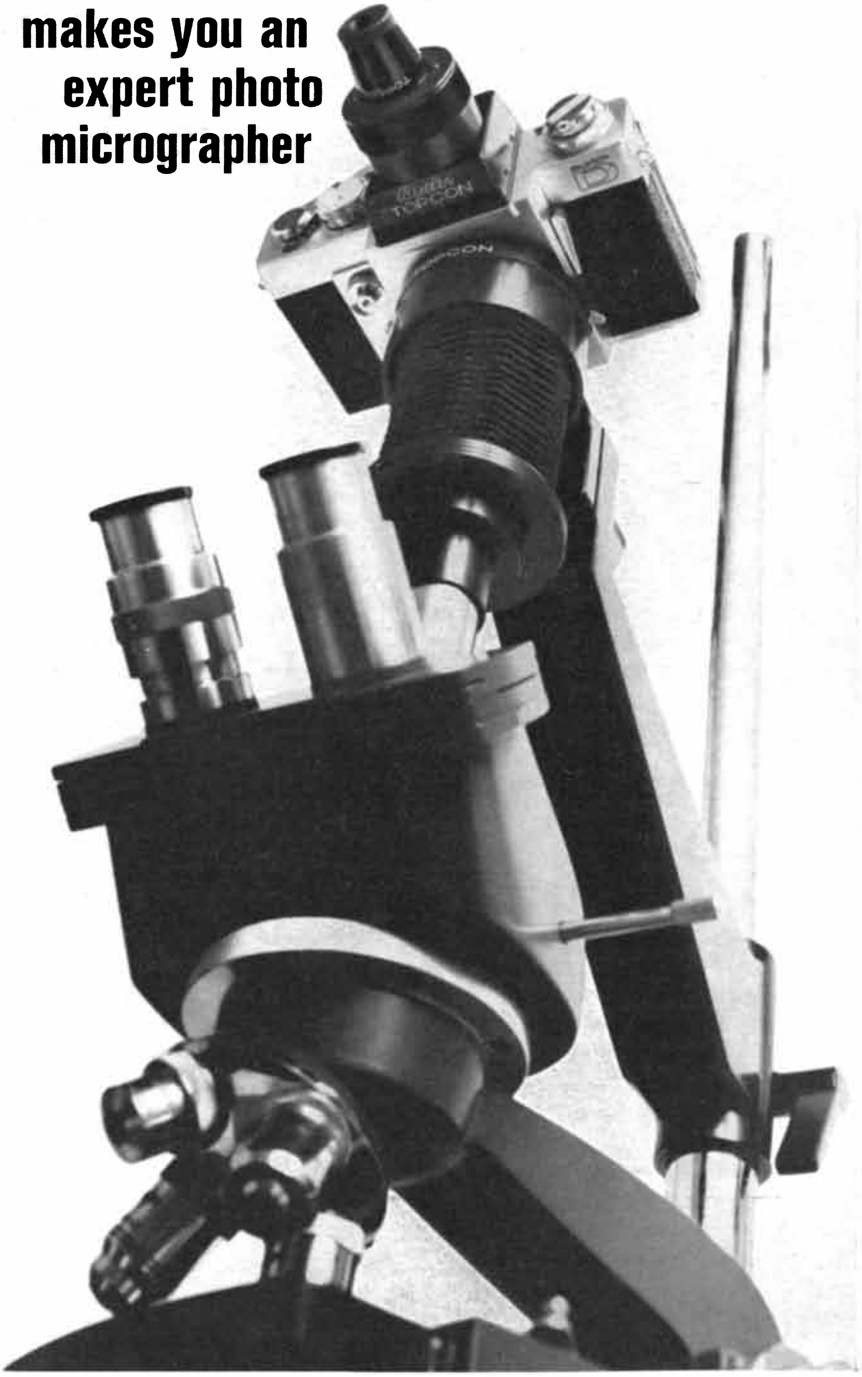
Change of address (or other subscription correspondence) should be addressed to Circulation Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017. Please notify us four weeks in advance of change. If possible, kindly furnish an address imprint from a recent issue. Send both old and new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



The Emancipator Becomes a Show-off!

The HP 9100A Calculator gave you

Freedom from waiting to get on the BIG computer;

Freedom from translating your problems into foreign computer languages;

Freedom from starvation-level computing with under-developed calculators;

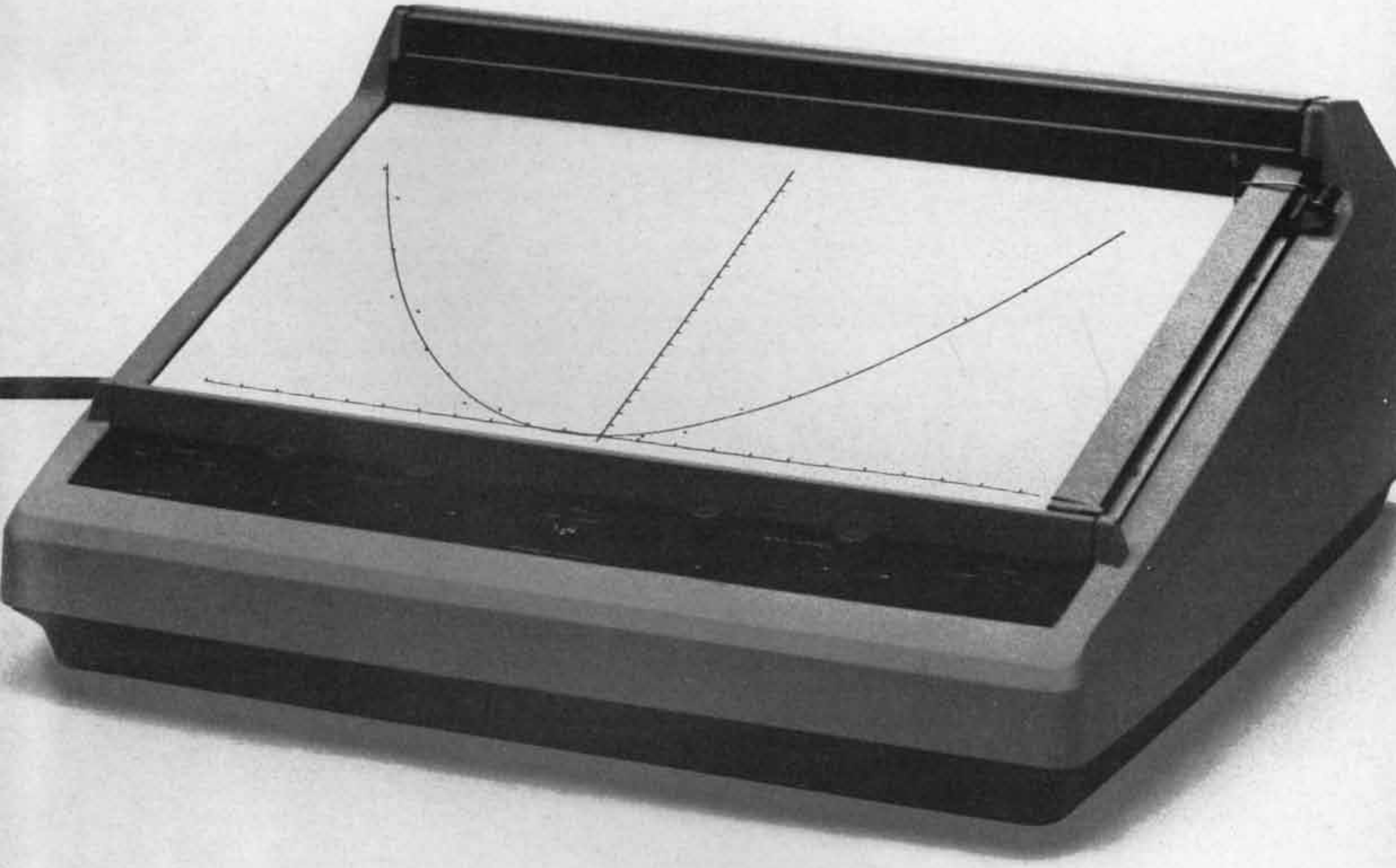
Freedom from the drudgery of manual computation– All for the one-time price of $4900.

NOW–EVEN GREATER COMPUTING FREEDOM IS YOURS. GRAPHIC DISPLAY

Add the new HP 9125A X-Y Recorder. Make your HP 9100A desk-top computing calculator a show-off. You will be liberated from the tediousness of hand-plotting, from the omnipresent human errors and from inundating reams of data that go with point-to-point hand plotting.

Think back! Would a picture be worth a thousand figures, to you?...to your company? Be an innovator...be a producer...a picture producer.

The 9125A recorder automatically draws the answers as the calculator solves your problems–such as curve-fitting (photo shows parabolic regression), differential equations, roots of polynomials and calibration curves. It shows you the results of your calculations faster, easier, and more accurately than you can get with hand plots! It ends the gnawing frustration of interpreting the limbo between data points. And...it gives you a calibrated plot so you can ready your units directly!



50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

JUNE, 1919: "We are gratified to note that in his testimony recently given before the House Naval Affairs Committee the Secretary of the Navy announced that the second three-year naval construction program had been abandoned. SCIENTIFIC AMERICAN was strongly opposed to this program on the grounds that it was both inexpedient and unnecessary—inexpedient because it would tend to destroy the excellent feeling that had sprung up between the U.S. and the Allies as the result of our participation in the war, unnecessary because the joint effect of the construction of our first three-year program and the elimination of the German fleet had enabled us to reach that strong position as second naval power, which the great national movement for preparedness had set as its ultimate goal. The fact that the second program has been abandoned suggests that the Government is satisfied that the coöperation of our Allies both in the matter of the proposed League of Nations and in the matter of naval disarmament is well assured. If these great results have been achieved, we believe the whole nation will join with us in recognizing that the President has performed a great and noble service in the interests not merely of the sorely stricken world of today but also of humanity at large for all time to come."

"Scarcely have we caught our breath at the announcement that the U.S. Navy seaplane NC-4, piloted by Commander Read, had flown from America to Europe by way of the Azores and the Spanish Main than the message comes that a British Vickers-Vimy bomber, with Captain Alcock at the controls, has swept across the Atlantic from Newfoundland to the Irish coast in one wild flight of 16½ hours. The outstanding facts of this non-stop flight are the astonishing speed of about 120 miles an hour and the unerring precision with which Lieutenant Brown, the navigator, held his way over the course of 1,950 miles."

"Perhaps the greatest significance of Major Hoke's recent invention of a precision block gage accurate within millionths of an inch is that it has forced the adoption by the Bureau of Standards of the wave-length as the measuring element. At present the standard yard is defined as the length between two marks on a certain metal rod at a certain temperature. This is not ideal, for several reasons. In the first place, why should one have to go to Washington or Paris to get a standard length? Would it not be more rational to define the yard, the meter, etc., in terms of something that could be reproduced independently anywhere? Would it not be a good idea at the same time to seek a standard that had no relation to temperature? It surely would, and the whole procedure in the Hoke case fairly thrusts such a standard upon our attention. Why should we not define the yard and the meter as so many wave-lengths of such a particular light? For testing Hoke gages cadmium light has been found to possess several notable advantages. Major Hoke looks forward to the day when the standard of length will be defined as 'so many wave-lengths of such-and-such a light'—cadmium perhaps, or something else if anything better can be found."

SCIENTIFIC AMERICAN

JUNE, 1869: "The rate at which immigration to the U.S. is now progressing is quite unprecedented. On the Atlantic side Europe is pouring in vast numbers of people of all ages who, fleeing from the pressure of want, anticipate a life of comparative ease and plenty in the less crowded industries of the American continent. On the Pacific slope the Asiatic races are beginning to form a large element of the population, and a useful element too, if report speaks truly. We are of those who think this influx of population will eventually be a benefit to the country, after the proper process of assimilation has been effected, and provided that a proper policy on the part of the general Government provides a home market for the increased production, consequent upon the increased number of producers."

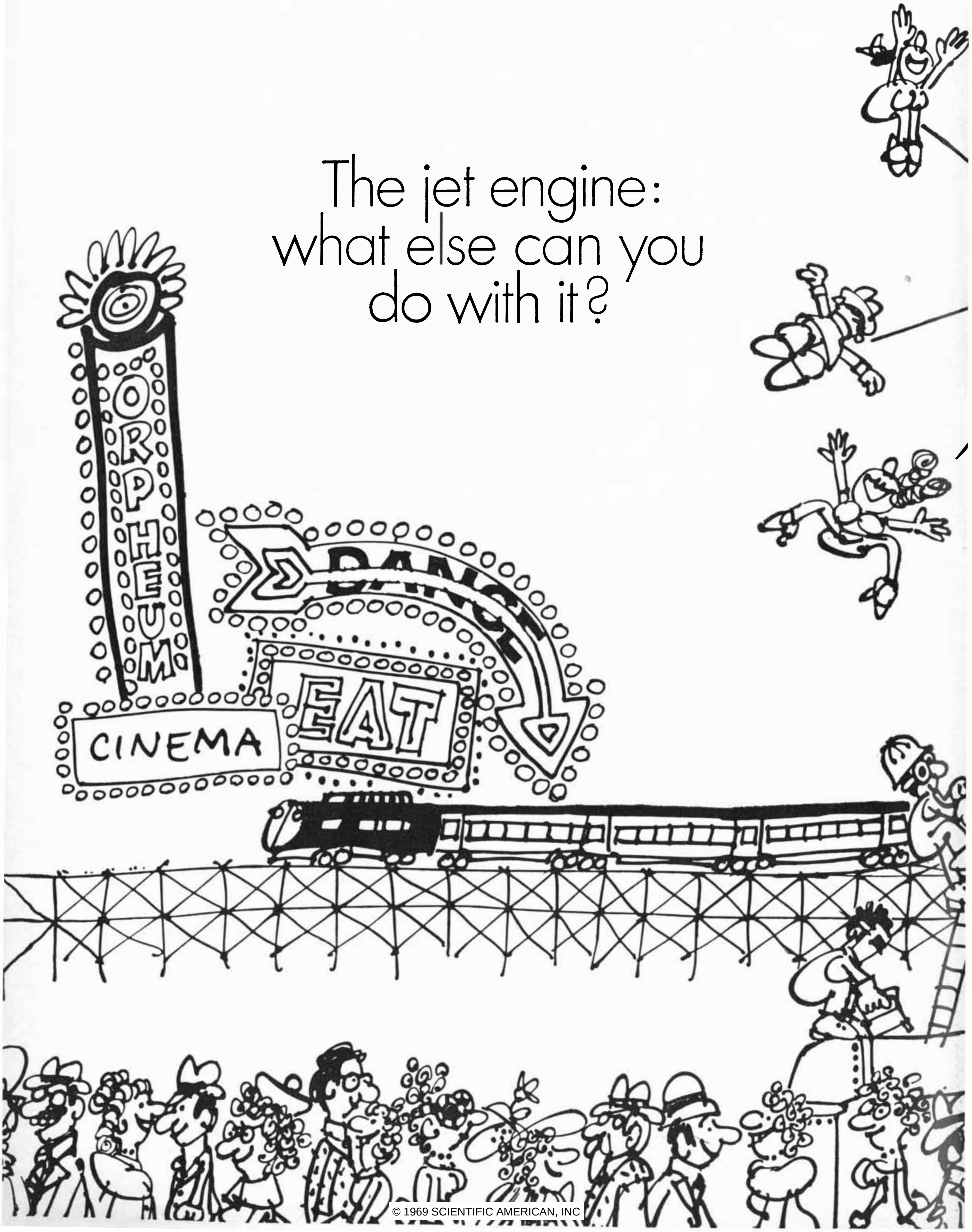
"The question of the eight-hour workday was taken up at the cabinet meeting on May 25, and after an extended discussion it was decided that President Grant should issue a proclamation, or an executive order, declaratory of the effect of the law upon wages, which is simply that the Act of Congress declares that eight hours' labor for the United States shall be a day's work, instead of 10 hours, *without affecting the rate of wages.*"

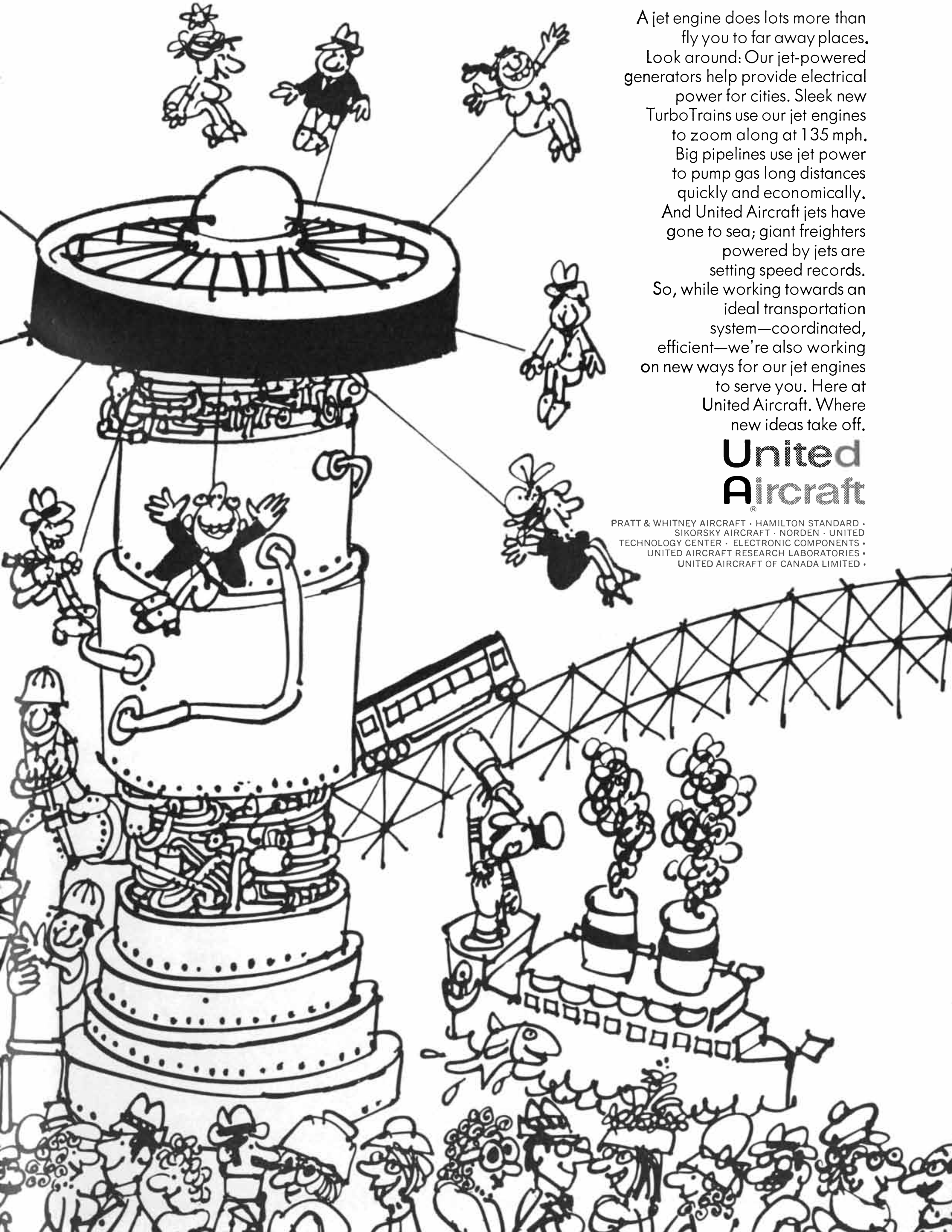
"The celebrated Darwinian theory is making proselytes rapidly. In fact, it may be said that its advocates no longer admit that it is simply a theory but hold it to be demonstrated by the scientific researches it has called forth. Among the most prominent scientific men who have fully embraced the theory is the veteran Sir Charles Lyell, whose past record does not indicate a man likely to be led to hasty conclusions or apt to adopt any theory unsupported by ample evidence of its truth."

"Professor Bunsen of Heidelberg recently met with a serious accident. He had received a large quantity of the metals of the platinum group and was engaged in the preparation of pure rhodium. He had precipitated a large quantity of the finely divided metal and had placed it to dry in a vessel heated by a bath of boiling water. Someone had then carelessly let the water boil out of the bath, so that when Bunsen went alone into his laboratory at midnight, he found that the heat of the vessel had risen to 300 degrees Fahrenheit, instead of 212 degrees, as the temperature would have been if water had remained in the bath. He approached the vessel, put down his light and put one finger in to mark the condition of things. Suddenly there was a fearful explosion. Both of his eyes were severely burned, and his hands were torn into a mass of open wounds. He had the presence of mind not to drop the platinum capsule containing the rhodium; he put it back upon the furnace before he called for help. The explosion and the call for assistance were fortunately heard by the servants, and he was immediately carried to his dwelling, which is in the same building as the laboratory. It is known that some years ago Bunsen lost the use of one eye in a similar explosion, and it was now feared that the other eye had been destroyed, but upon closer examination the physician expressed the hope that the injury was not incurable. Upon hearing this, the hero of science exclaimed: 'Thank God! I can now ascertain what was the condition of the metal when it blew up.' "

"The Erie Railroad Company have commenced selling through tickets by the Union Pacific Railroad from New York to San Francisco for $197.35."

The jet engine: what else can you do with it?





A jet engine does lots more than fly you to far away places. Look around: Our jet-powered generators help provide electrical power for cities. Sleek new TurboTrains use our jet engines to zoom along at 135 mph. Big pipelines use jet power to pump gas long distances quickly and economically. And United Aircraft jets have gone to sea; giant freighters powered by jets are setting speed records. So, while working towards an ideal transportation system—coordinated, efficient—we're also working on new ways for our jet engines to serve you. Here at United Aircraft. Where new ideas take off.

United Aircraft®

PRATT & WHITNEY AIRCRAFT · HAMILTON STANDARD · SIKORSKY AIRCRAFT · NORDEN · UNITED TECHNOLOGY CENTER · ELECTRONIC COMPONENTS · UNITED AIRCRAFT RESEARCH LABORATORIES · UNITED AIRCRAFT OF CANADA LIMITED ·

THE AUTHORS

R. W. DAVIES and R. AMANN ("Science Policy in the U.S.S.R.") are at the Centre for Russian and East European Studies of the University of Birmingham; Davies is director of the center and professor of Soviet economic studies and Amann is assistant lecturer. Davies was graduated from the University of London in 1950 and obtained his Ph.D. from Birmingham in 1954, whereupon he went to the University of Glasgow for two years before returning to Birmingham. He has collaborated with E. H. Carr in a forthcoming study, *Foundations of a Planned Economy in the U.S.S.R.* Amann was graduated from the University of Birmingham in 1964. He is a coauthor of part of a report by the Organisation for Economic Co-operation and Development, *Science Policy in the U.S.S.R.* In preparing the present article the authors were assisted by their colleague, M. J. Berry. The article is based on the O.E.C.D. study of Soviet science policy, which was made by an international team including the authors and Eugène Zaleski, Joseph P. Kozlowski and Helgard Wienert.

R. L. FLEISCHER, P. B. PRICE and R. M. WALKER ("Nuclear Tracks in Solids") began their collaboration at the Research and Development Center of the General Electric Company. Fleischer and Price are still there, although Price will be leaving soon to become professor of physics at the University of California at Berkeley. Walker was named McDonnell Professor of Physics and director of the Laboratory for Space Physics at Washington University in 1967. Fleischer received his Ph.D. in applied physics from Harvard University in 1956 and spent four years as assistant professor of metallurgy at the Massachusetts Institute of Technology before joining General Electric. Price, who received his doctorate in solid-state physics from the University of Virginia in 1958, joined General Electric in 1960 after study in England. As a visiting professor he organized a research group studying nuclear tracks in solids at the Tata Institute of Fundamental Research in Bombay. Walker joined General Electric in 1954 after receiving his Ph.D. in elementary-particle physics from Yale University. He is a founder of VITA, an association of more than 1,000 engineers and scientists who contribute their skills in working on technical problems of underdeveloped nations. He is also a member of the National Aeronautics and Space Administration's lunar-sample-analysis planning team.

RUSSELL ROSS ("Wound Healing") is associate professor of pathology at the University of Washington School of Medicine; his promotion to professor will take effect on July 1. He was graduated from Cornell University in 1951 with a degree in chemistry and received his Ph.D. in pathology from the University of Washington in 1962. He writes: "Beyond the work in the area of wound healing and fibrogenesis I have been working extensively for the past several years on the elastic fiber. Our knowledge of this important connective-tissue element has been limited and at this stage of development is somewhat akin to our knowledge of collagen 20 years ago. As for my other interests, they relate to music and sports, in particular tennis and squash."

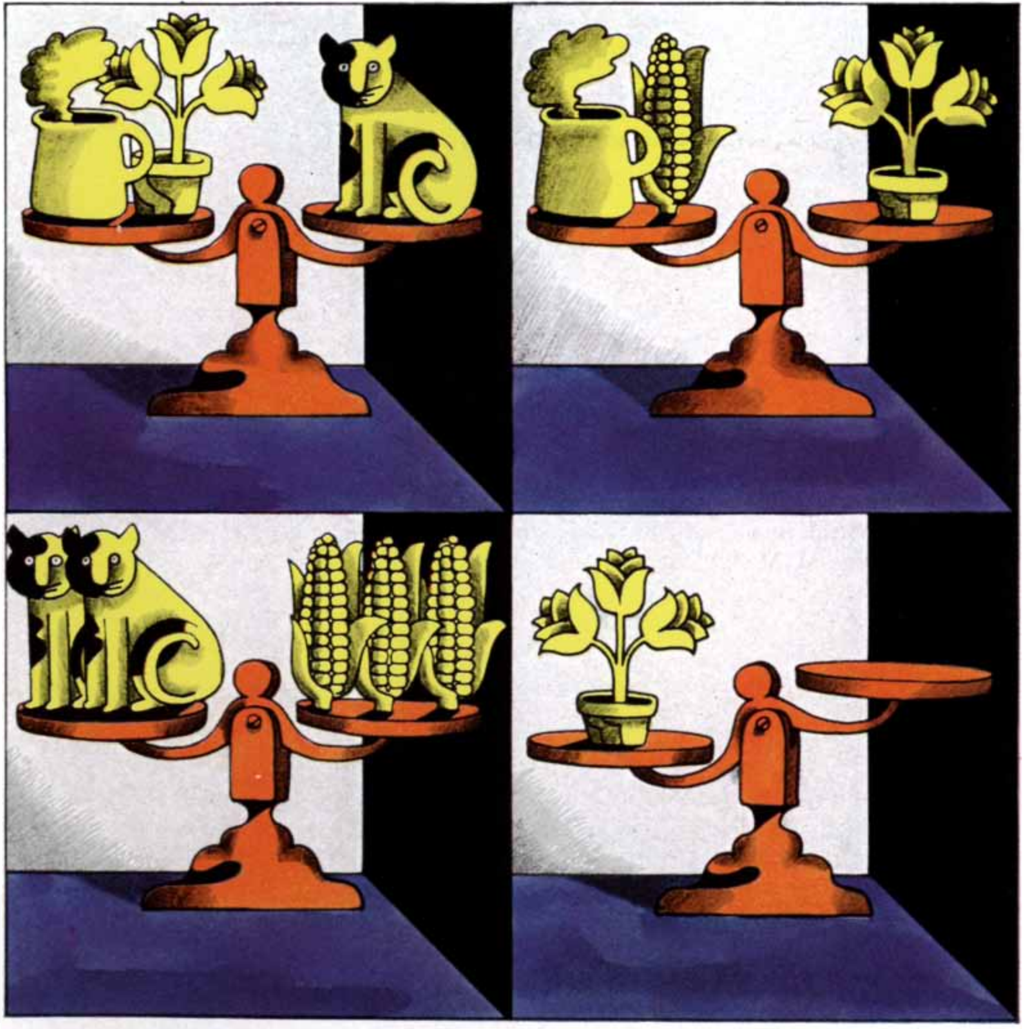
ALFRED TARSKI ("Truth and Proof") is professor of mathematics at the University of California at Berkeley. Born in Poland, he received his Ph.D. from the University of Warsaw in 1924 and taught there until he came to the U.S. in 1939. He was a research associate at Harvard University for two years and a member of the Institute for Advanced Study for one year before going to the University of California. Among his many writings is the book *Introduction to Logic and to the Methodology of Deductive Sciences,* which has appeared so far in 10 languages and is scheduled to be published in two more. The present article is based on the Faculty Research Lecture given by Tarski at Berkeley in 1963 and on the invited address he gave at the University of London in 1966.

JOHN TYLER BONNER ("Hormones in Social Amoebae and Mammals") is George M. Moffett Professor of Biology and chairman of the department of biology at Princeton University. A graduate of Harvard College in 1941, he received his master's degree at Harvard in 1942 and his Ph.D. there in 1947 (after four years of service in the Army Air Forces). He has been at Princeton since 1947. Bonner has worked extensively on the biology of such organisms as slime molds, algae and fungi. Among his books are *The Cellular Slime Molds,* now in its second edition, and *The Scale of Nature.*

LEO GOLDBERG ("Ultraviolet Astronomy") is Higgins Professor of Astronomy and chairman of the department of astronomy at Harvard University and director of the Harvard College Observatory. He was a student at Harvard from 1930 to 1938, receiving his bachelor's degree in 1934, his master's degree in 1937 and his Ph.D. in 1938. For 19 years before returning to Harvard in 1960 he was at the University of Michigan, serving from 1946 to 1960 as director of its observatory and chairman of the department of astronomy. He describes "the application of atomic physics to the interpretation of astrophysical problems" as his "deepest interest." He also writes: "I have served on almost every conceivable committee in universities, in professional societies and in Government agencies. Only a very few of these have been genuinely satisfying."

E. OTTO HÖHN ("The Phalarope") is associate professor of physiology at the University of Alberta. He writes: "Born March 14, 1919, in Basel, Switzerland, of Swiss parentage. Moved to England with parents in 1932 and gained entrance scholarship in modern languages of Guy's Hospital Dental School (University of London), obtaining the M.B.B.S. [bachelor of medicine and bachelor of science] degree in 1943. Then worked as hospital physician and house surgeon in hospital of Guy's Emergency Medical Service sector until 1944, when offered full-time demonstratorship in physiology at Guy's Medical School. Married the same year, wife English; two children, born 1949 and 1954. In 1947 obtained the position of assistant professor of physiology at the University of Alberta and acquired British nationality by naturalization before leaving for Canada. In 1951 obtained the Ph.D. in physiology from the University of London."

F. W. KARASEK ("Analytic Instruments in Process Control") is associate professor of chemistry at the University of Waterloo in Ontario. "When I got my doctorate at Oregon State University in 1952," he writes, "I went to the Phillips Petroleum Research Center for a short interlude of valuable industrial experience before entering on a university teaching career. That short interlude stretched to 17 years before I managed to begin the teaching career. I was attracted to the University of Waterloo not only because of the lovely new university but also because of the opportunity to set up new analytical and instrumentation courses. At home my main activities are related to the raising of seven sons."



Torinos

Cougar Convertibles

Mercury Montegos

Continental Mark III's

door sedans alone.

Mach I's. And not just Fords but other makes as well.

And—much to the joy of your company bookkeeper—they don't cost much more to rent than a regular standard sedan.

After all, while we understand that man cannot live by four-door sedans alone, we also understand that bookkeepers are not noted for their humanitarian instincts.

The biggest should do more. It's only right.

Hertz®

Now you can afford your own time-sharing computer: the GE starter set.

Information Systems

Serving Customer Needs-Worldwide

Watch it grow.

It's inevitable. Anyonewho nowbuys a little time-sharing, will soon be wanting far more.

General Electric makes your transition a lot easier, and a lot less expensive. With our new GE-410 system you can start out with a small computer (to serve a few people) and grow it into a very big one (to serve hundreds of people).

You can save thousands because you have no extra programming costs. Just change the heart of the computer — give it more power to handle more remote terminals — and you're back in business. All the programs you've already developed will still work — only faster.

You'll also save money because General Electric time-sharing systems do regular processing as well. Essentially, that's two computers for the price of one.

Only General Electric gives you such a wide choice in time-sharing. From a single terminal to the largest, most flexible time-sharing systems yet devised. One is right for you.

No wonder more time-sharing is done on General Electric computers than any other kind. Contact your nearest GE Information Systems Sales Representative. Or write:
Section 290-38, General Electric,
1 River Road,
Schenectady, N.Y. 12305.

290-38

GENERAL ELECTRIC

Established 1845 SCIENTIFIC AMERICAN June 1969 Volume 220 Number 6

Science Policy in the U.S.S.R.

A new study shows that the U.S.S.R. operates a research and development establishment comparable in size to that of the U.S. The differences between the two are getting fewer

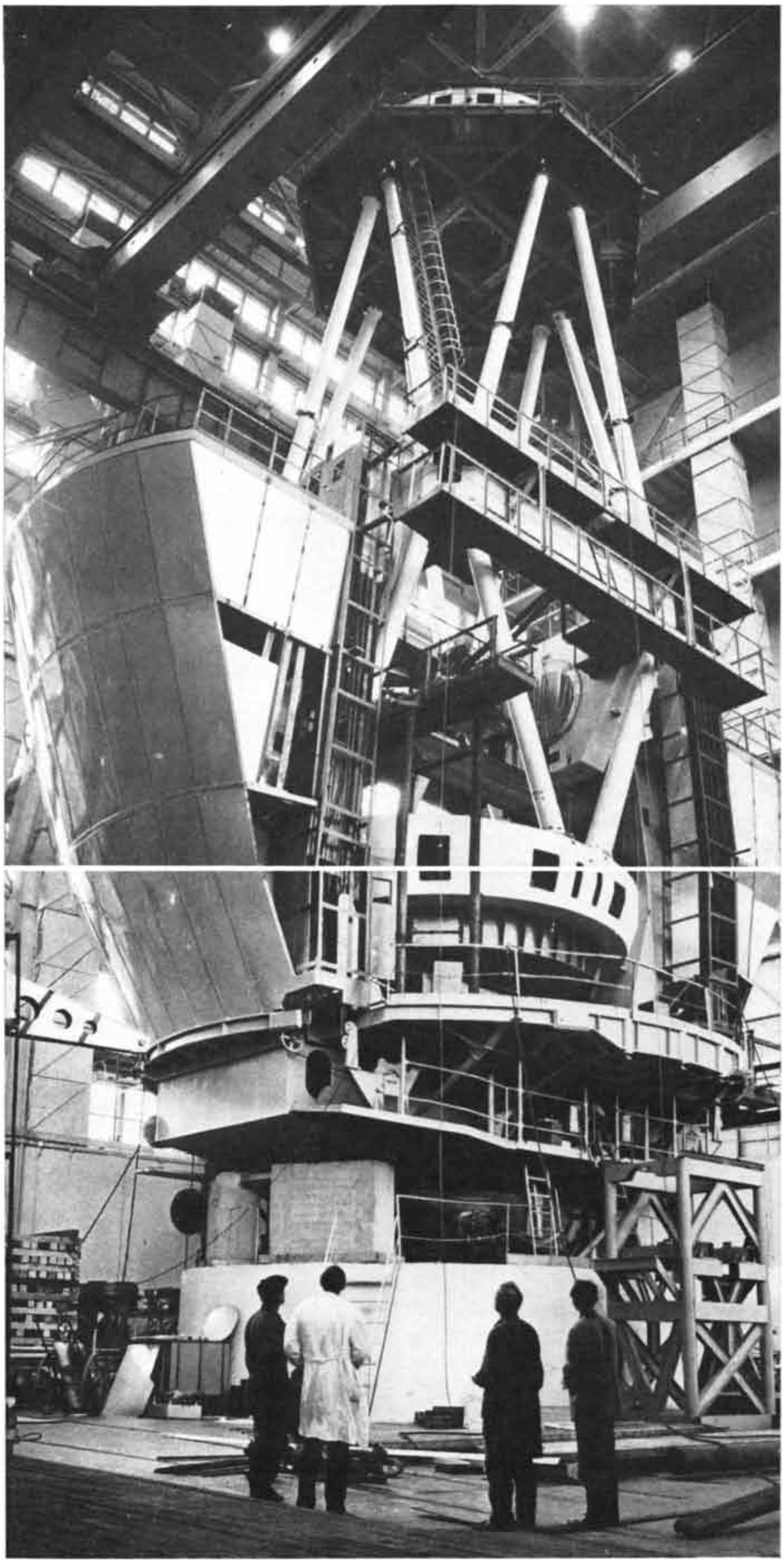
by R. W. Davies and R. Amann

The competition between the U.S. and the U.S.S.R. in space technology and advanced weapons is only one aspect of a major effort that has become the focus of attention in all industrial nations. This effort is "research and development." Research (the search for new knowledge) and development (the entire process of converting new knowledge into new processes and products) play a key role in industrial growth, public health and military power. There is a growing awareness, however, that the size of a nation's R. and D. budget is only one factor in the successful pursuit of scientific knowledge and sound technology. A great deal seems to depend on good management, and at least as much depends on providing adequate incentives, not only for discovery and invention but also for innovation: the task of carrying new ideas from the feasibility stage to commercial application. As a result there has been much critical discussion lately of how R. and D. is organized and applied in the two countries where it has reached its highest state.

It is widely supposed, particularly since the launching of the first Soviet artificial satellite, that the central planning of scientific and technical effort gives the U.S.S.R. a built-in advantage over the Western nations. This is not to say that the U.S.S.R. is at this stage ahead of the Western nations in science and technology but that to many observers it has seemed to be moving faster and with greater responsiveness to national and social needs. A closer examination of the Soviet R. and D. effort makes it clear, however, that this is an oversimplification. In the U.S.S.R. as elsewhere there is much questioning of current policies toward science and technology. An extensive study of Soviet science policy, just completed by an international team in which the authors of this article participated, shows that the U.S.S.R. is undertaking a reorganization of its R. and D. efforts and searching for new forms of incentive. The study was conducted for the Organisation for Economic Co-operation and Development by a group consisting of Eugène Zaleski of the National Center for Scientific Research in France, Joseph P. Kozlowski of the U.S. National Science Foundation, Helgard Wienert of the O.E.C.D. secretariat and the authors, assisted by our colleague M. J. Berry, of the Centre for Russian and East European Studies at the University of Birmingham. Similar O.E.C.D. surveys had previously been made of research and development policies in western Europe and North America. It is hoped that the comparative studies of the various strategies and experiences will be of value to all the countries concerned.

The starting point of our work on Soviet science was an attempt to estimate the magnitude of the R. and D. effort compared with that of the U.S. The comparison cannot be exact, because the two countries differ in their definitions of research and development. We can, however, make a rough estimate of the manpower engaged in these activities. Taking only those personnel who have completed a higher education, our Birmingham group calculates the number employed in R. and D. in the U.S.S.R. (for the year 1965) to be somewhere between 454,000 and 631,000, as compared with 497,000 in the U.S. (for the fiscal year 1963–1964). These figures do not include professionals doing research in the social sciences and humanities. Kozlowski puts the Soviet figure somewhat higher than we do; he calculates that in 1965 there were 588,000 engineers (not to speak of natural scientists, economists and so on) engaged in R. and D. in the U.S.S.R. On the whole it appears reasonable to conclude that in terms of manpower the U.S.S.R. is devoting at least as much effort as the U.S. to R. and D.

Of course this does not necessarily mean that the level of scientific and technical *output* in the U.S.S.R. is comparable to that in the U.S. The available evidence indicates that R. and D. productivity per man is distinctly lower in the U.S.S.R., except possibly in high-priority sectors such as military and space research. The lower productivity is mainly a reflection of the degree to which the U.S.S.R. lags behind the U.S. in economic development, education and technology. The total national income of the U.S.S.R., with its larger population, is no more than 60 percent of that of the



LARGEST SOVIET TELESCOPE will also be the world's largest when completed. It will have a mirror six meters (236 inches) in diameter, nearly a meter larger than the 200-inch reflector on Palomar Mountain. The new instrument will be located near Zelenchukskaya in the Caucasus at an altitude of 6,800 feet. In this picture, which consists of two separate exposures, the telescope is being assembled at the Optico-Mechanical Works in Leningrad.

U.S. More than 30 percent of the Soviet population is employed in agriculture, as against only 5 percent in the U.S. The median number of years of schooling that have been received by the adult population (age 16 or over) in the U.S.S.R. is only 6.6; in the U.S. it is 11.9. In the R. and D. sector, as in the Soviet economy as a whole, there is a relative shortage of the abundant equipment and lavish research facilities found in the more advanced U.S. economy.

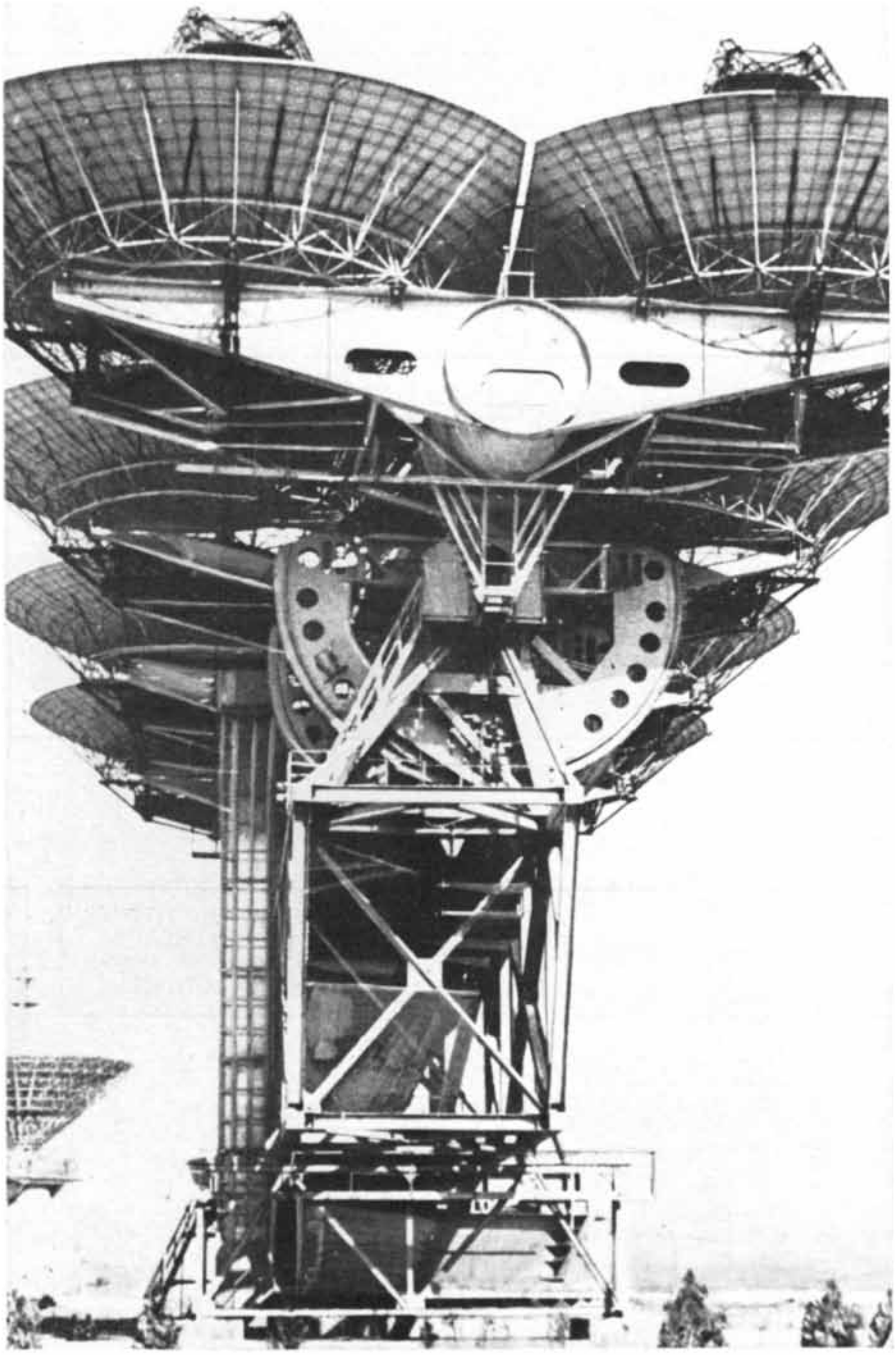
Just because the Soviet economy is less advanced than the U.S. economy the Soviet rivalry with the U.S. in advanced weapons and space programs has caused the Russians numerous difficulties. The pressure on material and qualified-manpower resources is considerable, particularly as the U.S.S.R. is currently applying an increased share of its resources to the production of consumer goods and the buildup of less advanced industries. It is therefore not surprising that, as in the U.S., its outlays on R. and D. have slowed markedly in the past few years. In the five years from 1957 to 1962 the total manpower employed in Soviet R. and D. more than doubled; in the four-year period from 1962 to 1966 it increased by only 28 percent [*see bottom illustration on page 26 and top illustration on page 27*]. There has been a similar slowing of the expenditure for R. and D. from an annual increase of 15 or 20 percent in the late 1950's to about 10 percent in recent years [*see bottom illustration on page 27*].

This is the background to the first of the critical issues now facing the Soviet policy-makers: How much of the national resources should be devoted to R. and D.? Officials of the State Committee for Science and Technology are pressing for a big rise in expenditure. Vadim A. Trapeznikov, first deputy chairman of the committee and a leading authority on automation, recently proposed an intensive program, involving an annual expenditure increase of 20 to 25 percent for R. and D., for the five-year period 1971–1975. He presented calculations suggesting that every ruble spent on research and development produced an annual increase of 1.49 rubles in national income, as against a return of only .39 ruble on each ruble invested in the economy as a whole. On the other hand, some officials of the State Planning Commission (Gosplan) have taken the opposite position. They are arguing against major increases in R. and D. expenditure and have cited figures purporting to show that the return on R. and D. is considerably less than Trapeznikov stated; they contend that the return fell short of the .74 ruble per ruble of expenditure that was expected in 1966. Thus a vigorous contest over the level of research expenditures seems to be under way at the top levels of the Soviet government.

The second major issue now troubling the Soviet policy-makers concerns the relative allocations to fundamental research and to development. We were not able to obtain precise figures on what proportions of the Soviet R. and D. resources are spent respectively on basic science, on applied research and on development. According to some Soviet statements, 10 to 12 percent of the R. and D. budget goes into fundamental research (about the same as the proportion devoted to science in the U.S.) and somewhat less than 50 percent goes into development (as against 66 percent in the U.S.). If these figures are accurate, about 40 percent of R. and D. in the U.S.S.R. is allocated to applied research, compared with 22 percent in the U.S. There would seem to be an odd paradox here: the U.S.S.R., which has always stressed the practical purposes of science, apparently spends considerably less of its R. and D. funds on translation of discoveries into the development of new processes and products than the U.S. A striking example of this relative skimping on development can be seen in the field of nuclear physics. The U.S.S.R. has poured huge investments into machines for basic studies in high-energy physics; it has built the world's most powerful particle accelerator, a proton machine of 76 billion electron volts (BeV), and is now considering whether to construct a 1,000-BeV accelerator (already designed in outline) or to go directly to one of 2,000 or even 10,000 BeV, perhaps in cooperation with other countries. Yet the U.S.S.R. has failed in the recent past to provide adequate funds for pilot facilities to develop nuclear power, according to a complaint voiced by the chairman of the committee responsible for the building of nuclear power plants.

There has been much public debate among Soviet scientists and officials on the need to correct this "imbalance" between research and development. All agree on the necessity of more attention to development; the discussion turns on whether the needed funds required should be found within the research budget or transferred from the general national capital budget. Leading scientists have objected to any invasion of the appropriations for basic science. At the general meeting of the U.S.S.R. Academy of Sciences last year the president opposed a proposal that the Academy enter on certain projects in applied research, on the ground that "this would divert resources intended for fundamental research."

The conflict over basic versus applied science is bringing to the fore important questions concerning the system of organization of R. and D. in the U.S.S.R. In certain respects this system differs sharply from the arrangements in the U.S. and Britain. Whereas in the Western countries basic research is conducted mainly in universities, in the U.S.S.R. most of it is carried on by academies that are not primarily educational institutions. A second important difference is that applied research and development in the U.S.S.R. are conducted principally in establishments directly under the control of the government departments responsible for particular industries, in contrast to the situation in the U.S., where industrial corporations carry on most of the applied research and

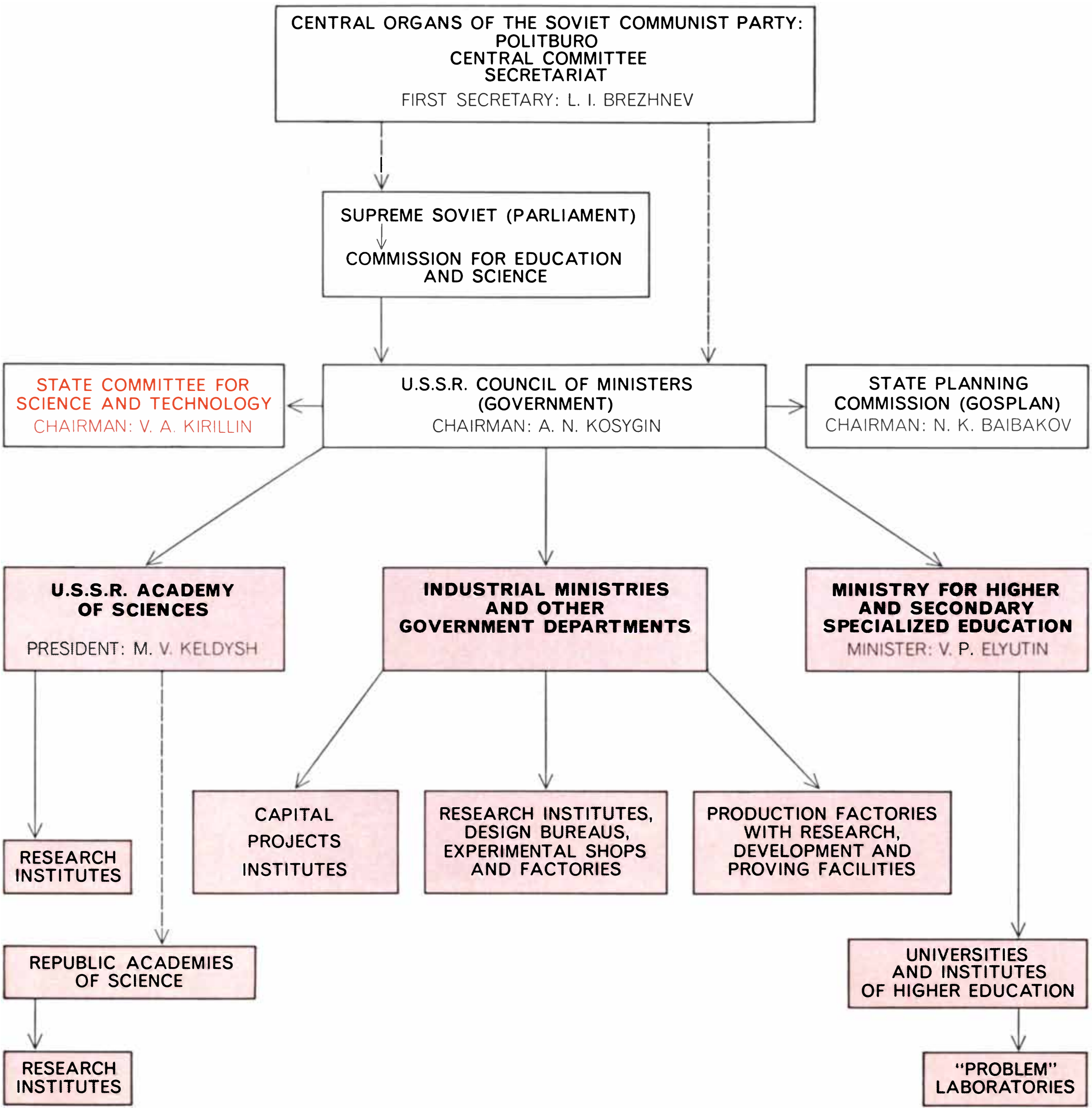


ARRAY OF ANTENNAS at a Soviet deep-space facility is used to track vehicles sent on interplanetary missions. The array provided two-way communication with *Venera 5* and *Venera 6*, which ejected instrumented capsules into the atmosphere of Venus last month. The capsules made measurements of the composition and pressure of the planet's atmosphere.

development, even that paid for by the government.

To understand the problems confronting the U.S.S.R. we need to take a more detailed look at its research organization. At the center of the system, but not in control of it, stands the Academy of Sciences [*see illustration below*]. In Czarist Russia the Imperial Academy of Sciences was primarily a club of distinguished scientists like the Royal Society of London. The Bolshevik revolution and the subsequent five-year plans transformed the Academy into a national center for fundamental research that now maintains nearly 200 scientific establishments spread over the U.S.S.R. from Leningrad all the way to eastern Siberia. In addition there are special research academies for medicine, agriculture and education, and each of the principal Soviet republics has an Academy of Sciences of its own. The network of academies as a whole employed in 1965 some 74,000 scientists, all of them engaged in research rather than teaching [*see top illustration on page 26*].

The various departments or ministries of the central government operate a



ORGANIZATION OF SOVIET SCIENCE fans downward from the central organs of the Communist party into three groups that administer all scientific activity in the U.S.S.R. They are the U.S.S.R. Academy of Sciences (together with the loosely affiliated academies in the separate Soviet republics), the industrial and governmental ministries and the ministry for higher and secondary specialized education. The last of the three groups is concerned primarily with education; first two groups sponsor the lion's share of research and development. The "Academy system" employs about 20 percent of all scientific workers in research and development, and the "Ministry [or branch] system" 80 percent. The State Committee for Science and Technology is responsible for the coordination of research and development. It has recently been pushing for a larger budget than the State Planning Commission thinks can be justified.

separate network of research institutes, known as the "branch" system, that is devoted mainly to research and development for the branches of industry and government. This system employs more than 316,000 scientists—about four-fifths of all the Soviet scientists who work in R. and D.

Characteristic of the difference between the research setup in the U.S. and that in the U.S.S.R. is the fact that, whereas the first nuclear reactor in the U.S. was built on a university campus (in the squash court of Stagg Field at the University of Chicago), the first reactor in the U.S.S.R. was constructed in a laboratory attached to the Academy of Sciences. In the U.S.S.R. creative research workers in the sciences and the humanities are attracted to the academies rather than to the institutions of higher education (VUZy), which are largely preoccupied with education. The VUZy do conduct some research, and in recent years the ministry in charge of them has increased their research facilities, so that they now have more than 700 research institutes and laboratories attached to them. The main lecturing staff, however, has such a heavy teaching schedule (generally 18 hours a week for 40 weeks a year) that the lecturers are left with little time for serious research. At a recent confrontation between the Minister of Higher Education and the presidium of the Academy of Sciences the president of the Academy tried to discourage proposals that research resources be shifted toward the VUZy network. He spoke approvingly of the supposed arrangements at the universities of Oxford and Cambridge, where, according to his account, "there is a category of 'tutors' who simply teach, nothing else being required of them."

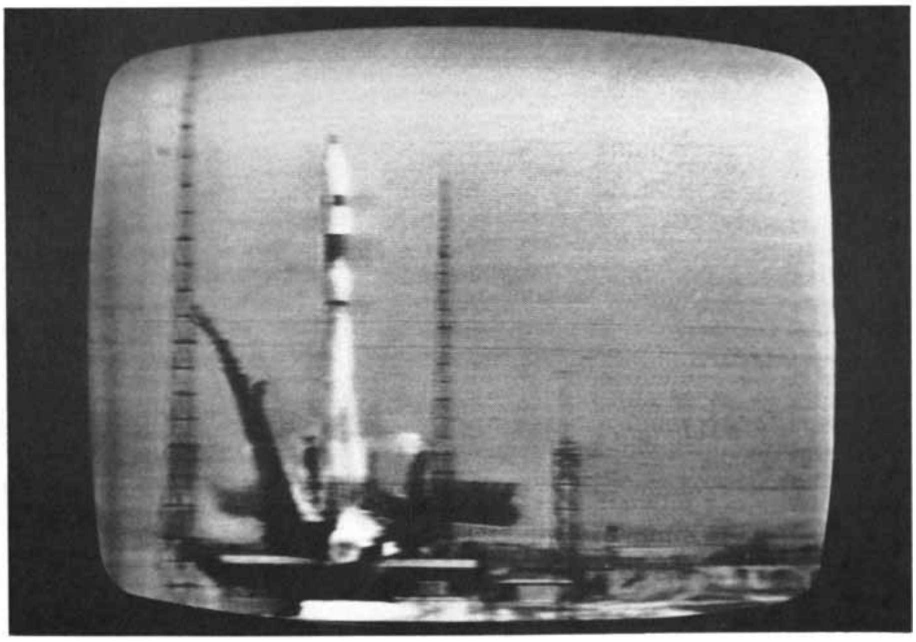
The branch system for industrial research and development is similarly under question and undergoing reexamination. The issue there turns on giving industrial firms or enterprises a larger role in R. and D. Some of them, notably in the heavy-electrical-equipment and machine-tool industries, already have developed this function in a major way, with well-equipped research laboratories and design departments attached to their factories. In general, however, most industrial research in the U.S.S.R. remains centralized in facilities operated by the government ministries. We found this reflected, for instance, in the awards of patents (called "certificates of authorship" in the U.S.S.R.); an analysis of those issued to identifiable inventors in a sample six-month period showed that about 52 percent went to R. and D. establishments under the ministries and only about 14 percent to industrial firms.

The centralization of industrial R. and D. is a heritage of the high-pressure drive for industrialization that was launched by the U.S.S.R. four decades ago. It enabled the government to mobilize the nation's scarce scientific and technical manpower, to channel in know-how from abroad, and to proceed rapidly with the planning and design of power stations, factories and products on a nationwide scale. In the past 10 years, however, it has become increasingly clear to the Soviet scientific community and to the government that a system designed to push a largely undeveloped economy through an elementary industrial revolution is not necessarily appropriate for the U.S.S.R. in its present state of industrialization, when it must cope with rapid technological change and is anxious to outstrip the U.S. technologically and economically. In hundreds of self-critical articles and speeches leading industrialists, scientists and politicians in the U.S.S.R. have recently urged that there is a pressing need to close the big technological gaps that exist between the U.S.S.R. and the U.S. in many industries.

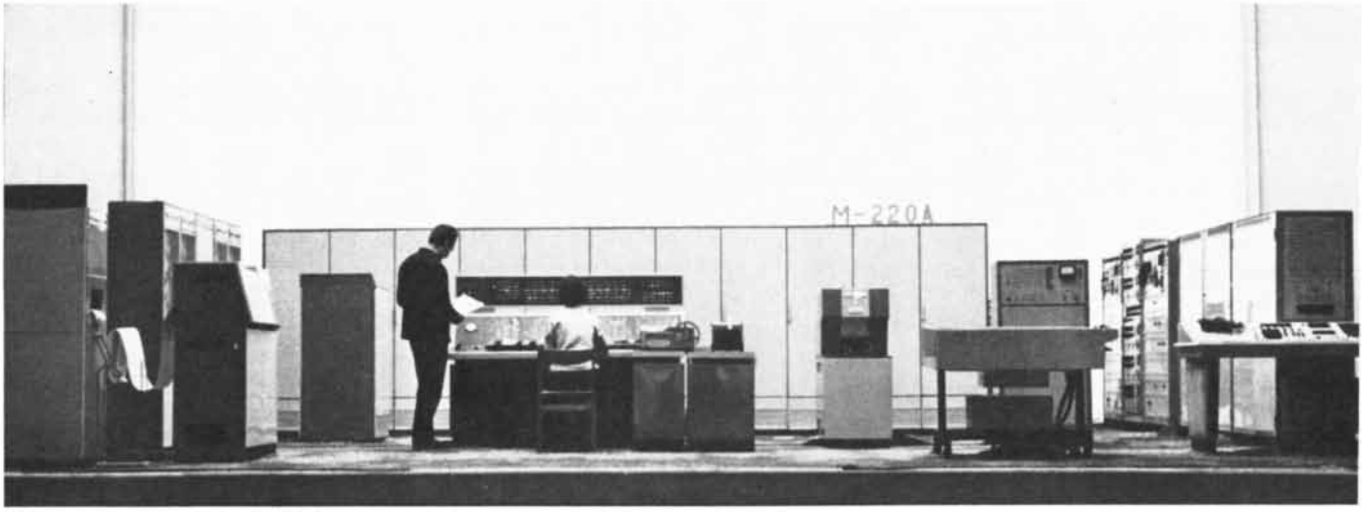
It is in this general context that Soviet policy-makers have undertaken in the past 10 years a massive reorganization of the R. and D. system. The reconstruction has two main aims: to increase the efficiency of research by improving its coordination, and to improve the system for applying research findings to industrial development.

Lack of coordination has been a persistent complaint of Soviet research workers for more than three decades. Writers on the subject have deplored the existence of considerable duplication of effort and poor communication between the different parts of the research system. Responsibility for research has been divided between the Academy system and the branch system; it has also been split up further among many different authorities within each system. Science expenditure has lacked unified control: the funds have come from diverse sources, mainly from the central state budget but also in part from industry, and within the state budget R. and D. has been divided among a number of heads.

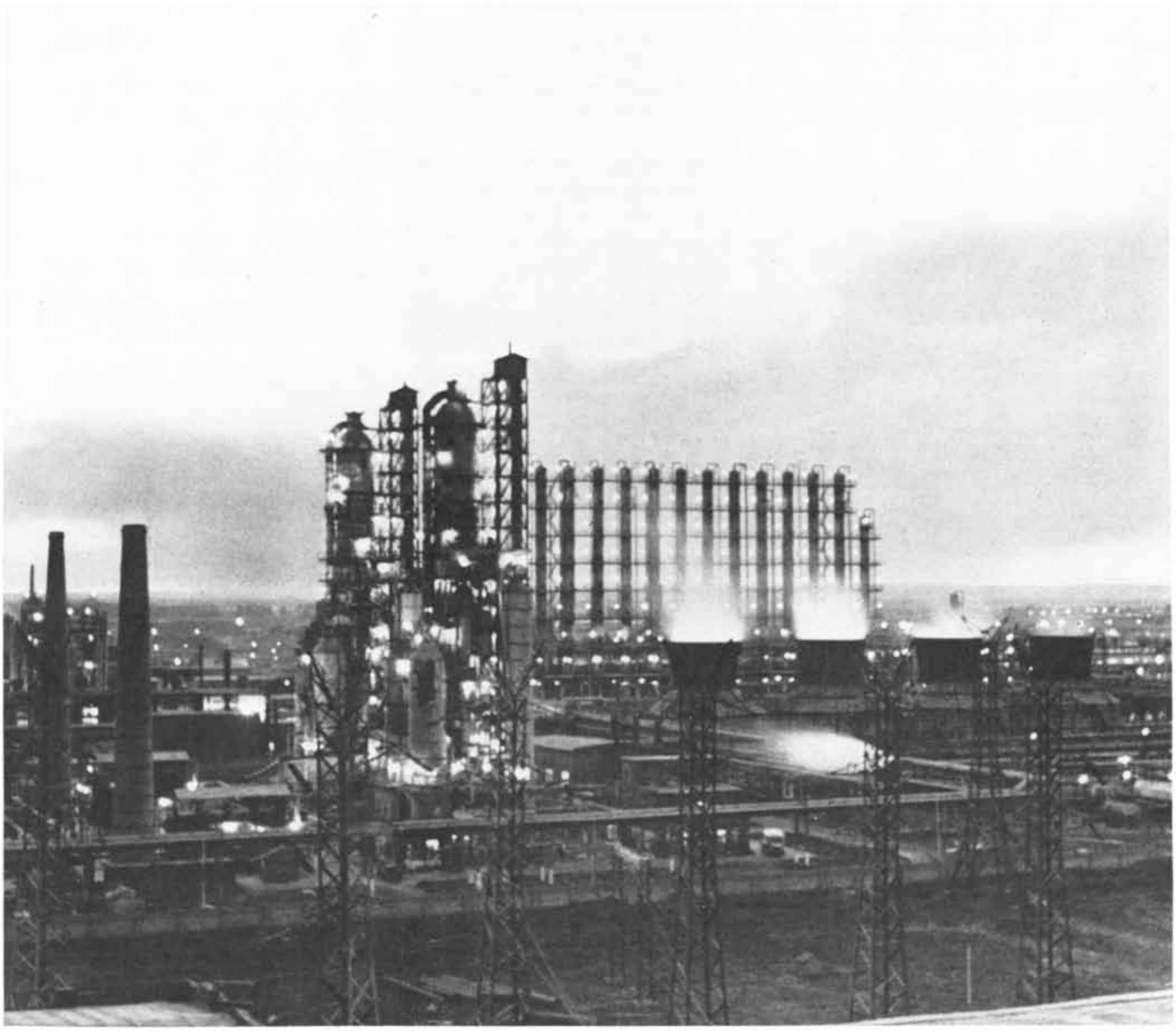
An important first step toward greater coordination was taken during World War II, when the U.S.S.R., like the U.S., set up large agencies, supported by its defense department, for certain major projects. The most important of these was the Soviet equivalent of the Manhattan project in the U.S., headed by Major General Leslie R. Groves; in 1943 the U.S.S.R. organized an atomic-bomb program under a Scientific and Technical Council with a well-known industrial administrator as chairman and a leading nuclear physicist, Igor V. Kurchatov, as deputy chairman. Kurchatov's role was comparable to that of J. Robert Oppenheimer. After the war the U.S.S.R. made



LAUNCHING OF *SOYUZ 4* on January 14 was shown on Soviet television. *Soyuz 4*, containing one cosmonaut, later docked in orbit with *Soyuz 5*, containing three cosmonauts. Two of the men in *Soyuz 5* transferred to *Soyuz 4*, after which the two spaceships landed.



ELECTRONIC COMPUTER M-220A, one of the latest Soviet models, resembles machines built in the West. The U.S.S.R. is now making plans to install a nationwide computer system for controlling the flow of raw materials, intermediate products and finished goods.

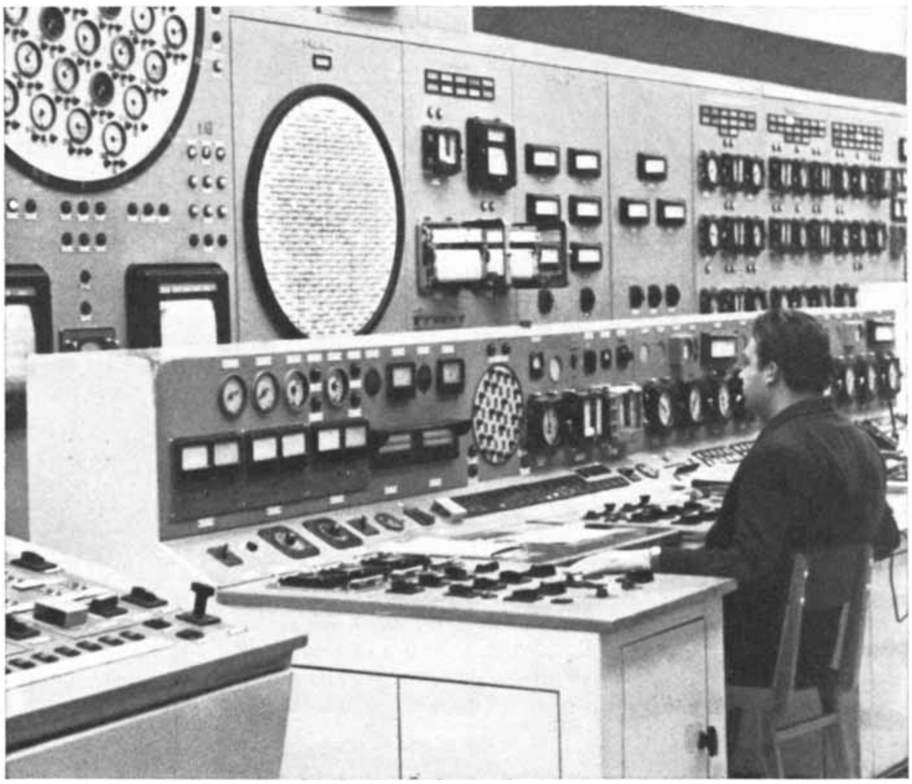


PETROCHEMICAL PLANT at Sterlitamak in the Bashkir Republic represents one of the largest Soviet construction projects of the past decade. One of the plant's main products is synthetic "natural" rubber (polyisoprene). For the past dozen years the U.S.S.R. has been giving high priority to the production of synthetic rubber, plastics and synthetic fibers. In the period 1959–1965 its output of these materials grew at an average rate of 19 percent a year compared with a growth rate of about 9 percent for industry in general.

several attempts to organize a coordinating agency for its R. and D.; the problem proved to be so difficult, however, that it did not succeed in forming a viable organization until 1961. The State Committee for Coordination of Scientific Research, established that year, was reorganized in 1965 as the State Committee for Science and Technology, and that body is now in general charge of all applied research and development.

Headed by Vladimir A. Kirillin, an outstanding thermodynamicist who carries much influence with political authorities as well as with his fellow scientists, this agency has a number of departments, dealing with specific branches of the economy, with coordination, with planning and with the financing of R. and D. It also has a number of advisory councils, each concerned with a broad scientific or technological problem. The committee has found it very difficult, however, to influence the direction of scientific activities or to institute a system of general coordination. Consequently, instead of attempting to plan all expenditures on research it has since 1965 been concentrating on a fairly small number of major scientific and technical problems. It is also sponsoring assessments of long-range scientific and technological planning.

Among the long-range projects are studies of the national resources of fuel and power and of the development of an integrated transport system, made by commissions whose reports have now been approved in principle by the government. For the short run the State Committee for Science and Technology has tried to plan for five-year periods or longer rather than for each year. Its plan for the 1966–1970 period dealt with about 250 problems, including the design of power plants, the construction of a 1,500-kilovolt transmission line from Kazakhstan to the center of European Russia and a start on the construction of a nationwide automatized system for the recording and control of supplies of materials and machines. Each "coordination plan" covers the whole cycle of creation of a facility from research to production and secures the necessary financing for the project.



ATOMIC POWER CONTROL ROOM in Novo-Voronezh regulates a water-cooled reactor of 365,000 kilowatts. The U.S.S.R. placed the first atomic power plant in operation in 1954.

It is not yet clear how effective these recent changes have been. Since the Soviet economic system is operated on a year-to-year basis by a detailed central plan, the committee's attempt to disregard minor projects and to cut down on annual planning has run into considerable criticism. Officials of Gosplan complain that the policy has "resulted in substantial deterioration in the planning of research" and has brought about a decline in the R. and D. performance of industrial ministries.

Fundamental research remains under the jurisdiction of the Academy of Sciences, which has been assigned full responsibility for the coordination and planning of work in the natural and human sciences. The presidium (governing body) of the Academy is charged with the task of preparing a comprehensive national plan for all research in these fields, wherever undertaken. The main responsibility here rests with the four sections into which the Academy is now divided, each of which is headed by a strong vice-president (the sections are physics and mathematics, chemistry and biology, the social sciences and the earth sciences).

In practice the efforts of the Academy to coordinate fundamental research have met many difficulties. Some of the formal procedures for preparing research plans are excessively bureaucratic, and there is much evidence that its influence on the other agencies involved in fundamental science (the VUZy and the other academies) is rather weak. Nevertheless, there are good signs of improvement in coordination. In recent years the U.S.S.R. has increasingly gone in for what is known in the West as "mission-oriented" fundamental research: the deliberate search for basic knowledge that is specifically required for a particular purpose. An increasing proportion of the support for science is being allotted to funds that are earmarked for special projects in which various disciplines are enlisted to work on a particular problem. A typical example is an extensive project now under way for study of food resources in the ocean. The presidium of the Academy first instructed two of its divisions to prepare, in collaboration with the Ministry of Fisheries, a program of needed studies. After receiving a preliminary report from this survey, the Academy instructed a number of its own establishments to undertake specified research projects and at the same time also requested three independent authorities (the ministries of fisheries and higher education and the Ukrainian Academy of Sciences) to carry out certain related studies. The central Academy assigned responsibility for coordinating all these research activities to one of its advisory councils.

The introduction of mission-oriented programs does not seem to have seriously cramped the general independence or autonomy of Academy scientists who are pursuing basic research. In the budget grant to every Academy institute a definite part is allocated to entirely free research. Even within the grants earmarked for specific projects the institute

	1956		1961		1965	
	SCIENTIFIC ESTABLISHMENTS		SCIENTIFIC ESTABLISHMENTS		SCIENTIFIC ESTABLISHMENTS	
	NUMBER	SCIENTISTS	NUMBER	SCIENTISTS	NUMBER	SCIENTISTS
U.S.S.R. ACADEMY OF SCIENCES	195	15,716	167	19,068	193	25,471
REPUBLIC ACADEMIES OF SCIENCES	295	8,673	401	21,846	349	26,709
ACADEMIES DEPENDING ON MINISTRIES OR STATE COMMITTEES	248	7,761	319	16,676	416	21,738
TOTAL FOR ACADEMY SYSTEM	738	32,150	887	57,590	958	73,918
PERCENT	26.8	30.2	21.3	24.1	20.3	18.9
MINISTRIES, STATE COMMITTEES, DEPARTMENTS AND ECONOMIC COUNCILS	2,018	74,250	3,285	180,910	3,766	316,482
PERCENT	73.2	69.8	78.7	75.9	79.7	81.1
TOTAL	**2,756**	**106,400**	**4,172**	**238,500**	**4,724**	**390,400**

DIVISION OF R. AND D. EFFORT between the Academy system and the branch system has been shifting. In 1956 26.8 percent of the scientific establishments and 30.2 percent of the scientists were under the jurisdiction of the academies, with the remainder under the ministries. By 1965 these percentages had dropped to 20.3 and 18.9 percent respectively as the number of ministry scientists more than quadrupled. Within the Academy system the republican academies grew much more rapidly than the U.S.S.R. Academy itself.

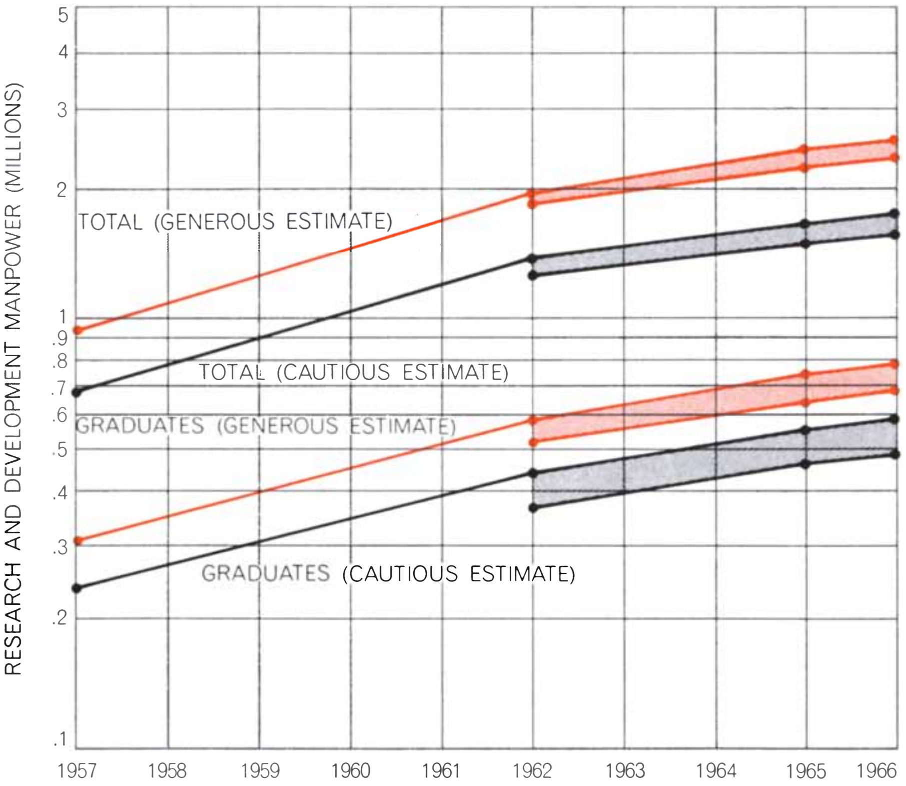
	1962		1966	
	CAUTIOUS ESTIMATE (1,000)	GENEROUS ESTIMATE (1,000)	CAUTIOUS ESTIMATE (1,000)	GENEROUS ESTIMATE (1,000)
A. SCIENTIFIC WORKERS IN R. AND D. (INCLUDING SOCIAL SCIENCES AND HUMANITIES)				
1. SCIENTIFIC ESTABLISHMENTS, PROJECTS AND DESIGN ORGANIZATIONS NOT SERVING CONSTRUCTION				
a. SCIENTISTS	299	299	418	418
b. SUPPORTING PERSONNEL	780	1,140	950	1,406
c. TOTAL	1,079	1,439	1,368	1,824
2. PROJECT AND PROJECT-SURVEY ORGANIZATIONS	45	45	53	53
3. HIGHER-EDUCATION ESTABLISHMENTS				
a. SCIENTISTS	45	45	59.5	59.5
b. SUPPORTING PERSONNEL	45	45	59 5	59.5
c. TOTAL	90	90	119	119
4. PERSONNEL ENGAGED IN R. AND D. AT ENTERPRISES	150	300	198	396
5. RESEARCH STUDENTS	49	49	73	73
TOTAL	**1,413**	**1,923**	**1,811**	**2,465**
B. SCIENTIFIC WORKERS IN SOCIAL SCIENCES AND HUMANITIES				
1. a. SCIENTISTS	41	41	65	65
b. SUPPORTING PERSONNEL	27	39	37	55
c. TOTAL	68	80	102	120
3. a. SCIENTISTS	19	19	26	26
b. SUPPORTING PERSONNEL	10	10	13	13
c. TOTAL	29	29	39	39
5. RESEARCH STUDENTS	12	12	15	15
TOTAL	**109**	**121**	**156**	**174**
SCIENTIFIC WORKERS IN R. AND D. (EXCLUDING SOCIAL SCIENCES AND HUMANITIES)	*1,304*	*1,802*	*1,655*	*2,291*

TOTAL SOVIET R. AND D. MANPOWER in 1966 was 1,811,000 according to a cautious estimate or 2,465,000 according to a more generous estimate, both figures being some 28 percent above comparable estimates for 1962. The number of workers formally classified as "scientists" is the same according to both estimates; the difference is in the number of supporting personnel. Comparable U.S. figures are not available, but the National Science Foundation estimates that the number of U.S. scientists and engineers engaged in R. and D. in 1966 was 520,500. Between 7 and 9 percent of Soviet R. and D. workers are in the social sciences and humanities.

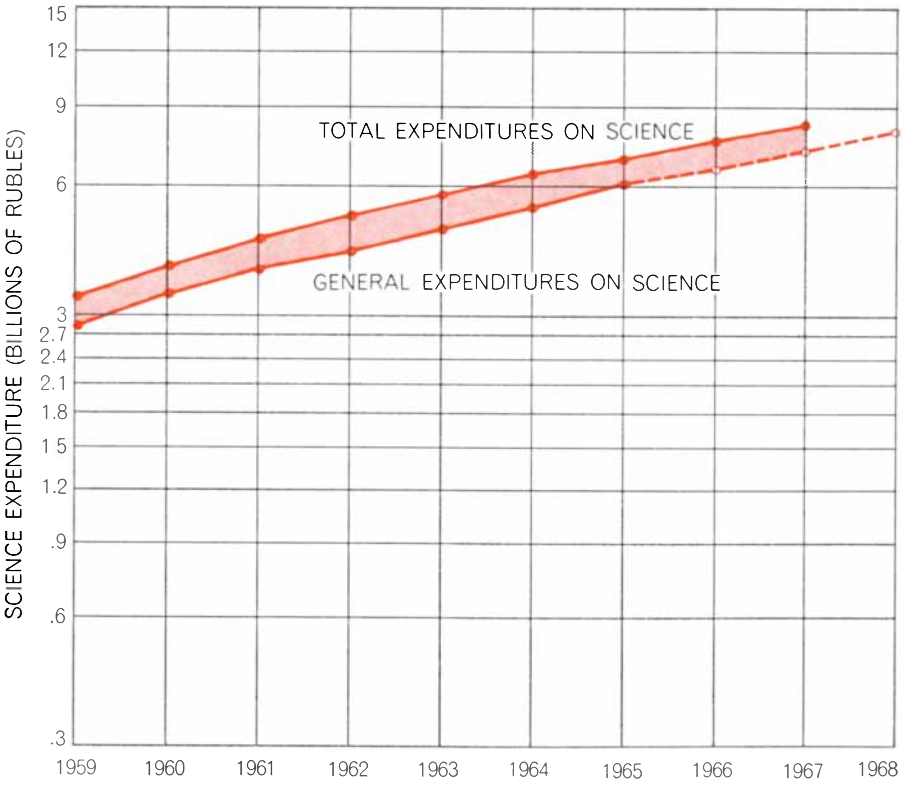
director and the investigators assigned to the research have considerable latitude in practice to decide on the actual program of studies for themselves. Moreover, the scientific community in the U.S.S.R. exercises a strong influence on the government's science policies. The central committee of the Communist party includes a number of members of the Academy of Sciences, and key administrative posts throughout the R. and D. system are held by scientists rather than by professional administrators or politicians. Norman Kaplan, an American sociologist who studied the administration of medical research in the U.S.S.R., found that the research institutes in this field are less subject to control by nonscientist administrators than similar institutes in the U.S.

What progress has the U.S.S.R. made on the second major problem, namely the practical application of its research results in the form of new products and industrial processes? This has received even more attention than the coordination of research. Everyone agrees that the development side of R. and D. is weak, and that even after industrial innovations are developed they diffuse into the economy at large too slowly.

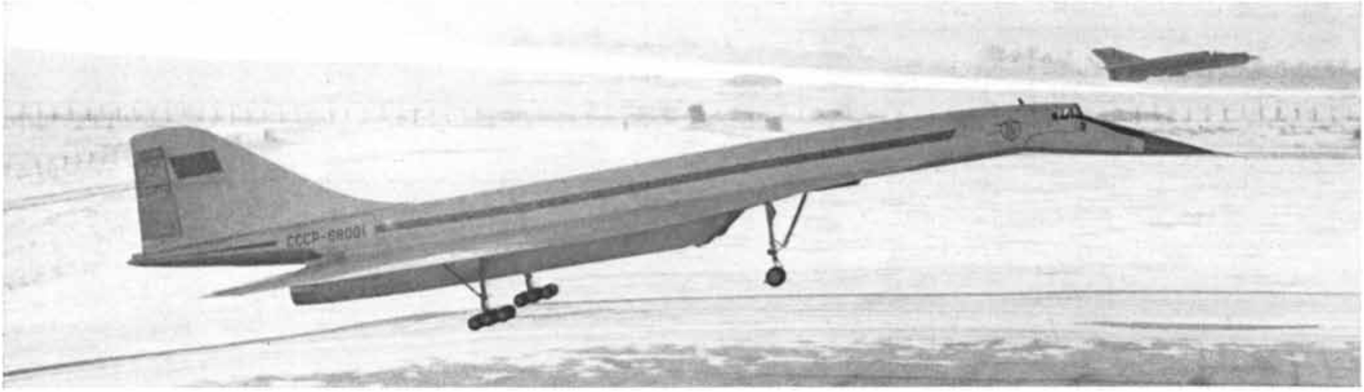
Two main factors appear to have contributed to this situation. The first is the organizational disunion we have already described: the separation between the academies' research establishments and the industrial R. and D. network, the fragmentation of responsibility in the latter network among a number of different ministries and the divorce of the factories in turn from the ministries' central institutes. The managers of Soviet industrial enterprises frequently complain that Academy institutes do not appreciate the needs of industry and that research and design staffs in the ministry institutes give too little regard to the requirements of the factory floor or the needs of the consumer (similar complaints are, after all, not unheard of in Western countries). The second factor standing in the way of the application of research results has been the strong orientation of Soviet economic planning toward production rather than innovation. The industrial ministries and factories have been judged primarily by their ability to meet the production goals set up for them and to stay within the limits placed on costs. As a result the ministries have skimped on expenditures for development and commercial application of new products and processes and the factory managements have re-



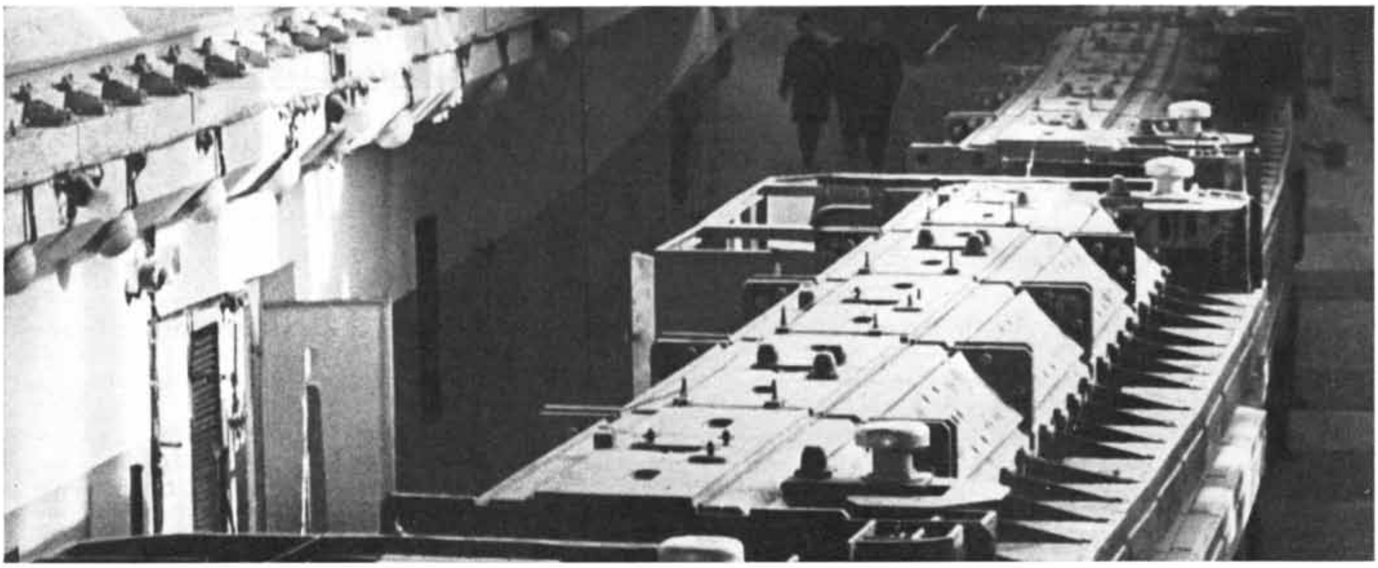
SOVIET SCIENTIFIC MANPOWER has been growing since 1957 at an annual rate of 11.5 percent, using either "generous" or "cautious" estimates. These logarithmic plots show, however, that the rates have been declining slightly; if constant, the lines would be straight. The shaded areas indicate the manpower fraction in the social sciences and humanities.



GROWTH OF SOVIET SCIENCE EXPENDITURES has been at the rate of about 12 percent a year, but again the rate has been slackening. The shaded area represents capital investment in facilities. The broken lines indicate planned expenditures. The actual total expenditure in 1967 was 8.2 billion rubles. It is estimated that the ruble buys between $2.50 and $3.50 worth of R. and D. at U.S. prices, so that the Soviet outlay on "science" fell between $20.5 billion and $28.7 billion. In 1967 the U.S. spent $23.8 billion on R. and D.



SUPERSONIC TRANSPORT, TU-144, made its first flight on December 31, 1968, several weeks ahead of the *Concorde*, the British-French supersonic transport. The TU-144, designed to carry 120 passengers at 1,500 miles an hour, may be in regular service by 1971.



ACCELERATOR AT SERPUKHOV near Moscow is more than twice as powerful as any in the West. Placed in service last year, the machine accelerates protons to an energy of 76 billion electron volts. Only a small part of the accelerator track appears here.

sisted innovations whose introduction might interrupt or slow down the flow of production.

Serious efforts to break through these obstacles have been made in the present decade. Since 1961 an annual plan for the introduction of research results into industry has been incorporated in the annual national economic plan. This has forced an increase in attention to the need for innovation but has proved to be at best a clumsy instrument, running into difficulties that we have already mentioned. The Soviet authorities have therefore experimented with changes in the basic structure of the R. and D. establishment that are designed to bring research, development and production more closely together.

One of these new organizational arrangements is a "factory center" where research institutes, design bureaus and a production enterprise are all combined in a single administrative unit. This scheme has found particular application in the heavy-engineering and machine-tool industries. Another system that may be more widely applied is the bringing together of research institutes working in related fields to form a research and teaching complex that is closely associated with industries in the area. Notable examples of institutions of this type are the long established Leningrad Polytechnical Institute and the relatively new Novosibirsk Science Center and University. A third experimental form might be called a "research corporation"; this works for industry as a kind of service organization, providing it on a fee basis with a variety of services from expert counseling on improvement of production processes to the design and construction of a completely new plant.

These Soviet experiments obviously resemble arrangements that are familiar in the West, particularly in the U.S.: the large industrial firm with its own elaborate R. and D. organization, the university with laboratories that perform applied research in close association with industry, the research firm that contracts its services to various industrial producers, the engineering company that designs and builds plants. In the U.S.S.R. these arrangements are being applied in a cautious way, with attention to retaining the outstanding features of Soviet R. and D.: large-scale operation, with its attendant economies, and the ability to focus on major national objectives. In the next decade or so the U.S.S.R. is likely to develop an interesting combination of its own traditional approaches and Western arrangements for R. and D.

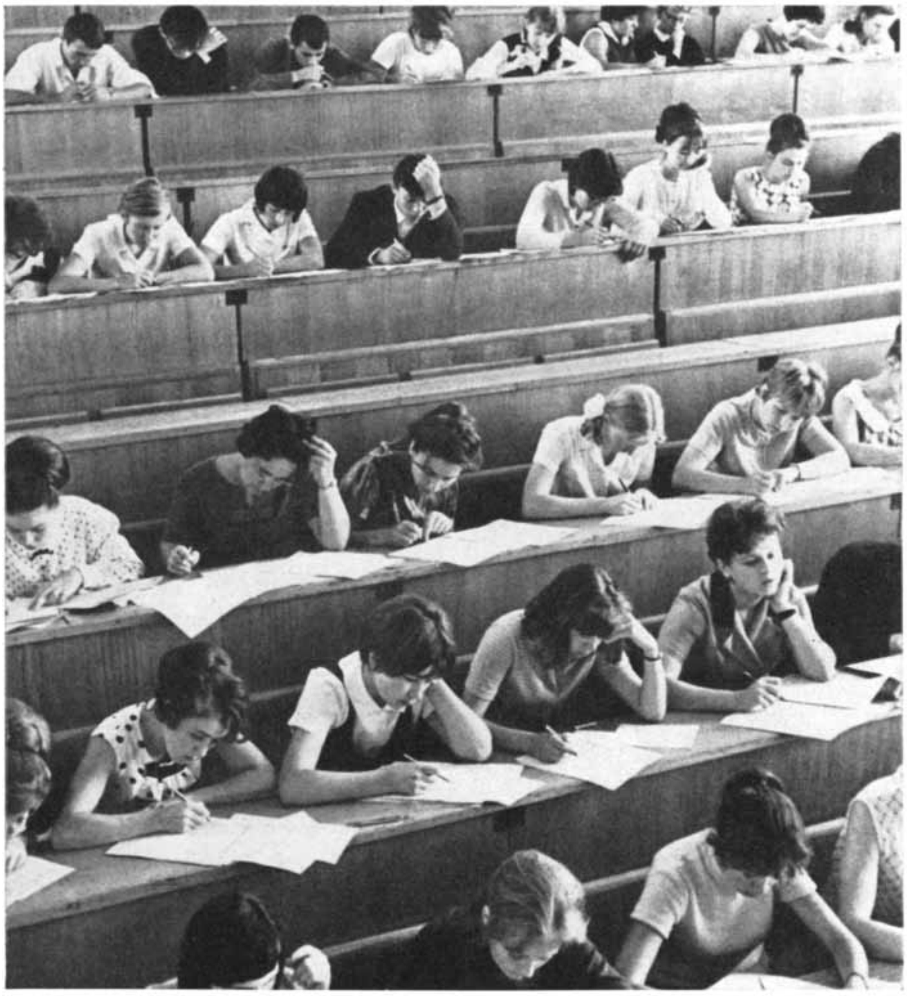
These ingenious and promising experiments to bring about a closer organizational union between research and production still leave the U.S.S.R. with the need to provide incentives to innovation throughout the economic system and to tailor products more directly to the requirements of industry.

Two proposals toward these ends have been approved and are now being put into effect. One is a system of contracts that will give the users of the products a greater voice in the direction of research. A large part of R. and D. resources will no longer be paid direct to the research organization from the state budget; instead it will be made available to ministries or to individual enterprises, which will place orders for the development or design of specified products. The other proposal is that more competition be introduced into R. and D. No one establishment, it is argued, should have a monopoly on a particular problem or a particular field. It is anticipated that the element of competition among R. and D. organizations will lead to a greater fertility of new ideas and products. Competition would, however, be restricted to the mock-up (or at most the prototype) stage; the Soviet system, it is argued, must avoid the wasteful duplication of production facilities characteristic of private capitalism.

These proposals, although important, still represent only a small step toward the effective promotion of innovation. They still leave the ultimate responsibility for approving and financing new products and processes with government planners and industrial ministries; they deal with the early stages of innovation but not with the crucial issue of the encouragement of commercial production. In the long run the Soviet government intends to deal with these difficulties by a much greater use of the profit motive to encourage industrial enterprises to innovate, but progress in this direction has been hindered by the Soviet pricing system. The prices assigned by the state to new products often fail to provide for the initial costs of tooling up for and introducing the products and are certainly not sufficient to encourage industrial firms to introduce new products on their own initiative. A number of Soviet economists have therefore proposed that prices for new products, instead of being fixed by a central state agency, be negotiated between the producer and the consumer, thereby allowing the market force of demand to operate in encouraging innovation.

Side by side with the search for economic incentives for R. and D. organizations and industrial firms, government leaders and planners in the U.S.S.R. are now making strenuous efforts to provide personal financial incentives to scientists, engineers and production workers, designed to foster inventions and their rapid application in industry. Considerable use is being made, for example, of bonuses promised in advance to scientists and designers for the prompt fulfillment of research projects and bonuses to factory workers for the successful development of new processes or products. Attempts are being made to relate the size of the bonus to the gain the economy will realize from the innovation.

The theory, or assumption, that seems to lie behind the great stress on economic incentives to organizations and individuals is that the economic gains derived from inventions and their application should be used to provide an automatic mechanism for stimulating innovation. Paradoxically, the Soviet leaders appear to be prepared to accept, within a framework of government planning, at least as great a reliance on the financial incentive to the individual and the profit motive as is now regarded as realistic among most economists in the private-enterprise economies of the West. Western experience would seem to indicate that personal bonuses cannot easily be related to the fate of particular projects. It is also fairly clear that the great emphasis on R. and D. and innovation by large corporations in the U.S. cannot be explained entirely by the higher prices to be obtained from new products. Much of the motivation for successful R. and D. flows from the competitive situation: companies must continually develop new processes and products to forestall rivals and to maintain their share of the market. On the other hand, most authorities in the Western countries believe that in many fields the spur of competition is no longer effective in promoting innovation, and that consequently there is a need for new economic relationships between government and industry and new methods of government management of research. Hence the efforts in the coming years on both sides—in the U.S.S.R. and in the Western nations—to devise satisfactory organizational forms for R. and D. and to develop satisfactory incentives for innovation should be of considerable mutual interest.



"SCIENCE CITY," Akademgorodok near Novosibirsk in western Siberia, was created several years ago to provide advanced training for young people with a special aptitude for mathematics and physics. This photograph shows some of the 3,200 students who took examinations at Akademgorodok last summer for admission to Novosibirsk state university.

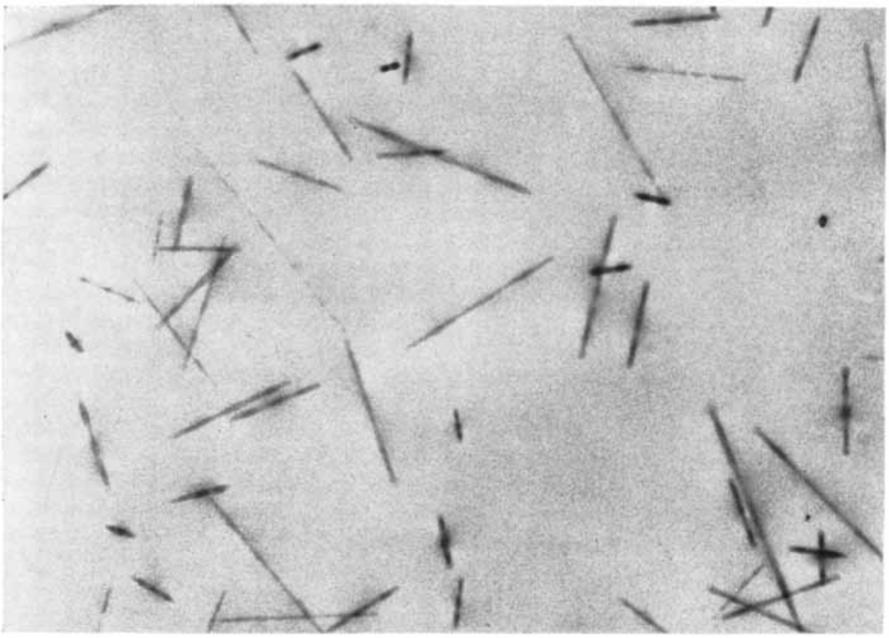
Nuclear Tracks in Solids

Charged particles leave damage trails in nonconducting solids such as mineral crystals, glasses and plastics. The etched tracks provide information on the age of the solid or the nature of the particles

by R. L. Fleischer, P. B. Price and R. M. Walker

In the 20th century physicists have devised ingenious systems for studying invisible subnuclear particles by visualizing and analyzing the tracks they leave in cloud chambers, photographic emulsions, bubble chambers and spark chambers. There is an older system of particle detection—one almost as old as the solar system. For billions of years the tracks of cosmic rays and of various products of radioactive decay have been recorded in naturally occurring solids.

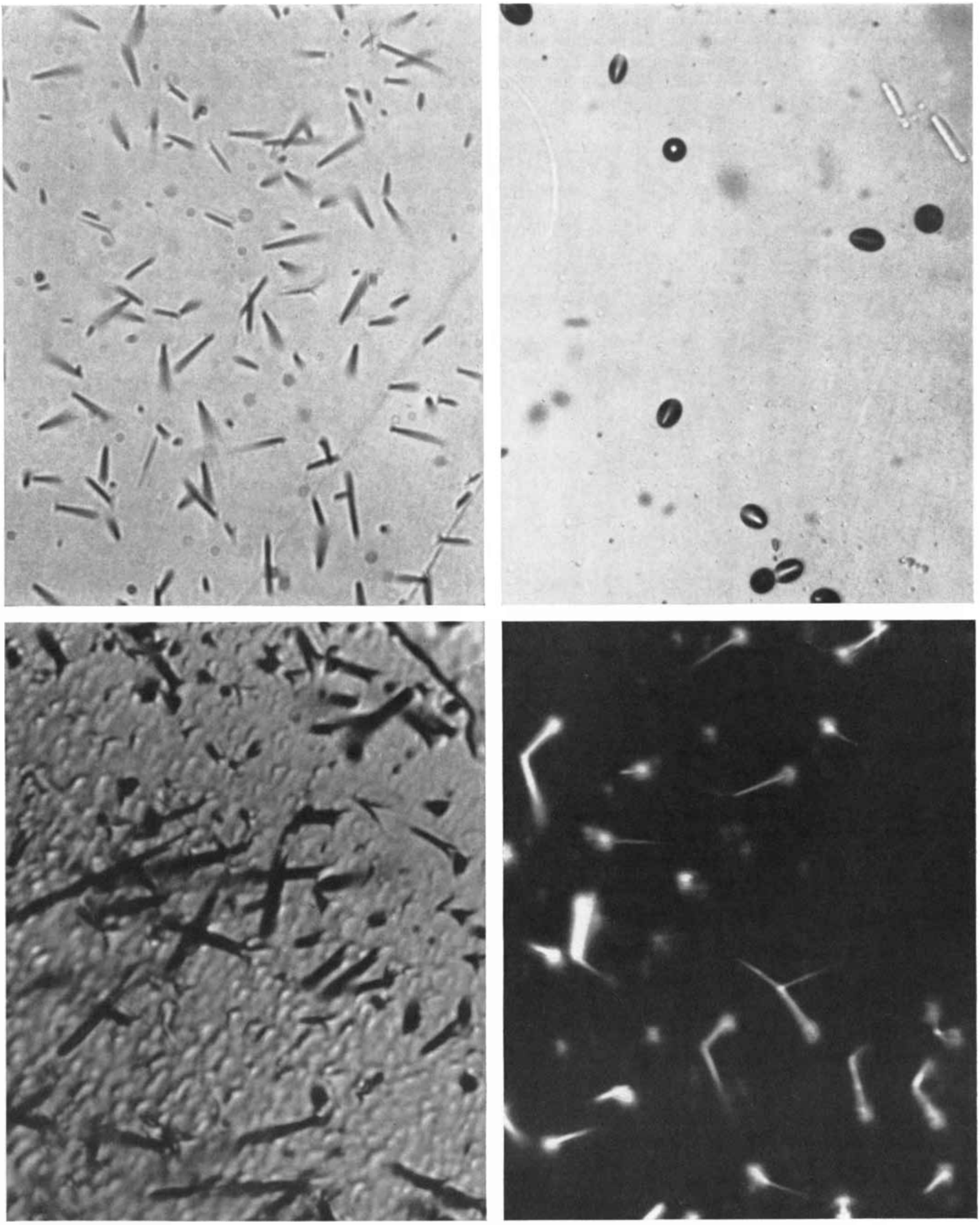
Seven years ago, at the General Electric Research and Development Center, we learned how to make these tracks readily visible. Since then, both in the same laboratories and at Washington University, we have examined them in rocks from the earth's crust, in ancient man-made objects, in meteorites and in lava from the ocean floor. The study of these fossil nuclear tracks yields information about the history of the earth and the solar system and about cosmic radiation. Contemporary particles also register tracks in insulating solids such as mica, glass and plastics. The unique characteristics of solid-state particle detectors, which are now in general use in laboratories throughout the world, have already produced significant results in physics and chemistry. We have found unexpected practical applications for track-registering solids in such diverse areas as biomedical technology, environmental research, microanalysis and nuclear technology. In this article we shall first tell how tracks are formed and revealed in solids and then discuss some of the scientific and technological applications of the phenomenon.



NUCLEAR TRACKS are enlarged 125,000 diameters in an electron micrograph, made by the authors, of a very thin sheet of mica that was irradiated with fission fragments. The tracks are visible because the crystal lattice was distorted along the lines of damage caused by the particles. The apparent length of the tracks is related to their angles of inclination.

When a massive charged particle passes through an electrically nonconducting solid, it produces a narrow trail of radiation-damaged material: the track. In the case of a heavy positive ion—a nucleus stripped of many or all of its orbiting electrons—passing through a crystalline solid, we believe the damage is caused by an "ion explosion." The positively charged particle knocks electrons out of atoms in the crystal lattice, leaving a wake of positive ions. These ions thereupon repel one another violently, disturbing the regular lattice. The process is slightly different in the case of a noncrystalline polymer such as a plastic [*see illustrations on page 32*]. We have found that virtually all insulating solids register the tracks of heavily damaging particles and under normal conditions retain them indefinitely. On the other hand, intensive search has so far not revealed any metals or good semiconductors that record such continuous tracks. We believe the reason is that electrons move so fast in conductors that they neutralize the atoms of an incipient track before it can form.

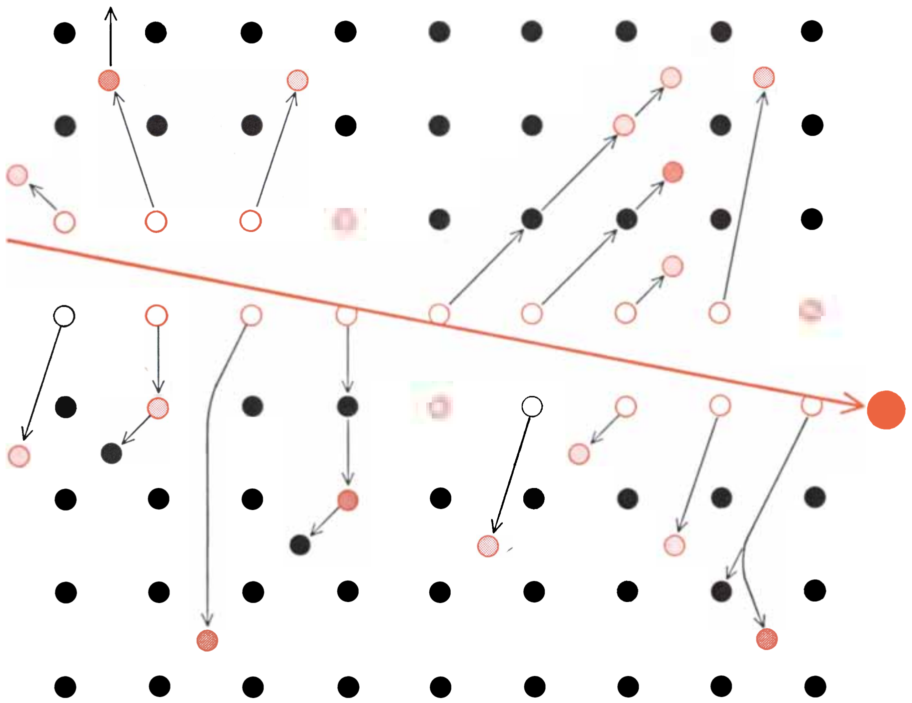
Tracks in solids were first described in 1959 by E. C. H. Silk and R. S. Barnes of the Atomic Energy Research Establishment at Harwell in England. They irradiated thin sheets of mica with fission fragments and observed the tracks with the electron microscope. In the mica they were using the tracks faded, however. Moreover, electron microscopy is a powerful tool for studying structural detail on a near-atomic scale but not for finding out about rare events (such as the



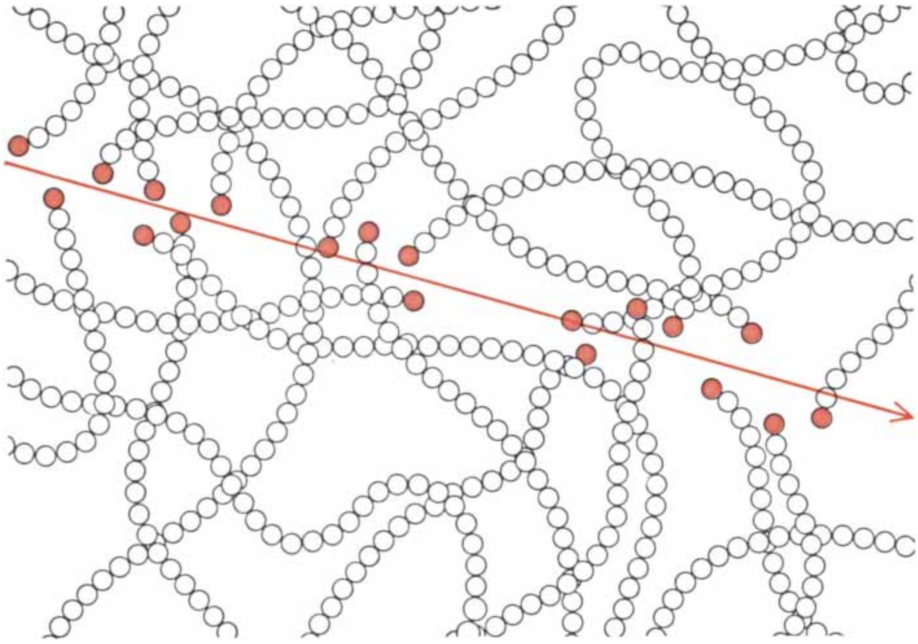
ETCHED NUCLEAR TRACKS are seen in photomicrographs made by the authors. Tracks of fission fragments in mica were etched with hydrofluoric acid for 30 minutes (*top left*). Fossil tracks from the spontaneous fission of uranium in obsidian are revealed as broad, cone-shaped etch pits after immersion in hydrofluoric acid for one minute (*top right*). Ancient tracks of heavy cosmic rays in a crystal from near the surface of a meteorite were etched for 50 minutes in concentrated sodium hydroxide; the longer tracks were left by nuclei heavier than iron (*bottom left*). When a thorium silicate crystal was bombarded with argon ions, compound nuclei were formed that decayed by fission; etching for one minute in phosphoric acid revealed a few three-pronged ternary fissions (*bottom right*). Enlargements are 700 diameters in the upper micrographs, 2,000 at bottom left and 1,400 at bottom right.

passage of specified nuclear particles) on a grand scale. In 1961 we began where Silk and Barnes had left off, attempting to stabilize the tracks and then to make them more readily visible.

We first found that the damage trails in mica could in effect be "developed" and fixed by immersing the crystal in hydrofluoric acid. The reason is that the damage trails are highly reactive regions. Any trail that intersects the surface is rapidly dissolved by the acid, leaving a hollow, cylindrical tube some 50 angstroms in diameter. (One angstrom is one ten-millionth of a millimeter.) If the exposure to the acid is prolonged, we then discovered, the etching continues at a slower rate. The undamaged wall of the tube is attacked and the diameter of the tube increases into the micron (thousandth of a millimeter) range. This is comparable to the wavelength of visible light, so that the enlarged tubes scatter transmitted light and show up clearly as dark lines when the sample is viewed with an ordinary light microscope (or as bright lines in a dark-field microscope).



TRACK IS FORMED in a crystalline solid when a highly charged particle (*dark color*) knocks electrons out of atoms along its path. The atoms, now similarly charged ions (*light color*), repel one another, leaving their original positions in the lattice (*open circles*). This creates a cylindrical region of disorder and distortion, easily attacked by an etching reagent.

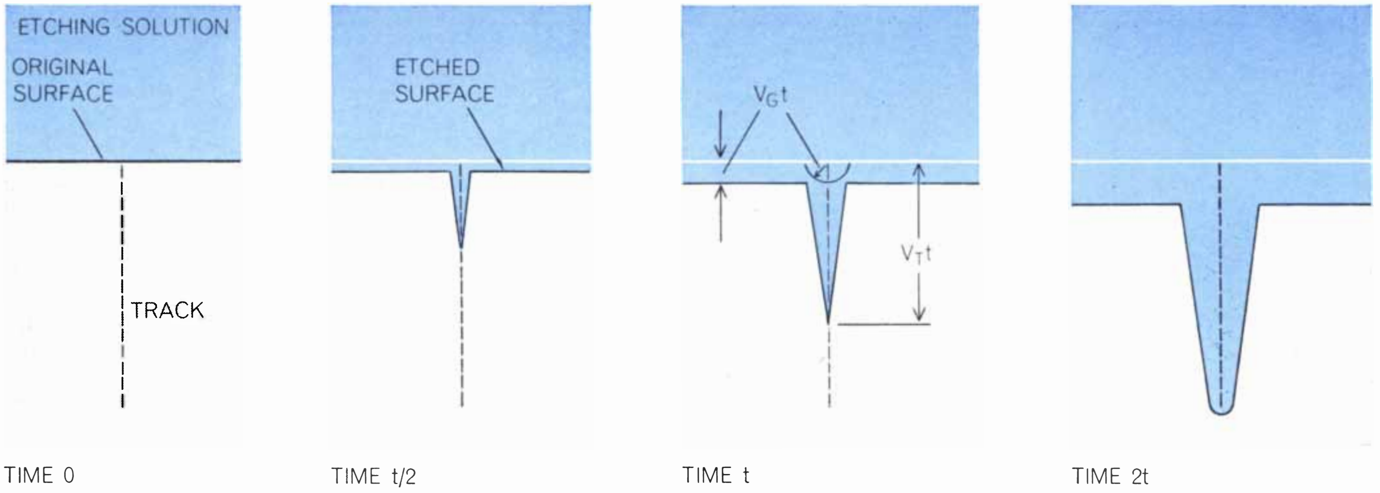


IN AN ORGANIC POLYMER the charged particle ionizes and excites molecules, breaking the chains. The chain ends form new species (*color*) that are highly reactive chemically.

The tracks take various forms, depending primarily on the solid, how it is etched and the nature of the particle [*see top illustration on opposite page*]. In some materials the radiation-damaged regions are attacked far more rapidly than the rest of the sample; the enlarged tracks that result taper very slightly and appear nearly cylindrical. In other solids the rate of attack is not much greater along the trail of damage than it is on undamaged surfaces, including the previously etched walls; the resulting tracks are conical in shape, the cone angle being governed by the ratio of the etching rates in damaged and undamaged material. Choosing the best etching reagent has been largely a matter of trial and error, although we have developed a few guidelines. Some easily attacked minerals such as mica and apatite can be etched with common acids at room temperature; more resistant minerals such as the feldspars require strong bases and high temperatures, and zircon responds to phosphoric acid heated to more than 900 degrees Fahrenheit. Many plastics, on the other hand, can be etched with such oxidizers as household bleach.

The density of the ionization damage caused by any moving particle is directly proportional to the square of the electric charge of the particle, which varies with its atomic number. It is also approximately inversely proportional to the square of the particle's velocity. These relations can be expressed by a family of curves for a number of different nuclei [*see bottom illustration on opposite page*] ranging from hydrogen (whose ion, the proton, has one unit of electric charge) up to the transuranium element curium (which has 96 units of charge when all its electrons have been removed). As the curves indicate, a particle does the least damage when it is traveling with nearly the speed of light; as its velocity decreases, its damage density increases. The reason is that the particle is more effective the more time it spends in the vicinity of each atom it encounters—until it is going so slowly that it again accumulates the electrons it had lost and thus becomes less heavily ionized.

Different solids have different sensitivities to nuclear particles. The most sensitive are the organic polymers, or plastics, some of which undergo detectable chemical changes when they are irradiated by slow protons and alpha

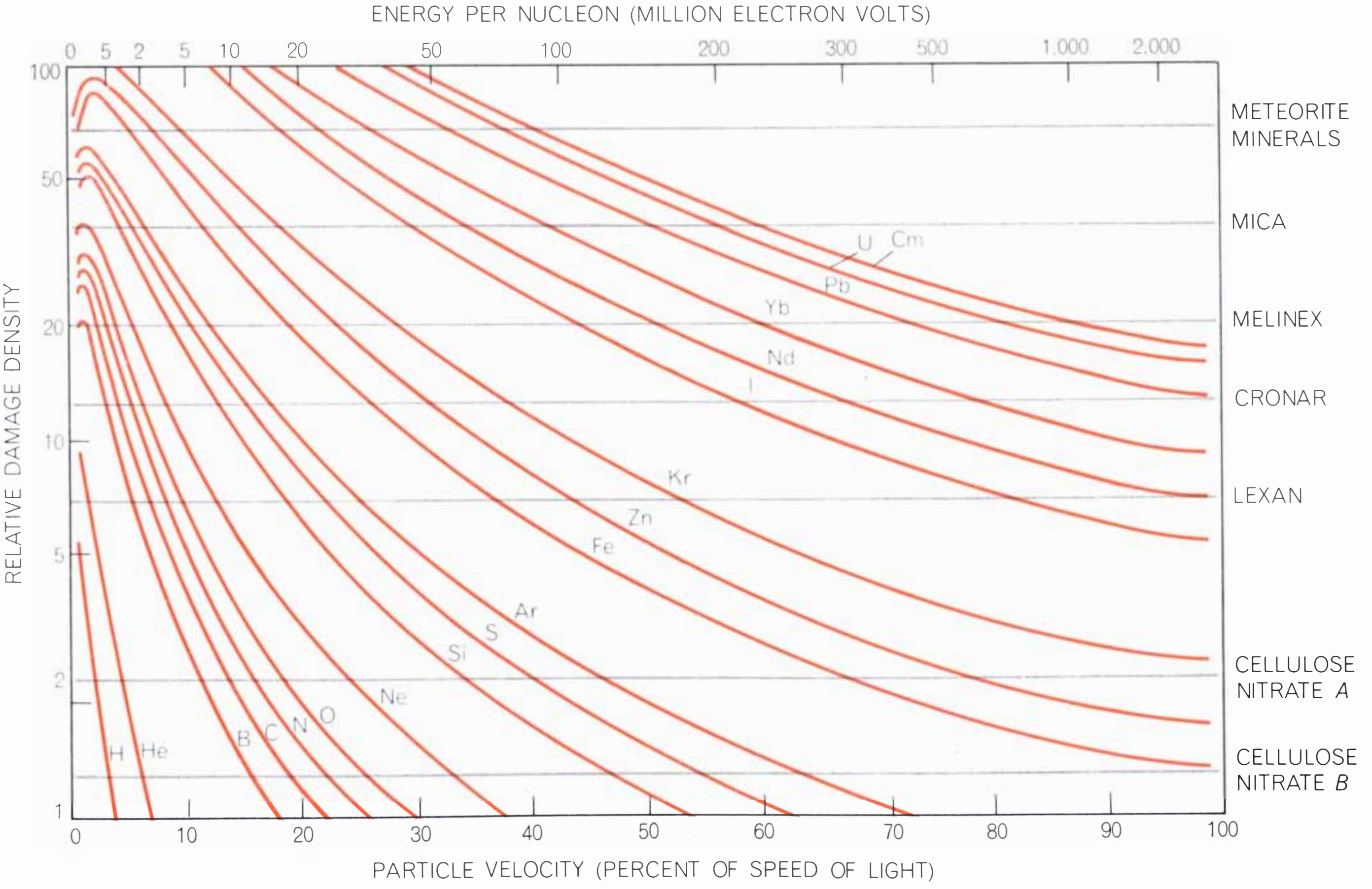


SHAPE OF ETCHED TRACK is governed by the ratio of the etching rate V_T along the track to the rate V_G on undamaged surfaces. When V_T is much the greater, the enlarged track is nearly cylindrical; when the rates are similar, the track is a shallow pit.

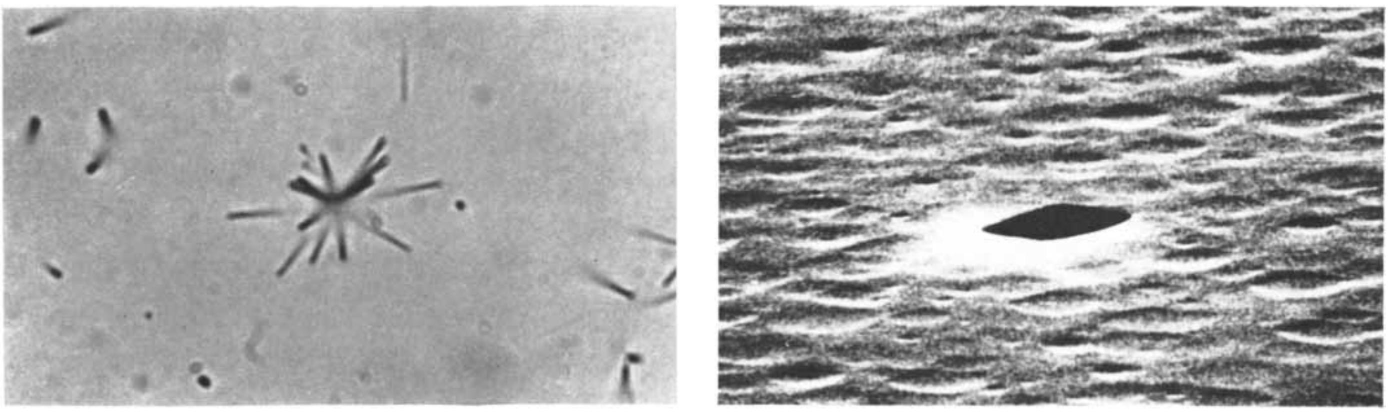
particles (helium ions). Some minerals, in contrast, are so insensitive that they only record the passage of nuclei as heavily damaging as low-energy iron ions. In experiments with heavy ions from particle accelerators and with cosmic rays we have found that each insulating solid has a well-defined threshold damage density below which no tracks are produced. This threshold characteristic, as we shall see, contributes significantly to the value of such solids as track detectors.

The starting point for our investigation of tracks in solids was the hope of finding stored nuclear records of natural radioactive processes that occur on the earth and in space. In the case of terrestrial materials, tracks result from the decay of radioactive atoms within the sample itself. In the case of meteoritic material, tracks also record the impact of cosmic rays.

Of all the decay processes the easiest to observe in solids is the spontaneous fission of uranium. From time to time a



DAMAGE DENSITY in various nonconducting solids is given as a function of velocity and of energy per nucleon for a number of nuclei. The damage density increases with increasing charge, or atomic number, from hydrogen (*H*) to curium (*Cm*); it also increases as the particle slows down (until it is going so slowly that it becomes less ionized). The gray horizontal lines represent the thresholds for track recording in materials ranging from sensitive plastics (*bottom*) to typical constituents of stony meteorites (*top*).

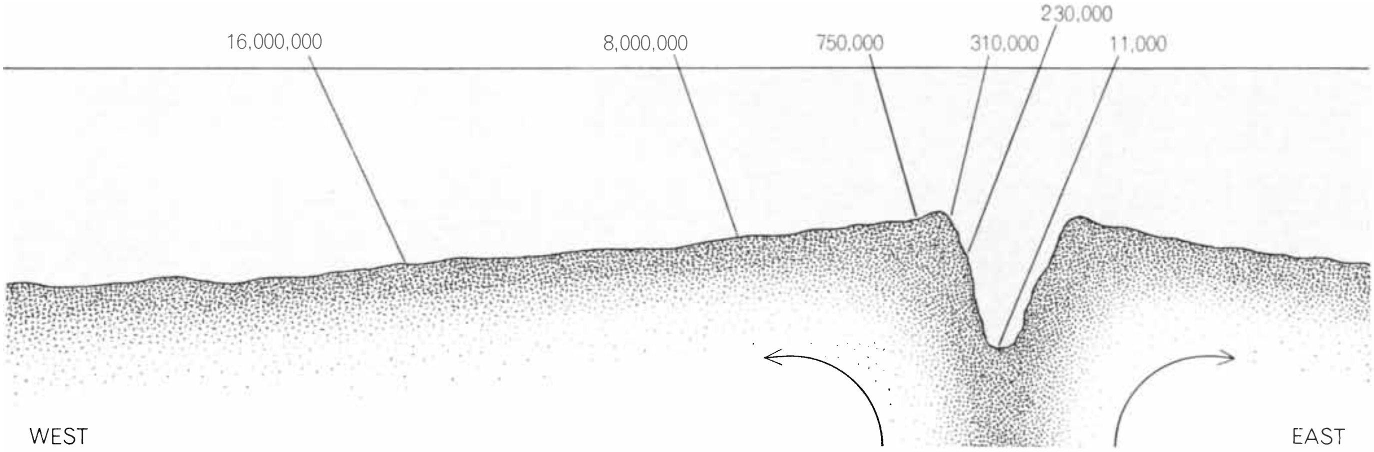


FOSSIL TRACKS of two kinds are seen in mica. Fission tracks are enlarged 1,350 diameters (*left*). Alpha-recoil tracks are enlarged 10,000 diameters (*right*). The latter, registered by nuclei recoiling as they emit an alpha particle, are more abundant than fission tracks (*black hole, center*) but so short that they are only visible under special illumination or (as here) in a scanning electron micrograph.

nucleus of uranium 238 splits into two heavy fragments that recoil in opposite directions and travel about 10 microns. Most natural solids contain uranium, in concentrations ranging from a few parts per billion to many parts per million. Within the solid the tracks of spontaneous fissions accumulate like the ticks of a clock, preserving a record of the time since the material solidified. To read the clock one need only etch and count the fossil tracks and then determine the uranium content of the sample. This is done by exposing the crystal to a dose of neutrons from a reactor that will induce fission in a known fraction (say one in a million) of the uranium atoms; a count of the new fission tracks gives the uranium concentration and then, on the basis of the known rate of spontaneous fission, one can calculate the age. (Although spontaneous fission occurs in a few atoms other than uranium 238, the rate is very low and such fissions can be ignored.)

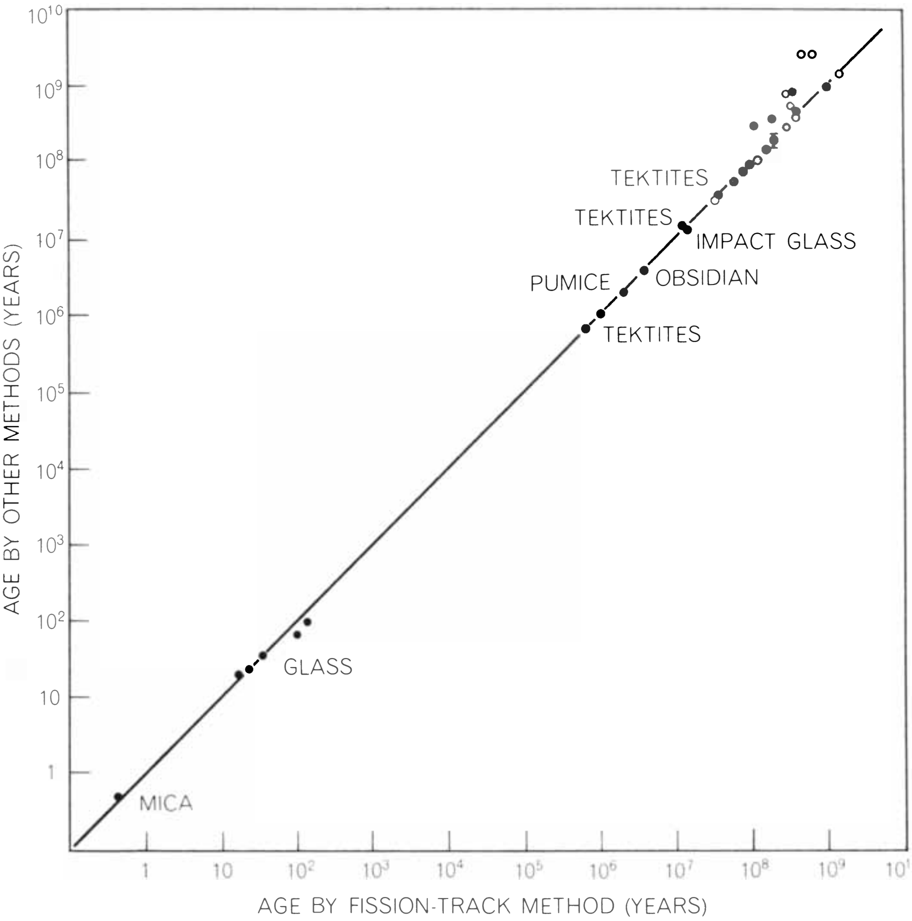
Fission-track dating is direct and visual and has proved to be applicable to many materials and over an enormous range of ages. The critical factor is the uranium content of the material to be dated. A concentration of one part per million—which is common in rocks—provides enough tracks to date an object older than some 100,000 years easily. By selecting minerals with a higher uranium content, such as zircon, apatite and certain volcanic glasses, we can date much younger samples, and some man-made glass objects from as late as the 19th century can be dated if they are unusually rich in uranium. Applying the fission-track method to human prehistory, we undertook to date a pumice deposit with which early hominid remains were associated at Olduvai Gorge in Tanzania. L. S. B. Leakey, G. H. Curtis and J. F. Evernden had established an age of 1.8 million years for another rock in the same bed, using a method based on the relative content of argon and radioactive potassium. Our result—obtained by a method subject to different kinds of error—was 2.03 million ± 280,000 years, a reasonably good agreement.

Some of our most difficult and most useful age determinations were made with samples of basaltic glass dredged up from the vicinity of the Mid-Atlantic Ridge by Fabrizio Aumento of the Geological Survey of Canada. The ridge is believed to be the line along which deep convection currents are continuously pushing molten material up to the ocean floor, causing it to spread symmetrically toward Europe and North America [see "Sea-Floor Spreading," by J. R. Heirtzler; SCIENTIFIC AMERICAN, December, 1968]. The spreading hypothesis was confirmed on the basis of geomagnetic reversals recorded in strips on the ocean floor and by dating samples at some distance from the ridge, but until the fission-track method was developed there was no way to date the most recently solidified material near the axis of the ridge. With J. R. M. Viertl, we were able to date such samples. They were remarkably young, ranging from only about



SEA-FLOOR SPREADING is confirmed by the steady increase in the age of samples dredged up at increasing distances from the rift valley at the axis of the Mid-Atlantic Ridge. Fission-track dating established the ages (11,000 and 230,000 years) of two samples near the valley too young to be dated by other methods. The next three samples, too altered to be track-dated, were dated by conventional methods. The oldest sample (16 million years) was protected by a ferromanganese crust and was dated on the basis of fission tracks.

11,000 years in the central "rift valley" to 230,000 years farther out. Beyond that most of the glassy ocean-floor rock has become too altered chemically for fission tracks to be preserved, but Aumento found one older sample that was sufficiently preserved to date by the fission-track method at about 16 million years [*see bottom illustration on opposite page*]. The regular sequence of ages increasing with distance makes it possible to establish a spreading rate of about an inch per year.



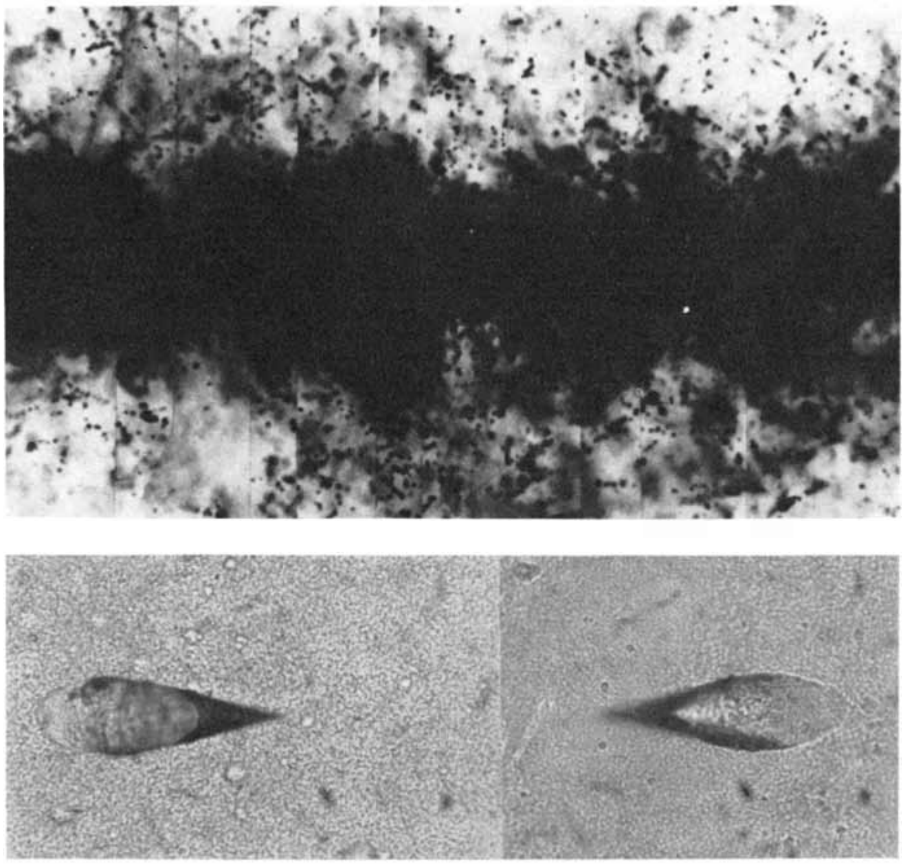
FISSION-TRACK DATING gives ages that compare well, over a wide time span, with known ages or ages determined by other methods. A mica sample (*bottom left*) was held against uranium foil for six months; the age of man-made glass was known. Fossil-track samples include mica (*open circles*) and such minerals as zircon, apatite, hornblende, allanite, epidote and sphene (*colored dots*). Objects dated in the 1,000-to-100,000-year range by the fission-track method were not graphed because they cannot be dated by other methods.

Recently a whole new vista for nuclear-track dating has opened up with the discovery that fossil tracks are formed by a second radioactive-decay process, one that leaves much more plentiful records than the rare spontaneous-fission events. For every uranium-238 nucleus that undergoes spontaneous fission there are some two million nuclei that decay by emitting an alpha particle. (Other nuclei, including those of thorium and certain rare-earth elements, also undergo alpha decay.) The alpha particles are not heavy enough to register tracks directly, but with each alpha emission the emitting nucleus recoils a little. These recoil nuclei are much heavier and leave permanent tracks. (The nuclei are electrically almost neutral, and we believe they produce damage not through ionization but by displacing atoms bodily.) The tracks are only about a two-thousandth as long as those of spontaneous fission and can be observed only with the phase-contrast light microscope or the electron microscope. Now that we know they are present, however, we are finding them in many samples. Because of their abundance they hold great promise for dating very young objects and also for providing higher age resolution.

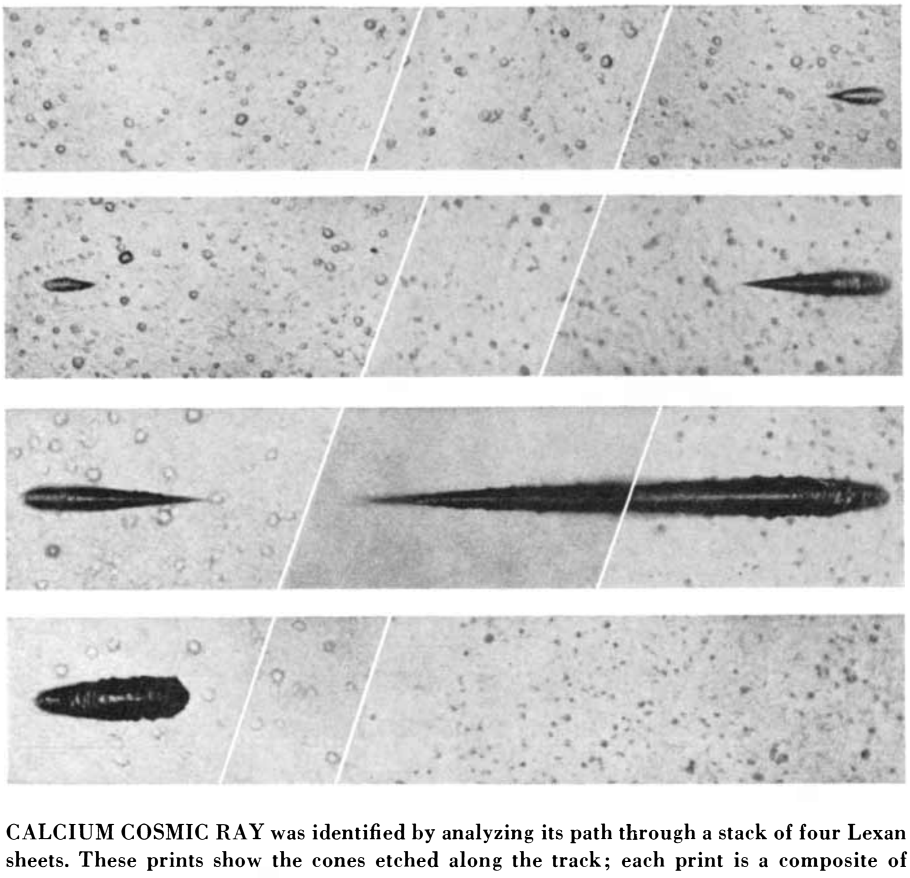
Meteorites are prized as investigative tools because they contain the oldest solid matter yet found in the solar system and because they have functioned as long-term probes of the interplanetary medium. Fossil tracks have yielded new information about both the composition of meteorites and the particles that have impinged on them. Consider first the material more than a foot below the surface of a large meteorite, where few cosmic ray nuclei can have penetrated; any particle tracks must date from the time when, as part of a much larger body, the matter first cooled to a temperature low enough for tracks to be permanently recorded.

When we examined crystals from deep within a number of large iron meteorites, we found far more tracks than could be accounted for by spontaneous uranium fission, even assuming an age of 4.5 billion years, the age of the solar system. Detailed calculations indicate that most of the tracks must be from the spontaneous fission of atoms of plutonium 244, an isotope of an extinct element that, because of its short half-life of 82 million years, has decayed away during the lifetime of the solar system; man knows of it only because it has been manufactured in nuclear reactors and accelerators. The spontaneous-fission rate is 400,000 times higher for plutonium 244 than for uranium 238, and so the plutonium leaves a clear record of its relatively short life. Its fission tracks serve as an "interval clock"; it measures the interval between the time when the solar system was formed from interstellar matter and a later time when the meteorites had cooled enough to retain tracks. The elapsed time turns out to be a few hundred million years. Studies of plutonium-244 fission tracks should be useful not only in establishing the chronology of meteorite formation and cooling but also in confirming current theories on the nuclear synthesis of very heavy elements.

For many millions of years regions near the surface of meteorites (and of other bodies such as the moon) have been bombarded by a heavy flux of the high-energy bare atomic nuclei we call cosmic rays, which are prevented by the atmosphere from reaching the earth. One of our original goals was to detect cosmic ray tracks in meteorites, and we first achieved that goal in 1964 with Michel Maurette of the University of Paris and Paul Pellas of the Paris Museum. In almost every meteorite we have examined we have found the tracks of true galactic cosmic rays—particles that originated beyond the solar system. Their number increases with the proximity of the sample to the outside of the meteorite and with the duration of the meteorite's exposure as an unprotected small body. (Particles emitted sporadically in solar flares, which might produce many more tracks of a similar nature, are lower in energy and would penetrate no more



COSMIC RAY TRACK of a superheavy nucleus (atomic number about 82) was recorded in a photographic emulsion (*top*) and in a sheet of Lexan polycarbonate plastic (*bottom*) lifted to about 130,000 feet by balloon. In the plastic what is seen are the cones etched inward along the track from the surface points at which the particle entered and left the sheet.



CALCIUM COSMIC RAY was identified by analyzing its path through a stack of four Lexan sheets. These prints show the cones etched along the track; each print is a composite of three exposures at different depths in the clear plastic. For interpretation see opposite page.

than about half an inch in a meteorite; that much is lost by ablation as the meteorite heats up in passing through the earth's atmosphere. We hope soon to see solar-particle tracks in material brought back from the crust of the moon.)

Cosmic rays consist of protons and a wide range of heavier particles, but meteorite crystals are quite insensitive and record only the tracks of extremely heavy nuclei. Analysis of the reduction in track density with increasing depth demonstrates that most of the cosmic rays we see in meteorites are due to the nuclei of iron, a fairly abundant element in the cosmic radiation; calibration studies with iron nuclei in accelerators confirm this conclusion. The tracks indicate that the rate of arrival and the energy of the iron nuclei have remained essentially constant for the past 20 million years. We can also calculate the rate at which the meteorites have been ground away by collisions with interplanetary dust: less than a millimeter every million years. Satellite data on impinging dust particles confirm this low rate.

Perhaps the most significant aspect of our cosmic ray studies in meteorites has been what one of our colleagues refers to as "breaking the iron curtain." Through studies with high-altitude balloons investigators had established the relative abundance in cosmic rays of the various elements—up to iron. The abundance of elements beyond iron was a mystery because such extremely heavy cosmic rays are exceedingly scarce. The damage-density curves [*see bottom illustration on page 33*] show that an iron nucleus will produce tracks in a meteorite mineral only when its velocity drops to about 6 percent of the speed of light, or in the last few microns of its trajectory in the solid. It is also clear that heavier nuclei will record tracks over a greater range. The problem of detecting extremely heavy nuclei therefore comes down to finding long tracks against a background of short iron tracks. With Maurette and then with D. Lal, R. S. Rajan and A. S. Tamhane of the Tata Institute of Fundamental Research in Bombay, we found that elements heavier than iron are indeed present in cosmic radiation; taken together, they are about a five-thousandth as abundant as the iron nuclei. The effort is now to establish the relation of track length to atomic number, so that the composition of the extremely heavy cosmic rays can be determined with high resolution.

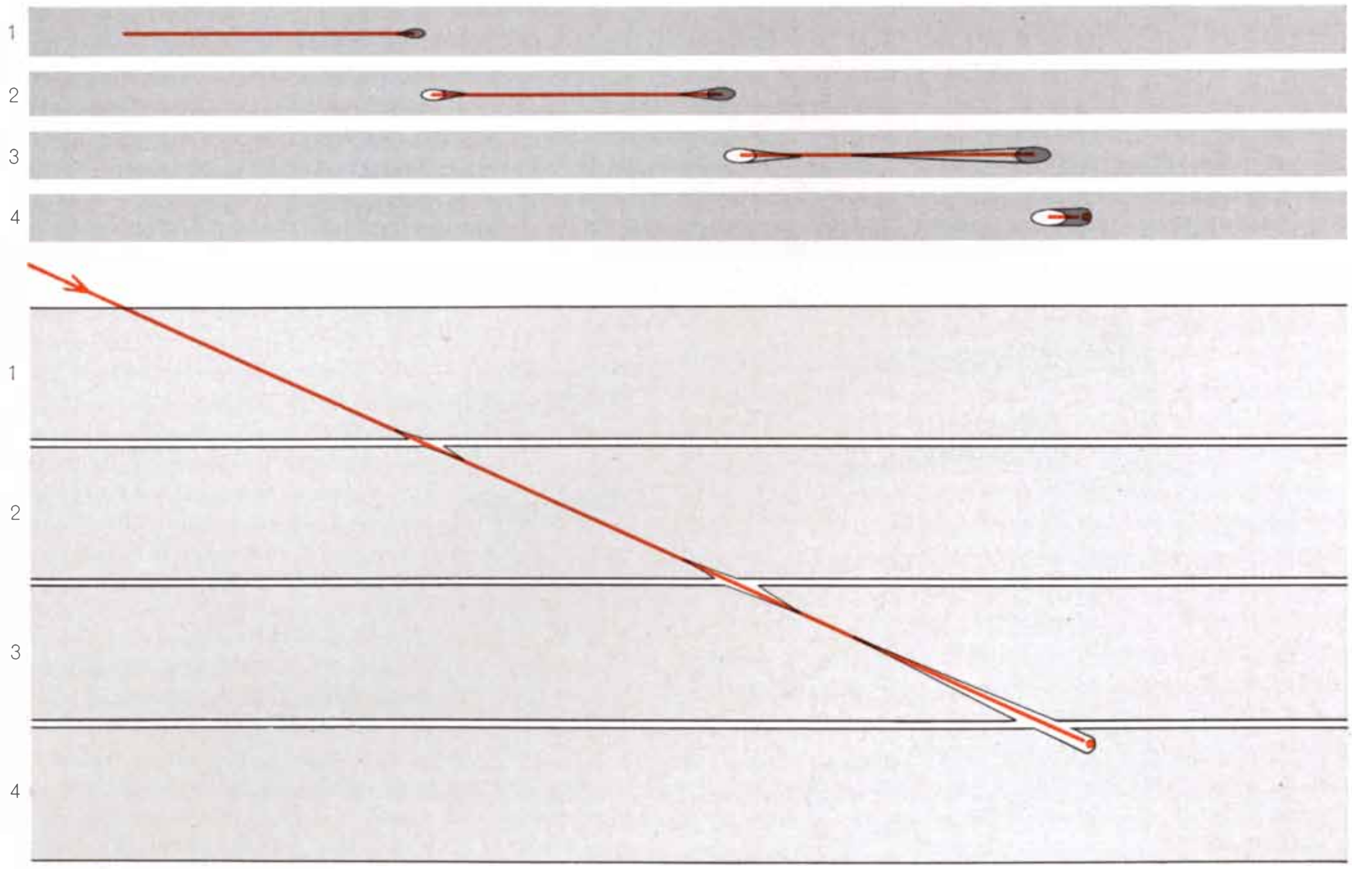
Having discovered the fossil tracks of extremely heavy nuclei in meteorites, we were anxious to search, with plastic detectors and nuclear emulsions, for these rare particles in present-day cosmic radiation. Before we could get our experiment off the ground (literally, since these are high-altitude balloon experiments), Peter H. Fowler of the University of Bristol succeeded in finding the first contemporary extremely heavy cosmic ray nuclei. He and his colleagues sent large packages of nuclear photographic emulsion to high altitudes. The tracks they recorded established the existence of cosmic rays as heavy as lead and uranium and suggested an exciting possibility: the detection of a transuranium nucleus.

This caused quite a stir among astrophysicists and nuclear physicists alike. First of all, no one is sure whether the cosmic rays that reach the solar system were born in our galaxy, and therefore have an average lifetime measured in millions of years, or are part of an all-pervasive universal radiation with a far greater lifetime and an awesome total energy. The existence of transuranium nuclei, with their short half-life, implies comparatively nearby (galactic) sources where heavy nuclei are built up by rapid neutron capture. Quantitative studies of the transuranium radiation's composition could provide the key to the age and origin of cosmic rays.

To nuclear theorists, moreover, the presence in cosmic rays of transuranium nuclei would offer a chance to test the "shell" model of the nucleus. The model predicts that the trend toward decreasing half-life with increasing atomic number beyond uranium should be reversed in the vicinity of atomic number 114 [see "The Synthetic Elements: IV," by Glenn T. Seaborg and Justin L. Bloom; SCIENTIFIC AMERICAN, April]; there should therefore be stable "superheavy" elements. Estimates of the half-life of the most stable superheavy nucleus range up to as much as 100 million years. That is too short for it to have survived in the earth's crust, but it may be long enough for such nuclei to be detectable in cosmic rays.

Fowler's experiments showed that the flux of superheavy cosmic rays at the top of the earth's atmosphere is no more than about one per 10 square meters of detector per day. The flux of iron nuclei is a million times as high and that of all cosmic rays is more than a billion times as high, so that it is not easy to find a superheavy nucleus. Plastic detectors lend themselves to a simple method we worked out for locating heavy-particle tracks. From a stack of plastic and photographic-emulsion detectors that have been exposed to cosmic radiation we remove certain plastic sheets and etch them until the cones etched inward along tracks meet and form holes through the plastic. When the sheets are fed through an Ozalid copying machine, ammonia diffuses through the holes and makes spots on the copying paper; we can then locate the tracks of the same particles in other sheets of plastic from the stack and in the emulsion for detailed study.

With George E. Blanford, M. W. Friedlander, Joseph Klarmann, J. P. Wefel and William C. Wells of Washington University, we have exposed five sets of such stacks, each about 15 square meters in total area, on balloon flights to an altitude of about 130,000 feet. (The last three flights were carried out in collaboration with Fowler's group.) So far we have studied the tracks of some 40



PARTICLE IDENTIFICATION is based on the fact that a cosmic ray nucleus slows down in the plastic sheets, and the consequent rate of change of the etched cone length with distance is a unique function of the atomic number and the mass of the particle. Here the path (*color*) of the calcium nucleus (*see bottom illustration on opposite page*) is diagrammed as seen from above (*top*) and from the side (*bottom*). In sheet *1* the entering ion's damage rate is below the threshold for track recording in Lexan. The rate increases as the particle's velocity decreases and a short cone is etched at the bottom of the sheet. In sheets *2* and *3* the cones get longer and the particle comes to rest in sheet *4*. Measurements of cone length establish a damage-density curve that identifies the particle.

nuclei with an atomic number above 60, of which half a dozen may be heavier than uranium. Our aim is precise identification of these superheavy particles, and for this purpose we have recently developed a new technique in collaboration with D. D. Peterson of the Rensselaer Polytechnic Institute and C. O'Ceallaigh, D. O'Sullivan and A. Thompson of the Dublin Institute for Advanced Studies. The method (which has a unique property: its resolution improves with increasing atomic number) is based on the fact that in plastic the etching rate along a track increases exponentially, rather than linearly, with damage density. Measurements of cone lengths at several points along a particle's trajectory enable us to trace out its damage-density curve with high precision and thus determine the particle's charge and mass. At this stage we think we can resolve individual atomic numbers and we have some hope of eventually resolving isotopes with the same atomic number.

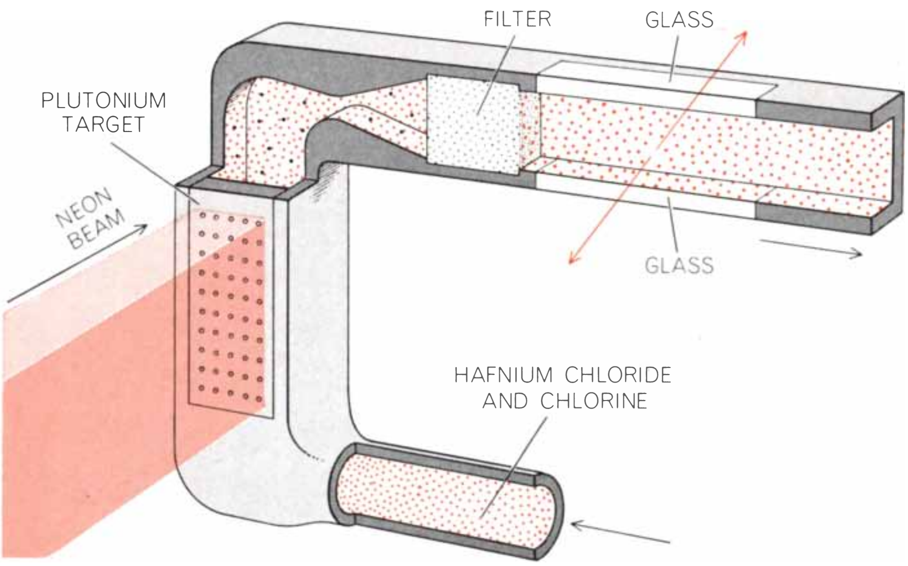
Since solid-state detectors are adapted to a search for rare, highly ionizing particles, it occurred to us to look for magnetic monopoles: hypothetical elementary particles of magnetism whose existence has been postulated on the basis of symmetry principles but that have never been detected [see "Magnetic Monopoles," by Kenneth W. Ford; SCIENTIFIC AMERICAN, December, 1963]. Because a moving magnetic monopole should have a strong electric field and consequently a high ionizing-damage rate, solid-state detectors should record their passage. In his dredging for ocean-floor samples to be dated, Aumento had brought up some ferromanganese material that because of its long period of formation and ability to trap magnetic particles seemed a likely place to hunt for magnetic monopoles. With H. R. Hart, Jr., W. M. Schwarz and I. S. Jacobs, we put about 18 pounds of the material through a high-field superconducting magnet with plastic track detectors at each end in an effort to extract, accelerate and then detect any trapped monopoles. None were found. As in the case of all previous monopole experiments, this does not quite prove that the particles do not exist. It does suggest, however, that the maximum number of magnetic monopoles reaching the earth since it was formed has been no more than .5 per square centimeter!

We realized early in our investigations that the threshold property of solid-state detectors would allow us to see the tracks of heavy particles emitted in nuclear reactions without being distracted by a dense background of light-particle tracks. In collaboration with workers at the Lawrence Radiation Laboratory of the University of California at Berkeley we undertook several experiments, one of them a search for ternary fission. W. J. Swiatecki of Berkeley had predicted that whereas uranium fissions into two fragments, superheavy nuclei should tend to fly apart into three fragments under the same circumstances. We bombarded a crystal of thorium silicate with the heaviest nuclei then available in an accelerator: argon ions with energies of 400 million electron volts. About one in a million of the argon ions collided and fused with a thorium nucleus, forming a highly unstable compound nucleus that usually fissioned immediately—and about one in 30 times we saw three fragment tracks instead of two. The key to the success of this experiment was the high threshold of the thorium silicate. The bombarding ions produced no tracks; tracks were formed only by heavier particles emitted in nuclear reactions, including the ones we were looking for.

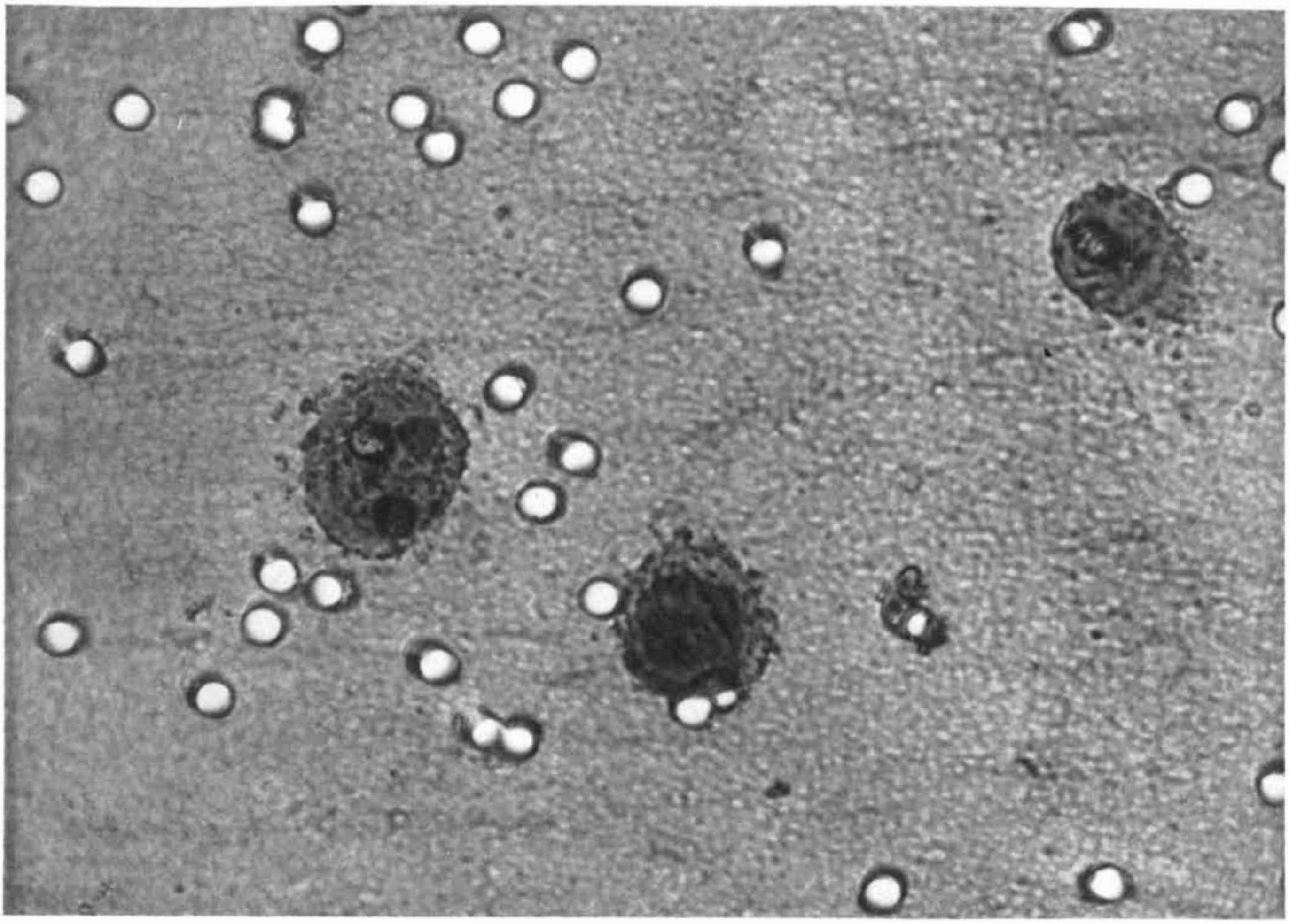
At the Joint Institute for Nuclear Research at Dubna, Russian physicists exploited the threshold property to detect one or two fissioning transuranium atoms a day in a cyclotron seething with radioactivity. The unstable elements beyond uranium had been synthesized as far as atomic number 103 (lawrencium); the Dubna group was trying to synthesize element 104. They bombarded a plutonium target with high-energy neon atoms and deduced, with glass detectors, that they were creating a few atoms a day of element 104, which appeared to have a half-life of .3 second and to decay by spontaneous fission. They called the element kurchatovium, after the late Russian physicist Igor V. Kurchatov.

To establish that the new element was indeed number 104 (an identification as yet unconfirmed in other laboratories) the Dubna workers devised a chemical test. It was based on the fact that element 104 falls below element 72, hafnium, in the periodic table of the elements and should therefore have similar chemical properties. They built a fast chemical separation system right into their cyclotron and proceeded to show that the chloride of their element had a boiling point near 300 degrees Celsius—comparable to that of hafnium chloride and much lower than chlorides of elements 93 through 103, the synthetic elements of the "actinide" series.

In the Dubna workers' system atoms of any kurchatovium (or actinides) that were produced in the neon-plutonium reaction would enter a stream of hafnium chloride and chlorine, be converted into chlorides and be swept into a heated filter within .1 second; any fissioning atoms that passed through the filter would register tracks in glass detectors [*see illustration at left*]. Control runs showed that the high-boiling-point actinide chlorides were stopped by the



KURCHATOVIUM, element number 104, was identified by Russian physicists with the aid of a chemical separation system built into their cyclotron. Atoms of a new element formed by the fusion of neon and plutonium atoms were chlorinated and carried through a filter along with hafnium chloride ions (*colored dots*). Passage through the filter showed that the new element was number 104, the chloride of which should have a boiling point like that of hafnium chloride; any actinide chlorides (*black dots*), with higher boiling points, would have been stopped. The presence of element 104 was established by tracks in glass detectors.



CANCER CELLS (*large gray structures*) **were filtered out of a blood sample by a Nuclepore filter that permitted blood cells to pass through. The filter, which has uniform cylindrical holes five microns (.005 millimeter) in diameter, is made by exposing a thin polycarbonate sheet to fission fragments and etching the tracks. The filter can be manufactured with various hole sizes and densities.**

filter. When the filter was kept at 250 degrees C., any chlorides made in the neon-plutonium reaction were also stopped. When the filter was heated to 350 degrees, however, the hafnium chloride could pass through—and so, apparently, did kurchatovium chloride: the detectors recorded about two spontaneous fissions a day, the rate predicted for the volatile kurchatovium chloride! The enormous amount of background radiation in the cyclotron during the required operating time of some 600 hours made glass or some other solid-state detector indispensable in this experiment.

New fundamental knowledge often has unexpected technological consequences. One practical device based on nuclear-track recording is the Nuclepore filter we developed at the General Electric Research and Development Center when we realized that the track-etching process could be regulated to make sieves with almost any desired density of precisely sized holes, meeting the increasing demand for filters that can discriminate between objects that are very similar in size. Plastic sheeting is exposed to collimated fission fragments in a small reactor, producing nearly parallel tracks through the plastic; the tracks are then etched to produce holes ranging in diameter from a few atom spacings up to many microns. In our laboratory C. P. Bean and R. W. DeBlois have made a device that sizes and counts viruses or other minute objects. A plastic membrane with just one hole a few tenths of a micron in diameter separates two electrically conducting solutions in which the particles to be sized are suspended. As a virus or other object passes through the hole, the electrical conductivity across the membrane suddenly drops—by an amount proportional to the decrease in the hole's volume resulting from the presence of the particle.

Track-etch plastics are being used by H. W. Alter and his colleagues at General Electric as image-recorders in neutron radiography and alpha-particle autoradiography and for bulk-tracing, a technique whereby the movement of water through a river system or of discharged gases through the atmosphere can be determined. A known concentration of tracer particles is introduced into the fluid system and later the particle concentration is measured at various points. The tracer can include an element that is not itself radioactive but that fissions or emits alpha particles on exposure to neutrons. Each sample collected is placed in contact with a plastic track-detector and exposed to neutron irradiation, and the resulting track density is a measure of the particle concentration at each site.

Bulk-tracing was one of the subjects discussed last month at the first international conference, in France, on track registration in solids. So was the information one can expect to derive from fossil-track analysis of the first samples from the surface of the moon, and so also were sea-floor spreading and trans-uranium cosmic rays. That is a wide spectrum for a scientific conference, but it is typical of the early stages of a new scientific development, when those who apply a new tool to widely differing tasks in very different disciplines still have much to say to one another.

WOUND HEALING

In man and other higher animals an intricate, three-step process involving a variety of cells renews injured tissue. Studies of this process could yield ways of controlling several diseases

by Russell Ross

When a salamander loses a leg, it can grow a new one, but man has not retained this ability to duplicate injured tissue. The liver and the surface layer of the skin are among the few mammalian tissues that can regenerate themselves; otherwise man must rely on wound healing. This is an intricate physiological process in which several different kinds of cells appear at successive intervals in order to absorb foreign matter, destroy bacteria and repair the injury.

Clearly wound repair is essential to health and comfort. It also represents a biological adaptation without which complex multicellular organisms could neither survive nor evolve. Because of its medical importance the general character of the process has been known for years. Recently, however, the advance of biochemistry and the development of the electron microscope and other research tools have enabled investigators to observe and understand in greater detail the molecular events that manifest themselves on the macroscopic level as inflammation, scar formation, restoration of the skin's surface and the other stages of wound repair. Furthermore, because the study of wound repair involves examining some of the most basic features of living tissue, progress in this field of inquiry could yield information that might lead to new understanding of such illnesses as cancer, rheumatoid arthritis and rheumatic heart disease.

The process of wound repair differs little from one kind of tissue to another and is generally independent of the form of injury. Wound repair can therefore be somewhat arbitrarily divided into three overlapping and probably related stages, each characterized by the activities of particular populations of cells.

These stages are clearly seen in a deep skin cut. Initially blood flows into the gap created by the cutting instrument, fills the space and clots, uniting the edges of the wound. Within several hours the clot loses fluid and the surface becomes dehydrated, so that it forms the hard scab that serves to protect the wound. Inflammation begins as fluids enter the wound around the clot. The injury at this stage may become swollen and painful. Then, about six hours after wounding, various kinds of white blood cells start to migrate into the wound and begin removing and breaking down cellular debris, bacteria and other foreign material. Subsequently in the dermis, or subsurface layer, the cells called fibroblasts enter the wound and build scar tissue by manufacturing collagen fibers and other proteins. Meanwhile the epidermis, or surface layer, creates a new surface similar to the old one. When this layer is almost completely formed, the scab sloughs off.

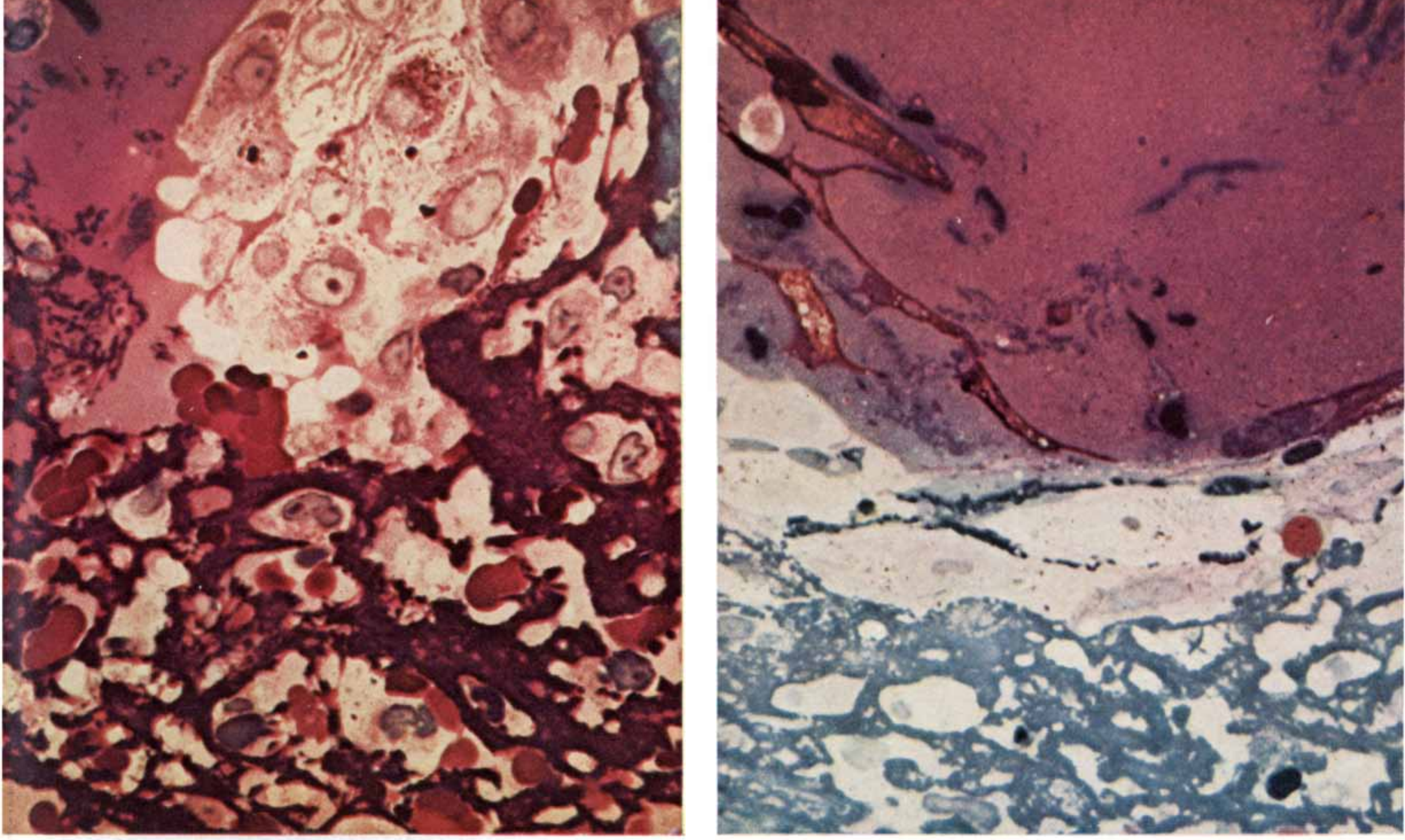
At the University of Washington School of Medicine we have been investigating the repair of both the surface and subsurface layers. Our tools are the light microscope, the electron microscope, autoradiography and substances labeled with radioactive isotopes. Medical-student volunteers provide some of the material for our study. To collect specimens we make one-centimeter incisions on the inside surface of the upper arm of each student and remove the wounded areas at successive intervals with a rotating-punch biopsy tool. This painless procedure enables us to study each stage of wound repair in detail, just as biology students observe the development of a chicken embryo by incubating eggs and opening them at different intervals.

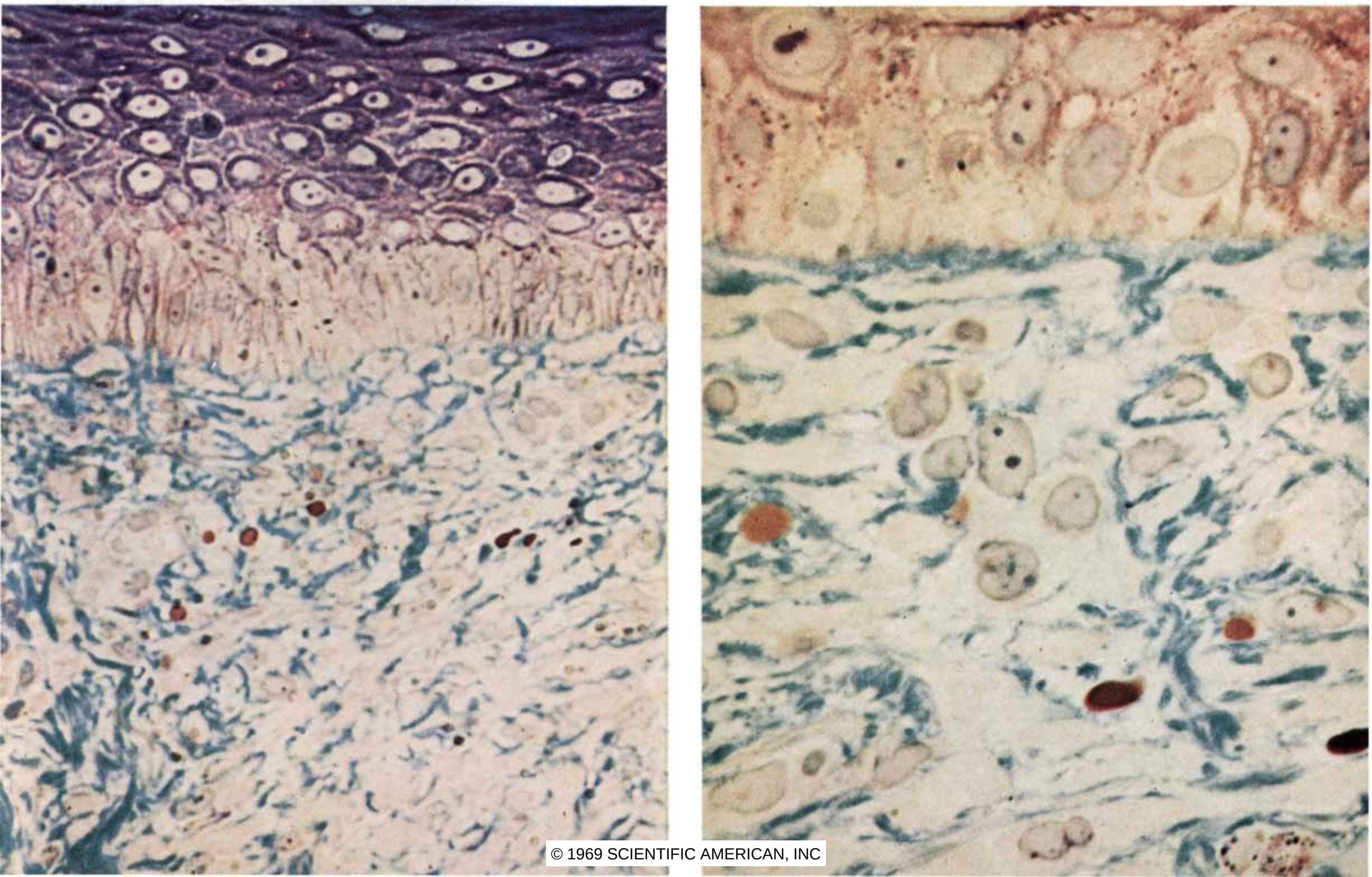
Healing in the dermal layer begins when a clot forms from the blood that flows into the wound immediately after injury. As the blood coagulates, fibrinogen molecules from the blood quickly link up into interconnected strands of fibrin. These strands provide a network throughout the wound defect that somewhat tenuously unites its edges. Meanwhile at the surface fibrin and other proteins in the blood serum dehydrate and form the scab.

After the clot has formed, the dying and broken tissue produces substances that cause blood vessels in the nearby uninjured tissue to leak, perhaps by altering the intercellular structure of the vessel walls so that the cells do not fit together as tightly as they did before. The resulting flow of serum contains a number of proteins such as globulin, albumin and antibodies. If the wound contains viruses (or is a lesion caused by viruses), the globulin and antibodies may attack them, but normally this fluid merely provides a sustaining environment for the white cells that begin to follow it into the wound about six hours later.

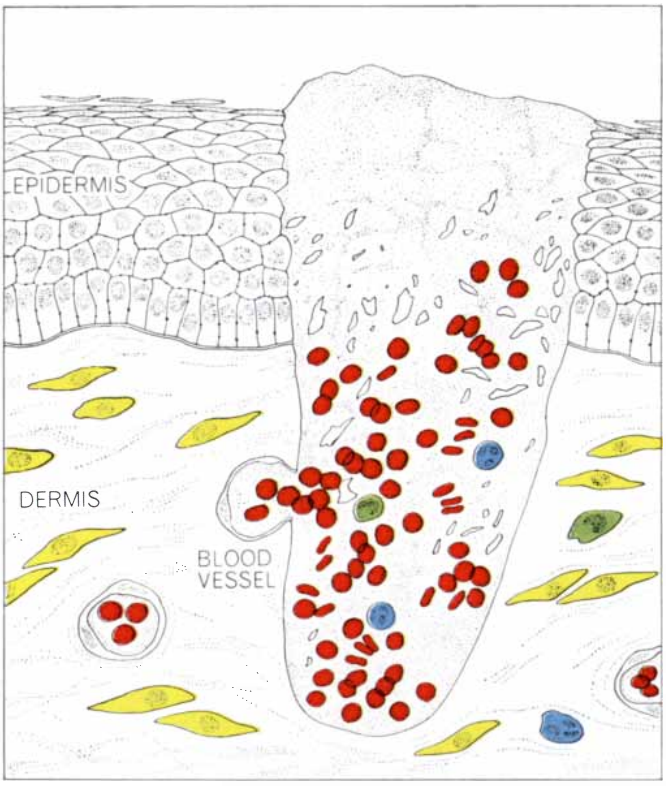
The first of these cells are the neutrophils. To get into the wound the neutrophil and the other white cells that follow it must migrate through the walls of the blood vessels in the nearby tissue. In order to get through the wall of a blood vessel a white cell inserts a bit of its cytoplasm between the cells lining the vessel and forces them apart. The blood cell can then slip through from the bloodstream into the adjacent tissue [*see illustration on page 43*].

Once inside the wound, a neutrophil can provide two kinds of defense. First, the cell can ingest organisms, by the process termed phagocytosis. In many cases the neutrophils kill the bacteria and digest most of their remains. In a

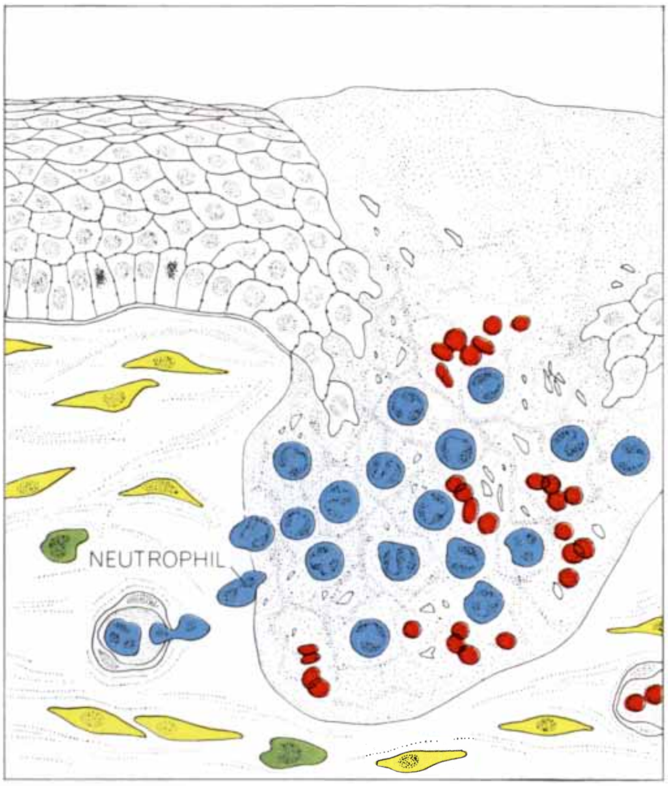




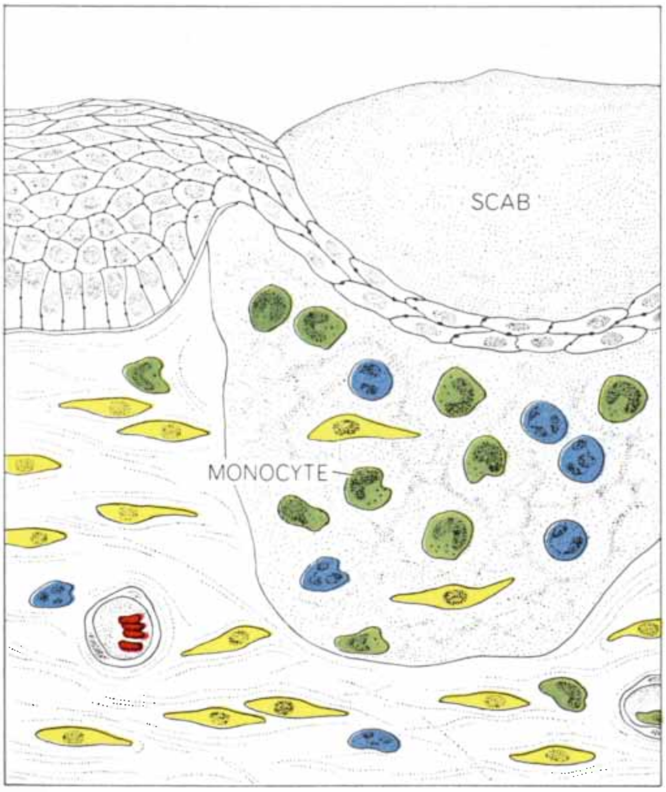
STAGES OF WOUND HEALING are shown in these photomicrographs. At top left, 24 hours after wounding, a wine red scab has already formed and purple fibrin strands unite the edges of the gap. Surface cells have intruded into the wound and begun to ingest fibrin and red and white blood cells. At top right, 24 hours later, surface cells have formed a layer under the scab. Inflammatory cells are lodged among the fibrin strands. At bottom left, 14 days after wounding, surface layers are complete and white connective-tissue cells have restored the bluish subsurface layer. At bottom right, six days later, high magnification reveals a subsurface layer of parallel collagen bundles and a few cells. Micrographs were made by James D. Huber of University of Washington.



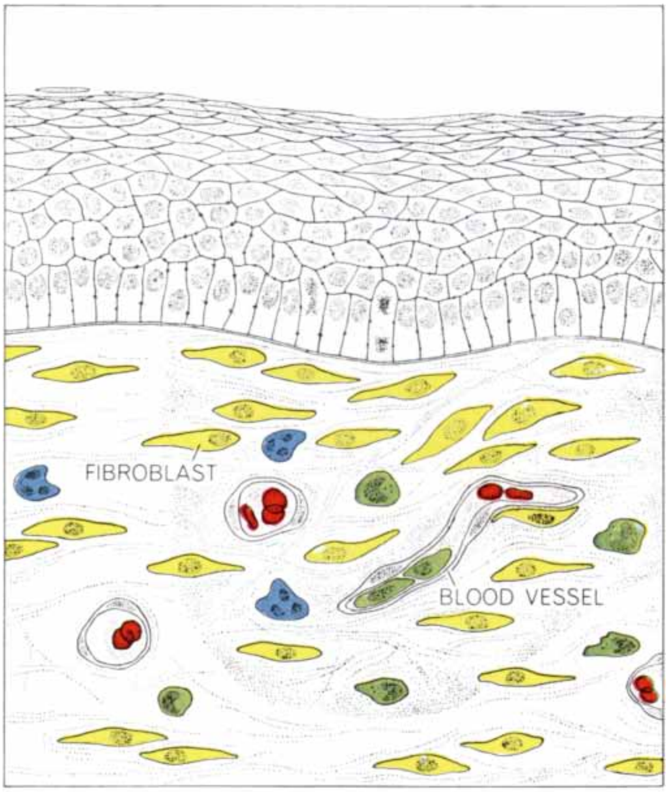
AT THE TIME OF WOUNDING gap fills with blood from broken vessels and cellular debris from the surface (epidermal) and subsurface (dermal) layers. Blood contains red cells and fibrinogen molecules, which form into strands that unite the wound's edges. Yellow cells of the dermis called fibroblasts surround the wound.



A DAY LATER bluish neutrophils enter the wound from surrounding blood vessels and begin to ingest bacteria and debris. The epidermal cells, held together by desmosomes, have also intruded. Other epidermal cells begin to reproduce (*note dividing nuclei*) so that those that migrate into the wound will be replaced.



TWO DAYS LATER wedges of epidermis have met under the scab, and the next generation of inflammatory cells, called monocytes, have entered the wound to complete the scavenging process. Meanwhile the fibroblasts that make the collagen and other connective-tissue proteins of the dermis have migrated into the defect.



SEVEN DAYS LATER the scab has sloughed from the restored epidermis. A few monocytes and neutrophils remain in the wound. At this stage the fibroblasts actively make collagen and other dermal-tissue proteins. The new connective tissue is denser than the original, unwounded dermis and is permeated by blood vessels.

sterile wound such as the one made by a surgeon a neutrophil has no bacteria to ingest. Under these circumstances the cell appears to degenerate and die. At this point the neutrophil's outer membrane ruptures, pouring enzyme-containing granules of different sizes into the wound. As the enzymes are released from the granules, they may attack the extracellular debris at the site of injury; such material can then be more easily removed by the cells that subsequently appear in the wound. Earl P. Benditt and I first observed this dissolution of neutrophils and the release of their granules when we studied healing skin wounds in guinea pigs in 1960. George F. Odland and I have since made identical observations in man. The mature neutrophil seems particularly suited for its specific task. Its cytoplasm contains only a few of the organelles found in most other cells. It has a life-span of only a few days and can neither synthesize proteins nor divide to replicate itself.

Within the first 12 hours after injury a second kind of white blood cell, the monocyte, begins to migrate into the wound. On entering the wound a monocyte becomes a macrophage: a phagocytic cell that removes most of the debris from the injured area by ingesting and then partially digesting it. Unlike the neutrophil, the monocyte has a fairly long life-span, it can synthesize proteins (particularly the enzymes it uses in phagocytosis) and it remains extremely active in the latter part of the inflammatory phase of wound repair. There is no evidence thus far that the monocyte can divide.

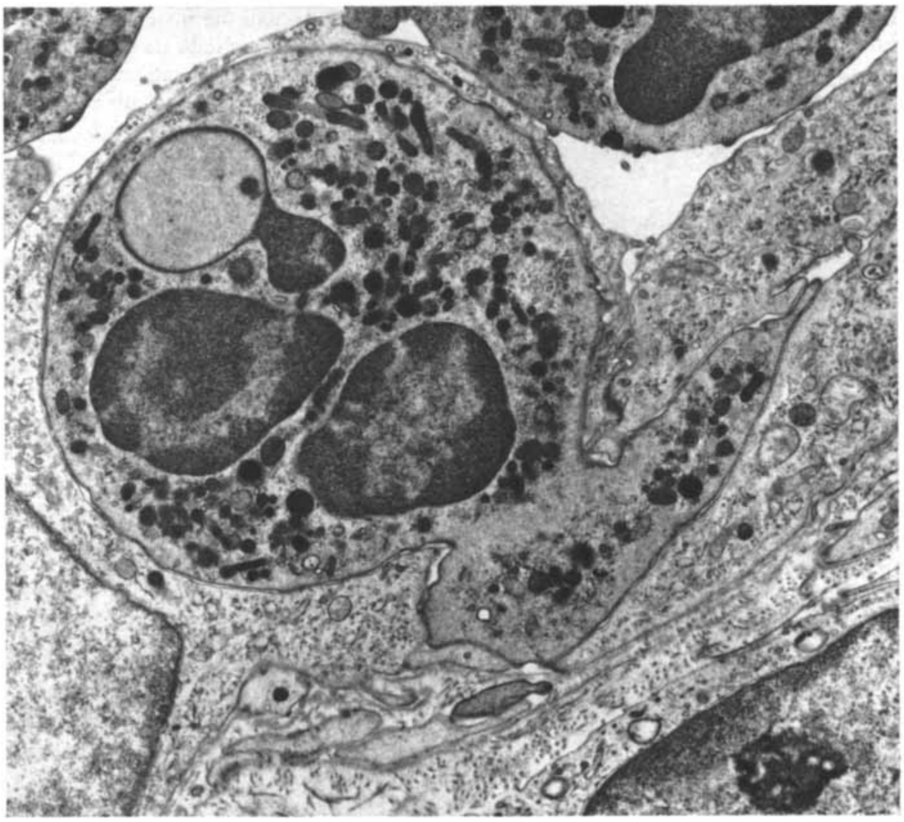
Electron-microscope studies of wounds in both guinea pigs and humans show that the monocyte goes through a series of changes before it begins to envelop and digest bacteria and debris. At first the cell contains comparatively few of the organelles responsible for the synthesis of protein within its cytoplasm. After the cell appears in the wound, however, it forms organelles. Subsequently small granules appear within the cytoplasm. This organelle system is called the rough endoplasmic reticulum, and it consists of membrane-lined channels with small, dense particles attached to the outer surface of the membrane. These particles are ribosomes; they synthesize the enzymes that degrade the debris ingested by the monocyte. It is remarkable to see how a mature monocyte becomes engorged with large vacuoles containing various forms of material, including whorls and membrane-like structures. It is believed that these structures are the remnants of the debris the cell is unable to digest completely [*see illustration at bottom right on page 45*].

Macrophages were once thought to originate in the connective and lymphoid tissues near the wound. Alvin Volkman and James L. Gowans of the University of Oxford recently demonstrated in a definitive series of experiments, however, that they stem from a small pool of rapidly dividing precursor cells in the bone marrow. Then they travel to the wound in the bloodstream.

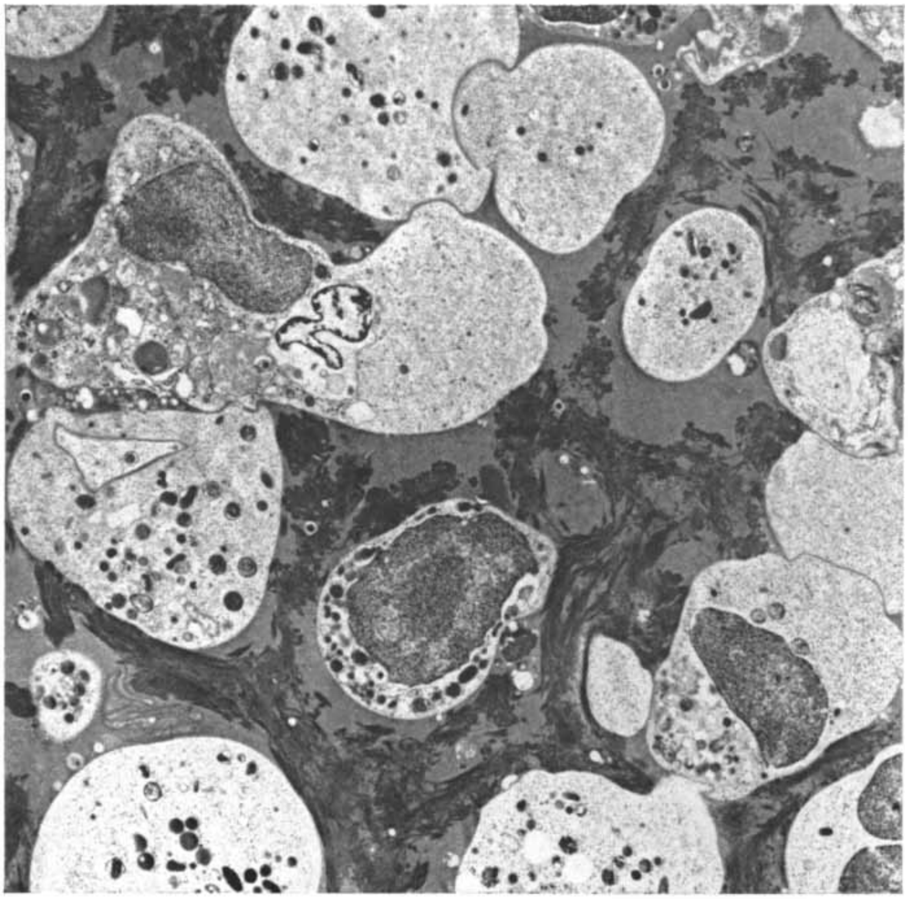
Although neutrophils and monocytes play a vital part in wound repair, there are times when they can be destructive. This is particularly true when exaggerated inflammation persists for abnormally long periods of time, as it does in some diseases such as rheumatoid arthritis. In this debilitating disease there is often marked destruction of the synovial membranes of the joints, which secrete the fluids that lubricate the joints, and in later stages of the disease the cartilage and bones of the joints may be partially destroyed and may subsequently adhere to each other. Repetitive inflammation, with its overabundance of inflammatory cells and release of their enzymes, may be partially responsible for much of the damage. It is for this reason that cortisone and other drugs that suppress inflammation have been used to treat arthritis. A better understanding of the role played by neutrophils and monocytes could lead to further improvement in the methods for controlling the disease.

Toward the end of the inflammatory response another kind of cell, the fibroblast, appears and begins to repair the injury by secreting the collagen and the protein polysaccharides that form scar tissue. Collagen endows the scar, which eventually serves to replace the wound defect, with great tensile strength. The most abundant protein in the animal kingdom, it is the major fibrous constituent of skin, tendon, ligament, cartilage, bone and the stroma (a form of connective tissue that supports most glands and organs).

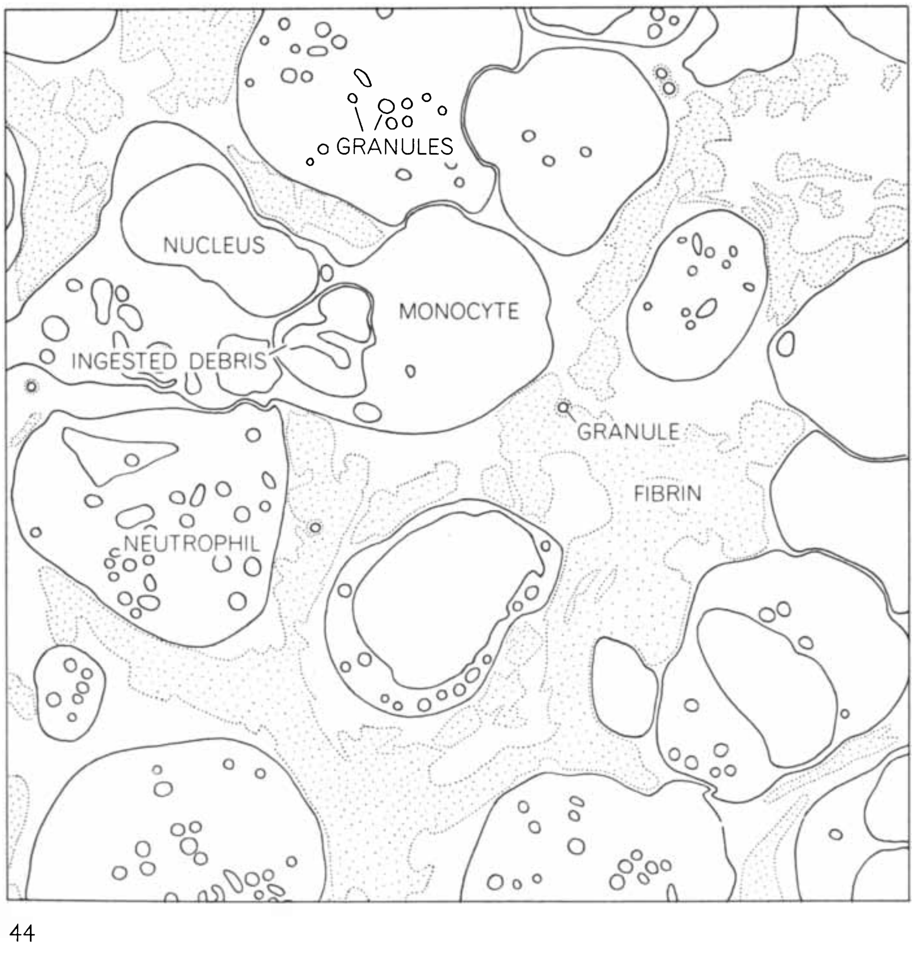
Scar tissue begins to form when the collagen molecules aggregate into fibrils that in the electron microscope exhibit a regular pattern of bands. The fibrils in turn are bound by the protein polysaccharides into fibers, which can be seen



NEUTROPHIL, enlarged by 14,000 diameters in this electron micrograph, begins its trip to the wound from inside a blood vessel embedded in connective tissue near the wound. To escape from the vessel into the connective tissue the granule-filled neutrophil squeezes a process, or extension of itself (*below center*), between the cells of the vessel wall and forces them apart. At upper left sections of two more neutrophils can be seen in the vessel's hollow.



AT THE CENTER OF A WOUND 24 hours old neutrophils are cleaning up fibrin strands, foreign matter and cellular debris. Some have already burst, pouring their protein-digesting granules into the gap. Fuzzy "halos" around some granules may indicate that they have already started to break down the plasma protein. A monocyte has also entered the wound and has begun to ingest material. Map below identifies cells, granules and other features.

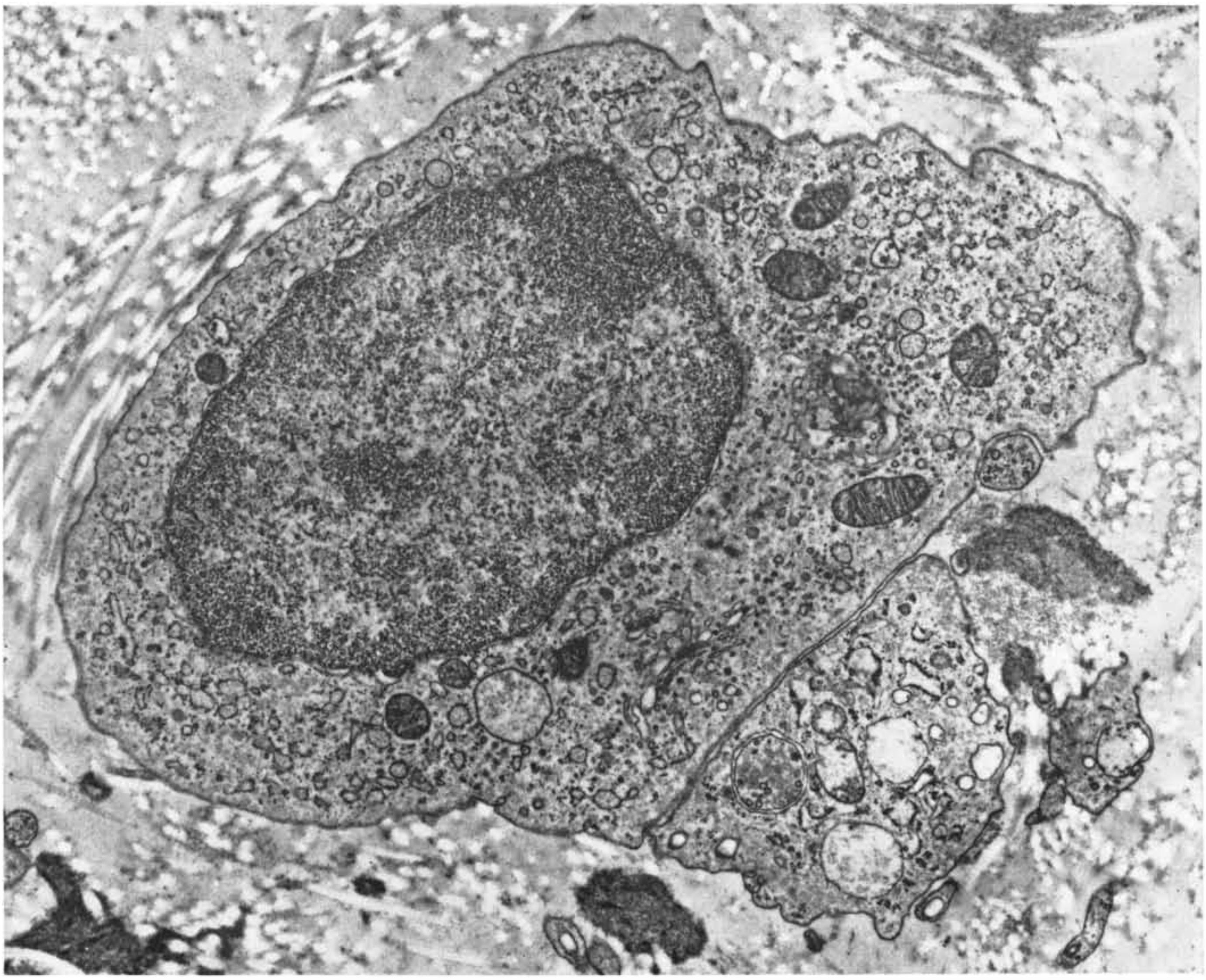


with the light microscope. After about two weeks (in the case of a linear incision) the synthesis of connective-tissue proteins begins to decrease and a process of remodeling takes place; it is this process that gives scar tissue its great strength. During remodeling many of the randomly oriented smaller collagen fibers that had formed earlier are broken down and reassembled into thick bundles. As the bundles form they orient themselves along the lines of stress in the healing wound. Eventually they gather into large fibers that look much like those in unwounded tissue. In fact, the difference between the original dermis and the scar tissue is one of degree rather than kind. Scar tissue is denser, its strands tend to lie more parallel to one another, and it contains fewer cells and blood vessels.

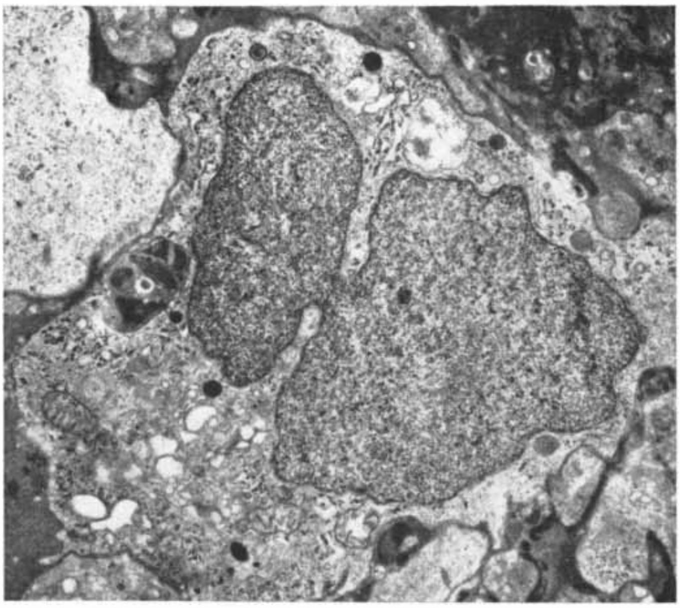
Just as the direction of stress determines how collagen bundles orient themselves in a wound, so the amount of scar tissue that forms depends in part on how much stress the wounded area normally receives as the body moves. More scar tissue may be needed to strengthen an injury on the hand or the leg than is needed for one on a comparatively quiet surface such as the chest or the abdomen. Heredity also plays a role. Negroes and Orientals, for instance, are more likely than whites to produce the raised scars called keloids.

The rough endoplasmic reticulum of the fibroblast synthesizes the collagen and other proteins that form scar tissue. In the fibroblast this organelle is highly developed, and the ribosomes attached to its membranes have a characteristic spiral or curved configuration [*see upper illustration on page 46*]. The rough endoplasmic reticulum of the monocyte is much less developed, which probably reflects the fact that the small amounts of protein synthesized by this cell are enzymes that work only within the cell to digest material in the wound. In contrast, the fibroblast synthesizes proteins for the construction of scar tissue and secretes them into the extracellular space.

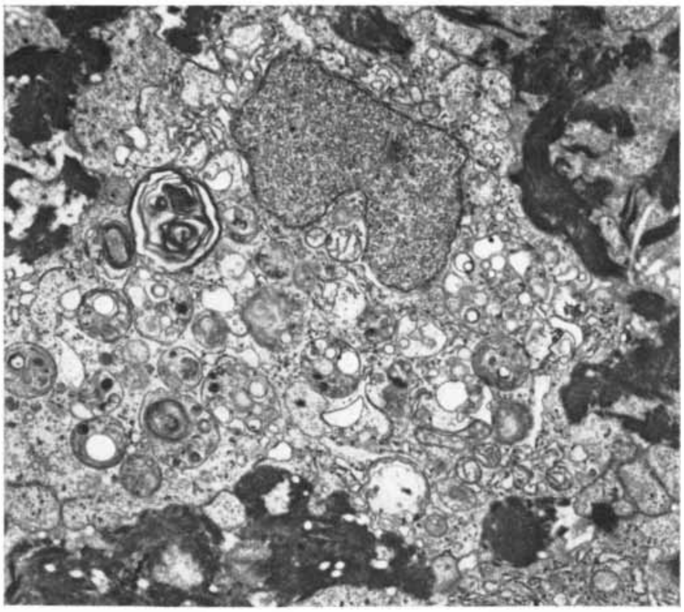
In a series of studies that confirmed this finding Benditt and I traced the manufacture of collagen through the organelles of individual fibroblasts. In order to follow the process we had to label some amino acid constituent of the protein with a radioactive atom. We chose the amino acid proline because it has the advantage of being the source of both proline and hydroxyproline, which together account for some 25 percent of the collagen molecule. We then followed labeled proline by making a series of electron-microscope autoradiographs.



MONOCYTE does not develop the capacity to ingest and digest foreign material until it has entered the wound. This cell in the dermis outside a wound displays features typical of the young monocyte: a large, round nucleus and a vesicle-filled cytoplasm.

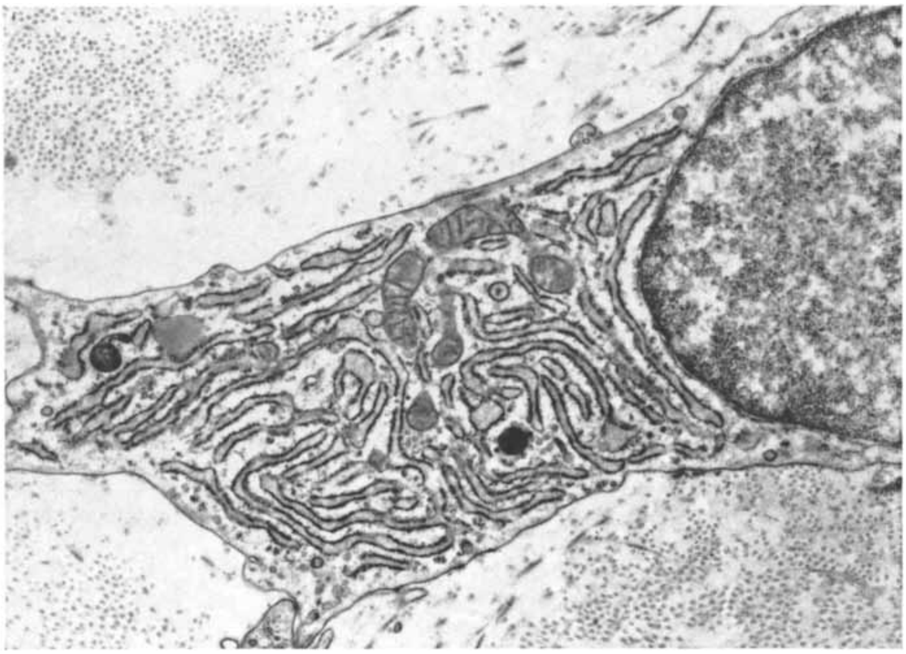


MONOCYTE BECOMES PHAGOCYTIC soon after it enters wound. The monocyte ingests foreign material by enclosing it in a vacuole, such as the one near the cell's nucleus. This particular vacuole contains strands of fibrin and serum protein from the clot.

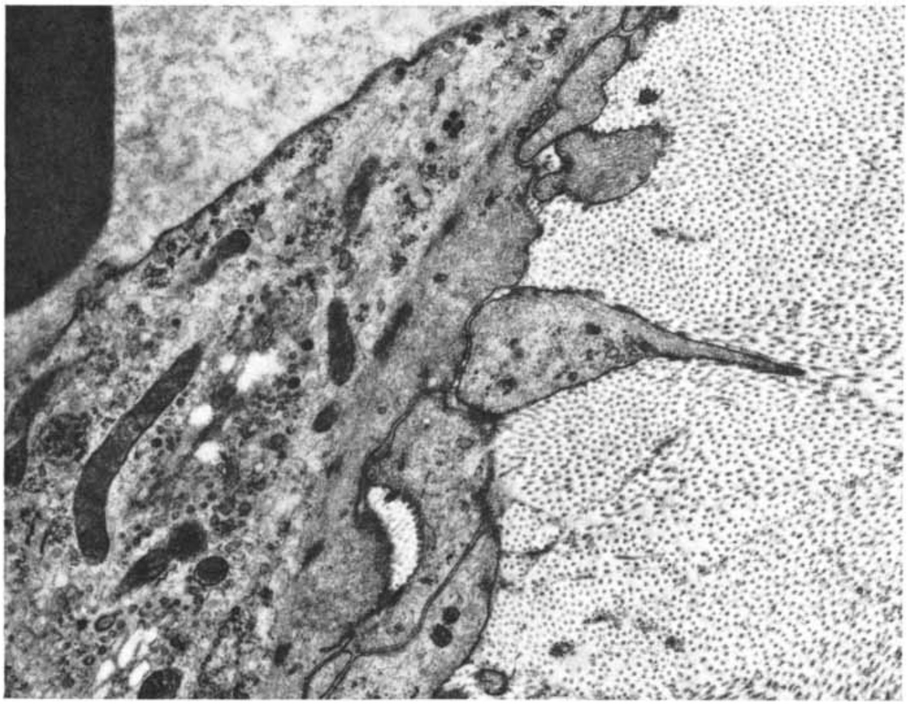


MATURE MONOCYTE, surrounded by dense fibrin strands, has a distended cytoplasm containing many vacuoles that are filled with ingested material, and whorl-like structures that may consist of undigested lipoprotein. This cell has been enlarged 11,000 diameters.

An autoradiograph is made by placing a layer of photosensitive emulsion on a section of tissue in which cells are synthesizing protein from amino acids, in this case proline. Since the silver grains of the emulsion lie close to the radioactively labeled proline they are exposed just as the grains on an ordinary photographic negative are. Under the high resolution of the electron microscope the exposed silver grains thus reveal which organelles in an individual fibroblast use proline. By quantitatively analyzing a large number of these autoradiographs made in sequence we have traced the pathways followed by the amino acid and the substances into which it became incorporated through the rough endoplasmic reticulum and the other organelles of the fibroblast [*see bottom illustration on opposite page*].



FIBROBLAST, enlarged 15,000 diameters, manufactures collagen molecules and other proteins in its mazelike "rough endoplasmic reticulum." Proteins are synthesized at the dark sites called ribosomes and are then transferred into the elongated channels. Many collagen fibrils, from which scar tissue forms, surround the cell. Part of nucleus can be seen at right.



BLOOD VESSEL REGENERATION begins, as shown in this electron micrograph, when a sprout branches from the vessel wall into the connective tissue. The sprout can be seen at center jutting out to the right, away from the grayish wall of the vessel. Sprouts eventually divide and form new series of vessels until a network extends throughout the wound. Part of a red blood cell can be seen in the old vessel's interior, in the upper left-hand corner.

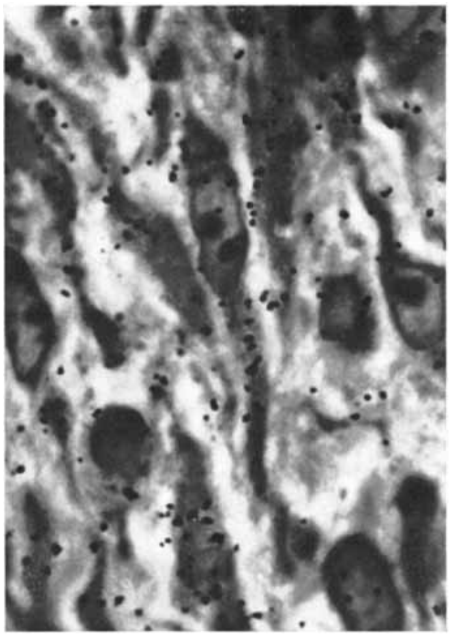
The fibroblast manufactures proteins as other cells do, but it appears that it may secrete one of the proteins, collagen, differently. For example, George E. Palade and his colleagues at Rockefeller University have demonstrated that the digestive enzymes synthesized by the exocrine cells of the pancreas are sequestered in the rough endoplasmic reticulum. Then the proteins are passed along to the region of the cell called the Golgi complex, which consists of a series of vesicles (cavities) and sacs that contain no ribosomes. Here the proteins are packaged into structures called condensing vacuoles. Subsequently these vacuoles form zymogen granules, in which the characteristic secretions of the exocrine cell are stored before they are released from the cell.

Although fibroblasts also have a Golgi complex, the process of secretion in these cells may be much more complex than it is in the exocrine cells. Benditt and I have reasoned that if a fibroblast stores its proteins after manufacture and transports them to special secretory mechanisms and sites as the exocrine pancreatic cell does, considerable time would elapse between the moment a fibroblast picked up an amino acid and the moment the amino acid, now part of a protein molecule, would be secreted back into the intracellular space. In order to measure the time a fibroblast needs to incorporate proline and other amino acids into a finished protein molecule we made a series of light-microscope autoradiographs of tissue in which fibroblasts were utilizing labeled proline and secreting collagen into the extracellular space. We observed that within an hour after the labeled amino acid had been injected into the abdominal cavity of a guinea pig it was localized predominantly within the fibroblasts, and that after an hour it had already begun to appear in the space adjoining the cells [*see top illustrations on opposite page*]. If the fibroblast has an appreciable storage phase in its manufacturing and secretion cycle, we would have expected the labeled proteins to remain for a longer period of time in some compartment within the cell, as they do in the exocrine pancreas.

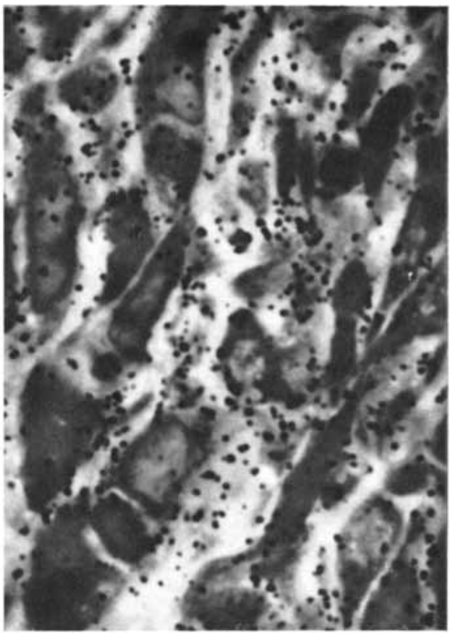
Fibroblasts, therefore, may secrete collagen directly from the rough endoplasmic reticulum, probably by forming small vesicles that pinch off from the sacs of the reticulum and migrate a comparatively short distance to the outer membrane of the cell. The vesicles would then fuse with the membrane and release their contents into the extracellular space [*see illustration on next page*]. There may also be a second secretion channel for collagen. We have observed sites that could be potential openings between the cavities of the rough endoplasmic reticulum and the extracellular space. Such a site may evolve into a passage through which material sequestered within the cavities could be released directly to the exterior.

Since autoradiography is a somewhat imprecise technique, this view of protein synthesis in the fibroblast cannot be stated as a definite fact. The difficulty is this: fibroblasts make several proteins simultaneously, including collagen, and proline is a constituent of all of them. Therefore in an autoradiograph it is not possible to tell whether collagen or some other labeled protein is represented by a particular silver grain. We will be able to remove some of the uncertainty, however, by tracing the passage of proteins through the fibroblast, and using biochemical information to determine which labeled proteins are present in each cell compartment at a given time. Current evidence of this kind seems to support the view that the rough endoplasmic reticulum synthesizes both collagen and the proteins that will later be joined with polysaccharides. These latter proteins then travel in vesicles to the Golgi complex, where polysaccharides are added, whereas the collagen may be secreted directly from the rough endoplasmic reticulum into the space outside the cell.

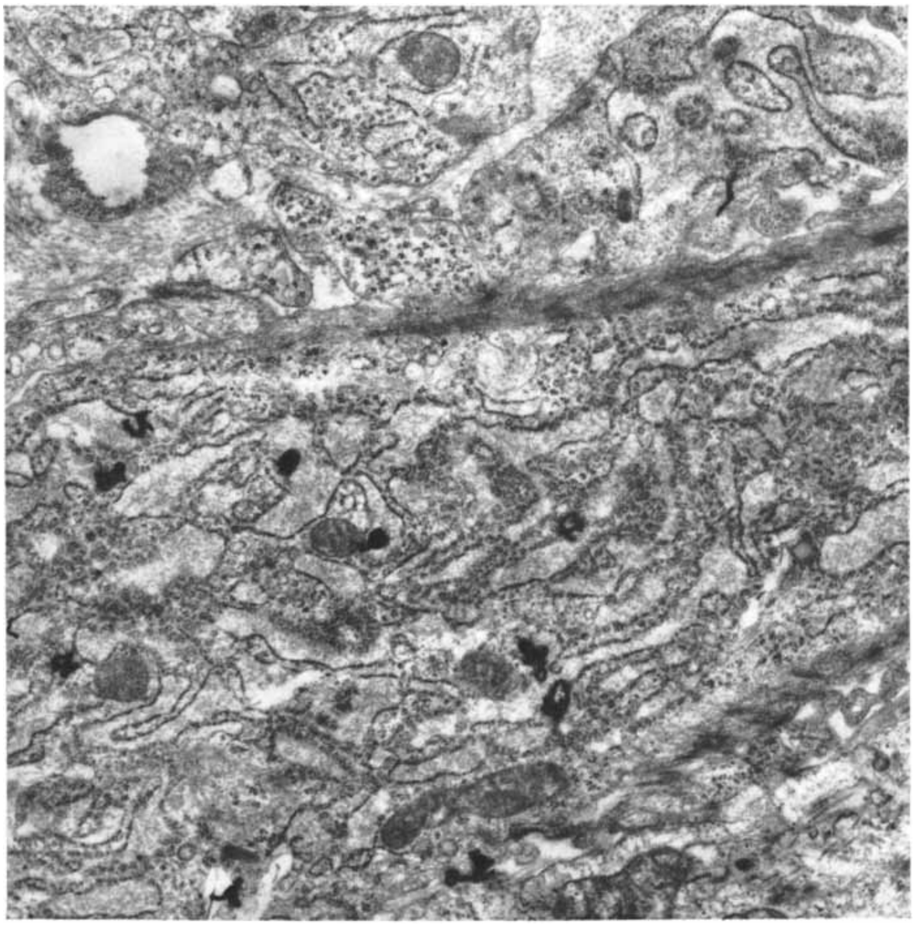
It will be interesting to determine definitely which intracellular pathway or pathways the fibroblast can use. Such information would help in the interpretation of pathological changes and suggest ways of studying the regulation of synthesis and secretion by fibroblasts. If clinicians could control the synthesis and secretion of collagen, they might be able to ward off some of the disabling effects of rheumatic heart disease. Patients with rheumatic fever become disabled because the heart continuously forms a new connective-tissue scar in order to repair a valve that is being repeatedly injured by infection. Eventually much of the valve is replaced by the dense scar tissue, and the thin, pliable valve becomes a thick, rigid structure that can-



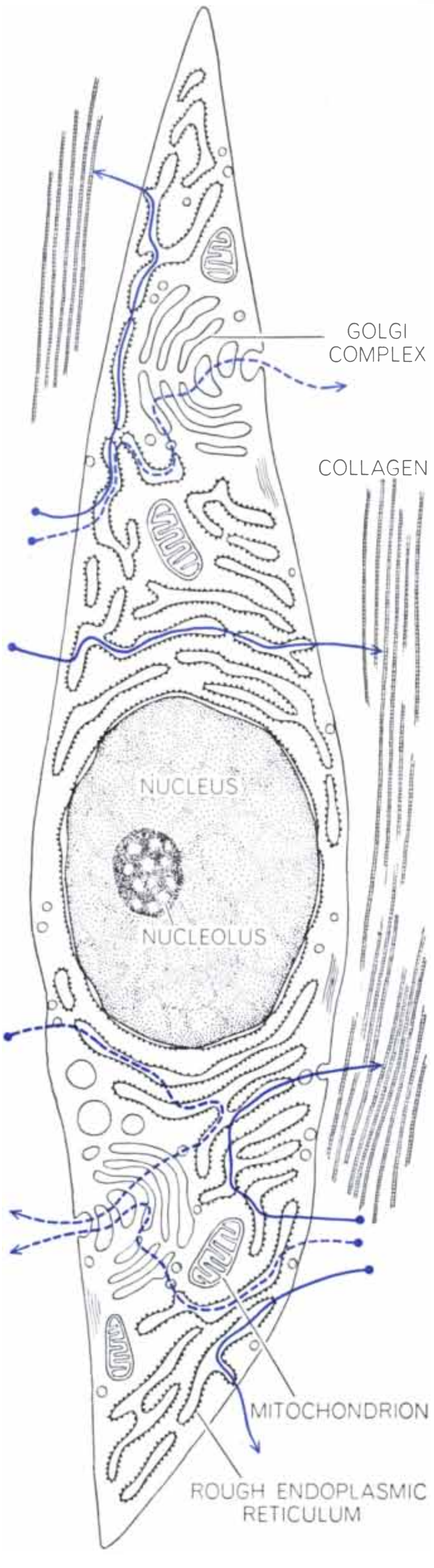
ONE HOUR after radioactively labeled proline has been administered to a guinea pig all this amino acid has been taken up by the dark fibroblasts, an event indicated by the localization of the exposed silver grains that appear as black dots around the dark fibroblasts in this light autoradiograph.



FOUR HOURS LATER most of the labeled proline has returned to the light gray space between the cells as part of the collagen they manufacture. This experiment shows that the fibroblasts secrete collagen immediately without storing it for long periods internally in organelles such as the Golgi complex.



ROUGH ENDOPLASMIC RETICULUM of a fibroblast is enlarged by 14,000 diameters in this electron-microscope autoradiograph. Exposed silver grains mark sites in this part of the cell where ribosomes are synthesizing collagen and other proteins. This autoradiograph is one of a series taken so that the pathways of protein synthesis in the cell could be traced.



PROTEIN PATHWAYS are shown in this idealized fibroblast. The evidence suggests that collagen (*solid arrows*) is secreted in two ways: a vesicle may detach from the rough endoplasmic reticulum and empty the collagen through the cell's membrane wall, or a channel may open that connects the reticulum directly with the space outside the cell. Other proteins (*broken arrows*) are synthesized in the reticulum, transferred to smooth-chambered Golgi complexes where polysaccharides are added, and secreted through the cell wall by speckled vesicles.

not close in the normal fashion. If the production of scar tissue could be regulated at the fibroblast level, hearts affected by rheumatic fever might be saved. There are other diseases in which too little scar tissue forms. They might be controlled if fibroblasts could be induced to manufacture more collagen.

Many early workers postulated that wound fibroblasts, like monocytes, arise from white blood cells. Together with Jack W. Lillywhite we have conducted several experiments using white cells isolated from the blood of experimental animals to test this hypothesis, and we have found it wanting. We obtained the white blood cells by two different methods. In one method the blood was withdrawn by puncturing either the heart or a vein, so that the needle had to pass through multiple layers of connective tissue before reaching the blood. In the other method the blood was secured by carefully isolating an artery and inserting a plastic tube into it. Here only the layers of the vessel wall were penetrated.

White blood cells obtained by these two methods were then allowed to grow in special chambers, which were planted under the skin or in the abdominal cavity of an experimental animal. The cells taken from the heart or a vein divided actively, formed large amounts of collagen and looked somewhat like fibroblasts or modified smooth-muscle cells. The white cells obtained from an artery did not divide, and no collagen formed in the chamber. Special filters with a mesh too fine to pass cells eliminated the possibility of fibroblast precursors entering the chambers once they were sealed. Therefore the connective-tissue cells found within those chambers in which collagen had formed must have appeared as a result of contamination of the blood by a few connective-tissue cells that were picked up when the needle punctured the heart or a vein, a result suggesting that the fibroblasts in a wound originate in connective tissue nearby.

As the fibroblasts synthesize the collagen and the protein polysaccharides of the healing wound's matrix of connective tissue, large numbers of small blood vessels form throughout the wound. According to the classic observations of Eliot R. Clark and Eleanor L. Clark of the University of Pennsylvania, these capillaries originate as budlike structures on nearby vessels, penetrate the wound and grow into loops. The loops then ramify throughout the wound by the division of their cells. As the capillaries from different sites migrate through the wound, they meet and form an interconnecting network of vessels.

In the early stages of wound repair this network of capillaries provides comparatively large quantities of oxygen for the cells that are actively synthesizing protein in the wound. John P. Remensnyder and Guido Majno of the Harvard Medical School have demonstrated that before the new capillary network forms there is a marked gradient of oxygen within the wound, the center of the wound being the most deficient in oxygen. This gradient, they say, may be partially responsible for the branching of new vessels into the region. Once the continuity of the connective tissue has been reestablished many of the new capillaries regress. Thus the wound changes from a tissue that is rich in blood vessels and actively dividing cells into one that has a much simpler cellular structure.

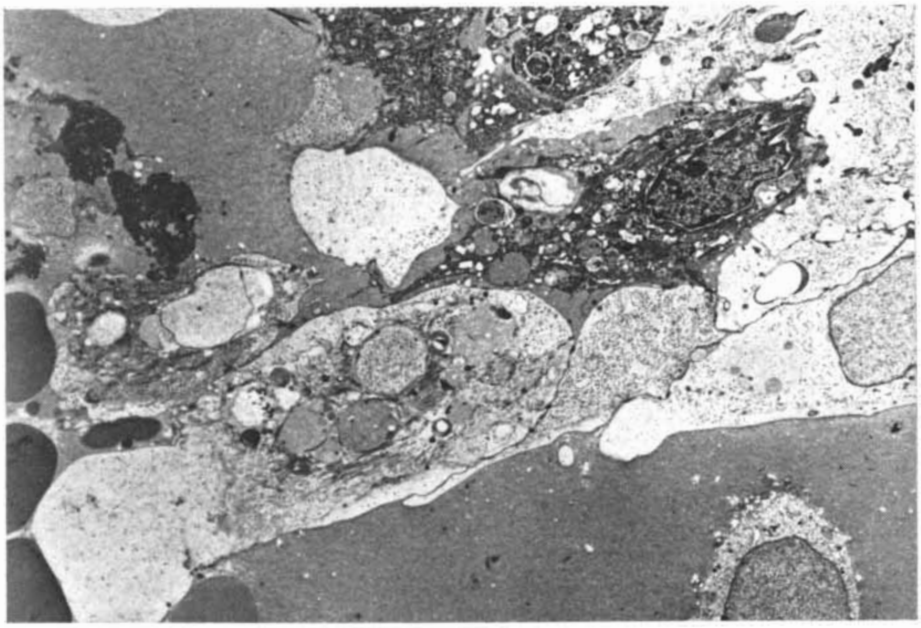
As scar tissue renews the dermis the epidermal cells begin to close the surface of the wound. In uninjured skin these ordered cells normally form a series of layers, each of which arises from the layer next to the dermis. The disruption of the epidermal layers causes some of the cells nearest the wound to degenerate, and most of the others lose the orderly, oriented appearance they had in their undisturbed state. Within 12 hours after wounding these cells become amorphous and develop ruffled borders and blisters, much like those seen in the actively moving cells grown in tissue culture.

The cells at the outer surface of the epidermis, the epithelial cells, are also highly mobile, and they travel quickly into the wound. The migration of the epithelial cell seems to be well organized. Normally these cells are attached to one another at sites called desmosomes, each of which is somewhat like a spot weld. In the course of migration the cells probably break up these attachments and re-form them. There are two reasons for believing that this happens; first, we have observed migrating cells attached to one another by desmosomes and, second, it would be difficult to explain the remarkable mobility of these cells if they were not able to rearrange their sites of attachment as they move into the wound.

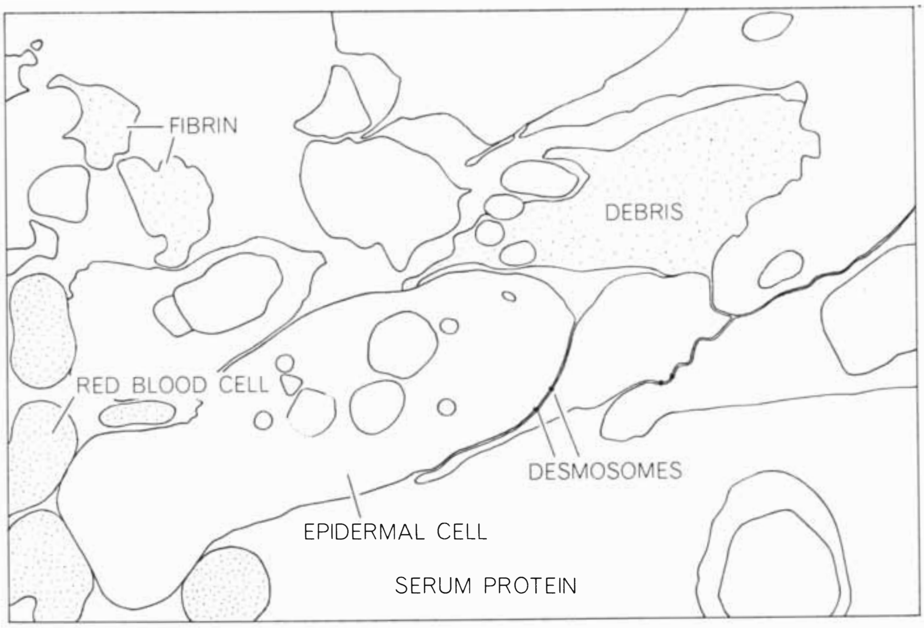
The scaffolding of fibrin derived from the clotted blood that forms the first provisional "patch" in the wound also guides the migrating epidermal cells as they slide glacially from all sides under the scab and into the wound. As the cells migrate through this milieu they exhibit a feature quite unusual for epidermis.

They actively participate in the ingestion and digestion of the strands of serum protein and fibrin lying in their path [*see upper illustration on next page*]. This behavior is similar to the phagocytosis accomplished by the white blood cells found deeper in the wound. In both instances such activity serves to clear the wound of debris and of the fibrin scaffolding that united the wound margins. When the leading edges of the sheets of epidermal cells meet and form a continuous layer beneath the scab, each cell regains its normal identity, and as the thickness of this layer of cells is restored the rate of division among the epidermal cells returns to what it was before the injury.

Once the epidermal cells have closed the wound, how do they "know" that the time has come to resume their normal shape and growth rate? It has been suggested by many investigations, in particular those conducted in the laboratory of Michael Abercrombie at University College London, that one of the factors in the control of cell division is "contact inhibition." Abercrombie noted that cells in tissue culture continue to divide until they establish contact, at which point they stop dividing. Perhaps the same process occurs in man and other animals, but there is no evidence that this is so. Moreover, no one knows how contact inhibition works. Some investigators have suggested that as the cells touch each other the distribution of electric charge on their surface changes, and that the change serves as a signal that halts growth, but here again we have little evidence. The problem is an interesting and important one. If the process through which normal cell growth stops and starts can be learned, investigators might be able to figure out what has happened in tumor cells that continue to divide even though they repeatedly come in contact with one another.



SURFACE HEALING BEGINS as the leading cell in a wedge of epidermis migrates into the wound. Fibrin strands and serum protein surround the cells. The cell's blunt, boxing-glove-like edge abuts a red blood cell. Desmosomes, which attach the epidermal cells to one another (*see map below*), can be seen at the cell's trailing edge, near the center of the picture. The cells and other bodies in this electron micrograph are enlarged by 4,500 diameters.

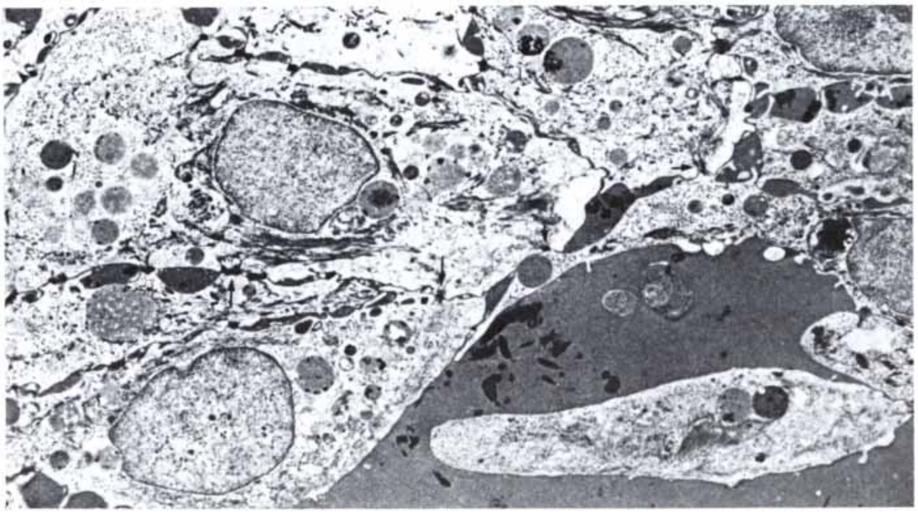


Since epidermis and dermis heal more or less simultaneously, it seems reasonable to ask if the two processes interact and are in any way dependent on each other. Hermes C. Grillo, who works in the laboratory of Jerome Gross at the Harvard Medical School, has found that the answer may well be affirmative. The connective tissue, it will be recalled, undergoes a process of remodeling in which bundles of collagen fiber form along the lines of stress running through the wound. This remodeling probably occurs as a result of the repeated dissolution and re-formation of collagen fibrils. What agent dissolves the fibrils? According to one hypothesis, the agent is the enzyme collagenase, which comes primarily from the epidermis, perhaps in response to chemical signals from the dermis. Grillo's experiments provide evidence supporting this idea. He found that when he added epidermal cells to dermal cells in a tissue culture, they greatly enhanced the production of collagenase. As in the case of contact inhibition, however, someone must demonstrate that epidermis and dermis behave this way together in the living animal.

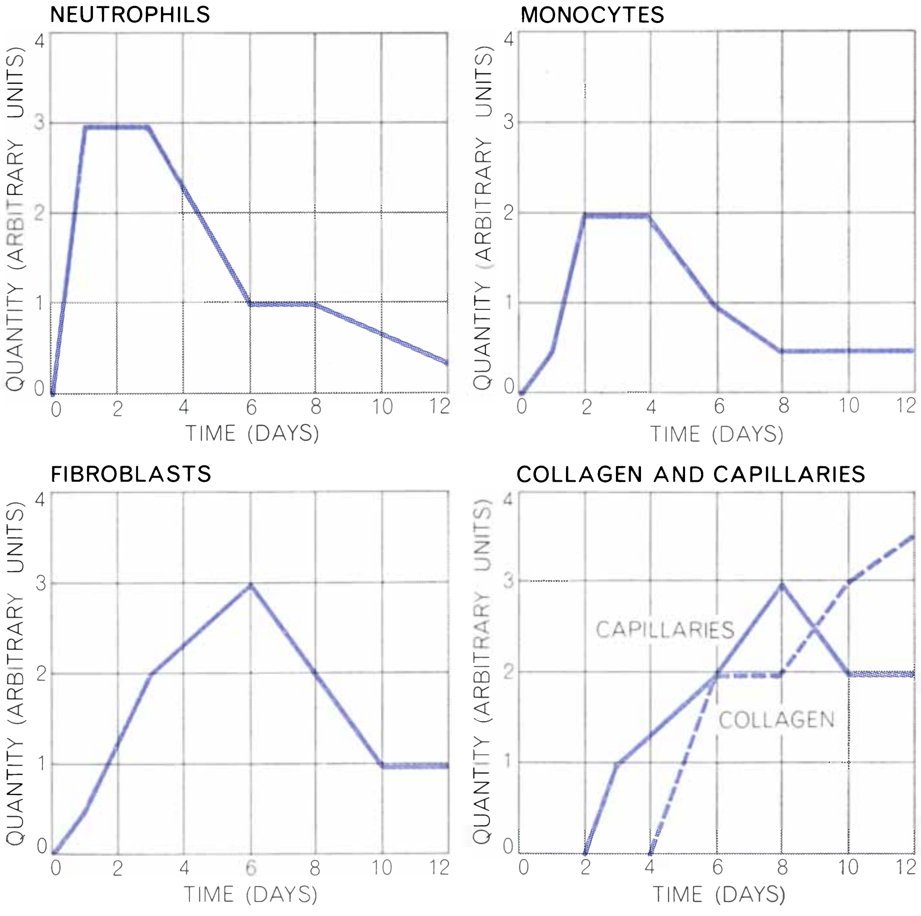
If the chemical events in simultaneous processes in the dermis and epidermis affect one another, it also seems possible that each stage in wound repair is an essential antecedent of the one following it. There is some evidence to support this hypothesis. As early as 1920 Alexis Carrel of the Rockefeller Institute succeeded in inhibiting the inflammatory response so that relatively few neutrophils appeared in the wound, and he found that the subsequent formation by fibroblasts of connective-tissue proteins was much delayed. Even earlier it was observed that the interruption of the formation of connective tissue can prevent wounds from healing. In his book *A Voyage around the World,* published in 1748, Richard Walter, chaplain to

George Anson, commander of the squadron that made the voyage, described the plight of seamen suffering from scurvy whose wounds failed to heal. Their scurvy was due to a lack of fresh fruit or vegetables in their diet, and it was later discovered that without such foods scurvy sufferers cannot form connective tissue because they have no source of vitamin C.

Burton Wolbach of the Harvard Medical School confirmed the importance of vitamin C to collagen formation in 1926. In the microscope he observed that fibroblasts from guinea pigs with scurvy produced a gelatinous, collagen-like substance with none of collagen's structural strength. It was only recently that the action of vitamin C in collagen formation has been clarified. Normally vitamin C allows hydroxyl groups (OH) to be added to the amino acids proline and lysine through the action of an enzyme, converting them into hydroxyproline and hydroxylysine. Hydroxylation, as the process is called, takes place after the amino acids have been synthesized into almost complete collagen molecules in the rough endoplasmic reticulum of the fibroblast. Without vitamin C the hydroxyl groups cannot be added, the collagen molecules remain incomplete and may not be secreted by the cell. Therefore collagen fibrils cannot form.



EPIDERMAL CELLS slightly to the rear of the leading edge of the epidermis contain round, dense vacuoles, filled with the serum protein and fibrin they have cleared from the wound by ingestion. Serum protein and somewhat darker patches of fibrin separate some of the cells from one another. These epidermal cells are strikingly different from those in undamaged tissue, which contain no vacuoles and have a different cytoplasmic structure.



CELL POPULATIONS that occupy the wound at different times can be divided into three groups, each with its own special task to perform. The stages tend to overlap one another and may well be interdependent, since suppression of one can delay the start of the next.

Our electron-microscope studies reveal that this chemical flaw is accompanied by a breakdown of the endoplasmic reticulum of the fibroblast. The walls of the cavities in the reticulum appear to be broken down, so that large, vacuole-like structures form. In addition, the ribosomes attached to the membrane of the endoplasmic reticulum are no longer arranged in the characteristic configuration seen in normal cells. We do not yet know how the morphologic changes observed in the fibroblasts are related to the biochemical changes that have been described in scurvy. The morphologic changes can be readily reversed, however, by feeding the experimental animal vitamin C; the cells resume their normal appearance within 24 hours, and recognizable collagen fibrils begin to appear in the extracellular space within 12 hours. It will be interesting to determine whether an increased rate of synthesis, or simply hydroxylation of a stored intracellular precursor of collagen, causes this rapid appearance of collagen.

Studies of wound repair have yielded much information about this important defense mechanism of the body. Like all research efforts, however, this one has generated new questions as well as new answers. For example, we should like to determine if there are important relationships between each of the different kinds of cells that migrate into the wound during the healing process, and to identify the factors that cause cells to start or stop multiplying. It may also be possible to modify with drugs the various protein-synthesizing activities of the cells in wounds in order to provide an optimal amount of connective tissue. These and other important problems remain to be solved, perhaps by further study of the healing wound.

In hope of doing each other some good

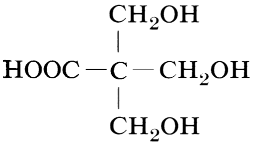
Sara and her bacillus

Sara Goodhue's daddy, Dr. Charles T., is a microbiologist who works at a company that makes a lot of film and cameras. You might wonder why a microbiologist would work at a place like that, but you would stop wondering when you remembered that molecular biology does not deal only with giant molecules of profound philosophical significance. It also deals with small, simple ones such as are needed in great big quantities because light-sensitive silver salts keep failing to level off their growth rate in the economy.

Sara's daddy was getting a little discouraged. Neither he nor Dr. James R. Schaeffer nor Mr. Hugh Risley nor Mr. Richard Stevens, who all worked together very hard, was able to find a bacterium to eat $C(CH_2OH)_4$. It's too darned artificial a molecule, but one some factories find easy to make.

Since Sara was already nine years old, her daddy decided it was time for her to help. So she put on her boots, went out to the little creek behind the house, and brought back some mud in which lived this bacillus, along with many other kinds:

And this bacillus really did eat $C(CH_2OH)_4$. Goodness knows why. Since nobody in his right mind truly loves a bacillus, Dr. Goodhue and Dr. Schaeffer were not interested in feeding this one just for its own sake. They named it *Flavobacterium oxydans* and enslaved its descendants, who, being naturally highly prone to mutation, were taught in only a couple of years to eat $C(CH_2OH)_4$ down only as far as



$$\begin{array}{ccccc} & & \mathrm{CH_2OH} & & \\ & & | & & \\ \mathrm{HOOC} & - & \mathrm{C} & - & \mathrm{CH_2OH} \\ & & | & & \\ & & \mathrm{CH_2OH} & & \end{array}$$

and then give it all up, nearly 100% of it, while subsisting on nothing but acetic acid.

The product which *F. oxydans* has been taught to give for its keep, known as tris(hydroxymethyl) acetic acid, used to be expensive and dangerous to make on a large scale. We like it very much.

If you think you'd like it too, we'd like to know. Please drop a note to Eastman Kodak Company, Eastman Organic Chemicals, Rochester, N.Y. 14650.

Omit gin

A certain well known brand of quinine water can be made to lase. It lases much better than its equally well known competitor. We have no intention of disclosing which is which, nor the circumstances that inspired this particular study, nor the precise ultimate objectives of our main study of organic dye lasers. When the work turns up something that ought to be published in a serious way, we present it to a proper forum.

Meanwhile if some ambitious youngster is asking for your guidance in making a liquid laser, refer him or her to Bruce Burdick, Mail Code 918, Eastman Kodak Company, Rochester, N.Y. 14650.

On American soil

At a proper forum entitled "New Horizons in Color Aerial Photography" under the auspices of the American Society of Photogrammetry and the Society of Photographic Scientists and Engineers at New York's Coliseum on Monday morning, June 9, a man from the Kodak Research Laboratories explains why he has been assembling samples of as many different categories of American soil and sand as a concerted effort can collect, and to what purpose the light reflectance of all these samples has been measured—wet and dry—over a broad spectral range.

If you see why you should attend but are reading this too late to do so, contact can be established through Aerial Product Planning, Eastman Kodak Company, Rochester, N.Y. 14650.

What do you want to do with your money?

We keep trying to sell things. Perhaps, as hinted above, color film *could* make a soil scientist's working days more productive. Perhaps super 8 film *can* teach better than a book. Perhaps someone *had* better keep the electron microscopes supplied with reliable film. How much are these things worth?

Through the city where the things are made, a river flows into a large lake. Far fewer species of animals and plants probably inhabit that river than when the Seneca tribe of the Iroquois confederacy ruled its banks. How many Senecas today are pedologists, teachers, or electron microscopists we do not know, nor how abundantly sturgeon spawned in the coves back when no Seneca could or would aspire to such a life. Since nobody expected to pick up a can of beans after a hard day at the electron microscope, very little forest was cleared for bean acreage. Plenty of tree root hairs kept the soil from silting the river.

Enough persuasive words have been spilled. It is now dimly seen that convenience sometimes carries too high a price in tawdriness. Let us ask ourselves how far to reverse our mighty engines. Let us give ecologists credit for being as smart as our engineers, chemists, and geologists.

Applied ecology probably cannot restore the exact pre-Columbian state of the river and the lake but might come closer than you think, given sufficient backing. How much do you want to pay?

If you want beans or film at the lowest possible price, it might still be economically feasible to seed the river banks several times a year with mud turtles from a commercial mud-turtle hatchery. You would be able to show the kids a turtle sunning himself on the side of a car that some older kid ran over the bank on a summer night in '62. And that would be that, if you want the lowest possible price.

The cost of reducing and reversing environmental degradation could be put on the same financial basis as the more familiar and traditional price-determining factors. Do you buy that idea?

Now that we are all agreed, next question is how much applied ecology to apply. The United Nations General Assembly has unanimously (!) agreed to hold a conference in 1972 to pick some winning answers. We shall try to anticipate them. Why wait for 1972?

EASTMAN KODAK COMPANY An equal-opportunity employer

There's no such thing as

last year's Xerox machine.

There are no '67, '66, '65, '64 or '63's, either.

Instead of making your Xerox machine go obsolete by plan, we're forever looking for ways to make it go better. By plan.

Even while the machine's in an office somewhere, running off copies, we're continually improving things inside it. Changing this, reworking that.

You don't have to request it. We just do it. At no extra charge.

Keeping your Xerox machine "this year's" doesn't stop here, though.

Suppose your needs should change next month, and you prefer a different Xerox model.

Fine. There's a clause in your Xerox lease agreement that lets you have it—on just 15 days' notice.

Business generosity? No.

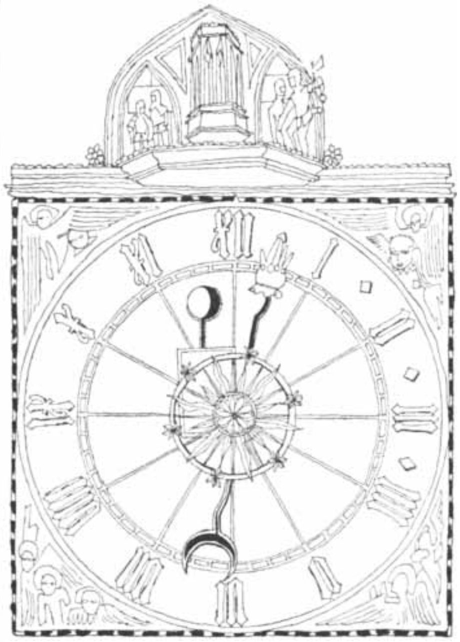
We figure even if you signed with Xerox back in '63, you'd better get results that are '69.

Or maybe you'll turn to somebody else's '70.

XEROX

XEROX CORPORATION, ROCHESTER, NEW YORK 14603. OVERSEAS: SUBSIDIARIES THROUGHOUT LATIN AMERICA; AND IN ASSOCIATION WITH RANK ORGANISATION LTD., RANK XEROX LTD., LONDON; AND THROUGH RANK XEROX LTD., FUJI-XEROX CO., LTD., TOKYO. XEROX IS A REGISTERED TRADEMARK OF XEROX CORPORATION.

SCIENCE AND



No Secrets

The Massachusetts Institute of Technology, which receives more research money from the Department of Defense than any other educational institution, is accepting no more contracts for secret research pending a report from an 18-man panel. The panel, created by Howard W. Johnson, the president of M.I.T., consists of faculty members, students, alumni, members of the M.I.T. Corporation and members of the two laboratories affiliated with M.I.T. that now conduct secret research: the Lincoln Laboratory and the Instrumentation Laboratory. The two laboratories receive nearly $100 million a year from the Department of Defense.

The decision to create the panel followed several weeks of protest by students and faculty members who are opposed to M.I.T.'s doing secret research of any kind. They particularly object to two projects of the Instrumentation Laboratory involving guidance systems for helicopters and for Poseidon missiles, the multiwarhead rockets that will replace Polaris missiles. The Instrumentation Laboratory, which earlier developed the guidance systems for the Thor and Polaris missiles, also conducts nonsecret research, such as the development of the guidance and navigation system used by the Apollo spacecraft.

In announcing the formation of the panel Johnson said: "In the last few years [there has been] deep concern about the essential nature of a university and its relations to the society and to government. . . . At M.I.T. this concern has been directed specifically at the operation and programs of the Instrumentation and Lincoln laboratories. . . . A large part of the M.I.T. community would like to see the roles, procedures and programs objectively reexamined. The function of the panel will be to evaluate the benefit and the implications that the laboratories have for [M.I.T.] in its prime responsibility for education and research and in its responsibility for service to the nation." Johnson added: "The panel must recognize that the two laboratories have such highly developed technical capability that they have become important national resources, which could not be reproduced without tremendous effort and a long period of time."

Of the more than $5 billion that the U.S. is spending for research in the current fiscal year, $1.5 billion goes to colleges and universities, and of this amount $247 million comes from the Department of Defense. According to Lee A. DuBridge, the President's science adviser, about 4 percent of the $247 million goes to secret research. Two years ago the figure was 8 percent.

The Rubella Vaccines

Live-virus vaccines that have produced immunity against rubella ("German measles") in more than 95 percent of the test subjects who have received them will probably be licensed in the U.S. within a matter of weeks. Widespread immunization with these vaccines could prevent another wave of infection, anticipated for the early 1970's, like the one that swept across the U.S. in 1964, causing at least 8,000 fetal deaths and afflicting 15,000 to 20,000 infants with deafness, heart disease, cataracts, glaucoma, psychomotor retardation and blood disorders.

The vaccines are based on the HPV-77 strain, a live, attenuated form of the rubella virus produced by Harry M. Meyer, Jr., and Paul D. Parkman of the National Institutes of Health. Regulations setting standards for the effectiveness and safety of the vaccines, developed from this strain and produced by Merck Sharp & Dohme Research Laboratories and Philips Roxane, have been published in *The Federal Register*. The vaccines will be available when manufacturers, investigators and clinicians have commented on the regulations and final standards are published.

Physicians and public health agencies are counseled by the Communicable Disease Center's advisory committee on immunization practices to use the vaccines in an attempt to eradicate the disease among children, who are its major carriers. Some children will receive the vaccine from family physicians, and others will be immunized in mass programs financed by public and private funds. Since 85 percent of the adults in the U.S. have already had rubella and are therefore immune for life, the advisory committee sees little point in routine vaccination of adults. There will also be no mass immunization of women of childbearing age because it is not known how the live-virus vaccine will affect the fetus. Instead women will be immunized by their physicians on an individual basis. Physicians must be careful, the advisory committee warns, to avoid inoculating women who are pregnant or may soon be.

Clinicians will ultimately have a variety of vaccines with which to protect their patients. HPV-77 was only the first discovered of the three attenuated rubella strains from which vaccines are being developed. Industry and the NIH concentrated on it to get a vaccine on the market as soon as possible. In Belgium, Recherche et Industrie Thérapeutiques, now a subsidiary of Smith Kline & French Laboratories, has produced a vaccine from the Cendehill strain, a variety of virus raised in rabbit-kidney tissue. Switzerland has already licensed the vaccine, and Smith Kline & French is now completing its application for a U.S. license. At the Wistar Institute of Anatomy and Biology in Philadelphia, Stanley A. Plotkin has produced another attenuated strain by growing rubella virus in a tissue culture of embryonic human lung cells. The Institut Merieux in France is currently developing a vaccine from this strain.

Whether the next rubella outbreak is averted depends on how soon it occurs, on whether enough vaccine can be produced and on the number of children inoculated. If the epidemic begins next year, even the most intense effort could fail. An outbreak may not, however, come that soon. Sometimes a high incidence of the disease during the preced-

ing year foreshadows an epidemic. During the spring of 1969 no major wave of rubella was detected, so that it would seem that rubella will soon cease to be a major threat.

Heavyweight

The synthesis of element 104, the heaviest man-made element, has been announced by Albert Ghiorso and his colleagues at the Lawrence Radiation Laboratory of the University of California at Berkeley. Ghiorso described the latest addition to the periodic table of the elements at the Mendeleev Centennial Symposium organized by the American Chemical Society to commemorate the origination of the table by Dmitri Mendeleev in 1869. Ghiorso said that he and his co-workers have discovered two isotopes of the element: 104-257 and 104-259. They have less conclusive evidence for the isotope 104-258, he said, and so far have been unable to confirm the existence of the isotope 104-260, which was reported by a group of Russian workers in 1964.

An element's atomic number reflects the number of protons in its nucleus; an isotope's mass number reflects its total number of protons and neutrons. Thus the isotope 104-257 has 104 protons and 153 neutrons. Ghiorso said that his group created 104-257 by bombarding a target of californium 249 (an isotope of element 98) with nuclei of carbon 12; the combination of the two nuclei was followed by the emission of four neutrons. The isotope has a half-life of four to five seconds and decays into the previously known nobelium 253 by emitting an alpha particle (two protons and two neutrons). The same reaction, except for the emission of three neutrons, yielded the isotope 104-258, which has a half-life of about .01 second and decays by spontaneous fission, so that it is difficult to assign a particular atomic number to the fissioning nucleus. The isotope 104-259 is formed by the combination of carbon 13 and californium 249, followed by the emission of three neutrons. The isotope has a half-life of three to four seconds and decays by alpha radiation into the previously known nobelium 255.

Element 104 is the first element in a region of the periodic table that Glenn T. Seaborg, the codiscoverer of elements 94 through 102 who is now chairman of the Atomic Energy Commission, has called the transactinide region (see "The Synthetic Elements: IV," by Glenn T. Seaborg and Justin L. Bloom; SCIENTIFIC AMERICAN, April). He predicted in 1944 that a number of elements would show a family-like relation, namely that elements 90 through 103 would resemble the series of rare-earth elements from element 58 through element 71, which had been named the lanthanide series because they have chemical properties similar to those of the immediately preceding element, lanthanum (element 57). By analogy the series from element 90 through element 103 was named the actinide series after actinium (element 89), which immediately precedes the first element in the series (thorium). Seaborg has predicted that following the transactinide elements (104 through 121) there should be another rare-earth-like series of elements, from about 122 through 153, that can be called the superactinide series.

Old Australopithecines

Ever since the first discovery of "man-ape" fossils in South Africa during the 1920's, students of human evolution have faced a dilemma. The extinct hominid genus, named *Australopithecus,* was an excellent candidate for the role of man's immediate ancestor—except that the fossils were not old enough. Their apparent age made some man-apes contemporary with *Homo erectus,* the primitive but fully human species that first appeared about a million years ago. How then could *Australopithecus* have given rise to *Homo?* This temporal barrier has now been removed by the discovery of *Australopithecus* fossils that are three million years old.

The new finds were made in Ethiopia (in the fossil-rich formations of the valley of the Omo River, a tributary of Lake Rudolf) in 1967 and 1968 by members of a joint U.S.-French expedition. They consist of four incomplete jaws and more than 50 individual teeth. The fossil formations contain volcanic ash; potassium-argon isotope dating shows that its age ranges from less than two million years, fairly early in the Pleistocene epoch, to more than four million years, when the preceding epoch, the Pliocene, was drawing to a close.

Opposition to the acceptance of *Australopithecus* as man's immediate ancestor began to lessen when remains of the genus uncovered at Olduvai Gorge by L. S. B. Leakey in 1959 proved to be some 1.75 million years old. It further diminished when a 2.5-million-year-old fragment of hominid arm bone, discovered by Bryan Patterson of Harvard University in 1965 near Lake Rudolf in Kenya, was identified as another *Australopithecus* fossil.

Announcing the Omo finds recently,

one of the discoverers, F. Clark Howell of the University of Chicago, pointed out that the two jaws found by his group evidently belong to *Australopithecus robustus,* the sturdier of the two species that comprise the genus. Both jaws are about two million years old. One is the largest *Australopithecus* mandible yet found. In Howell's opinion, most of the teeth discovered by his group and perhaps the jaws and some of the teeth unearthed by French workers belong to the smaller species, *A. africanus.* Some of the teeth are about three million years old, making them the oldest hominid remains known. In addition to giving *Australopithecus* ample time to fit into man's family tree, the Omo discoveries support the view that the two species had begun to develop along divergent lines long before Olduvai times.

Legal Remedy

In the first action by a state to end the use of DDT, Michigan has moved to cancel permission for the sale of products containing the pesticide at the end of June. The State Agriculture Commission, which had stopped the sale of DDT for mosquito control a year ago, took the new action on the basis of a recommendation from the state university's agricultural experiment station. A major consideration was apparently concern for commercial and sport fishing in Lake Michigan, into which much of the surface water in the state eventually drains. High concentrations of DDT have been reported in the fat of whitefish, trout, perch and salmon.

DDT, like other chlorinated hydrocarbons, is a persistent chemical, not readily broken down. It accumulates in soil and water and is concentrated as it moves up the aquatic food chain from algae to increasingly large fishes and carnivorous birds (see "Toxic Substances and Ecological Cycles," by George M. Woodwell; SCIENTIFIC AMERICAN, March, 1967). Both its persistence and its lethal effect across a broad spectrum of animal life have spurred the development of newer pesticides that are more specific in their effect and less long-lasting, and now DDT is coming under increasing attack. In Sweden a two-year moratorium on the use of DDT and a ban on some other chlorinated hydrocarbons will go into effect in 1970. A trial one-year moratorium is in force in Arizona. The Wisconsin Department of Natural Resources has been holding intensive hearings on a citizens' petition to declare DDT a pollutant and end its use. Legislation to the same effect is being considered in several states and has been introduced in the U.S. Senate.

Two-Planet Solar System

Six years ago Peter van de Kamp, director of the Sproul Observatory of Swarthmore College, announced that photographic evidence of minute wobbles in the path of Barnard's star had led him to the conclusion that a planet-sized dark companion was revolving around this nearby star once every 24 years. Since then the motion of Barnard's star across the sky has been further plotted with the aid of the Sproul 24-inch refracting telescope. These observations have confirmed the existence of the perturbation observed in the earlier photographs; moreover, calculations based on the new photographs have indicated a very elongated orbit for the companion. This fact, plus signs of a small secondary perturbation, have recently led van de Kamp and his colleagues to present an alternative interpretation, "namely that the observed perturbation is the resultant of two perturbations caused by *two* unseen companions of Barnard's star."

According to van de Kamp, the two objects, which he calls *B*1 and *B*2, appear to be revolving around Barnard's star in approximately the same plane in circular orbits with periods of 26 and 12 years respectively. He has calculated the respective masses of *B*1 and *B*2 to be 1.1 and .8 times the mass of Jupiter. The radii of their orbits around Barnard's star are closely comparable to the orbits of Jupiter and an average asteroid in our solar system. Therefore van de Kamp concludes these objects can be considered planets.

The Heart and Hard Water

Several studies in the past decade or so have suggested that the death rate from coronary disease is inversely correlated with the hardness of the local water supply: the harder the water, the lower the coronary rate. Because there was no correlation with the presence or absence of calcium, magnesium or any other substance commonly reported in municipal water supplies, epidemiologists began to look for an unknown "water factor" in coronary deaths.

A study recently published in *The New England Journal of Medicine* reports evidence that the excess coronary deaths in soft-water areas are almost entirely sudden deaths outside the hospital. In such cases the cause of death is usually certificated by a coroner rather than by a physician. T. W. Anderson, W. H. le Riche and J. S. MacKay of the University of Toronto School of Hygiene reviewed the death certificates of 55,000 people who died in the province of Ontario in 1967 and classified the deceased individuals according to whether their local water supply had a hardness of less than 100 parts per million, of between 100 and 200 p.p.m. or of more than 200 p.p.m. They found that the standardized death rate from ischemic, or coronary, heart disease declined from 416 per 100,000 in the first category (softest water) to 390 per 100,000 in the second category (medium-hard) and to 365 per 100,000 in the third category (hardest). There was no correlation for any other disease.

When the deaths from ischemic heart disease were classified into deaths reported by coroners (sudden deaths) and those not reported by coroners, the correlation with type of water supply was even more striking. The sudden-death rates per 100,000 were 195 (soft water), 164 (medium-hard) and 120 (hardest). There was no correlation between water hardness and nonsudden deaths from ischemic heart disease. The Toronto workers conclude: "The main effect of water hardness may be on the mechanism causing death rather than on the underlying process of myocardial infarction." They suggest that something present (or absent) in soft water may favor cardiac arrhythmias that often lead to sudden death.

In an accompanying editorial in *The New England Journal of Medicine* Henry A. Schroeder of the Dartmouth Medical School offers the hypothesis that soft water, which is often slightly acid, dissolves cadmium normally present as a contaminant in the zinc used to galvanize pipes. It has been shown that soft water passing through galvanized pipes can dissolve enough cadmium to exceed by two to eight times the limits allowed in food (10 micrograms per liter). In rats trace amounts of cadmium are known to produce the "pathological picture" of hypertensive disease. If this hypothesis is correct, says Schroeder, there should be a perceptible decline in coronary deaths in coming years because copper has been widely used in place of galvanized iron piping in dwellings built since World War II.

New Bond

A new chemical bond has been discovered in nature: the carbon-phosphorus bond. The enormously complex chemistry of living things is based essentially on carbon atoms stably bonded to one another and to atoms of a few other elements, notably hydrogen, oxygen, nitrogen and sulfur; when still other elements appear in biological compounds, they are typically bonded, rather unstably, by way of oxygen or nitrogen. The carbon-phosphorus bond was noted first some 10 years ago by Masaaki Horiguchi and Makoto Kandatsu of the University of Tokyo and then in 1962 at the City of Hope Medical Center in Duarte, Calif. Now the American discoverers, James S. Kittredge and Eugene Roberts, have reported in *Science* on what amounts to a new field of phosphorus biochemistry dealing with a group of compounds known as aminophosphonic acids.

The carbon-phosphorus bond came to light when Kittredge and Roberts noted a previously unidentified spot on a paper chromatogram of material from the sea anemone. It migrated like a phosphate ester, ethanolamine phosphate, but it resisted acid breakdown and it left a large amount of ash on combustion—both of which led to the discovery that the phosphorus in this compound was stably and directly bonded to carbon rather than being linked, as it is in a phosphate, by way of oxygen. They identified the substance as 2-aminoethylphosphonic acid (AEP). Later they learned that Horiguchi and Kandatsu had previously isolated AEP from one-celled animals called ciliates that exist as symbionts in the stomach of sheep. AEP and related compounds have since been found in a number of marine invertebrates, notably coelenterates and mollusks, and the route by which they are synthesized has been tentatively worked out. Carbon-phosphorus compounds have been shown to be incorporated in fats and enzymes.

Kittredge and Roberts speculate on the biological advantages that may be conferred by the carbon-phosphorus compounds. Phospholipids (with the usual oxygen-phosphorus bond) are normal components of many biological membranes; perhaps phosphonolipids, armed with the highly resistant carbon-phosphorus bond, serve as components of membranes that need to be resistant to certain enzymes. By the same token, enzymes that contain the carbon-phosphorus bond may act like those with phosphate bonds but be more resistant to the action of other enzymes they encounter. Similar reasoning would explain the adaptive value of the carbon-phosphorus bond to organisms in the sheep's stomach, which is rich in enzymes that break down ordinary phosphates.


Now pollution control can help pay for itself.

The greatest show on the moon.

Earth.


This mini-furnace could heat a 9-room house.


Computer on board new survey ship speeds exploration for offshore oil.

Looking through the window of Apollo 8, more than 240,000 miles deep in space, man first sees the inviting azure earth rising beyond the desolate horizon of the moon.

Guiding the spacecraft on each Apollo mission is Raytheon's on-board computer, the brain of the guidance and navigation system.

Raytheon is exploring new frontiers in a variety of fields – in space, beneath the floor of the ocean, in the research laboratory.

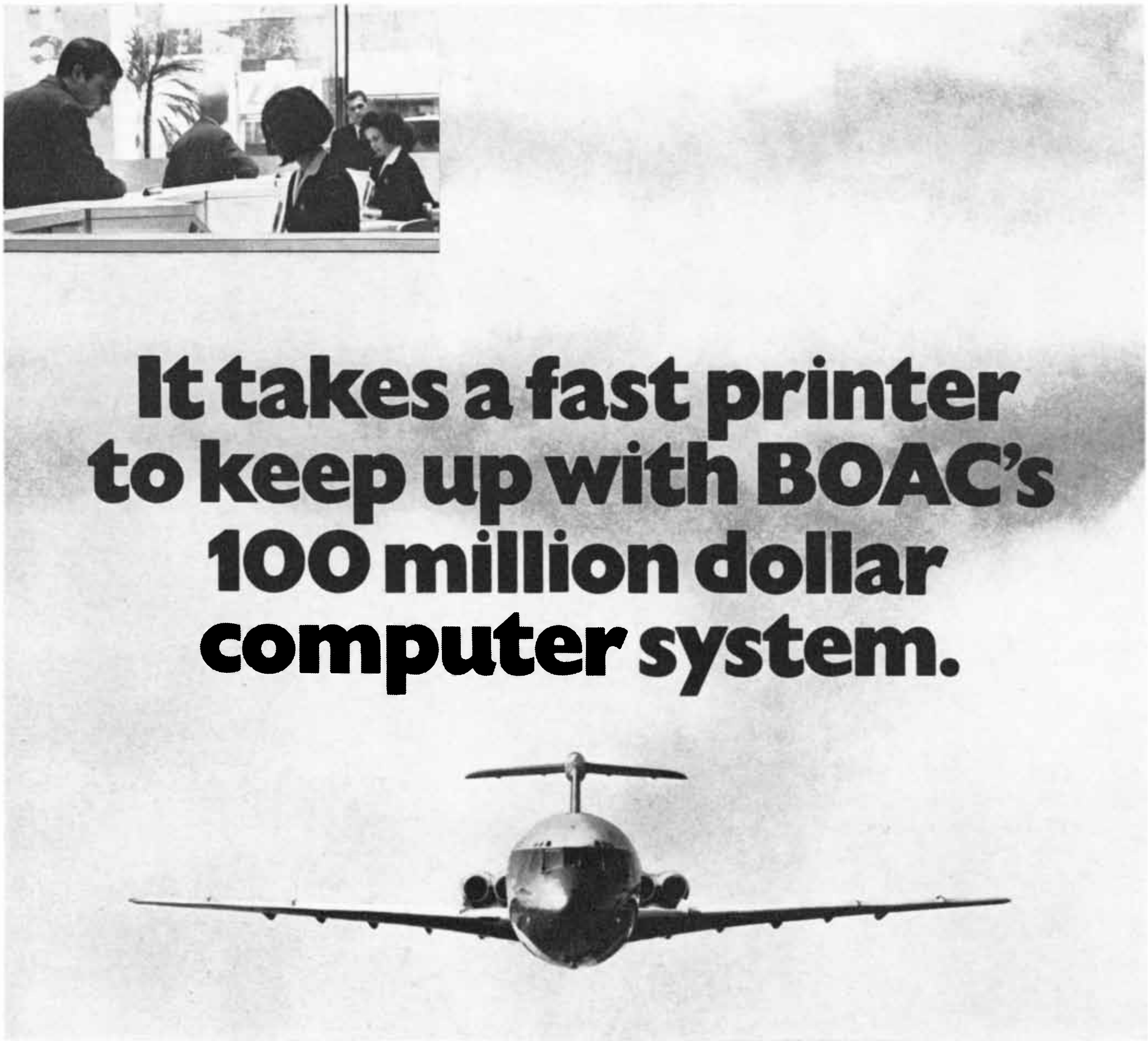
While exploring new methods for cooling high power microwave tubes, Raytheon scientists developed an experimental miniature furnace the size of a two-pound coffee can – so efficient it can heat a nine-room house.

Exploring for oil offshore, our Seismograph Service Corporation subsidiary is operating a new survey ship that employs a high-power, non-explosive acoustic energy source to probe beneath the ocean's floor.

At Raytheon, common threads of technology run through and weave together our diverse and growing businesses . . . in space and defense systems, natural resources, education, home appliances, electronic components, communications, and marine electronics. Write for our annual report. Raytheon Company, Lexington, Mass. 02173.

Raytheon makes things happen.

RAYTHEON



It takes a fast printer to keep up with BOAC's 100 million dollar computer system.

Kleinschmidt.

The amazing Kleinschmidt 311™ Teleprinter is part of BOAC's 100 million dollar computer system called Boadicea. This system puts all the BOAC offices in the United States right next door to the home office in London.

For example, if the BOAC office in Miami needs a passenger list for any BOAC flight, Kleinschmidt gets the answer from London in seconds. At 400 words a minute.

The Kleinschmidt 311 Teleprinter gives BOAC the ultimate in speed and reliability. Furthermore, the 311 Teleprinter can accept parallel information on demand at an average speed of 40 characters a second, which eliminates the formatting problems of serial printers.

What Kleinschmidt is doing for BOAC it can do for any business that requires quick reception of the printed word. If you would like to learn more, please write Kleinschmidt Division of SCM Corporation, Deerfield, Ill. 60015. Or call us at 312-945-1000.

SCM®

TRUTH AND PROOF

The antinomy of the liar, a basic obstacle to an adequate definition of truth in natural languages, reappears in formalized languages as a constructive argument showing not all true sentences can be proved

by Alfred Tarski

The subject of this article is an old one. It has been frequently discussed in modern logical and philosophical literature, and it would not be easy to contribute anything original to the discussion. To many readers, I am afraid, none of the ideas put forward in the article will appear essentially novel; nonetheless, I hope they may find some interest in the way the material has been arranged and knitted together.

As the title indicates, I wish to discuss here two different though related notions: the notion of truth and the notion of proof. Actually the article is divided into three sections. The first section is concerned exclusively with the notion of truth, the second deals primarily with the notion of proof, and the third is a discussion of the relationship between these two notions.

The Notion of Truth

The task of explaining the meaning of the term "true" will be interpreted here in a restricted way. The notion of truth occurs in many different contexts, and there are several distinct categories of objects to which the term "true" is applied. In a psychological discussion one might speak of true emotions as well as true beliefs; in a discourse from the domain of esthetics the inner truth of an object of art might be analyzed. In this article, however, we are interested only in what might be called the logical notion of truth. More specifically, we concern ourselves exclusively with the meaning of the term "true" when this term is used to refer to sentences. Presumably this was the original use of the term "true" in human language. Sentences are treated here as linguistic objects, as certain strings of sounds or written signs. (Of course, not every such string is a sentence.) Moreover, when speaking of sentences, we shall always have in mind what are called in grammar declarative sentences, and not interrogative or imperative sentences.

Whenever one explains the meaning of any term drawn from everyday language, he should bear in mind that the goal and the logical status of such an explanation may vary from one case to another. For instance, the explanation may be intended as an account of the actual use of the term involved, and is thus subject to questioning whether the account is indeed correct. At some other time an explanation may be of a normative nature, that is, it may be offered as a suggestion that the term be used in some definite way, without claiming that the suggestion conforms to the way in which the term is actually used; such an explanation can be evaluated, for instance, from the point of view of its usefulness but not of its correctness. Some further alternatives could also be listed.

The explanation we wish to give in the present case is, to an extent, of mixed character. What will be offered can be treated in principle as a suggestion for a definite way of using the term "true", but the offering will be accompanied by the belief that it is in agreement with the prevailing usage of this term in everyday language.

Our understanding of the notion of truth seems to agree essentially with various explanations of this notion that have been given in philosophical literature. What may be the earliest explanation can be found in Aristotle's *Metaphysics:*

> **To say of what is that it is not, or of what is not that it is, is false, while to say of what is that it is, or of what is not that it is not, is true.**

Here and in the subsequent discussion the word "false" means the same as the expression "not true" and can be replaced by the latter.

The intuitive content of Aristotle's formulation appears to be rather clear. Nevertheless, the formulation leaves much to be desired from the point of view of precision and formal correctness. For one thing, it is not general enough; it refers only to sentences that "say" about something "that it is" or "that it is not"; in most cases it would hardly be possible to cast a sentence in this mold without slanting the sense of the sentence and forcing the spirit of the language. This is perhaps one of the reasons why in modern philosophy various substitutes for the Aristotelian formulation have been offered. As examples we quote the following:

> **A sentence is true if it denotes the existing state of affairs.**
>
> **The truth of a sentence consists in its conformity with (or correspondence to) the reality.**

Due to the use of technical philosophical terms these formulations have undoubtedly a very "scholarly" sound. Nonetheless, it is my feeling that the new formulations, when analyzed more closely, prove to be less clear and unequivocal than the one put forward by Aristotle.

The conception of truth that found its expression in the Aristotelian formulation (and in related formulations of more recent origin) is usually referred to as the *classical,* or *semantic conception of truth.* By semantics we mean the part of logic that, loosely speaking, discusses the relations between linguistic objects (such as sentences) and what is expressed by these objects. The semantic character of the term "true" is clearly revealed by the explanation offered by Aristotle and by some formulations that will be given later in this article. One speaks sometimes of the correspondence theory of truth as the theory based on the classical conception.

(In modern philosophical literature some other conceptions and theories of truth are also discussed, such as the pragmatic conception and the coherence theory. These conceptions seem to be of an exclusively normative character and have little connection with the actual usage of the term "true"; none of them has been formulated so far with any degree of clarity and precision. They will not be discussed in the present article.)

We shall attempt to obtain here a more precise explanation of the classical conception of truth, one that could supersede the Aristotelian formulation while preserving its basic intentions. To this end we shall have to resort to some techniques of contemporary logic. We shall also have to specify the language whose sentences we are concerned with; this is necessary if only for the reason that a string of sounds or signs, which is a true or a false sentence but at any rate a meaningful sentence in one language, may be a meaningless expression in another. For the time being let us assume that the language with which we are concerned is the common English language.

We begin with a simple problem. Consider a sentence in English whose meaning does not raise any doubts, say the sentence "snow is white". For brevity we denote this sentence by "S", so that "S" becomes the name of the sentence. We ask ourselves the question: What do we mean by saying that S is true or that it is false? The answer to this question is simple: in the spirit of Aristotelian explanation, by saying that S is true we mean simply that snow is white, and by saying that S is false we mean that snow is not white. By eliminating the symbol "S" we arrive at the following formulations:

(1) "snow is white" is true if and only if snow is white.
(1′) "snow is white" is false if and only if snow is not white.

Thus (1) and (1′) provide satisfactory explanations of the meaning of the terms "true" and "false" when these terms are referred to the sentence "snow is white". We can regard (1) and (1′) as partial definitions of the terms "true" and "false", in fact, as definitions of these terms with respect to a particular sentence. Notice that (1), as well as (1′), has the form prescribed for definitions by the rules of logic, namely the form of logical equivalence. It consists of two parts, the left and the right side of the equivalence, combined by the connective "if and only if". The left side is the definiendum, the phrase whose meaning is explained by the definition; the right side is the definiens, the phrase that provides the explanation. In the present case the definiendum is the following expression:

"snow is white" is true;

the definiens has the form:

snow is white.

It might seem at first sight that (1), when regarded as a definition, exhibits an essential flaw widely discussed in traditional logic as a vicious circle. The reason is that certain words, for example "snow", occur in both the definiens and the definiendum. Actually, however, these occurrences have an entirely different character. The word "snow" is a syntactical, or organic, part of the definiens; in fact the definiens is a sentence, and the word "snow" is its subject. The definiendum is also a sentence; it expresses the fact that the definiens is a true sentence. Its subject is a name of the definiens formed by putting the definiens in quotes. (When saying something of an object, one always uses a name of this object and not the object itself, even when dealing with linguistic objects.) For several reasons an expression enclosed in quotes must be treated grammatically as a single word having no syntactical parts. Hence the word "snow", which undoubtedly occurs in the definiendum as a part, does not occur there as a syntactical part. A medieval logician would say that "snow" occurs in the definiens *in suppositione formalis* and in the definiendum *in suppositione materialis*. However, words which are not syntactical parts of the definiendum cannot create a vicious circle, and the danger of a vicious circle vanishes.

The preceding remarks touch on some questions which are rather subtle and not quite simple from the logical point of view. Instead of elaborating on them, I shall indicate another manner in which any fears of a vicious circle can be dispelled. In formulating (1) we have applied a common method of forming a name of a sentence, or of any other expression, which consists in putting the expression in quotes. The method has many virtues, but it is also the source of the difficulties discussed above. To remove these difficulties let us try another method of forming names of expressions, in fact a method that can be characterized as a letter-by-letter description of an expression. Using this method we obtain instead of (1) the following lengthy formulation:

(2) The string of three words, the first of which is the string of the letters Es, En, O and Double-U, the second is the string of letters I and Es, and the third is the string of the letters Double-U, Aitch, I, Te, and E, is a true sentence if and only if snow is white.

Formulation (2) does not differ from (1) in its meaning; (1) can simply be regarded as an abbreviated form of (2). The new formulation is certainly much less perspicuous than the old one, but it has the advantage that it creates no appearance of a vicious circle.

Partial definitions of truth analogous to (1) (or (2)) can be constructed for other sentences as well. Each of these definitions has the form:

(3) "*p*" is true if and only if *p*,

where "*p*" is to be replaced on both sides of (3) by the sentence for which the definition is constructed. Special attention should be paid, however, to those situations in which the sentence put in place of "*p*" happens to contain the word "true" as a syntactical part. The corresponding equivalence (3) cannot then be viewed as a partial definition of truth since, when treated as such, it would obviously exhibit a vicious circle. Even in this case, however, (3) is a meaningful sentence, and it is actually a true sentence from the point of view of the classical conception of truth. For illustration, imagine that in a review of a book one finds the following sentence:

(4) Not every sentence in this book is true.

By applying to (4) the Aristotelian criterion, we see that the sentence (4) is true if, in fact, not every sentence in the book concerned is true, and that (4) is false otherwise; in other words, we can assert the equivalence obtained from (3) by taking (4) for "*p*". Of course, this equivalence states merely the conditions under which the sentence (4) is true or is not true, but by itself the equivalence does not enable us to decide which is actually the case. To verify the judgment expressed in (4) one would have to read attentively the book reviewed and analyze the truth of the sentences contained in it.

In the light of the preceding discussion we can now reformulate our main problem. We stipulate that the use of the term "true" in its reference to sentences in English then and only then conforms with the classical conception of truth if it enables us to ascertain every equivalence of the form (3) in which "p" is replaced on both sides by an arbitrary English sentence. If this condition is satisfied, we shall say simply that the use of the term "true" is adequate. Thus our main problem is: can we establish an adequate use of the term "true" for sentences in English and, if so, then by what methods? We can, of course, raise an analogous question for sentences in any other language.

The problem will be solved completely if we manage to construct a general definition of truth that will be adequate in the sense that it will carry with it as logical consequences all the equivalences of form (3). If such a definition is accepted by English-speaking people, it will obviously establish an adequate use of the term "true".

Under certain special assumptions the construction of a general definition of truth is easy. Assume, in fact, that we are interested, not in the whole common English language, but only in a fragment of it, and that we wish to define the term "true" exclusively in reference to sentences of the fragmentary language; we shall refer to this fragmentary language as the language L. Assume further that L is provided with precise syntactical rules which enable us, in each particular case, to distinguish a sentence from an expression which is not a sentence, and that the number of all sentences in the language L is finite (though possibly very large). Assume, finally, that the word "true" does not occur in L and that the meaning of all words in L is sufficiently clear, so that we have no objection to using them in defining truth. Under these assumptions proceed as follows. First, prepare a complete list of all sentences in L; suppose, for example, that there are exactly 1,000 sentences in L, and agree to use the symbols "s_1", "s_2", . . . , "$s_{1,000}$" as abbreviations for consecutive sentences on the list. Next, for each of the sentences "s_1", "s_2", . . . , "$s_{1,000}$" construct a partial definition of truth by substituting successively these sentences for "p" on both sides of the schema (3). Finally, form the logical conjunction of all these partial definitions; in other words, combine them in one statement by putting the connective "and" between any two consecutive partial definitions. The only thing that remains to be done is to give the resulting conjunction a different, but logically equivalent, form, so as to satisfy formal requirements imposed on definitions by rules of logic:

(5) **For every sentence x (in the language L), x is true if and only if either**
s_1, and x is identical to "s_1", or
s_2, and x is identical to "s_2",
. . . .
. . . .
or finally,
$s_{1,000}$, and x is identical to "$s_{1,000}$".

We have thus arrived at a statement which can indeed be accepted as the desired general definition of truth: it is formally correct and is adequate in the sense that it implies all the equivalences of the form (3) in which "p" has been replaced by any sentence of the language L. We notice in passing that (5) is a sentence in English but obviously not in the language L; since (5) contains all sentences in L as proper parts, it cannot coincide with any of them. Further discussion will throw more light on this point.

For obvious reasons the procedure just outlined cannot be followed if we are interested in the whole of the English language and not merely in a fragment of it. When trying to prepare a complete list of English sentences, we meet from the start the difficulty that the rules of English grammar do not determine precisely the form of expressions (strings of words) which should be regarded as sentences: a particular expression, say an exclamation, may function as a sentence in some given context, whereas an expression of the same form will not function so in some other context. Furthermore, the set of all sentences in English is, potentially at least, infinite. Although it is certainly true that only a finite number of sentences have been formulated in speech and writing by human beings up to the present moment, probably nobody would agree that the list of all these sentences comprehends all sentences in English. On the contrary, it seems likely that on seeing such a list each of us could easily produce an English sentence which is not on the list. Finally, the fact that the word "true" does occur in English prevents by itself an application of the procedure previously described.

From these remarks it does not follow that the desired definition of truth for arbitrary sentences in English cannot be obtained in some other way, possibly by using a different idea. There is, however, a more serious and fundamental reason that seems to preclude this possibility. More than that, the mere supposition that an adequate use of the term "true" (in its reference to arbitrary sentences in English) has been secured by any method whatsoever appears to lead to a contradiction. The simplest argument that provides such a contradiction is known as the *antinomy of the liar;* it will be carried through in the next few lines.

Consider the following sentence:

(6) **The sentence printed in red on page 65 of the June 1969 issue of *Scientific American* is false.**

Let us agree to use "s" as an abbreviation for this sentence. Looking at the date of this magazine, and the number of this page, we easily check that "s" is just the only sentence printed in red on page 65 of the June 1969 issue of *Scientific American*. Hence it follows, in particular, that

(7) **"s" is false if and only if the sentence printed in red on page 65 of the June 1969 issue of *Scientific American* is false.**

On the other hand, "s" is undoubtedly a sentence in English. Therefore, assuming that our use of the term "true" is adequate, we can assert the equivalence (3) in which "p" is replaced by "s". Thus we can state:

(8) **"s" is true if and only if s.**

We now recall that "s" stands for the whole sentence (6). Hence we can replace "s" by (6) on the right side of (8); we then obtain

(9) **"s" is true if and only if the sentence printed in red on page 65 of the June 1969 issue of *Scientific American* is false.**

By now comparing (8) and (9), we conclude:

(10) **"s" is false if and only if "s" is true.**

This leads to an obvious contradiction: "s" proves to be both true and false. Thus we are confronted with an antinomy. The above formulation of the antinomy of the liar is due to the Polish logician Jan Łukasiewicz.

Some more involved formulations of this antinomy are also known. Imagine, for instance, a book of 100 pages, with just one sentence printed on each page.

On page 1 we read:

> **The sentence printed on page 2 of this book is true.**

On page 2 we read:

> **The sentence printed on page 3 of this book is true.**

And so it goes on up to page 99. However, on page 100, the last page of the book, we find:

> **The sentence printed on page 1 of this book is false.**

Assume that the sentence printed on page 1 is indeed false. By means of an argument which is not difficult but is very long and requires leafing through the entire book, we conclude that our assumption is wrong. Consequently we assume now that the sentence printed on page 1 is true—and, by an argument which is as easy and as long as the original one, we convince ourselves that the new assumption is wrong as well. Thus we are again confronted with an antinomy.

It turns out to be an easy matter to compose many other "antinomial books" that are variants of the one just described. Each of them has 100 pages. Every page contains just one sentence, and in fact a sentence of the form:

> **The sentence printed on page *00* of this book is *XX*.**

In each particular case "*XX*" is replaced by one of the words "*true*" or "*false*", while "*00*" is replaced by one of the numerals "1", "2", . . . , "100"; the same numeral may occur on many pages. Not every variant of the original book composed according to these rules actually yields an antinomy. The reader who is fond of logical puzzles will hardly find it difficult to describe all those variants that do the job. The following warning may prove useful in this connection. Imagine that somewhere in the book, say on page 1, it is said that the sentence on page 3 is true, while somewhere else, say on page 2, it is claimed that the same sentence is false. From this information it does not follow at all that our book is "antinomial"; we can only draw the conclusion that either the sentence on page 1 or the sentence on page 2 must be false. An antinomy does arise, however, whenever we are able to show that one of the sentences in the book is both true and false, independent of any assumptions concerning the truth or falsity of the remaining sentences.

The antinomy of the liar is of very old origin. It is usually ascribed to the Greek logician Eubulides; it tormented many ancient logicians and caused the premature death of at least one of them, Philetas of Cos. A number of other antinomies and paradoxes were found in antiquity, in the Middle Ages, and in modern times. Although many of them are now entirely forgotten, the antinomy of the liar is still analyzed and discussed in contemporary writings. Together with some recent antinomies discovered around the turn of the century (in particular, the antinomy of Russell), it has had a great impact on the development of modern logic.

Two diametrically opposed approaches to antinomies can be found in the literature of the subject. One approach is to disregard them, to treat them as sophistries, as jokes that are not serious but malicious, and that aim mainly at showing the cleverness of the man who formulates them. The opposite approach is characteristic of certain thinkers of the 19th century and is still represented, or was so a short while ago, in certain parts of our globe. According to this approach antinomies constitute a very essential element of human thought; they must appear again and again in intellectual activities, and their presence is the basic source of real progress. As often happens, the truth is probably somewhere in between. Personally, as a logician, I could not reconcile myself with antinomies as a permanent element of our system of knowledge. However, I am not the least inclined to treat antinomies lightly. The appearance of an antinomy is for me a symptom of disease. Starting with premises that seem intuitively obvious, using forms of reasoning that seem intuitively certain, an antinomy leads us to nonsense, a contradiction. Whenever this happens, we have to submit our ways of thinking to a thorough revision, to reject some premises in which we believed or to improve some forms of argument which we used. We do this with the hope not only that the old antinomy will be disposed of but also that no new one will appear. To this end we test our reformed system of thinking by all available means, and, first of all, we try to reconstruct the old antinomy in the new setting; this testing is a very important activity in the realm of speculative thought, akin to carrying out crucial experiments in empirical science.

From this point of view consider now specifically the antinomy of the liar. The antinomy involves the notion of truth in reference to arbitrary sentences of common English; it could easily be reformulated so as to apply to other natural languages. We are confronted with a serious problem: how can we avoid the contradictions induced by this antinomy? A radical solution of the problem which may readily occur to us would be simply to remove the word "true" from the English vocabulary or at least to abstain from using it in any serious discussion.

Those people to whom such an amputation of English seems highly unsatisfactory and illegitimate may be inclined to accept a somewhat more compromising solution, which consists in adopting what could be called (following the contemporary Polish philosopher Tadeusz Kotarbiński) "the nihilistic approach to the theory of truth". According to this approach, the word "true" has no independent meaning but can be used as a component of the two meaningful expressions "it is true that" and "it is not true that". These expressions are thus treated as if they were single words with no organic parts. The meaning ascribed to them is such that they can be immediately eliminated from any sentence in which they occur. For instance, instead of saying

> **it is true that all cats are black**

we can simply say

> **all cats are black,**

and instead of

> **it is not true that all cats are black**

we can say

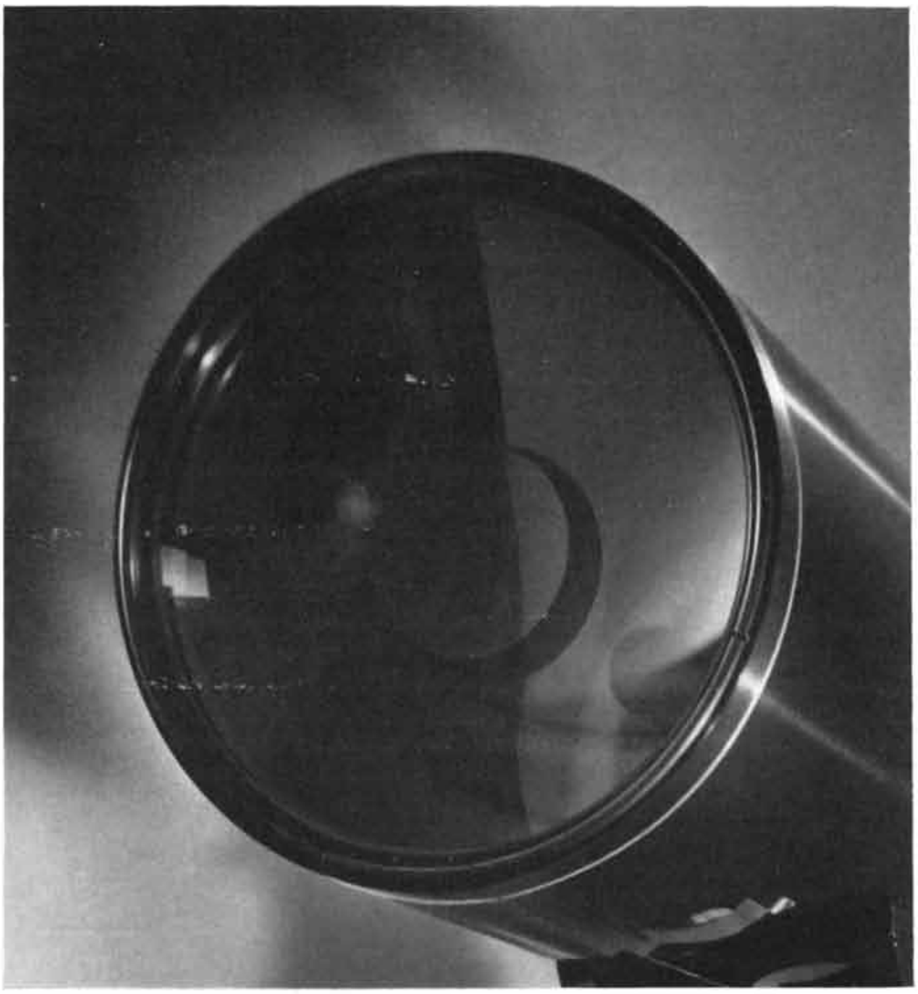
> **not all cats are black.**

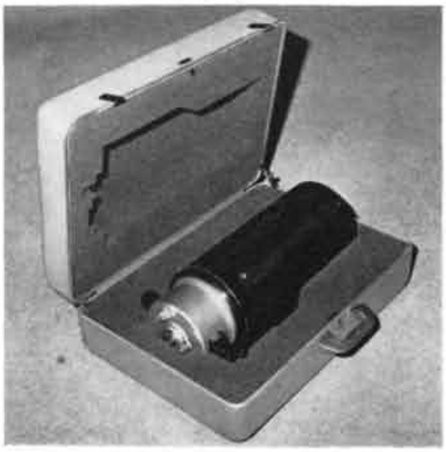
In other contexts the word "true" is meaningless. In particular, it cannot be used as a real predicate qualifying names of sentences. Employing the terminology of medieval logic, we can say that the word "true" can be used syncategorematically in some special situations, but it cannot ever be used categorematically.

To realize the implications of this approach, consider the sentence which was the starting point for the antinomy of the liar; that is, the sentence printed in red on page 65 in this magazine. From the "nihilistic" point of view it is not a meaningful sentence, and the antinomy simply vanishes. Unfortunately, many uses of the word "true", which otherwise seem quite legitimate and reasonable, are similarly affected by this approach. Imagine, for instance, that a certain term occurring repeatedly in the works of an ancient mathematician admits of several interpretations. A historian of science who studies the works arrives at the conclusion that under one of these interpretations all the theorems stated by the mathematician prove to be true; this leads him naturally to the conjecture that the same will apply to any work of this mathematician that is not known at present but may be discovered in the future. If, however, the historian of science shares the "nihilistic" approach to the notion of truth, he lacks the possibility of expressing his conjecture in words. One could say that truth-theoretical "nihilism" pays lip service to some popular forms of human speech, while actually removing the notion of truth from the conceptual stock of the human mind.

We shall look, therefore, for another way out of our predicament. We shall try to find a solution that will keep the classical concept of truth essentially intact. The applicability of the notion of truth will have to undergo some restrictions, but the notion will remain available at least for the purpose of scholarly discourse.

To this end we have to analyze those features of the common language that are the real source of the antinomy of the liar. When carrying through this analysis, we notice at once an outstanding feature of this language—its all-comprehensive, universal character. The common language is universal and is intended to be so. It is supposed to provide adequate facilities for expressing everything that can be expressed at all, in any language whatsoever; it is continually expanding to satisfy this requirement. In particular, it is semantically universal in the following sense. Together with the linguistic objects, such as sentences and terms, which are components of this language, names of these objects are also included in the language (as we know, names of expressions can be obtained by putting the expressions in quotes); in addition, the language contains semantic terms such as "truth", "name", "designation", which directly or indirectly refer to the relationship between linguistic objects and what is expressed by them. Consequently, for every sentence formulated in the common language, we can form in the same language another sentence to the effect that the first sentence is true or that it is false. Using an additional "trick" we can even construct in the language what is sometimes called a self-referential sentence, that is, a sentence S which asserts the fact that S itself is true or that it is false. In case S asserts its own falsity we





can show by means of a simple argument that S is both true and false—and we are confronted again with the antinomy of the liar.

There is, however, no need to use universal languages in all possible situations. In particular, such languages are in general not needed for the purposes of science (and by science I mean here the whole realm of intellectual inquiry). In a particular branch of science, say in chemistry, one discusses certain special objects, such as elements, molecules, and so on, but not for instance linguistic objects such as sentences or terms. The language that is well adapted to this discussion is a restricted language with a limited vocabulary; it must contain names of chemical objects, terms such as "element" and "molecule", but not names of linguistic objects; hence it does not have to be semantically universal. The same applies to most of the other branches of science. The situation becomes somewhat confused when we turn to linguistics. This is a science in which we study languages; thus the language of linguistics must certainly be provided with names of linguistic objects. However, we do not have to identify the language of linguistics with the universal language or any of the languages that are objects of linguistic discussion, and we are not bound to assume that we use in linguistics one and the same language for all discussions. The language of linguistics has to contain the names of linguistic components of the languages discussed but not the names of its own components; thus, again, it does not have to be semantically universal. The same applies to the language of logic, or rather of that part of logic known as metalogic and metamathematics; here we again concern ourselves with certain languages, primarily with languages of logical and mathematical theories (although we discuss these languages from a different point of view than in the case of linguistics).

The question now arises whether the notion of truth can be precisely defined, and thus a consistent and adequate usage of this notion can be established at least for the semantically restricted languages of scientific discourse. Under certain conditions the answer to this question proves to be affirmative. The main conditions imposed on the language are that its full vocabulary should be available and its syntactical rules concerning the formation of sentences and other meaningful expressions from words listed in the vocabulary should be precisely formulated. Furthermore, the syntactical rules should be purely formal, that is, they should refer exclusively to the form (the shape) of expressions; the function and the meaning of an expression should depend exclusively on its form. In particular, looking at an expression, one should be able in each case to decide whether or not the expression is a sentence. It should never happen that an expression functions as a sentence at one place while an expression of the same form does not function so at some other place, or that a sentence can be asserted in one context while a sentence of the same form can be denied in another. (Hence it follows, in particular, that demonstrative pronouns and adverbs such as "this" and "here" should not occur in the vocabulary of the language.) Languages that satisfy these conditions are referred to as formalized languages. When discussing a formalized language there is no need to distinguish between expressions of the same form which have been written or uttered in different places; one often speaks of them as if they were one and the same expression. The reader may have noticed we sometimes use this way of speaking even when discussing a natural language, that is, one which is not formalized; we do so for the sake of simplicity, and only in those cases in which there seems to be no danger of confusion.

Formalized languages are fully adequate for the presentation of logical and mathematical theories; I see no essential reasons why they cannot be adapted for use in other scientific disciplines and in particular to the development of theoretical parts of empirical sciences. I should like to emphasize that, when using the term "formalized languages", I do not refer exclusively to linguistic systems that are formulated entirely in symbols, and I do not have in mind anything essentially opposed to natural languages. On the contrary, the only formalized languages that seem to be of real interest are those which are fragments of natural languages (fragments provided with complete vocabularies and precise syntactical rules) or those which can at least be adequately translated into natural languages.

There are some further conditions on which the realization of our program depends. We should make a strict distinction between the language which is the object of our discussion and for which in particular we intend to construct the definition of truth, and the language in which the definition is to be formulated and its implications are to be studied. The latter is referred to as the metalanguage and the former as the object-language. The metalanguage must be sufficiently rich; in particular, it must include the object-language as a part. In fact, according to our stipulations, an adequate definition of truth will imply as consequences all partial definitions of this notion, that is, all equivalences of form (3):

"p" is true if and only if p,

where "p" is to be replaced (on both sides of the equivalence) by an arbitrary sentence of the object-language. Since all these consequences are formulated in the metalanguage, we conclude that every sentence of the object-language must also be a sentence of the metalanguage. Furthermore, the metalanguage must contain names for sentences (and other expressions) of the object-language, since these names occur on the left sides of the above equivalences. It must also contain some further terms that are needed for the discussion of the object-language, in fact terms denoting certain special sets of expressions, relations between expressions, and operations on expressions; for instance, we must be able to speak of the set of all sentences or of the operation of juxtaposition, by means of which, putting one of two given expressions immediately after the other, we form a new expression. Finally, by defining truth, we show that semantic terms (expressing relations between sentences of the object-language and objects referred to by these sentences) can be introduced in the metalanguage by means of definitions. Hence we conclude that the metalanguage which provides sufficient means for defining truth must be essentially richer than the object-language; it cannot coincide with or be translatable into the latter, since otherwise both languages would turn out to be semantically universal, and the antinomy of the liar could be reconstructed in both of them. We shall return to this question in the last section of this article.

If all the above conditions are satisfied, the construction of the desired definition of truth presents no essential difficulties. Technically, however, it is too involved to be explained here in detail. For any given sentence of the object-language one can easily formulate the corresponding partial definition of form (3). Since, however, the set of all sentences in the object-language is as a rule infinite, whereas every sentence of the metalanguage is a finite string of signs, we cannot arrive at a general definition simply by forming the logical conjunction of all partial definitions. Nevertheless, what we eventually obtain is in some intuitive sense equivalent to the imaginary infinite conjunction. Very roughly speaking, we proceed as follows. First, we consider the simplest sentences, which do not include any other sentences as parts; for these simplest sentences we manage to define truth directly (using the same idea that leads to partial definitions). Then, making use of syntactical rules which concern the formation of more complicated sentences from simpler ones, we extend the definition to arbitrary compound sentences; we apply here the method known in mathematics as definition by recursion. (This is merely a rough approximation of the actual procedure. For some technical reasons the method of recursion is actually applied to define, not the notion of truth, but the related semantic notion of satisfaction. Truth is then easily defined in terms of satisfaction.)

On the basis of the definition thus constructed we can develop the entire theory of truth. In particular, we can derive from it, in addition to all equivalences of form (3), some consequences of a general nature, such as the famous laws of contradiction and of excluded middle. By the first of these laws, no two sentences one of which is the negation of the other can both be true; by the second law, no two such sentences can both be false.

The Notion of Proof

Whatever may be achieved by constructing an adequate definition of truth for a scientific language, one fact seems to be certain: the definition does not carry with it a workable criterion for deciding whether particular sentences in this language are true or false (and indeed it is not designed at all for this purpose). Consider, for example, a sentence in the language of elementary high school geometry, say "the three bisectors of every triangle meet in one point". If we are interested in the question whether this sentence is true and we turn to the definition of truth for an answer, we are in for a disappointment. The only bit of information we get is that the sentence is true if the three bisectors of a triangle always meet in one point, and is false if they do not always meet; but only a geometrical inquiry may enable us to decide which is actually the case. Analogous remarks apply to sentences from the domain of any other particular science: to decide whether or not any such sentence is true is a task of the science itself, and not of logic or the theory of truth.

Some philosophers and methodologists of science are inclined to reject every definition that does not provide a criterion for deciding whether any given particular object falls under the notion defined or not. In the methodology of empirical sciences such a tendency is represented by the doctrine of operationalism; philosophers of mathematics who belong to the constructivist school seem to exhibit a similar tendency. In both cases, however, the people who hold this opinion appear to be in a small minority. A consistent attempt to carry out the program in practice (that is, to develop a science without using undesirable definitions) has hardly ever been made. It seems clear that under this program much of contemporary mathematics would disappear, and theoretical parts of physics, chemistry, biology, and other empirical sciences would be severely mutilated. The definitions of such notions as atom or gene as well as most definitions in mathematics do not carry with them any criteria for deciding whether or not an object falls under the term that has been defined.

Since the definition of truth does not provide us with any such criterion and at the same time the search for truth is rightly considered the essence of scientific activities, it appears as an important problem to find at least partial criteria of truth and to develop procedures that may enable us to ascertain or negate the truth (or at least the likelihood of truth) of as many sentences as possible. Such procedures are known indeed; some of them are used exclusively in empirical science and some primarily in deductive science. The notion of proof—the second notion to be discussed in this paper—refers just to a procedure of ascertaining the truth of sentences which is employed primarily in deductive science. This procedure is an essential element of what is known as the axiomatic method, the only method now used to develop mathematical disciplines.

The axiomatic method and the notion of proof within its framework are products of a long historical development. Some rough knowledge of this development is probably essential for the understanding of the contemporary notion of proof.

Originally a mathematical discipline was an aggregate of sentences that concerned a certain class of objects or phenomena, were formulated by means of a certain stock of terms, and were accepted as true. This aggregate of sentences lacked any structural order. A sentence was accepted as true either because it seemed intuitively evident, or else because it was proved on the basis of some intuitively evident sentences, and thus was shown, by means of an intuitively certain argument, to be a consequence of these other sentences. The criterion of intuitive evidence (and intuitive certainty of arguments) was applied without any restrictions; every sentence recognized as true by means of this criterion was automatically included in the discipline. This description seems to fit, for instance, the science of geometry as it was known to ancient Egyptians and Greeks in its early, pre-Euclidean stage.

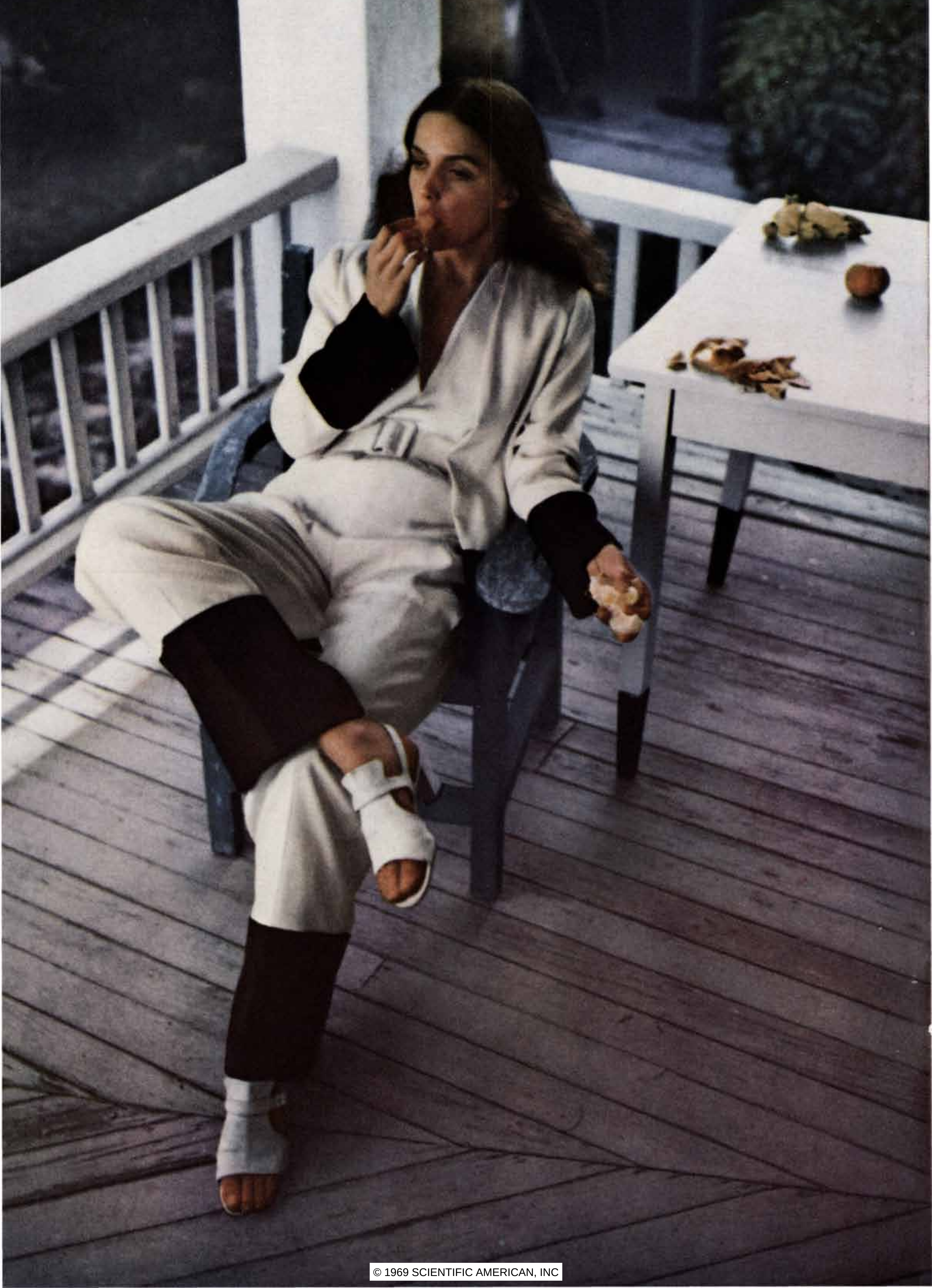
It was realized rather soon, however, that the criterion of intuitive evidence is far from being infallible, has no objective character, and often leads to serious errors. The entire subsequent development of the axiomatic method can be viewed as an expression of the tendency to restrict the recourse to intuitive evidence.

This tendency first revealed itself in the effort to prove as many sentences as possible, and hence to restrict as much as possible the number of sentences accepted as true merely on the basis of intuitive evidence. The ideal from this point of view would be to prove every sentence that is to be accepted as true. For obvious reasons this ideal cannot be realized. Indeed, we prove each sentence on the basis of other sentences, we prove these other sentences on the basis of some further sentences, and so on: if we are to avoid both a vicious circle and an infinite regress, the procedure must be discontinued somewhere. As a compromise between that unattainable ideal and the realizable possibilities, two principles emerged and were subsequently applied in constructing mathematical disciplines. By the first of these principles every discipline begins with a list of a small number of sentences, called axioms or primitive sentences, which seem to be intuitively evident and which are recognized as true without any further justification. According to the second principle, no other sentence is accepted in the discipline as true unless we are able to prove it with the exclusive help of axioms and those sentences that were previously proved. All the sentences that can be recognized as true by virtue of these two principles are called theorems, or provable sentences, of the given discipline. Two analogous principles concern the use of terms in constructing the discipline. By the first of them we list at the beginning a few terms, called undefined or primitive terms, which appear to be directly understandable and which we decide to use (in formulating and proving theorems) without explaining their meanings; by the second principle we agree not to use any further term unless we are able to explain its meaning by defining it with the help of undefined terms and terms previously defined. These four principles are cornerstones of the axiomatic method; theories developed in accordance with these principles are called axiomatic theories.

As is well known, the axiomatic method was applied to the development of geometry in the *Elements* of Euclid about 300 B.C. Thereafter it was used for over 2,000 years with practically no change in its main principles (which, by the way, were not even explicitly formulated for a long period of time) nor in the general approach to the subject. However, in the 19th and 20th centuries the concept of the axiomatic method did undergo a profound evolution. Those features of the evolution which concern the notion of proof are particularly significant for our discussion.

Until the last years of the 19th century the notion of proof was primarily of a psychological character. A proof was an intellectual activity that aimed at convincing oneself and others of the truth of a sentence discussed; more specifically, in developing a mathematical theory proofs were used to convince ourselves and others that a sentence discussed had to be accepted as true once some other sentences had been previously accepted as such. No restrictions were put on arguments used in proofs, except that they had to be intuitively convincing. At a certain period, however, a need began to be felt for submitting the notion of proof to a deeper analysis that would result in restricting the recourse to intuitive evidence in this context as well. This was probably related to some specific developments in mathematics, in particular to the discovery of non-Euclidean geometries. The analysis was carried out by logicians, beginning with the German logician Gottlob Frege; it led to the introduction of a new notion, that of a *formal proof,* which turned out to be an adequate substitute and an essential improvement over the old psychological notion.

The first step toward supplying a mathematical theory with the notion of



Sokolsky on Sokolsky on Polaroid Land Film

Mel Sokolsky, a major photographic talent and innovator, did the cover and feature photos for 3 issues of *The New York Sunday Times* women's fashion supplement. The most recent one was all in Polacolor film.

"I'm really interested in communications as a whole. All of it. In all forms. I'm a talker, as you see.

Photography is a means of communication, so I'm in it.

But I try new approaches so I don't get bored. That's one of the reasons I switched to Polaroid color film. And it works beautifully.

In commercial photography, the usual methods involve a loss of time that's really from the Stone Age. Yet everyone wants their stuff in a hurry. With Polaroid Land film, I can give it to them.

But to me, the big advantages are color quality, and what my own idea of realism is.

Most modern film emulsions are thin. The results are hard, sharp, slick.

But our eyes see things more softly, with more 3-dimensionality. The way things look in old-time photos with their thick emulsions.

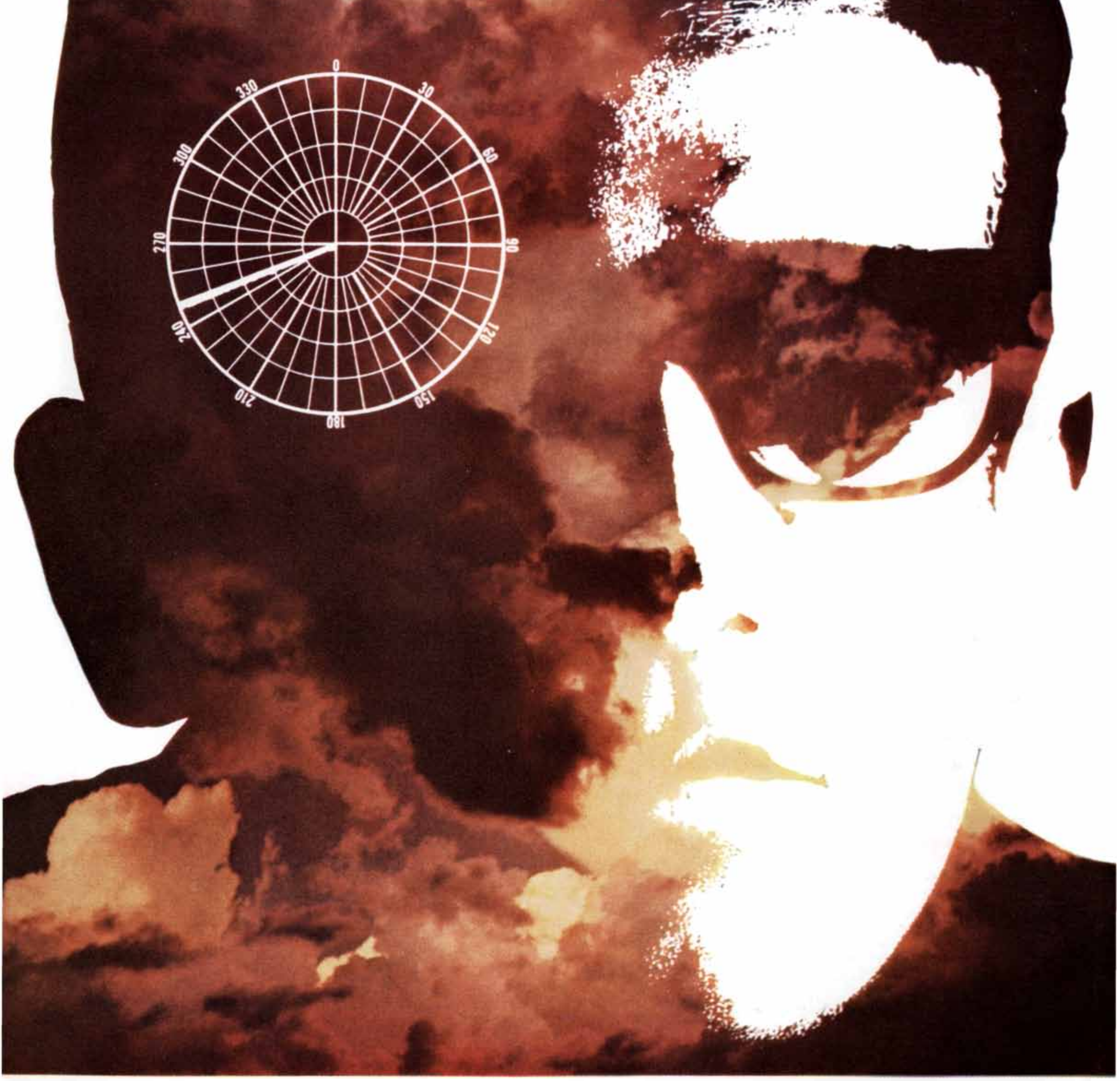
The nostalgia we feel when we look at them isn't just because of the subject matter, but because it's the way we really see things.

That's what's so good about Polaroid color film. It has a palette of its own, that gives me the quality I'm looking for. It's warmer, more human. And to me, more real.

Of course I try to get realism through candidness and naturalness too, as you can see in this picture which I did on the *Times* assignment. I don't go for the excessively posed and arranged fashion shots.

And with Polaroid film, I can see what I've just taken. If I don't get what I want the first time, I simply try again till I do get it.

I'd like to make a prediction. Before long, the engraving of Polaroid shots for commercial production will spread like wildfire. Because more and more professionals are seeing the advantages."



Extending man's senses:

Command joint forces from 20,000 feet.

Bounce a newscast around the world.

Navigate precisely from Cape Horn to the Cape of Good Hope.

Turn night into day with an airborne floodlight.

Set up a TACAN transceiver, anywhere, in minutes.

Airborne command and control stations. Communications satellite networks. The global Omega navigation system. Nighttime military and civilian operations. Portable TACAN.

All demand fail-safe technology—the finest electronics and systems available.

That's our business.

LTV Electrosystems has the scientific and engineering talent, fast-reaction capacity and the production facilities (15 nationwide) to build the sophisticated, new-generation systems our customers need to extend their senses and capabilities into every environment.

Why don't you join us?

We welcome scientists and engineers for careers in research, design, development or management. Contact: Bill Hickey, Supervisor of Professional Placement, P.O. Box 6118, Dallas, Texas 75222. Tel. (214) 276-7111. An equal opportunity employer.

LTV ELECTROSYSTEMS, INC.
PO BOX 6030, DALLAS, TEXAS 75222
A quality company of Ling-Temco-Vought, Inc. **LTV**

a formal proof is the formalization of the language of the theory, in the sense discussed previously in connection with the definition of truth. Thus formal syntactical rules are provided which in particular enable us simply by looking at shapes of expressions, to distinguish a sentence from an expression that is not a sentence. The next step consists in formulating a few rules of a different nature, the so-called rules of proof (or of inference). By these rules a sentence is regarded as directly derivable from given sentences if, generally speaking, its shape is related in a prescribed manner to the shapes of given sentences. The number of rules of proof is small, and their content is simple. Just like the syntactical rules, they all have a formal character, that is, they refer exclusively to shapes of sentences involved. Intuitively all the rules of derivation appear to be infallible, in the sense that a sentence which is directly derivable from true sentences by any of these rules must be true itself. Actually the infallibility of the rules of proof can be established on the basis of an adequate definition of truth. The best-known and most important example of a rule of proof is the rule of detachment known also as *modus ponens*. By this rule (which in some theories serves as the only rule of proof) a sentence "*q*" is directly derivable from two given sentences if one of them is the conditional sentence "if *p*, then *q*" while the other is "*p*"; here "*p*" and "*q*" are, as usual, abbreviations of any two sentences of our formalized language. We can now explain in what a formal proof of a given sentence consists. First, we apply the rules of proof to axioms and obtain new sentences that are directly derivable from axioms; next, we apply the same rules to new sentences, or jointly to new sentences and axioms, and obtain further sentences; and we continue this process. If after a finite number of steps we arrive at a given sentence, we say that the sentence has been formally proved. This can also be expressed more precisely in the following way: a formal proof of a given sentence consists in constructing a finite sequence of sentences such that (1) the first sentence in the sequence is an axiom, (2) each of the following sentences either is an axiom or is directly derivable from some of the sentences that precede it in the sequence, by virtue of one of the rules of proof, and (3) the last sentence in the sequence is the sentence to be proved. Changing somewhat the use of the term "proof", we can even say that a formal proof of a sentence is simply any finite sequence of sentences with the three properties just listed.

An axiomatic theory whose language has been formalized and for which the notion of a formal proof has been supplied is called a formalized theory. We stipulate that the only proofs which can be used in a formalized theory are formal proofs; no sentence can be accepted as a theorem unless it appears on the list of axioms or a formal proof can be found for it. The method of presenting a formalized theory at each stage of its development is in principle very elementary. We list first the axioms and then all the known theorems in such an order that every sentence on the list which is not an axiom can be directly recognized as a theorem, simply by comparing its shape with the shapes of sentences that precede it on the list; no complex processes of reasoning and convincing are involved. (I am not speaking here of psychological processes by means of which the theorems have actually been discovered.) The recourse to intuitive evidence has been indeed considerably restricted; doubts concerning the truth of theorems have not been entirely eliminated but have been reduced to possible doubts concerning the truth of the few sentences listed as axioms and the infallibility of the few simple rules of proof. It may be added that the process of introducing new terms in the language of a theory can also be formalized by supplying special formal rules of definitions.

It is now known that all the existing mathematical disciplines can be presented as formalized theories. Formal proofs can be provided for the deepest and most complicated mathematical theorems, which were originally established by intuitive arguments.

The Relationship of Truth and Proof

It was undoubtedly a great achievement of modern logic to have replaced the old psychological notion of proof, which could hardly ever be made clear and precise, by a new simple notion of a purely formal character. But the triumph of the formal method carried with it the germ of a future setback. As we shall see, the very simplicity of the new notion turned out to be its Achilles heel.

To assess the notion of formal proof we have to clarify its relation to the notion of truth. After all, the formal proof, just like the old intuitive proof, is a procedure aimed at acquiring new true sentences. Such a procedure will be adequate only if all sentences acquired with its help prove to be true and all true sentences can be acquired with its help. Hence the problem naturally arises: is the formal proof actually an adequate procedure for acquiring truth? In other words: does the set of all (formally) provable sentences coincide with the set of all true sentences?

To be specific, we refer this problem to a particular, very elementary mathematical discipline, namely to the arithmetic of natural numbers (the elementary number theory). We assume that this discipline has been presented as a formalized theory. The vocabulary of the theory is meager. It consists, in fact, of variables such as "*m*", "*n*", "*p*", ... representing arbitrary natural numbers; of numerals "0", "1", "2", ... denoting particular numbers; of symbols denoting some familiar relations between numbers and operations on numbers such as "=", "<", "+", "−"; and, finally, of certain logical terms, namely sentential connectives ("and", "or", "if", "not") and quantifiers (expressions of the form "for every number *m*" and "for some number *m*"). The syntactical rules and the rules of proof are simple. When speaking of sentences in the subsequent discussion, we always have in mind sentences of the formalized language of arithmetic.

We know from the discussion of truth in the first section that, taking this language as the object-language, we can construct an appropriate metalanguage and formulate in it an adequate definition of truth. It proves convenient in this context to say that what we have thus defined is the set of true sentences; in fact, the definition of truth states that a certain condition formulated in the metalanguage is satisfied by all elements of this set (that is, all true sentences) and only by these elements. Even more readily we can define in the metalanguage the set of provable sentences; the definition conforms entirely with the explanation of the notion of formal proof that was given in the second section. Strictly speaking, the definitions of both truth and provability belong to a new theory formulated in the metalanguage and specifically designed for the study of our formalized arithmetic and its language. The new theory is called the metatheory or, more specifically, the meta-arithmetic. We shall not elaborate here on the way in which the metatheory is constructed—on its axioms, undefined terms, and so on. We only point out that it is within the framework of this metatheory that we formulate and solve the problem of whether the set of provable sentences coincides with that of true sentences.

The solution of the problem proves to be negative. We shall give here a very rough account of the method by which the solution has been reached. The main idea is closely related to the one used by the contemporary American logician (of Austrian origin) Kurt Gödel in his famous paper on the incompleteness of arithmetic.

It was pointed out in the first section that the metalanguage which enables us to define and discuss the notion of truth must be rich. It contains the entire object-language as a part, and therefore we can speak in it of natural numbers, sets of numbers, relations among numbers, and so forth. But it also contains terms needed for the discussion of the object-language and its components; consequently we can speak in the metalanguage of expressions and in particular of sentences, of sets of sentences, of relations among sentences, and so forth. Hence in the metatheory we can study properties of these various kinds of objects and establish connections between them.

In particular, using the description of sentences provided by the syntactical rules of the object-language, it is easy to arrange all sentences (from the simplest ones through the more and more complex) in an infinite sequence and to number them consecutively. We thus correlate with every sentence a natural number in such a way that two numbers correlated with two different sentences are always different; in other words, we establish a one-to-one correspondence between sentences and numbers. This in turn leads to a similar correspondence between sets of sentences and sets of numbers, or relations among sentences and relations among numbers. In particular, we can consider numbers of provable sentences and numbers of true sentences; we call them briefly provable* numbers and true* numbers. Our main problem is reduced then to the question: are the set of provable* numbers and the set of true* numbers identical?

To answer this question negatively, it suffices, of course, to indicate a single property that applies to one set but not to the other. The property we shall actually exhibit may seem rather unexpected, a kind of *deus ex machina*.

The intrinsic simplicity of the notions of formal proof and formal provability will play a basic role here. We have seen in the second section that the meaning of these notions is explained essentially in terms of certain simple relations among sentences prescribed by a few rules of proof; the reader may recall here the rule of *modus ponens*. The corresponding relations among numbers of sentences are equally simple; it turns out that they can be characterized in terms of the simplest arithmetical operations and relations, such as addition, multiplication, and equality—thus in terms occurring in our arithmetical theory. As a consequence the set of provable* numbers can also be characterized in such terms. One can describe briefly what has been achieved by saying that the definition of provability has been translated from the metalanguage into the object-language.

On the other hand, the discussion of the notion of truth in common languages strongly suggests the conjecture that no such translation can be obtained for the definition of truth; otherwise the object-language would prove to be in a sense semantically universal, and a reappearance of the antinomy of the liar would be imminent. We confirm this conjecture by showing that, if the set of true* numbers could be defined in the language of arithmetic, the antinomy of the liar could actually be reconstructed in this language. Since, however, we are dealing now with a restricted formalized language, the antinomy would assume a more involved and sophisticated form. In particular, no expressions with an empirical content such as "the sentence printed in such-and-such place", which played an essential part in the original formulation of the antinomy, would occur in the new formulation. We shall not go into any further details here.

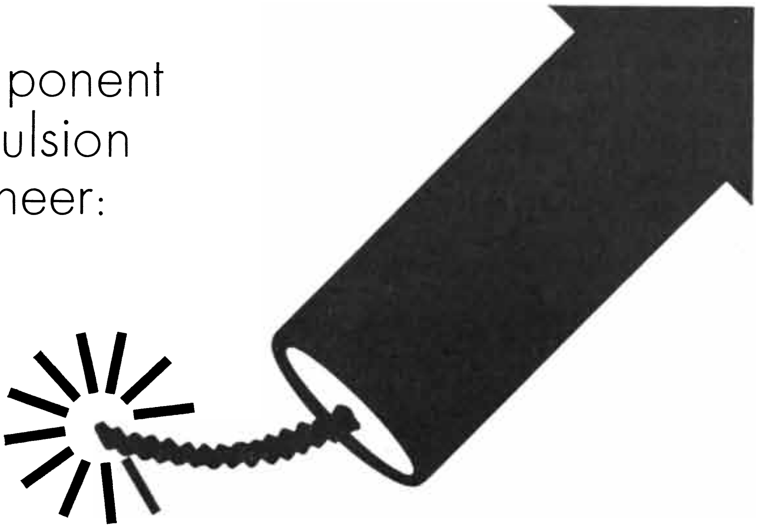
Thus the set of provable* numbers does not coincide with the set of true* numbers, since the former is definable in the language of arithmetic while the latter is not. Consequently the sets of provable sentences and true sentences do not coincide either. On the other hand, using the definition of truth we easily show that all the axioms of arithmetic are true and all the rules of proof are infallible. Hence all the provable sentences are true; therefore the converse cannot hold. Thus our final conclusion is: there are sentences formulated in the language of arithmetic that are true but cannot be proved on the basis of the axioms and rules of proof accepted in arithmetic.

One might think that the conclusion essentially depends on specific axioms and rules of inference, chosen for our arithmetical theory, and that the final outcome of the discussion could be different if we appropriately enriched the theory by adjoining new axioms or new rules of inference. A closer analysis shows, however, that the argument depends very little on specific properties of the theory discussed, and that it actually extends to most other formalized theories. Assuming that a theory includes the arithmetic of natural numbers as a part (or that, at least, arithmetic can be reconstructed in it), we can repeat the essential portion of our argument in a practically unchanged form; we thus conclude again that the set of provable sentences of the theory is different from the set of its true sentences. If, moreover, we can show (as is frequently the case) that all the axioms of the theory are true and all the rules of inference are infallible, we further conclude that there are true sentences of the theory which are not provable. Apart from some fragmentary theories with restricted means of expression, the assumption concerning the relation of the theory to the arithmetic of natural numbers is generally satisfied, and hence our conclusions have a nearly universal character. (Regarding those fragmentary theories which do not include the arithmetic of natural numbers, their languages may not be provided with sufficient means for defining the notion of provability, and their provable sentences may in fact coincide with their true sentences. Elementary geometry and elementary algebra of real numbers are the best known, and perhaps most important, examples of theories in which these notions coincide.)

The dominant part played in the whole argument by the antinomy of the liar throws some interesting light on our earlier remarks concerning the role of antinomies in the history of human thought. The antinomy of the liar first appeared in our discussion as a kind of evil force with a great destructive power. It compelled us to abandon all attempts at clarifying the notion of truth for natural languages. We had to restrict our endeavors to formalized languages of scientific discourse. As a safeguard against a possible reappearance of the antinomy, we had to complicate considerably the discussion by distinguishing between a language and its metalanguage. Subsequently, however, in the new, restricted setting, we have managed to tame the destructive energy and harness it to peaceful, constructive purposes. The antinomy has not reappeared, but its basic idea has been used to establish a significant metalogical result with far-reaching implications.

Nothing is detracted from the significance of this result by the fact that its philosophical implications are essentially negative in character. The result shows indeed that in no domain of mathematics is the notion of provability a perfect substitute for the notion of truth. The belief that formal proof can serve as an adequate instrument for establishing truth of all mathematical statements has proved to be unfounded. The original triumph of formal methods has been followed by a serious setback.

Whatever can be said to conclude this discussion is bound to be an anticlimax. The notion of truth for formalized theories can now be introduced by means of a precise and adequate definition. It can therefore be used without any restrictions and reservations in metalogical discussion. It has actually become a basic metalogical notion involved in important problems and results. On the other hand, the notion of proof has not lost its significance either. Proof is still the only method used to ascertain the truth of sentences within any specific mathematical theory. We are now aware of the fact, however, that there are sentences formulated in the language of the theory which are true but not provable, and we cannot discount the possibility that some such sentences occur among those in which we are interested and which we attempt to prove. Hence in some situations we may wish to explore the possibility of widening the set of provable sentences. To this end we enrich the given theory by including new sentences in its axiom system or by providing it with new rules of proof. In doing so we use the notion of truth as a guide; for we do not wish to add a new axiom or a new rule of proof if we have reason to believe that the new axiom is not a true sentence, or that the new rule of proof when applied to true sentences may yield a false sentence. The process of extending a theory may of course be repeated arbitrarily many times. The notion of a true sentence functions thus as an ideal limit which can never be reached but which we try to approximate by gradually widening the set of provable sentences. (It seems likely, although for different reasons, that the notion of truth plays an analogous role in the realm of empirical knowledge.) There is no conflict between the notions of truth and proof in the development of mathematics; the two notions are not at war but live in peaceful coexistence.



Hormones in Social Amoebae and Mammals

The substance that attracts social amoebae to one another to form a sluglike mass has recently been identified. It turns out that the same substance also acts as a "messenger" in mammalian cells

by John Tyler Bonner

One of the pleasures of science is to see two distant and apparently unrelated pieces of information suddenly come together. In a flash what one knows doubles or triples in size. It is like working on two large but separate sections of a jigsaw puzzle and, almost without realizing it until the moment it happens, finding that they fit into one. I recently had this pleasure, although my own work directly concerned only one section of the puzzle. The assembling of the other section began some years ago, when a substance in the family of adenosine phosphates was found to be a "chemical messenger" intimately involved in the action of many mammalian hormones, including man's. The substance is cyclic AMP, or cyclic-3′,5′-adenosine monophosphate. ("Cyclic" refers to the fact that the atoms of the phosphate group form a ring.) The relatives of cyclic AMP are, among others, the more familiar, noncyclic ADP (adenosine diphosphate) and ATP (adenosine triphosphate), substances that play key roles in plant and animal metabolism.

Cyclic AMP was discovered by Earl W. Sutherland, Jr., now at the Vanderbilt University School of Medicine, and Theodore W. Rall of the Western Reserve University School of Medicine in 1958. They and their collaborators have since learned a number of remarkable facts about the substance, the bare bones of which I shall outline here. To deal with the basic biochemistry first, cyclic AMP is formed from ATP in a one-step reaction; the enzyme responsible for the conversion is adenyl cyclase. The substance is subject to further modification, being converted into 5′ AMP (5′-adenosine monophosphate) in another one-step reaction. The enzyme that carries out this step is a phosphodiesterase specific for the reaction.

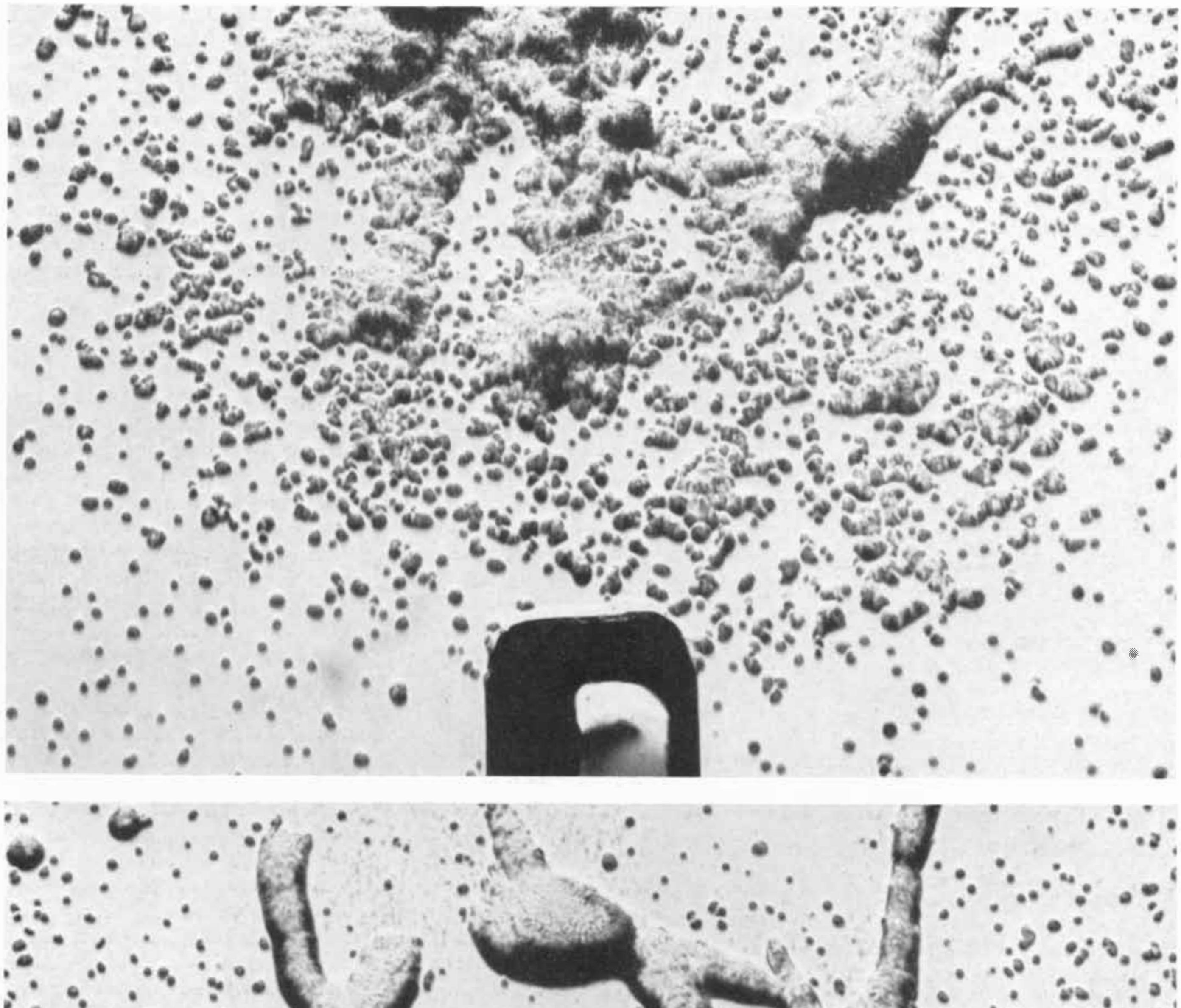
The hormone that Sutherland and Rall first showed to be involved with cyclic AMP was adrenalin, or epinephrine, as it is now usually called in the U.S. One is taught in elementary biology that the stress of anger, pain or fear produces a great surge of epinephrine in the blood, with a resulting quick rise in the amount of blood sugar available as energy for emergency action. Inevitably one develops the simpleminded notion that the epinephrine directly splits glycogen, or animal starch, into its subunits of glucose, or blood sugar. This is far from being the case. As Sutherland and Rall discovered, cyclic AMP is implicated in a much more complex process. I shall go into a small amount of detail here because the example of epinephrine illustrates the way cyclic AMP works with hormones generally.

Skipping the step that puts the epinephrine into the blood, we can start with the liver, which is one of the places where the hormone does its work. Cells in the liver are suddenly bathed in epinephrine, brought to them by neighboring blood vessels. Adenyl cyclase, the enzyme that converts ATP into cyclic AMP, is attached in some way to the surface membrane of liver cells. When the epinephrine reaches the cell surface, it specifically stimulates the enzyme, the conversion takes place and cyclic AMP is formed inside the cell.

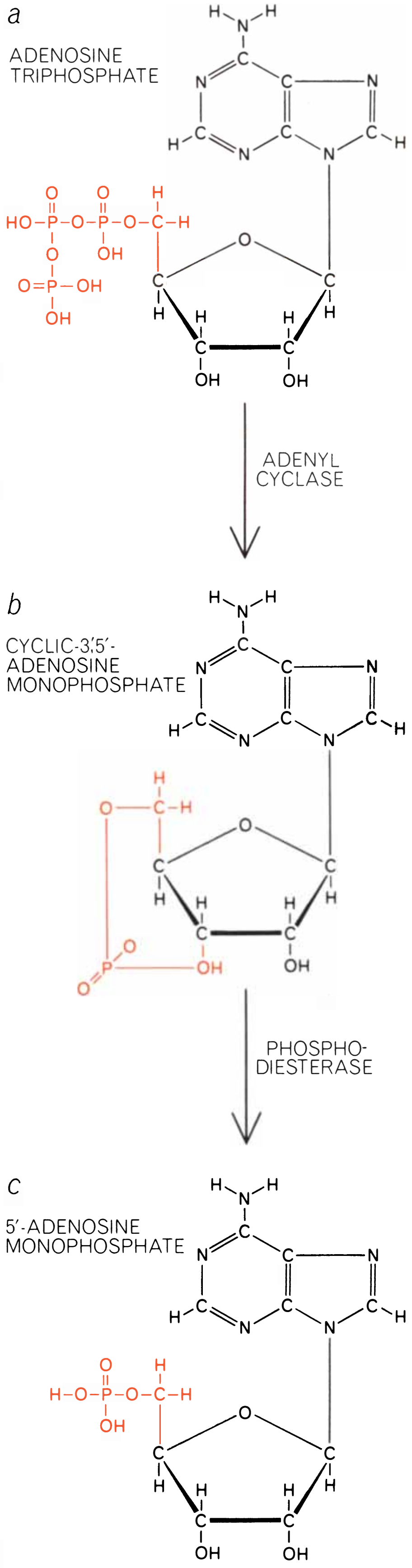
The cyclic AMP inside the cell now stimulates a second enzyme, changing it from an inactive form to an active one. Once activated, the second enzyme acts on a third to activate still another enzyme: phosphorylase. This fourth enzyme cleaves glycogen to produce glucose-1-phosphate. Before the sugar can escape from the cell, however, it must go through two more enzyme steps, one that converts it to glucose-6-phosphate and one that converts glucose-6-phosphate to glucose, the sugar that enters the blood [*see illustration on page 81*]. The entire sequence is remarkable for many reasons, not least of which is the extraordinary number of chemical steps involved in what originally was thought to be a simple process.

At the moment more than a dozen mammalian hormones are known to use cyclic AMP as this kind of "second messenger," to use the name Sutherland and his co-workers have given it. I shall not list them all here but prominent among them are glucagon and insulin, both of which are involved in blood-sugar levels, three hormones of the anterior pituitary and one of the posterior pituitary. In addition large amounts of the substance are present in brain tissue. It has been known for some time that the synaptic gaps between nerves are bridged by neurohormones such as acetylcholine. Bruce McL. Breckenridge of the Rutgers University Medical School has recently put forward the hypothesis that cyclic AMP is somehow involved in the bridging sequence.

One of the many unanswered questions about cyclic AMP is how the single enzyme adenyl cyclase can respond to so many different hormones and do so in such a specific way. There is some evidence that the enzyme, which always appears to be associated with cell membranes, differs in different cell systems: one type can respond only to one hormone, another type only to another hormone, and so on. This raises further questions of deep interest. For example, how could such a chemical system have evolved? It would almost seem as though the fundamental part of the hormone system is the cyclic AMP, and that the variety of methods for turning on the system had evolved subsequently. I shall return to some evolutionary considerations presently, but the main basis for such a



EFFECT OF ATTRACTANT on the social amoebae of the slime mold *Dictyostelium* is shown in these two photographs. In each, a small block of agar (*bottom*) contains the attractant, cyclic-3′,5′-adenosine monophosphate (cyclic AMP). In the upper photograph hundreds of amoebae have been placed near the block. In the lower photograph many have responded to the attractant by streaming toward the block. In normal development secretion of the attractant brings the amoebae together to form a multicellular organism.



CYCLIC AMP (*b*) is formed from the more familiar adenosine triphosphate, or ATP (*a*), in a one-step reaction catalyzed by the enzyme adenyl cyclase. The word "cyclic" refers to the ring shape of the molecule's phosphate group (*color*). The substance is converted into 5′ AMP (*c*) by a phosphodiesterase that is specific for the reaction.

contention is that cyclic AMP appears to be a common component of many cells and tissues and is even found in bacteria.

I should now like to ask the reader to forget momentarily what has gone before while we look at the other piece of the puzzle. This piece involves an entire series of experiments, including some of my own, in a quite specialized area of biology. Not until the very last step, however, was there the faintest suspicion that cyclic AMP played any part in the work.

The social amoebae known as the cellular slime molds were first described by Oskar Brefeld of Germany exactly a century ago. A few years later their peculiar life history was elucidated, and today a score or more species are known. Here I am going to concentrate mainly on the species *Dictyostelium discoideum,* which was discovered in the 1930's by Kenneth B. Raper of the University of Wisconsin. *Dictyostelium* amoebae live in the soil; they play a key role in the food chain because they are the immediate predators of the soil bacteria.

In appearance and size the amoebae are similar to human leukocytes, the white blood cells, and like leukocytes they engulf bacteria by surrounding them with protoplasm. After they clear an area of food and enter a period of starvation, the individual amoebae stream together into central collection points. In a reasonably well-populated area an aggregate may number 100,000 or more cells, although if food is scarce, the number will be fewer.

The assembled amoebae gather into a cartridge-shaped mass, one or two millimeters long, which has a distinct front and hind end. The mass crawls through the cavities in the soil like a slug and shows an ability to orient toward light and heat, an ability that is completely lacking in the separate amoebae before they aggregate. As I have argued elsewhere, spore dispersal must be of key importance in the natural selection of soil organisms, and the formation of a slug and its migration to "a better place in the sun" as it wanders through the soil or among rotting leaves in the humus is therefore an important adaptation for more effective spore dispersal.

After a period of migration the slug upends itself until it points at a right angle to the surface. Some of its cells will now form a stalk, while others will go to make up a spherical mass of spores. The stalk cells swell, become hollow like the pith of a plant and die. The result is a slender, tapering hollow rod that rises a few millimeters into the air and is filled with dead cells that serve as internal trusses. As the stalk grows the spore mass is carried aloft with it so that ultimately the delicate stalk bears a sphere of spores at its tip. Any one of the spores is capable of starting a new generation.

Because this organism is a relatively simple system of cells that differentiate (that is, take up specialized functions) it has been an object of interest to students of developmental biology for some years. One of the prime targets of study has been the social amoebae's mechanism of aggregation. How are the individuals oriented so that they stream into the central collecting points? The action is a rather pure example of morphogenesis, or formative movement, a process that is known to be most important in the initial stages of the development of all animal embryos.

At the turn of the century it was suggested that the aggregation might be due to the attraction of the individual amoebae by some kind of chemical stimulus, but there was no evidence to support this idea. Then in 1942 Ernest H. Runyon of Agnes Scott College showed that if he put an aggregation center on one side of a semipermeable membrane (such as a sausage casing) and put amoebae that were ready to aggregate on the other side, the amoebae would react. In spite of the membrane barrier they would swarm to a point on their side of the surface exactly opposite the aggregation center on the other side. Since it was known that large molecules, such as proteins, cannot readily penetrate such membranes, Runyon suggested that the chemical attractant might be a small molecule.

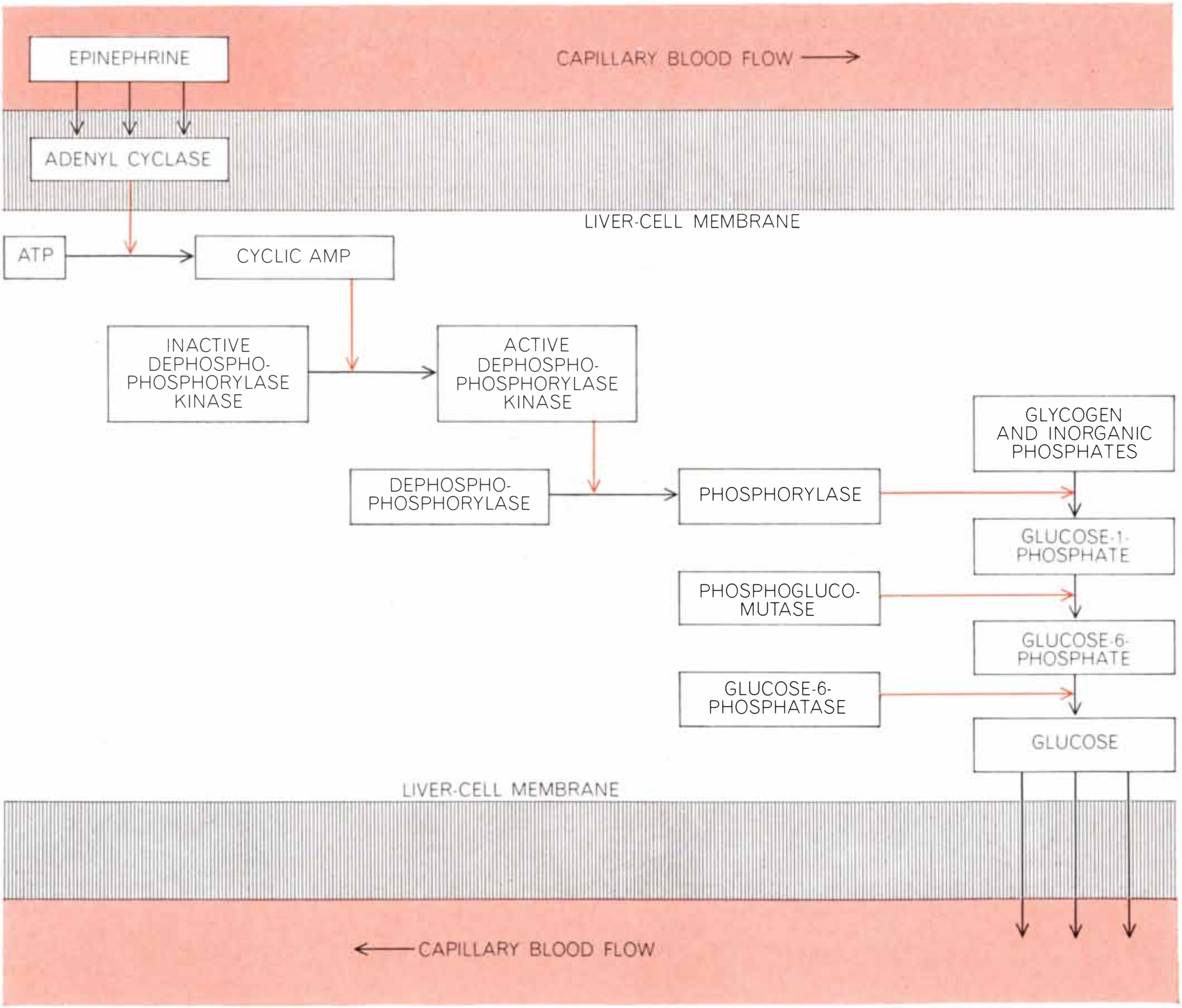
More evidence in support of the hypothesis of a chemical attractant was needed; it was clear that other forces, for example electric forces, could also penetrate a membrane. Studying the attractant problem further, I found in 1947 that if the amoebae began to aggregate underwater, and if one moved the water slowly past them (as though they were at the bottom of a brook), then the amoebae upstream became completely disoriented but the ones downstream oriented beautifully toward the aggregation center, even over large distances. The obvious conclusion was that aggregation was mediated by a free-diffusing agent; it might be heat or it might be a chemical substance, as Runyon had suggested. For various reasons I favored the idea of a chemical, and I gave the unidentified attractant the name acrasin because the proper name of the cellular slime molds is Acrasiales, and also because in Edmund Spenser's

Faerie Queene there is a witch named Acrasia who attracted men and transformed them into beasts.

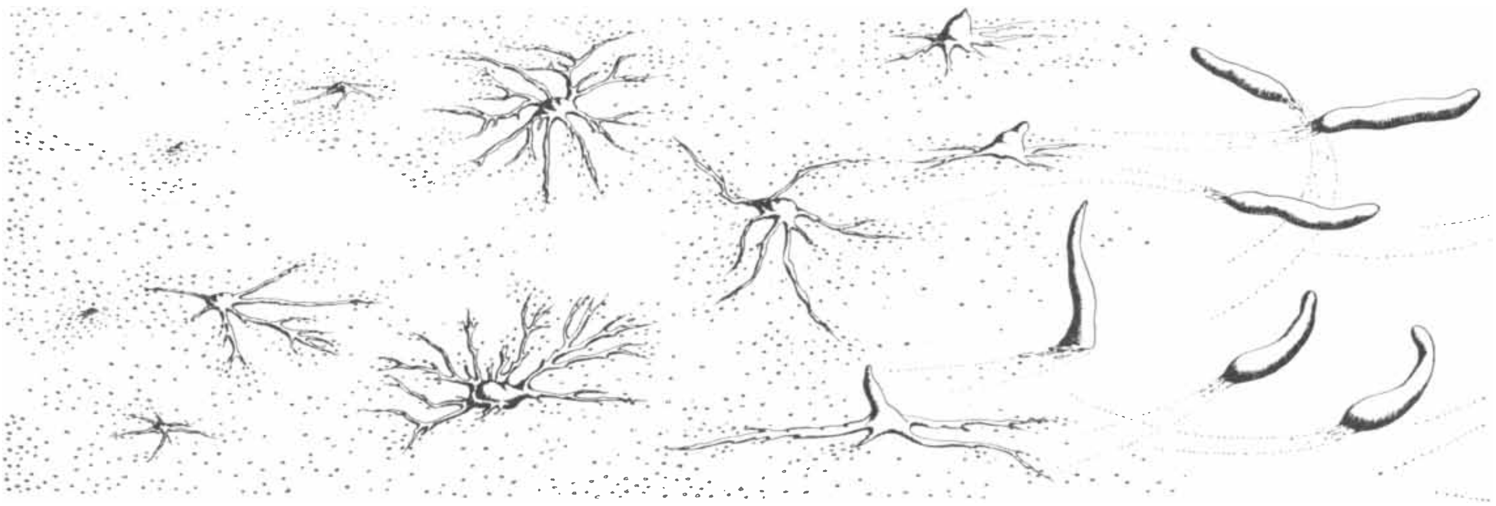
The final proof that acrasin really existed came from a remarkable set of experiments performed by Brian M. Shaffer of the University of Cambridge. At Princeton we had been unable to find any trace of the attractant at aggregation centers we had killed. Yet we felt that they must contain acrasin, because a few seconds earlier, when they were still alive, they had been actively attracting amoebae. Unknown to us, Shaffer was having the same difficulty; like us he suspected that for some reason the acrasin was rapidly disappearing.

To test this suspicion Shaffer sandwiched some active amoebae between a glass slide and a small block of agar [*see top illustration on page 88*]. To the edges of the sandwich he added drops of water that had been near an aggregation center (and therefore contained acrasin) every few seconds. In a few minutes all the amoebae under the block streamed toward the wet edges. If he let a minute or more pass between drops, there was no attraction; acrasin evidently disappeared quickly and, in order to attract, it had to be applied at short successive intervals. When Shaffer froze water that had been near aggregation centers in capillary tubes and three months later used the contents of the capillaries at short intervals, the amoebae were attracted; acrasin was apparently stable at low temperatures. In any case, there was no longer any doubt about the attractant's existence. As Shaffer wrote to me after these experiments, "I have bottled it."

The next problem was to find out why acrasin was so unstable. Shaffer showed conclusively that the reason was that a companion substance, an enzyme, rapidly destroyed it. He did this while visiting Princeton, so that I had a close view. His principal method was to plunge water containing acrasin into methanol, which denatured the accompanying enzyme; the remaining acrasin was stable. This was an economical experiment, but I always had special admiration for what might be called his Rube Goldberg demonstration. He suspended a long cylinder of sausage casing in a moist chamber; then he painted the outside of the casing with amoebae that were about to aggregate and dripped water along the inside of the casing. The water collected in a beaker at the bottom of the casing. The acrasin it contained was stable at room



ACTION OF SEVERAL HORMONES in mammals utilizes cyclic AMP as a chemical messenger; illustrated here is the role of the substance in the production of blood sugar by a liver cell following stimulation by the presence of epinephrine in the bloodstream. The enzyme adenyl cyclase, attached to the cell membrane (*top left*), is stimulated by epinephrine and changes ATP inside the cell into cyclic AMP. The messenger substance activates a second enzyme, beginning a five-step sequence that yields glucose (*bottom right*).



SOCIAL PHASE of the life cycle of the cellular slime molds begins when individual amoebae (*far left*) begin to stream together into a central collection point and gather into a cartridge-shaped mass that moves through the soil like a slug and is attracted to light.

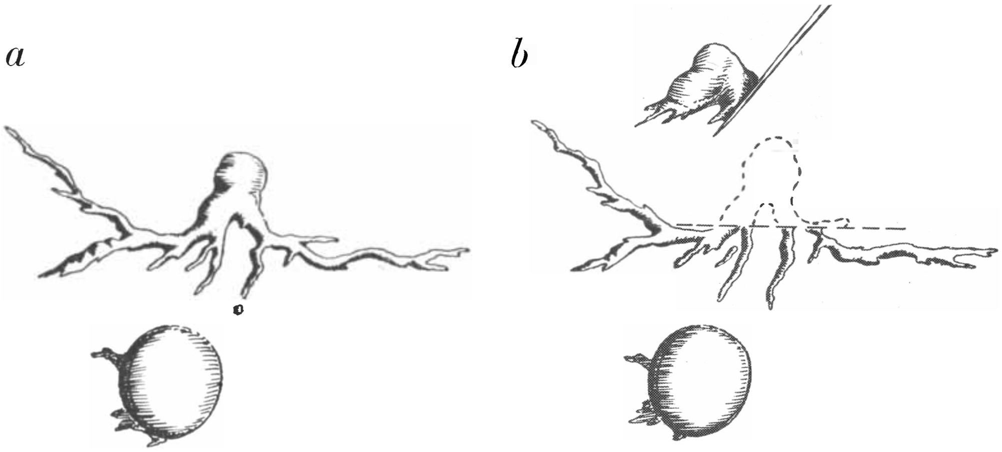
temperature for extended periods of time. The membrane had passed the small molecules of acrasin but barred the large molecules of the destructive enzyme, or acrasinase, that would ordinarily have destroyed the attractant.

Shaffer's experiment with the glass slide and the agar block constituted a test or an assay for the presence or absence of acrasin. The existence of such an assay, even though it was not a quantitative one, led to a flurry of activity as workers in various laboratories attempted to determine the chemical identity of acrasin. Various substances became suspects, and I shall mention a few of the attempts that were clearly going in the right direction. Barbara E. Wright of the Retina Foundation found that urine from pregnant women contained a component that attracted the amoebae of *Dictyostelium* (it does contain such a component), but she thought this had something to do with the steroids in the urine (it does not). She and her co-workers made the interesting discovery, however, that *Dictyostelium* synthesizes rather large amounts of one steroid with an unknown function. At Brandeis University, Maurice Sussman and his collaborators found attractive substances that absorbed light at wavelengths in the ultraviolet part of the spectrum (which was on the right track), and Shaffer, then working in conjunction with organic chemists at Cambridge, went so far as to suspect that acrasin was a member of the class of compounds called purines.

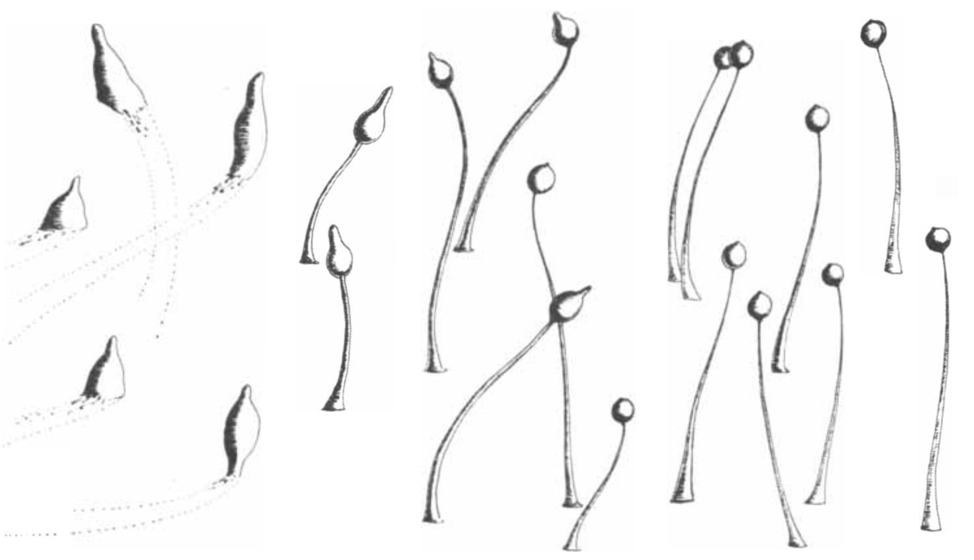
The ultimate result of these various efforts was that a time came when no one was working on the acrasin problem. Simultaneously, and without knowing about each other's activities until much later, Theo M. Konijn of the Hubrecht Laboratory in the Netherlands and I decided that since all was quiet the time had come to have a calm, noncompetitive look at the problem. We had both decided that the first necessity was to devise a truly quantitative assay, one that would not only show that an attractant was present but also measure just how attractive it was.

My test consisted in placing small cellophane squares on a dish of agar that contained the test substance; amoebae were then put on the cellophane squares and the rate at which they moved off the cellophane was recorded. At the time we thought the test had limitations (a concern that has recently proved groundless), but it taught us several things. It showed, for example, that urine from both pregnant and nonpregnant women was loaded with an active substance, and my colleague Ruth Reisberg found good evidence that the substance was a phosphate compound. It also proved that bacteria were loaded with an attractant of some kind. (This was a fact that a number of workers had learned before. Barbara Wright had even shown the attractiveness of bacterial extracts using the Shaffer assay.) Undoubtedly the greatest benefit of my test, however, was that when Konijn read our first report, he immediately wrote to tell us of his remarkable work, which led to his coming to Princeton for a year.

Because the Konijn test was the key to later successes I shall describe it in some detail. The basic idea is to put a very small droplet of saline solution that contains amoebae on an agar plate and then to put a water droplet containing the test substance nearby. If after a few hours the amoebae spread out of the confines of their droplet, the test is scored as positive. The assay is made quantitative simply by varying the distance between the two droplets. There are, however, two key factors in making the Konijn test work. One is that the agar must be washed repeatedly and must have exactly the right degree of rigidity. Unless these requirements are met the amoebae may not stay within the confines of their droplet even though



ABILITY OF BACTERIA to secrete the slime mold attractant, named acrasin by the author, was discovered during the search for the substance's chemical identity; one laboratory proof is illustrated here. The collection point (*a*) of a slime mold aggregation is re-



After a period of migration the slug upends itself and some cells form a stalk that rises upward, carrying aloft a spherical mass of other cells that have turned into spores (*right*).

the test droplet contains no attractant.

The other key factor in the Konijn test is that the amoebae must be at just the right stage in their development in order to respond. At first Konijn would stay up until the small hours waiting for that precise moment. Then matrimony intervened, with the result that he made an interesting discovery. When he stored the test plates overnight at five degrees Celsius, within minutes after the plates had returned to room temperature the next day all the amoebae were responsive. In addition, their movements were more precisely synchronized than they had ever been before.

Working with this test, Konijn and a colleague at the University of Utrecht, J. G. C. van de Meene, began an attempt to identify acrasin. We were making a similar attempt with my test, but they carried the matter much further. We had both shown that it was unaffected by heating and that it had a negative electric charge. They were able to add the fact that the attractant had a low molecular weight (between 200 and 400). When he came to Princeton, Konijn brought along some of his purified extract. He and David S. Barkley, who was then a graduate student, determined that the attractant also absorbed light in the ultraviolet part of the spectrum, with an absorption peak at a wavelength of 259 microns.

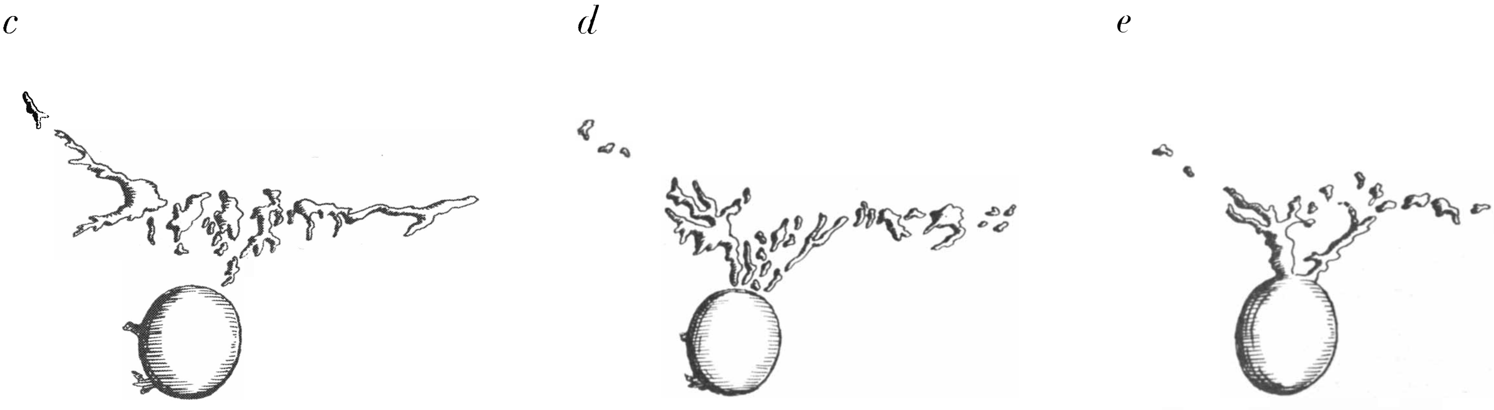
One evening Konijn and Barkley were sitting together reviewing all they knew about acrasin—the data just noted and the fact that the substance was found in bacteria and urine. Barkley had an inspiration: Why not see if cyclic AMP was an attractant? They quickly obtained some of the substance and in no time the hunch was verified. I was in Canada at the time and received an excited telephone call from the two; they had found that cyclic AMP was amazingly effective at attracting amoebae in the Konijn test. Eventually they were to learn that as little as .01 millimicrogram (10^{-11} gram) was enough to make the amoebae burst out of their droplet.

This important discovery was a wonderful bit of good fortune but a big question remained. Was cyclic AMP a naturally occurring acrasin, synthesized by the amoebae? By chance we did an incredibly stupid thing that turned out to be incredibly foresighted. At the time we happened to have another slime mold species, *Polysphondylium pallidum*, growing well in liquid culture. Without thinking, we decided to see if this species synthesized cyclic AMP. We found that it produced the substance in large quantities and that it evidently produced no other kind of attractant. This took us about a month.

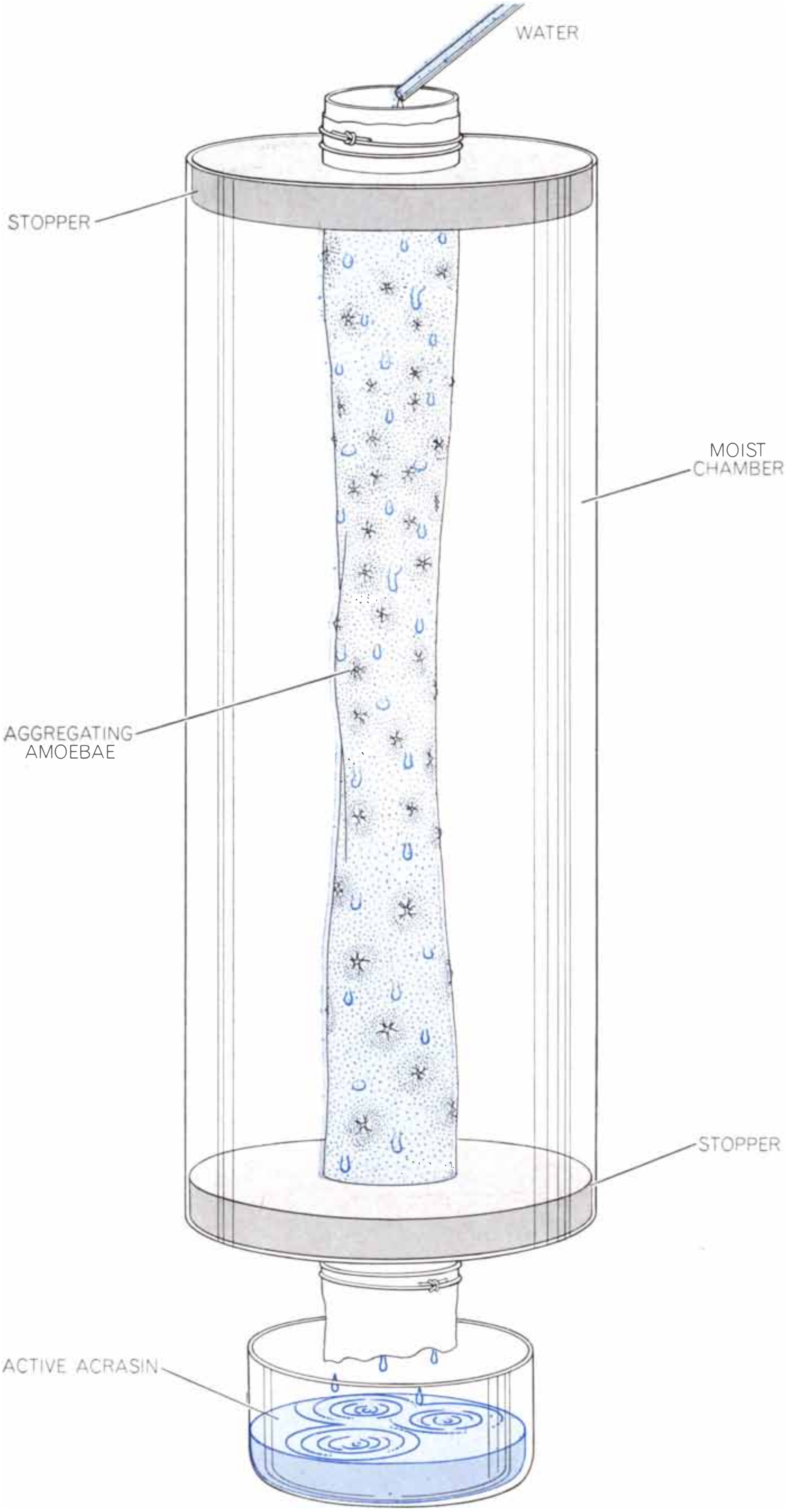
We decided to complete the story by showing that cyclic AMP also attracts *Polysphondylium*. To our horror it did no such thing: the species produces the attractant but does not respond to it. This discovery raised interesting questions that are still unanswered regarding the specificity of slime mold species and also presented a puzzle: We do not know the chemical mechanism whereby *Polysphondylium* amoebae form aggregates.

It seemed as though a whole month had been wasted. To make matters worse, when we attempted to extract cyclic AMP from *Dictyostelium* we not only found that it was totally absent but also could discover no trace of any attractant of any kind. While groping in the depths of depression, one of our collaborators, Ying-Ying Chang, decided to see if our negative results were related to those that had plagued Shaffer when he first tried to "bottle" acrasin. Shaffer had shown that *Dictyostelium* produced the destructive enzyme acrasinase, and he had kept the enzyme from making the acrasin disappear by separating the two.

There is an enzyme (a phosphodiesterase) that changes cyclic AMP into 5′ AMP. In mammals it had been isolated and partially purified by R. W. Butcher of Vanderbilt University, who was kind enough to send us some. In fact, Butch-



moved (*b*), leaving the stream of converging amoebae disoriented (*c*). A nearby clump of the common bacterium *Escherichia coli* now begins to exercise an attraction (*d*) and within 45 minutes the amoebae are streaming in the opposite direction (*e*). The ancestors of the bacteria-devouring social amoebae probably were helped in their search for food by their prey's secretion of the attractant.



"RUBE GOLDBERG" DEVICE for collecting acrasin, the attractant produced by slime molds, was constructed by Brian M. Shaffer of the University of Cambridge during his visit to Princeton University. A sausage casing was suspended in a moist chamber; aggregating amoebae were placed on the outside of the casing while water was dripped down the inside. The small molecules of acrasin passed through the membrane, entered the water and were collected (*bottom*). The molecules of an enzyme that is produced by the amoebae along with acrasin and that usually destroys the attractant, however, were too large to pass through the membrane. Acrasin collected by this method remained active for long periods.

er's enzyme was one of the substances we had used to prove that the attractant secreted by *Polysphondylium* actually was cyclic AMP; when we incubated the attractant with the enzyme, all the attractant's activity disappeared. If *Dictyostelium* was producing such an enzyme, our failure to find cyclic AMP would be explained.

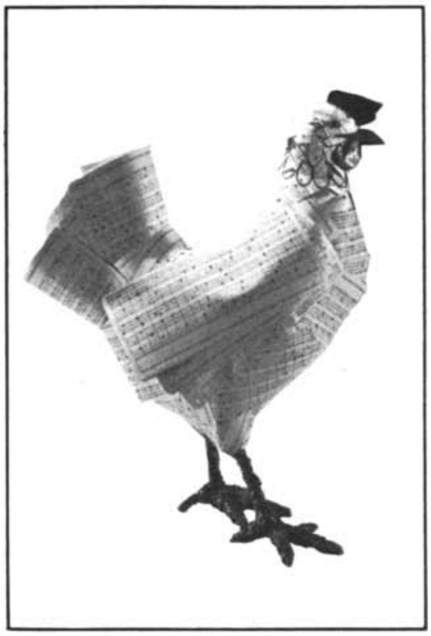
In a rather rapid series of experiments Chang showed that such was the case. *Dictyostelium* produces large amounts of a phosphodiesterase that has many (although not all) of the properties of Butcher's enzyme. The main similarity is that, like the mammalian enzyme, the slime mold enzyme breaks down cyclic AMP. The main difference is that the mammalian enzyme is found within the cell but the slime mold enzyme is almost entirely extracellular. I shall return to the significance of this fact.

Clearly Chang's work had pulled us out of our difficulties. Now all we needed was some way of proving that cyclic AMP was actively synthesized by *Dictyostelium*. We approached the problem in two ways. The first method was to grow the amoebae on dead bacteria that contained residual amounts of cyclic AMP. After a period the culture was assayed, great pains being taken to denature the destructive enzyme quickly. The assays showed that there was an increase in an attractant that had the characteristics of cyclic AMP.

The second method, devised by Barkley, depended on separating the destructive enzyme from the attractant. One way he did this was to add beads made of a special resin to a solution containing both substances. The resin had the property of adsorbing molecules with a negative charge, such as the molecules of cyclic AMP, thus putting them out of reach of the enzyme. Another approach was a modification of Shaffer's device: amoebae were kept on one side of a semipermeable membrane and water was kept on the other. Since the membrane would not pass the enzyme's big molecules, the attractant that escaped through the membrane and entered the water was stable. The material collected in both ways was subjected to chemical analysis. Once again the only attractant of low molecular weight to be isolated was identical in characteristics with cyclic AMP. Thus we now seem to have an understanding of the basic acrasin-acrasinase system, at least in the case of *Dictyostelium discoideum*.

We have recently examined the question of when during their life cycle *Dictyostelium* amoebae produce cyclic

The Electronic Chicken and how it grew.



"Improving the breed" of domestic fowls and animals has been a major concern of farmers since the ancient beginnings of agriculture. Today, modern techniques are accelerating the pace of genetic improvement as never before. And nowhere is the progress more rapid than in the raising of poultry where rapid life cycle and modern data processing have resulted in progress of amazing swiftness.

Arthur Heisdorf started out in the State of Washington in 1945 with fifteen acres and an idea. Today, Heisdorf and Nelson Farms produce more than 200 *million* chickens per year. The healthiest, lowest-maintenance, most highly productive chickens that ever laid an egg.

The genetic idea responsible for this amazing success is called Reciprocal Recurrent Selection. It's based on the maintenance of parent stock lines that blend or "nick" to produce hybrid chicks of maximum vigor and quality. Originally, H&N marketed its "Nick Chicks" directly to farmers. Today, the superiority of the H&N chicken is so widely accepted that the company markets just the parent birds in over 70 countries throughout the free world. The hatcheries in turn breed and sell the finished hybrid chicks.

At its home farms, the company maintains some 80 pure breeds of chickens, as well as many precious crossbred strains of parent stock. Careful research and evaluation of all breeding results is essential. But the most crucial object of attention is the finished "Nick Chick." Each hybrid test chicken is carefully examined for some 22 different genetic traits, all duly entered on punch cards and electronically evaluated. Researchers match performance ratings against family trees to determine which parent strains have crossed best to produce superior chicks.

The success of their white leghorn "Nick Chick" has led H&N research teams to apply the same sophisticated techniques with equal success to brown-egg layers, meat chickens and turkeys. But the real gleam in H&N's eye is for the chicken of the future—a miniature leghorn that weighs and eats about 30% less while producing as many (if slightly smaller) eggs, with far greater net return to the farmer.

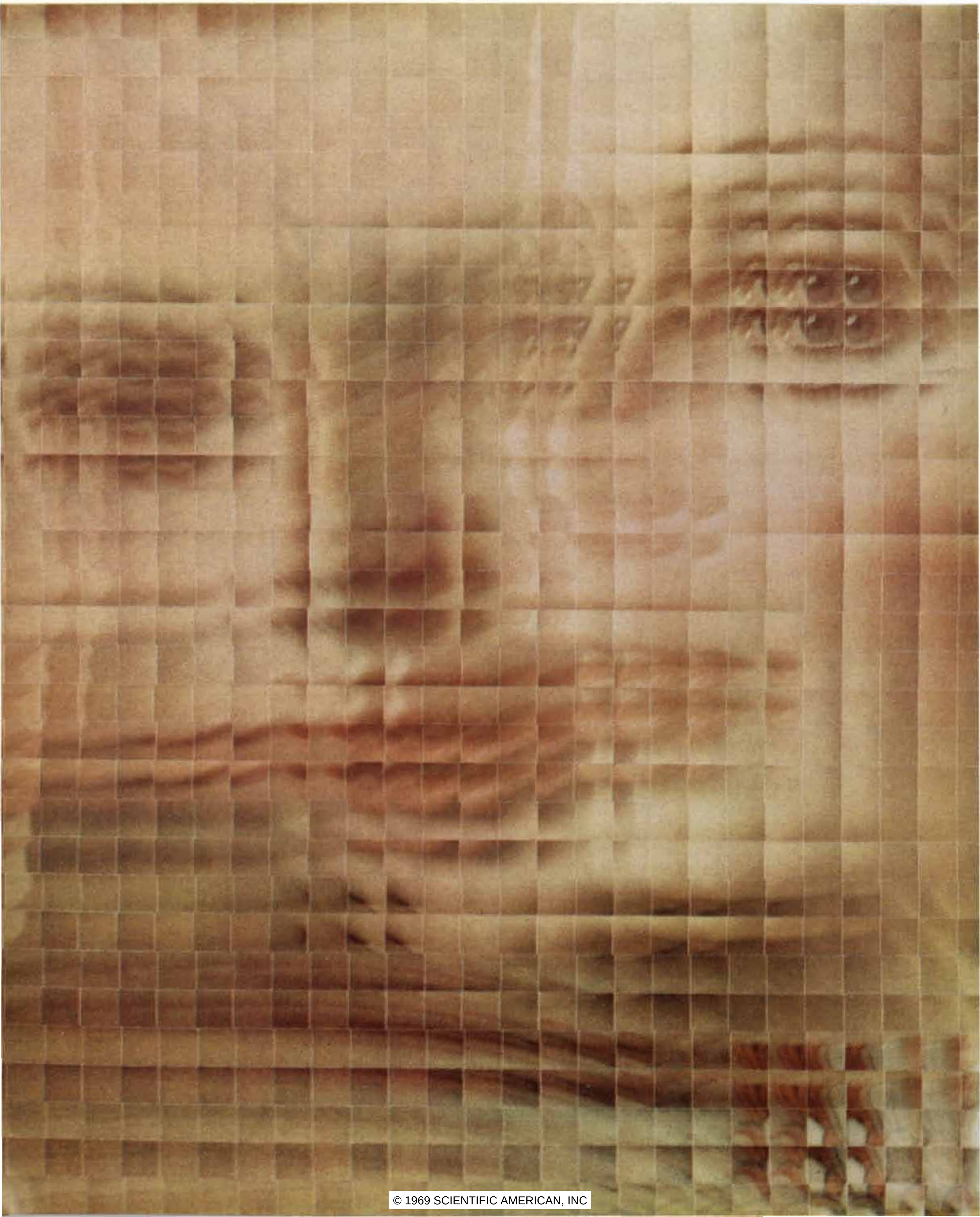
The State of Washington is a fertile field for the exciting break-through industries of the future. In agriculture, oceanography, aerospace, nuclear energy, fisheries and many other areas, the smart money is betting on the State of Washington.

Daniel J Evans
Governor

STATE OF WASHINGTON

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

You talk so much yo



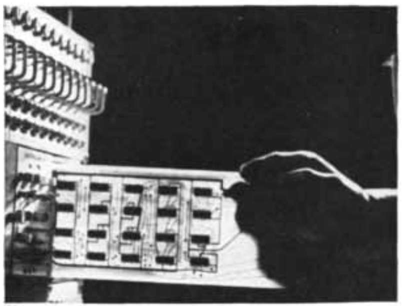
ur voice is changing.

The world is becoming a chatterbox. More and more people are talking to more and more people. We're talking on telephones, on television, on radio, via teletypewriter and computer. Our communications channels are getting as crowded as a one-telephone home with four teenage daughters.

How does the communications industry cope with the avalanche of words? One technique is called pulse code modulation. Before your voice is sent over the telephone line, it's chopped up and converted into a digital code. It is then translated at the other end and your "Hello" comes out just as you said it. That way twenty-four conversations can be transmitted simultaneously on wires that used to carry two.

As advanced electronic techniques bring the science of the computer to your telephone exchange, you'll be getting increasingly more convenient services, too. "Call Forwarding," for instance. Just dial the number where you'll be into your own phone and any incoming calls will be switched automatically. Or "Call Waiting" which alerts you to another incoming call and holds it until you're off the wire. "Speed Calling" lets you dial frequently called numbers by dialing just one or two digits.

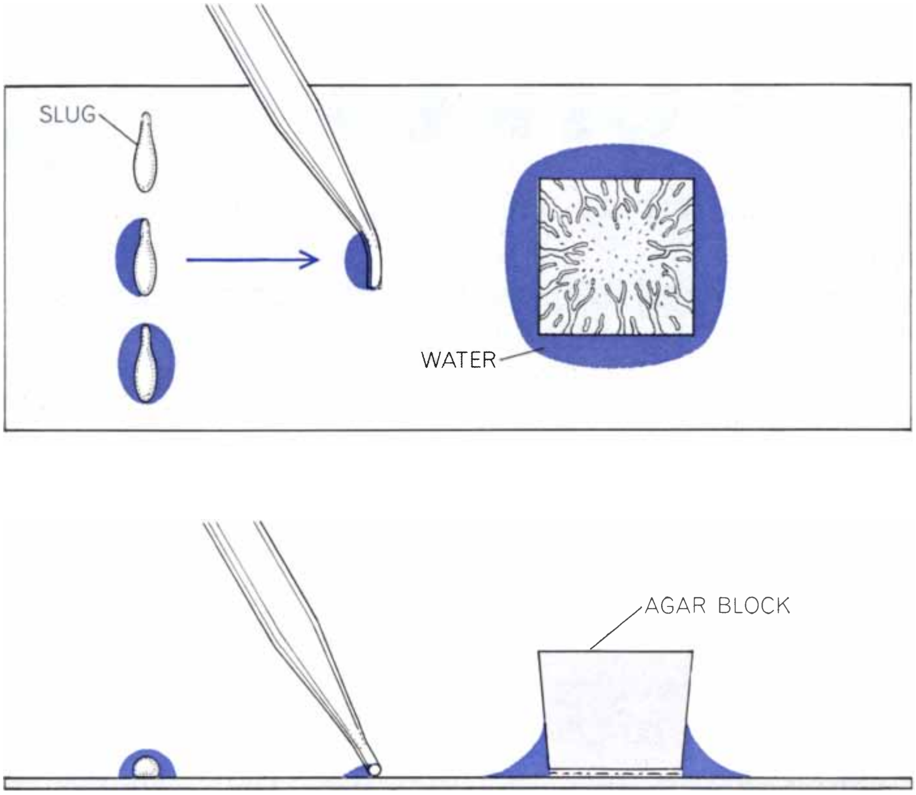
General Dynamics, through its Stromberg-Carlson subsidiary, supplies the independent telephone industry with telephones, exchanges and other communications systems. And it's an industry that's growing even faster than the gross national product.


Microcircuits help code phone transmission.

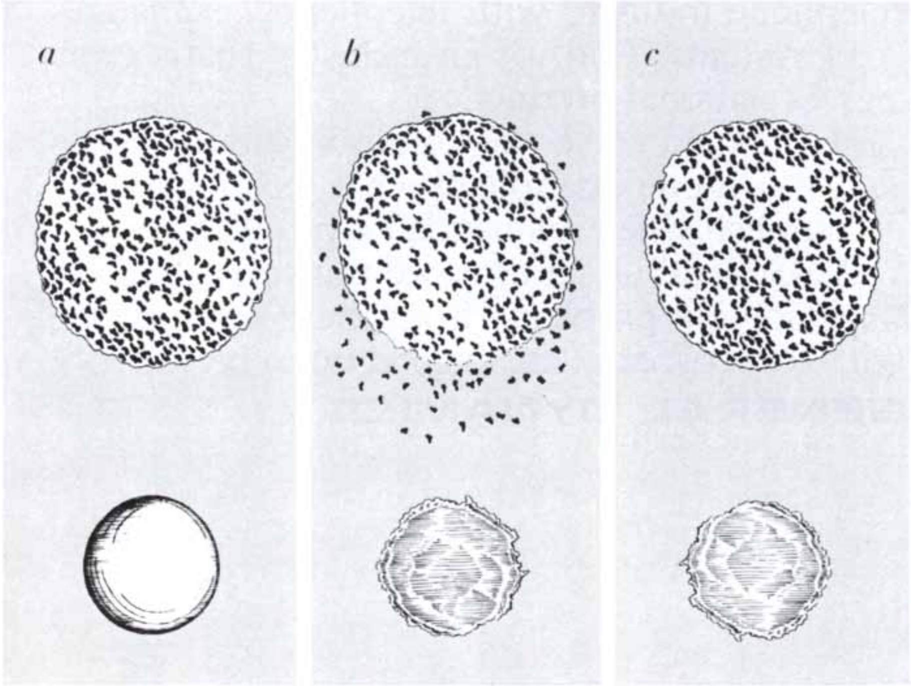
Helping to solve the communications explosion is just one example of what technology can accomplish when it's handed a problem.

At General Dynamics, we put technology to work solving problems from the bottom of the sea to outer space...and a good bit in between.

GENERAL DYNAMICS



QUALITATIVE ASSAY, proving that slime mold collection centers are rich in acrasin but that the attractant is short-lived, was devised when Shaffer sandwiched amoebae between a glass slide and an agar block. He then dampened the edges of the block every few seconds with water that had been in contact with cell masses that were actively secreting acrasin; the amoebae streamed toward the dampened edges. When the applications of water were minutes apart, however, the amoebae were no longer attracted; the acrasin had vanished.



QUANTITATIVE ASSAY, indicating the presence or absence of acrasin and also gauging the amount present, was devised by Theo M. Konijn of the Hubrecht Laboratory in the Netherlands. A droplet containing amoebae in saline solution is put on one part of an agar plate and a droplet of the substance being tested is put nearby (*a*). If the amoebae emerge from their droplet in an hour or so (*b*), the test is scored as positive. If the amoebae stay inside the droplet (*c*), the test substance is deemed lacking in acrasin. The strength of the attractant may be measured by repeating the test with increased distances between droplets.

AMP. We grew the amoebae on one side of a membrane and bathed the other side with water that was collected at two-hour intervals. The water samples were concentrated and tested for the presence of cyclic AMP, using the Konijn assay. The relative strength of each sample was determined by comparing it with solutions containing known amounts of commercial cyclic AMP.

We found that when the amoebae pass through the aggregation phase there is at least a hundredfold increase in their secretion of the attractant. Simultaneously there is a hundredfold increase in the amoebae's sensitivity to cyclic AMP. This means that the whole mechanism of chemical attraction is at least 10,000 times more effective during aggregation than during the earlier stages of the amoebae's development [*see illustration on page 91*].

We were also able to show that, just as the amoebae continuously produce small amounts of cyclic AMP during the early stages of their development, they also continuously produce the enzyme that destroys the attractant. This mechanism of simultaneous production and destruction evidently serves an important purpose in the amoebae's life cycle. If the individual amoeba's output of attractant were not steadily eliminated, the attractant would keep accumulating until the gradient in attractant concentration that guides the individual during aggregation would be drowned out. An example is the role the destructive enzyme plays in the Konijn assay. The fact that the amoebae are initially concentrated in one spot means that a considerable amount of the enzyme is also concentrated in the same area; it destroys any attractant in the immediate vicinity of the amoebae. The result is that if attractant is applied externally, it is destroyed in the immediate region of the amoebae, producing a steep outward gradient of attractant that stimulates the amoebae to move out radially beyond the confines of their droplet.

I have pointed out that soil amoebae feed on soil bacteria and also that bacteria secrete cyclic AMP. It is quite reasonable to suppose that a positive response to cyclic AMP, among other substances, helped the solitary predators that were the ancestors of the slime mold amoebae to track their prey. Cyclic AMP is a small molecule that diffuses readily and is remarkably stable. As evolution proceeded and slime molds with multicellular fruiting bodies began to enjoy the selective advantages of more effective spore dispersal, it is possible to

Basic Research at Honeywell
Research Center
Hopkins, Minnesota

The use of Infrared Detection in Industrial Measurement of Temperature

New processes and materials have increased the need for devices measuring temperatures from 100° to 500°F. Sensing Infrared energy may offer practical solutions.

Industry today, working with plastics, rubber, textiles, paper and glass, requires a new order of devices for temperature measurement, with greater precision, speed of response, and reliability. Requirements vary considerably for production processes, however most uses demand a device which can make a very rapid and accurate measurement of a small "target" portion of the product. Because the product may be moving rapidly or may be corrosive, the sensing must be accomplished without coming in contact with the surface.

One basic temperature measurement device, the thermocouple, is still used. Unfortunately, the sensing portion of the thermocouple is usually placed in a mold or tool, so it yields an *indirect* measure of average heat. Also, a thermocouple requires a large temperature differential between the sensor and the reference junction to produce a practical measureable voltage.

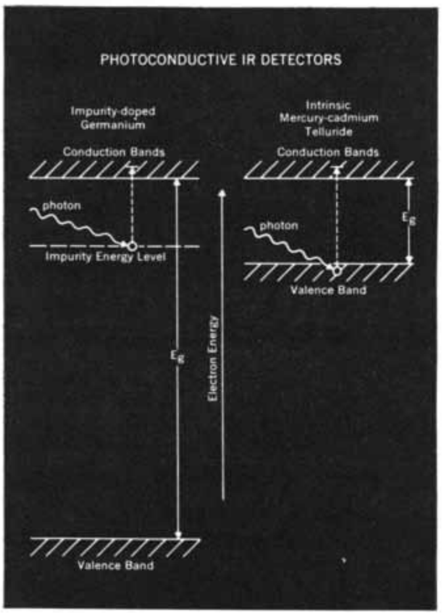
Connecting a number of thermocouples, in series, as in a thermopile, does increase the voltage output, but both the thermocouple and thermopile share the additional disadvantage of slow response.

If we could measure heat (radiant energy) rather than heat transfer directly, the delay in response would be eliminated. Radiant, or infrared energy, is directly proportional to temperature by the formula: $E = \epsilon\sigma T^4$ (σ = emissivity constant). The amount of infrared radiation is determined by the number of photons which are emitted by a hot object and by their energy or wavelengths. An instrument which counts the number of photons of the proper energy would give a signal which could be directly converted to temperature.

Semiconductors convert photons of infrared radiation into signals. One elemental semiconductor, germanium – to which impurities have been added to produce "doped" germanium – has been used in temperature measurement. When IR radiation impinges on a doped crystal of germanium, the photons are absorbed by electrons from the crystal impurities. The photon energy enables the electrons to escape the bonds holding them to the atoms. In this phenomenon, called photoconductivity, the photo-excited electrons carry a current freely through the crystal which can be measured electrically and converted to a measure of temperature.

However, doped germanium detectors are largely restricted to laboratory or military use because only small amounts of impurities can be added to pure germanium crystals, resulting in few potentially excitable electrons and a faint signal. These faint photoelectric signals are obscured by conduction from thermally excited electrons unless the sensor is housed in an elaborate liquid helium cooling apparatus.

To eliminate the low temperature requirement limiting the practical use of doped germanium detectors, scientists at Honeywell began, in 1960, to search for a compound with intrinsic excitability, on the theory that, if the electrons of the crystal material itself could be photoexcited, there would be a greater number of current-carrying electrons and a stronger signal.



Electron energy level diagram of photoconductors. The small impurity energy gap gives germanium its photoconductive property.

The Honeywell scientists assumed that an alloy combining a compound having a vanishingly small energy gap, with a wide energy gap semi-conductor compound, would have intrinsic excitability. After comparing a number of different alloys, they developed a photoconductive alloy of mercury telluride and cadmium telluride which could effectively sense lower energy, longer wavelengths, and which could be cooled either thermoelectrically or by liquid nitrogen, thus obviating the need for an elaborate apparatus.

The alloy mercury-cadmium telluride also proved to be designable. By varying the amounts of HgTe and CdTe, an alloy could be designed to yield an optimum response at the specific wavelength dictated by the temperature of the material being measured.

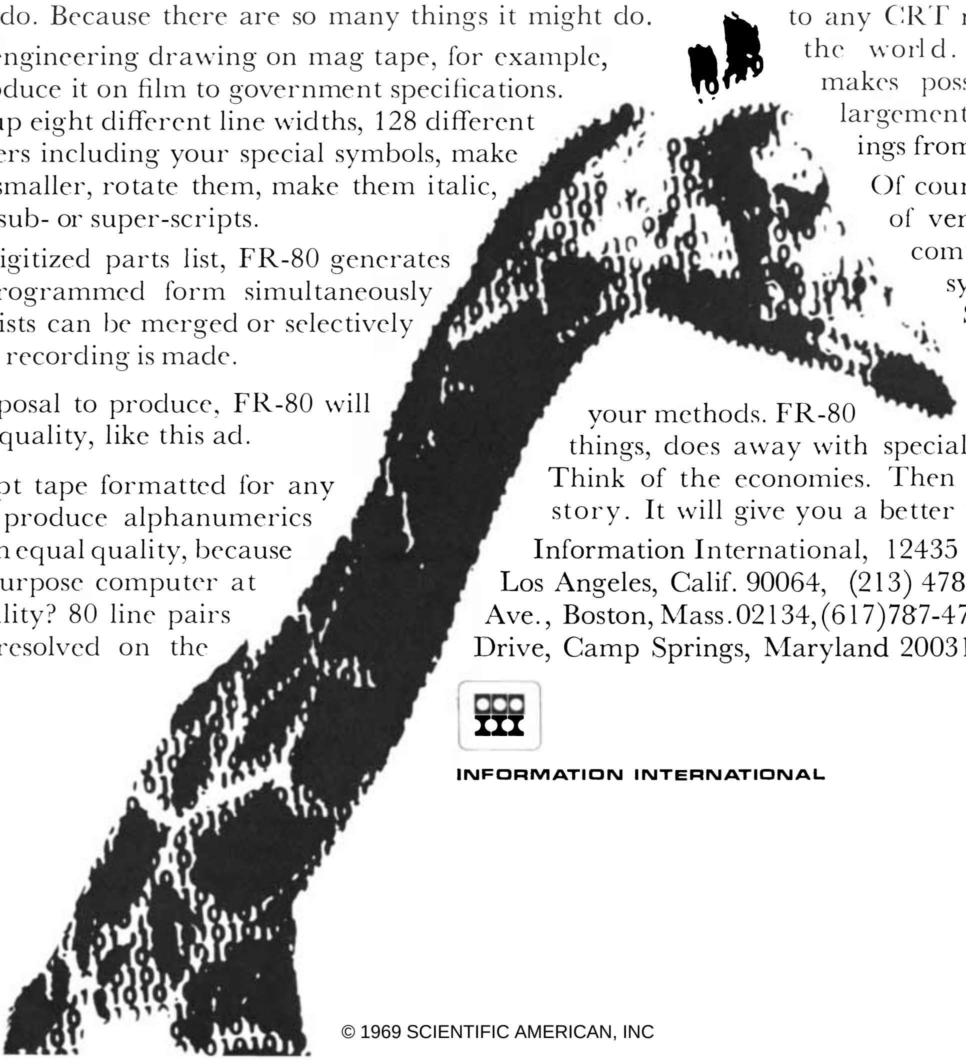
Since these materials are extremely sensitive to the presence of impurities, highly pure single crystals of the alloy had to be produced. By solidification from a molten mixture of HgTe and CdTe, homogeneous crystals were grown to fit specific industrial requirements. The crystals, combined with optical equipment to focus the radiation, plus a signal amplifier and a cooling system, produce a practical industrial instrument which is now being used at temperatures ranging from room temperature to greater than 2000° F. The Honeywell infrared pyrometer can measure temperature in less than 10 ms at temperatures from 100° F., with a target size of less than .25 sq. in. Error tolerances are ±2% and repeatability is 0.5%.

Future research will center on the problem of greater accuracy, particularly in room-temperature applications.

If you are working in the field of infrared pyrometers, and want to learn more about Honeywell's research in infrared detection, contact Dr. Donald Long, Honeywell Research Center, Hopkins, Minnesota. If you are interested in industrial applications for Honeywell's infrared pyrometer contact Mr. John Myers of the Honeywell Research Center. If you have an advanced degree and are interested in a career in research at Honeywell, please write to Dr. John Dempsey, Vice President Science and Engineering, Honeywell Inc., 2701 4th Avenue So., Minneapolis, Minn. 55408.

Honeywell
AUTOMATION

NO, I SAID GRAPH!



You have to be a little careful when you tell our new film recorder, FR-80, what to do. Because there are so many things it might do.

If you have an engineering drawing on mag tape, for example, FR-80 can reproduce it on film to government specifications. It will summon up eight different line widths, 128 different kinds of characters including your special symbols, make them larger or smaller, rotate them, make them italic, or place them as sub- or super-scripts.

If you have a digitized parts list, FR-80 generates your own preprogrammed form simultaneously with the data. Lists can be merged or selectively sorted as the film recording is made.

If you have proposal to produce, FR-80 will set type in book quality, like this ad.

FR-80 will accept tape formatted for any peripheral and produce alphanumerics and graphics with equal quality, because it is a general purpose computer at heart. What quality? 80 line pairs per millimeter resolved on the 35 mm film, which is superior to any CRT recorder system in the world. Such sharpness makes possible extreme enlargements, like E-size drawings from microfilm.

Of course, this is the kind of versatility a modern company's information system has to have. Special purpose graphics output systems freeze your methods. FR-80 lets you improve things, does away with special purpose systems. Think of the economies. Then write for the full story. It will give you a better view from the top.

Information International, 12435 W. Olympic Blvd., Los Angeles, Calif. 90064, (213) 478-2571; 89 Brighton Ave., Boston, Mass. 02134, (617) 787-4700; 7880 Coolridge Drive, Camp Springs, Maryland 20031 (301) 449-4248.

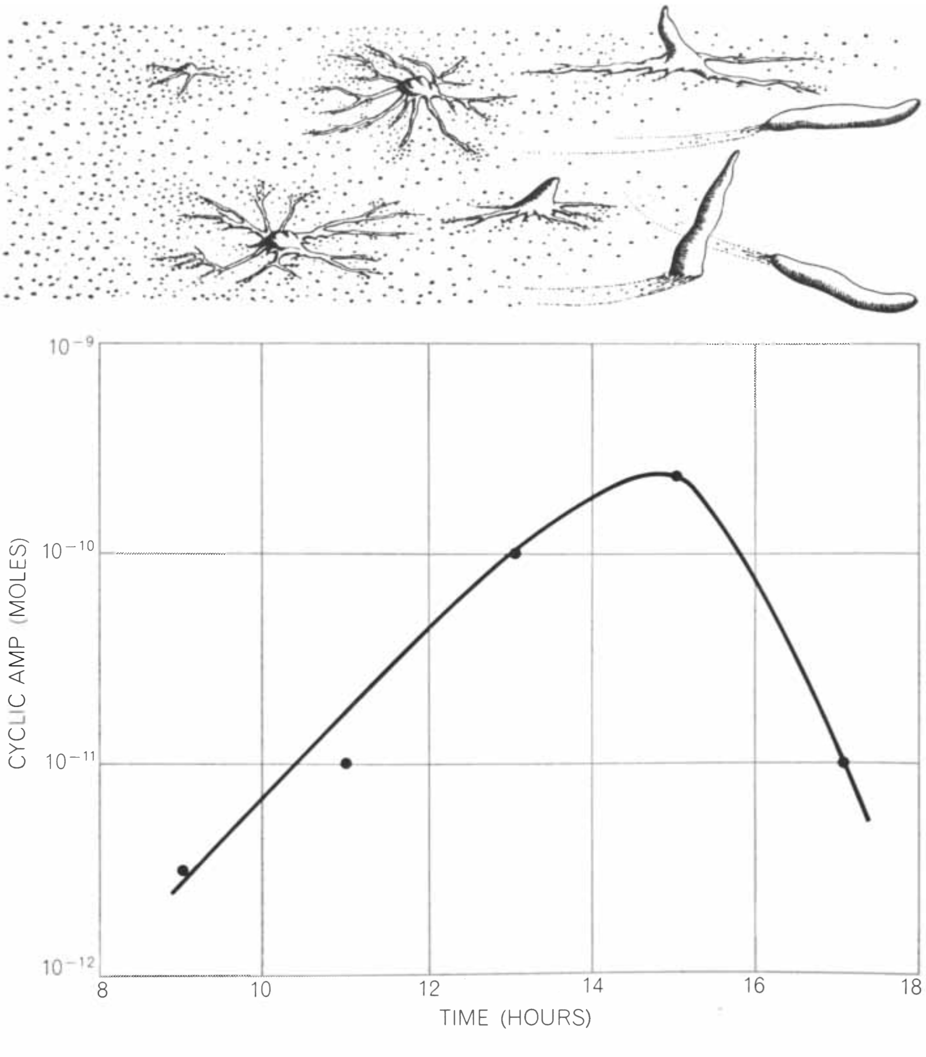
INFORMATION INTERNATIONAL

assume that social amoebae were helped to aggregate by the same mechanism of chemical attraction that already assisted their quest for food. The only required change would be to increase the sensitivity of the mechanism enormously, so that individuals would ignore their surroundings, stop hunting and swarm together instead.

Thus we find that the question of chemical attractants among social amoebae and the question of hormonal activity in mammals are connected by a single biochemical link, and that the two large pieces of the puzzle that we mentioned in the beginning do indeed fit together. There remains much to be done, however; the rest of the puzzle must be filled in.

If we look at the link, we see that in *Dictyostelium* cyclic AMP appears to be the main hormone responsible for the amoebae's social existence. We should like to know more about how it is controlled within the individual amoeba and how it orients amoebae, but its central role is now established. Like the principal mammalian hormones, it is extracellular and provides communication between cells that are separated from one another in space. In mammals, on the other hand, cyclic AMP is triggered by the extracellular hormones and acts inside the cells. It may possibly act externally as well, but such actions have not yet been elucidated. That the chain of chemical events involving cyclic AMP in mammals is longer and more complex than it is in social amoebae is something that is to be expected of more complicated organisms.

To me, from a biochemical and an evolutionary point of view, one of the most fundamental unsolved questions is the role of cyclic AMP in bacteria. There have been some interesting beginnings on this problem, and when it is fully elucidated, I feel it will provide an important insight into the basic role of the substance in all living systems, an insight that is very much needed at this moment.



HUNDREDFOLD INCREASE in the amount of attractant secreted by social amoebae takes place during the seven hours in which the individual amoebae begin to stream together at collection centers. The peak is reached as aggregation is completed (*top, right of center*); production falls sharply as the fully formed multicellular slug starts to migrate.

ULTRAVIOLET ASTRONOMY

Telescopes now in orbit are measuring the ultraviolet radiation from celestial objects that never penetrates the earth's atmosphere. The results may significantly alter man's view of the universe

by Leo Goldberg

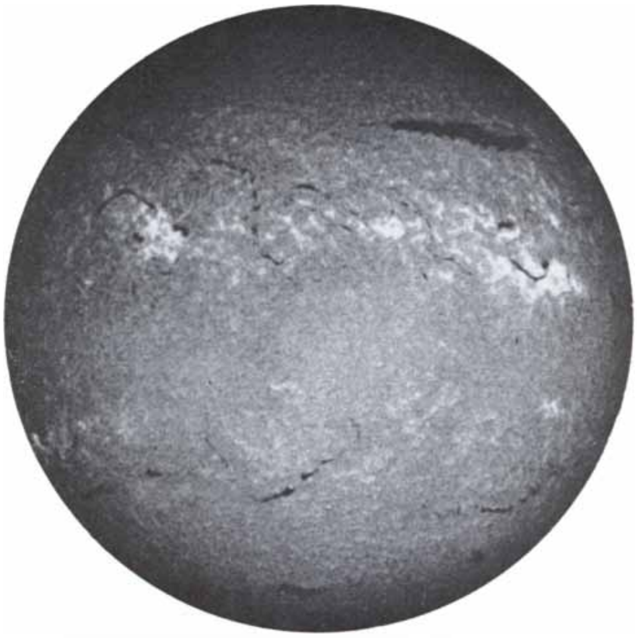
The ability to put telescopes and other instruments in orbit above the earth's atmosphere is giving astronomers an opportunity to observe the sky in a way they could only dream of a dozen years ago. It is well known that the atmosphere blocks out most of the shortwave radiation emitted by the sun and other stars. Thus the radiation reaching the average telescope on the earth has a spectrum only slightly broader in wavelength than the spectrum visible to the eye: 4,000 to 7,000 angstroms. A few infrared "windows" provide glimpses of wavelengths longer than 7,000 angstroms. At the short end of the spectrum the blockage is almost complete below 3,000 angstroms. Consequently astronomers could only guess about the character and intensity of radiation emitted by celestial objects between 3,000 and 300 angstroms in the ultraviolet part of the spectrum. (Wavelengths shorter than 300 angstroms but longer than .1 angstrom are arbitrarily defined as X rays, and wavelengths shorter than .1 angstrom are designated gamma rays.)

The Orbiting Astronomical Observatory (OAO-II), placed in orbit last December, is now providing a systematic sampling of the ultraviolet output of some 50,000 stars. In this article I shall report some of the preliminary findings of OAO-II and explain why astronomers have been looking forward to its results with such keen anticipation. In the broadest sense ultraviolet astronomy, now in its infancy, should help to answer some of the fundamental questions about the universe, including the processes by which stars, and even galaxies, are born, grow old and die. Ultraviolet data should also tell us much about the nature of quasars, the immensely bright starlike objects that seem to be among the most remote beacons in the universe, and about pulsars, the strange stars that flash on and off with precise regularity from once every few seconds to as rapidly as 30 times a second.

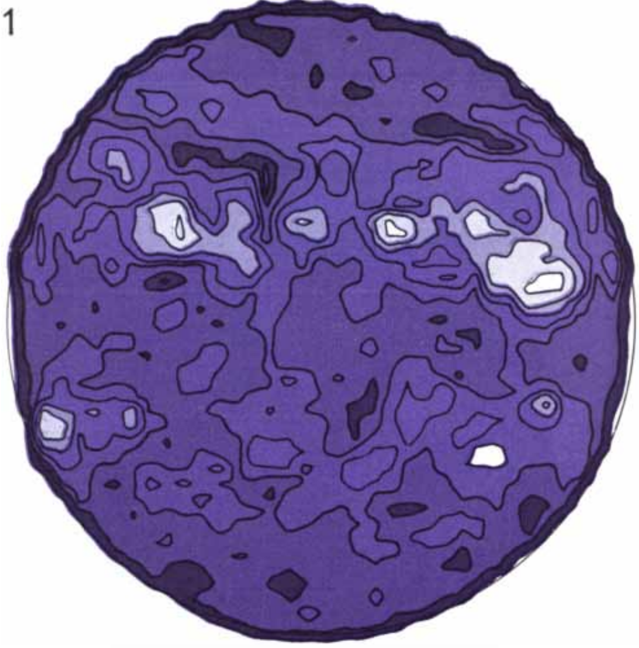
In the not too distant future, perhaps by 1980, astronomers hope to establish a great national observatory in space, or preferably an international one. Among the instruments being considered for such an Astronomical Space Observatory are a reflecting telescope at least three meters in diameter for observing planets, stars, nebulas and galaxies; a number of special telescopes for studying the sun, the largest of which might be 1.5 meters in diameter and nearly 40 meters long; large-aperture telescopes and arrays of detectors for monitoring sources of X rays and gamma rays in the universe, and a fantastic long-wave radio telescope shaped like a rhomboid 10 kilometers on a side. The observatory would be designed as a fully automatic permanent facility, but it would probably require intermittent visits by engineer-astronauts to maintain, repair and modernize the instruments. Measured by its potential for fundamental discovery, the proposed space observatory is by far the most rewarding scientific activity that can be envisioned for the U.S. space program in the next two decades.

The utility of telescopes in space goes beyond their capacity to intercept radiation that cannot penetrate the atmosphere. The air around and above us is in constant turbulent motion, with the result that the images produced by earth-based telescopes are severely degraded. In theory the angular resolving power of a telescope is directly proportional to the size of its aperture, but in practice there is little or no gain in resolving power when the aperture exceeds 12 inches or so. The 200-inch telescope on Palomar Mountain resolves features on the moon that are about 3,000 feet across, but a telescope of the same size above the atmosphere could distinguish craters as small as 100 feet in diameter. The brightness of the night sky places another serious limitation on earth-based telescopes. Even in locations where there is very little dust in the air, auroral emission in the upper atmosphere and the airglow ultimately fog photographic plates and drown out the radiation from the faintest and most remote celestial objects. Telescopes in orbiting observatories above the airglow will be able to collect this faint radiation by long exposures. A space telescope 120 inches in diameter should be able to detect stars 100 times fainter than the faintest detectable from the earth. Data

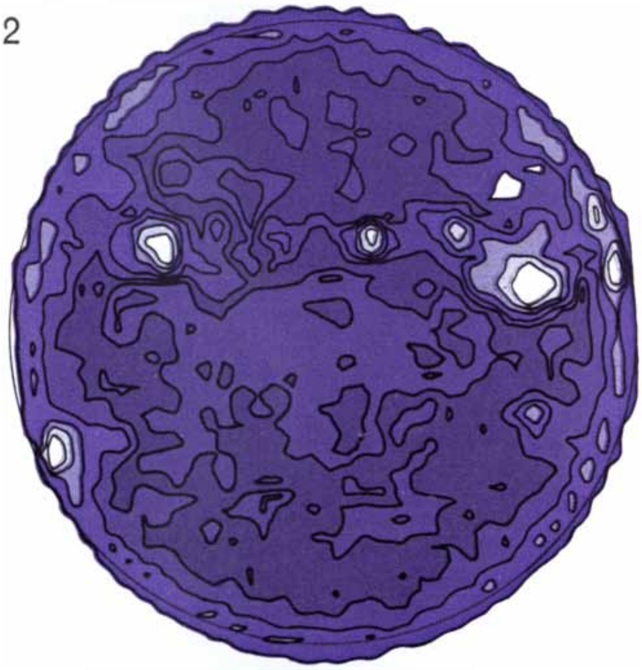
ULTRAVIOLET SOLAR IMAGES, constructed from measurements made by Orbiting Solar Observatory IV, can be compared with a photograph (*top left on opposite page*) taken on the same day at the Sacramento Peak Observatory of the Air Force. The photograph was made by recording a single wavelength: the red line of un-ionized hydrogen. It shows extensive sunspots on the northern hemisphere of the sun's surface and lesser activity in the southern hemisphere. The dark patches are clouds of comparatively cool gas. The five ultraviolet images are numbered in order of increasing temperature, and therefore height, in the solar atmosphere: 10,000 degrees (*1*), 100,000 degrees (*2*), 325,000 degrees (*3*), 1.4 million degrees (*4*) and 2.25 million degrees (*5*). Temperatures are on the absolute, or Kelvin, scale. The last two images respectively correspond to heights of 15,000 kilometers and 200,-000 kilometers. The hottest regions are in the palest color; the coolest regions are dark gray. It can be seen that the sunspots exert an influence throughout the solar atmosphere.



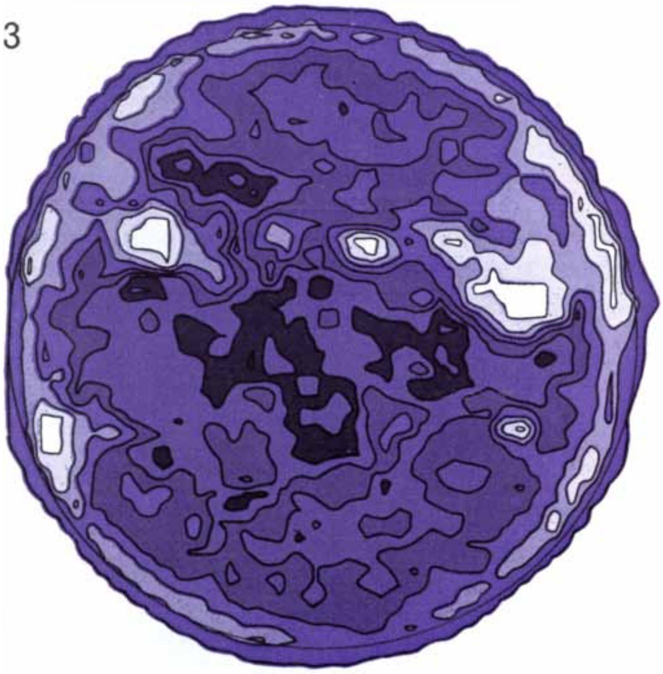
1



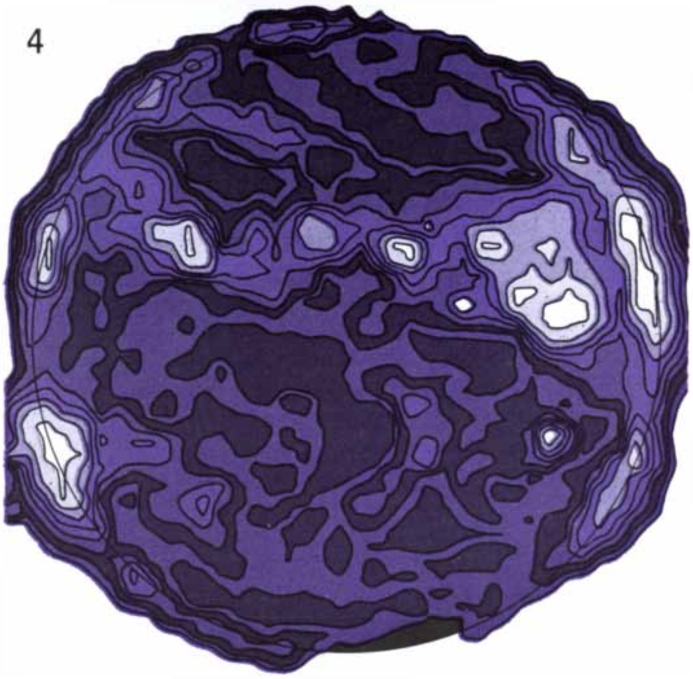
2



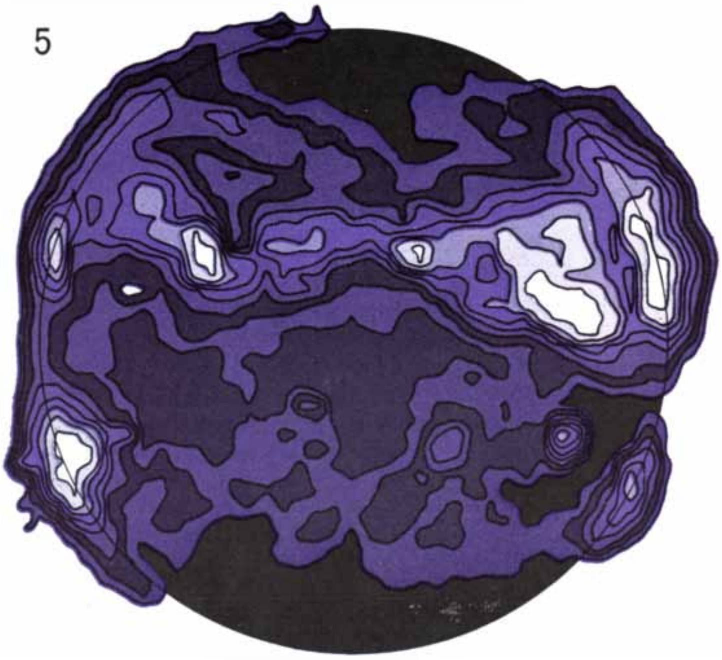
3

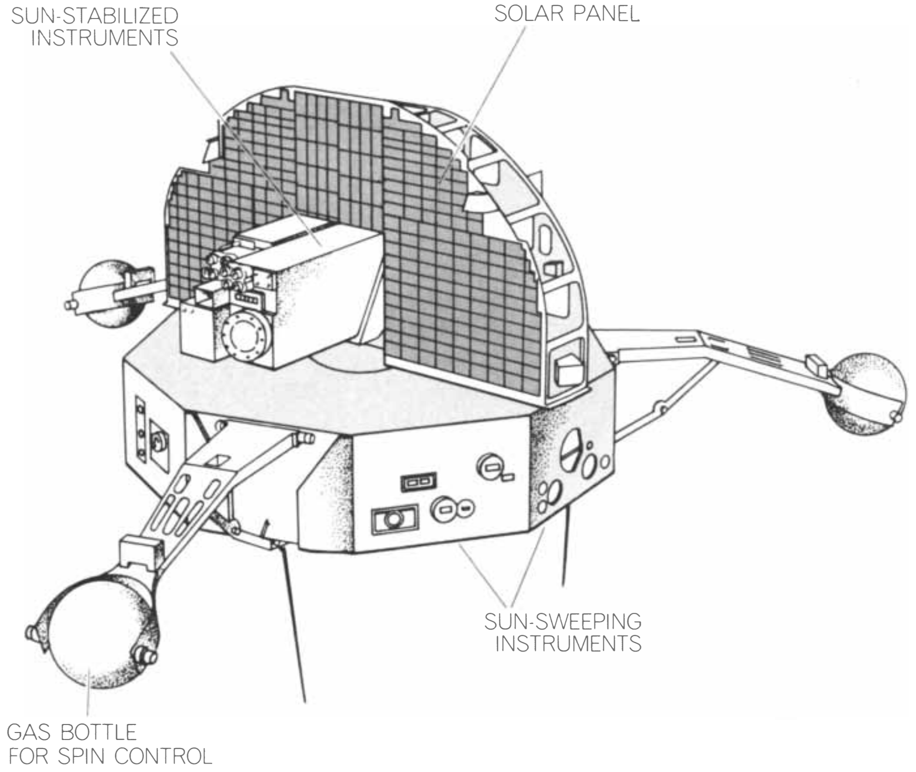


4

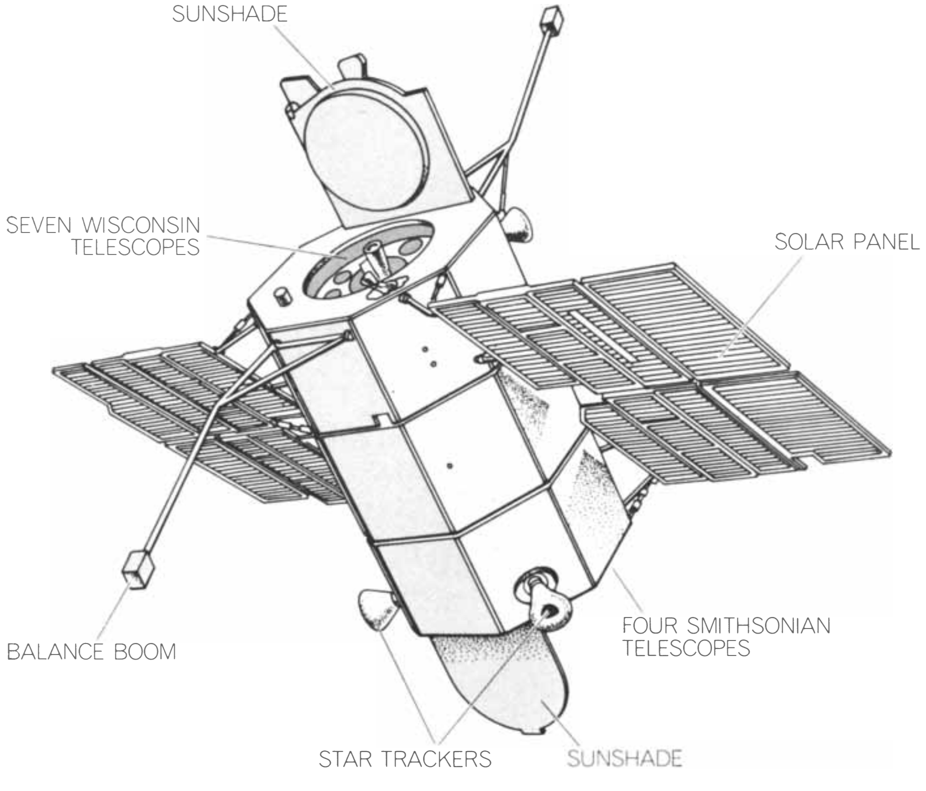


5





ORBITING SOLAR OBSERVATORY IV, launched October 18, 1967, carries some 250 pounds of instruments in a circular orbit 350 miles high. The base section, which rotates 30 times a minute, has instruments for six experiments. The upper section, which can be aimed at the sun, carries an X-ray telescope, X-ray spectrometer and ultraviolet spectroheliograph designed by the author and his colleagues at the Harvard College Observatory.



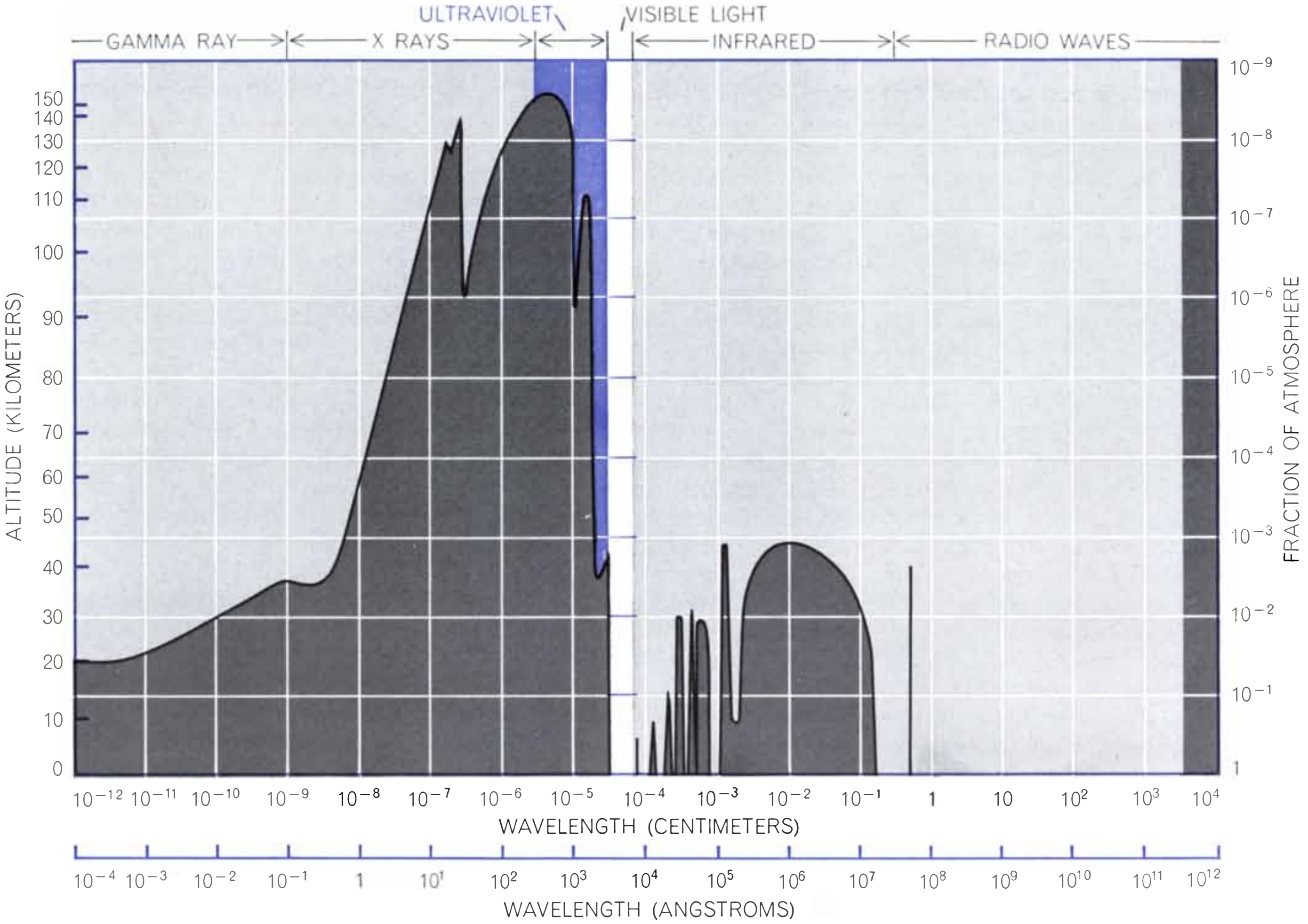
ORBITING ASTRONOMICAL OBSERVATORY II (OAO-II), with a scientific payload of 1,000 pounds, is the most complex unmanned satellite developed by the U.S. It was launched December 7, 1968, into a circular orbit 480 miles high. Designed expressly to measure ultraviolet radiation from stars, OAO-II is equipped with 11 telescopes: three of 16-inch aperture, four of 12-inch aperture and four of eight-inch aperture. Four of the telescopes were provided by the Smithsonian Astrophysical Observatory, seven by the University of Wisconsin.

on faint objects are critical for settling major questions in cosmology, such as whether the universe is infinite or not.

The atoms and molecules responsible for absorbing ultraviolet and infrared radiation in the earth's atmosphere and the heights at which they are found are fully known, so that one can estimate how high it is necessary to go to get above them [*see illustration on opposite page*]. The infrared is absorbed chiefly by water vapor and carbon dioxide, most of which is below 30 kilometers, a height that can be exceeded by stratospheric balloons. On the other hand, ultraviolet radiation is absorbed at much higher altitudes, which can be exceeded only by rocket propulsion. The band of wavelengths between 2,900 and 2,200 angstroms is absorbed by the layer of ozone that is formed by the action of solar ultraviolet radiation on oxygen molecules at an altitude of about 50 kilometers. The spectrum from 2,200 to 900 angstroms is absorbed by molecular oxygen, most of which lies below 100 kilometers. Molecular nitrogen, atomic oxygen and atomic nitrogen absorb ultraviolet radiation of still shorter wavelength, the maximum absorption occurring at about 500 angstroms. At this wavelength very little radiation penetrates below 240 kilometers.

The height of the ozone layer was not well known until the advent of sounding rockets in 1946, and was in fact usually much underestimated. Thus in the late 1920's unsuccessful attempts were made to photograph the solar ultraviolet spectrum from manned balloons at heights up to nine kilometers, and as late as the 1930's it was still thought by some that an unmanned balloon observatory could penetrate the ozone layer and photograph the solar spectrum as far down as 900 angstroms. Space astronomy finally became a reality in October, 1946, when a group at the Naval Research Laboratory headed by Richard Tousey photographed the solar spectrum down to a wavelength of 2,200 angstroms with a spectrograph carried to an altitude of 80 kilometers in the nose cone of a German V-2 rocket. From this modest beginning the frontier of the solar-radiation spectrum has been pushed steadily toward shorter wavelengths by instruments of increasing refinement and complexity mounted in sounding rockets and in artificial satellites.

More recently the development of precision pointing devices has also brought the stellar universe within the range of space telescopes, so that the



ATMOSPHERIC ABSORPTION of electromagnetic radiation is strong in the ultraviolet region of the spectrum, where the hottest stars emit most of their energy. The upper boundary of the dark gray areas specifies the altitude where the intensity of external radiation at each wavelength is reduced to half its original value. Radio waves beyond 50 meters in length are blocked completely.

possibilities for gaining basic new information from space astronomy are now virtually unlimited. Several new subdivisions of observational astronomy have been created, each concerned with its own region of the spectrum and requiring unique observing techniques and instrumentation. The subdivisions include X-ray and gamma-ray astronomy, ultraviolet astronomy and infrared astronomy (although much important work in the last category can still be accomplished from the ground). A fourth subdivision, particle astronomy, is concerned with the detection of cosmic rays from the sun, the galaxy and the extragalactic universe. I shall deal here only with ultraviolet astronomy, beginning with the sun.

The physical nature of the sun has been under investigation for more than 350 years, ever since the sun was first observed through a telescope by Galileo. There are a number of reasons why the sun continues to occupy a major share of the attention of astronomers. First of all, no other star is close enough to appear as more than a point of light in the most powerful telescopes. Thus the sun is the one star out of billions that we can hope to understand in much detail. In addition, the sun now assumes an important role in geophysics for the reason that the flow of certain types of solar radiation (notably X rays, ultraviolet radiation and charged particles) is not at all steady but often exhibits sudden large transients, which are frequently followed by correspondingly violent disturbances in the earth's upper atmosphere and magnetic field. Finally, the closeness of the sun makes it a unique physical laboratory where the behavior of matter can be studied under an enormous range of physical conditions that have not so far been duplicated on the earth. For example, the temperature at the center of the sun, where thermonuclear reactions convert mass into energy, is on the order of 15 million degrees Kelvin. The relatively new discipline of plasma physics owes much to astronomical studies of the motions of ionized and magnetized solar gases as observed with telescopes and spectroscopes.

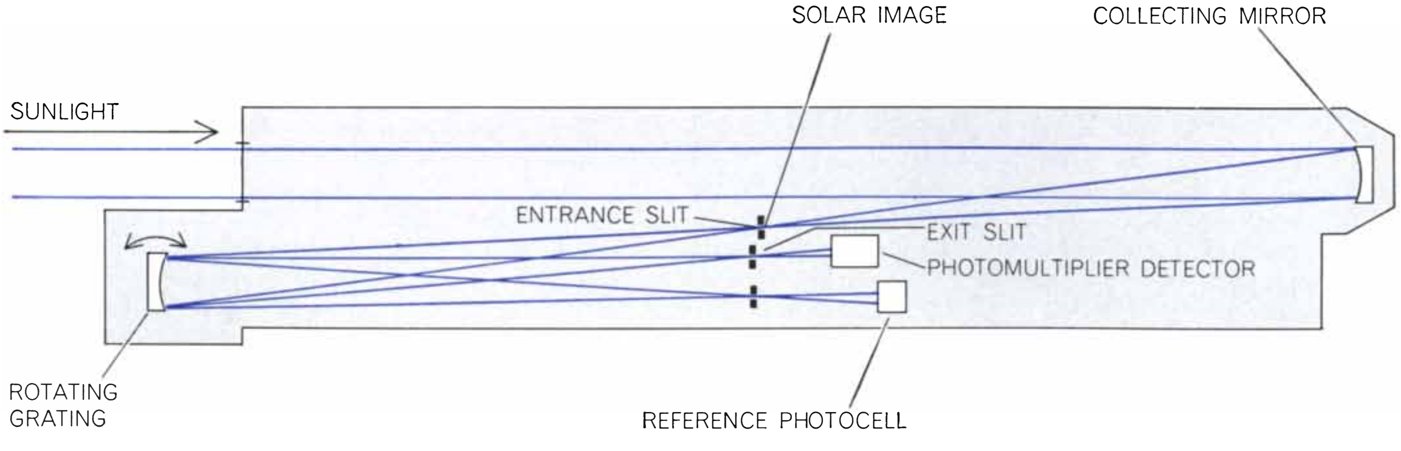
Although a great body of knowledge about the sun had been accumulated by 1940, the modern era of solar physics is only a little older than a quarter of a century, dating from the surprising discovery made by Bengt Edlén that the solar corona is an extremely hot ionized plasma with a temperature of about a million degrees K. Since the temperature of the sun's visible surface is only about 6,000 degrees K., the high temperature of the corona presented a puzzle. Studies over the past two decades have revealed that the temperature actually drops to about 4,500 degrees K. at the top of the photosphere, the layer adjacent to the surface that emits most of the visible radiation. The temperature remains roughly constant for a few hundred kilometers and then begins to rise in the region known as the chromosphere, slowly at first and then very rapidly, jumping from 50,000 degrees to 500,000 degrees in a distance of no more than a few hundred kilometers, and then rising more slowly to temperatures exceeding two million degrees [*see illustration on opposite page*].

For heat to flow outward from cool regions to much hotter regions would seem to violate the second law of thermodynamics. In reality, however, there is no violation because work is being done on the atmosphere by turbulent columns of gas moving upward from an unstable region below the photosphere. The evidence for such large-scale convection is provided by the granulated appearance of the sun's visible surface, which is caused by the contrast between the hot, bright material moving upward and the cooler, darker material moving downward [*see illustration on page 102*]. It is generally accepted that the motions of the turbulent gases are transformed into waves of several types that dissipate their energy in the form of heat as they travel up through the atmosphere, but a detailed theory of the heating process has not yet been verified. Such a theory should be able to predict the rates at which the temperature and the density change with height through the chromosphere and corona. Such predictions are difficult, if not impossible, to test by conventional ground-based observational techniques.

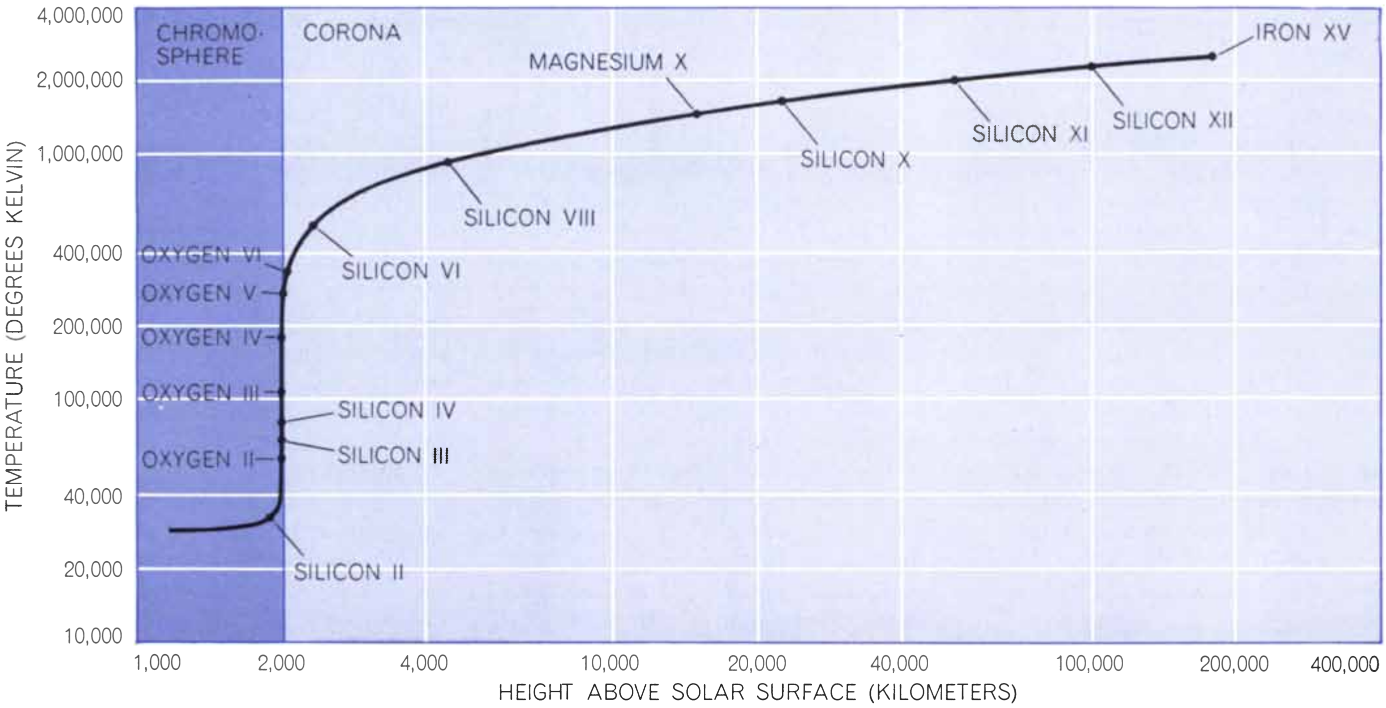
Data on the transition region between the chromosphere and corona are particularly elusive. This is the region where the temperature jumps suddenly from 50,000 degrees to a million degrees within a very short distance. In this zone the number of electrons stripped from atoms is determined by the local temperature; the higher the temperature, the higher the degree of ionization. Because the temperature increases rapidly with height, the total volume occupied by any one species of ion is extremely small and therefore only the strongest emission lines of each species can be detected. These lines fall in the far-ultraviolet region of the spectrum. The highly ionized atoms in the transition zone also emit visible light, but it is too feeble to be observed. Only in the corona itself is visible radiation from highly ionized atoms detectable. Even here comparatively few ions make their presence known in this way, and they do so only from the region peripheral to the solar disk. In order to learn more about the structure of the hot transition regions of the sun, where matter radiates chiefly in the ultraviolet part of the spectrum, it is necessary to mount ultraviolet detectors on space platforms.

Many exploratory investigations of the sun's far-ultraviolet spectrum have been carried out by laboratories in this country and abroad, chiefly with the aid of sounding rockets. These studies have paved the way for the much more advanced instruments carried aloft in spacecraft such as the series of five orbiting solar observatories (OSO's) that have been placed in orbit by the National Aeronautics and Space Administration since March, 1962 [*see top illustration on page 94*]. Each OSO provides a stabilized platform on which about 70 pounds of instruments can be mounted and aimed at the sun with an accuracy of within one minute of arc (about a thirtieth of the sun's diameter). In addition to the section that can be trained continuously on the sun the spacecraft includes a rotating section whose instruments survey the sun intermittently.

In 1891 George Ellery Hale of the U.S. and Henri A. Deslandres of France independently conceived the idea of the spectroheliograph, an instrument for constructing monochromatic images of the sun in individual spectral lines radiated by atoms and ions. The device has been one of the principal means for investigating the structure of the lower part of the chromosphere. When the fourth orbiting solar observatory (OSO-IV) was launched in October, 1967, it carried an ultraviolet spectroheliograph designed and constructed by the Harvard College Observatory to record the spectrum of radiation from the center of the sun's disk in the wavelength region from 300 to 1,400 angstroms, and to produce monochromatic solar images at any desired wavelength in this spectral range [*see illustration below*]. Some of the results of the Harvard experiment are presented here as an example of what is now being achieved by the techniques of solar ultraviolet astronomy. In the far-ultraviolet region of the spectrum, radiation from the photosphere is too weak to be detected. Thus instead of the familiar patterns of the visible spectrum, in which dark lines are superposed on a bright continuous spectrum, the spectrum consists entirely of bright emission lines and bands of continuous emission from the chromosphere and the corona [*see illustration on pages 98 and 99*]. Most of the strong bright lines in that part of the spectrum have been identified, but only a few are labeled in the illustration; they suggest the great variety of ionic species that contribute to the sun's ultraviolet emission. The approximate temperatures at which some of these emission lines are formed can be found in the illustration on the opposite page. The great bulge of continuous emission extending to shorter wave-

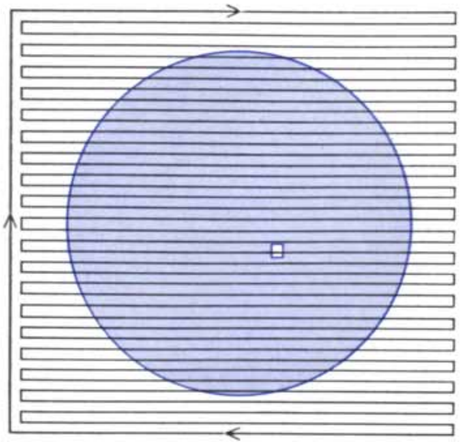


SPECTROHELIOGRAPH aboard OSO-IV supplied the measurements for the illustrations on pages 93, 98 and 99 and on the cover of this issue of SCIENTIFIC AMERICAN. Sunlight enters at the left, is brought to a focus at the entrance slit and strikes an optical grating that separates the radiation into its component wavelengths. By rotating the grating to various positions any narrow band of ultraviolet radiation between 300 and 1,400 angstroms can be centered on the photomultiplier detector. The instrument scans the sun in a



TEMPERATURE OF SUN'S ATMOSPHERE rises steeply some 2,000 kilometers above the photosphere, or visible surface of the sun, which itself is only about 6,000 degrees K. The region where the temperature jumps from 50,000 degrees K. to more than 500,000 divides the chromosphere from the corona. Temperatures are inferred from the spectral lines of various ions, atoms stripped of one or more electrons. The Roman numeral following the name of each element is one greater than the number of missing electrons.

lengths from 912 angstroms is formed when ionized hydrogen recombines with free electrons to form neutral hydrogen at a temperature of about 10,000 degrees K. At the other extreme, bright lines of Fe XV and Fe XVI (iron atoms with respectively 14 and 15 missing electrons) are emitted in highly localized superhot regions of the corona where the temperature may be as high as two million to four million degrees K. Most of the other emission lines are produced by the abundant elements (hydrogen, helium, carbon, nitrogen, oxygen, neon, magnesium, silicon, sulfur and iron) in various stages of ionization; each ion emits radiation at just the height in the atmosphere where the temperature is right for its formation and maintenance.



raster pattern (*far right*), providing a total of 1,920 measurements in a period of about five minutes. The area covered by each measurement is indicated by the white square.

When the Harvard instrument is used as a spectroheliograph, it is able on command from the ground to select radiation from any desired narrow band of wavelengths and by scanning the sun in a raster pattern to construct a picture of the solar regions emitting that particular wavelength. Since each emission line is emitted in a narrow height range, pictures of the sun made in this way outline the three-dimensional structure of the solar atmosphere. The pictures are not made by direct photography because the satellite cannot return the film to the ground. Instead images are reconstructed on the ground from a matrix of brightness measurements made at 1,920 points covering the sun's visible disk and the surrounding inner corona. The brightness measurements are made photoelectrically by a photon-counting system incorporated into the instrument. They are then recorded on magnetic tape and transmitted to the ground by radio at the end of each orbit when the satellite passes over one of the several ground stations along a north-south line between Rosman, N.C., and Santiago in Chile.

Approximately a third of the data were received at Fort Myers, Fla., and were then relayed almost immediately to Cambridge, Mass., over a leased telephone line. Hence astronomers in Cambridge were able to examine pictures of the sun 20 to 30 minutes after they had been received on the ground and to make changes in the observing program as desired. The distribution of brightness across the solar image can be represented in a number of different ways. A typical set of five ultraviolet images, recorded on November 22, 1967, are shown on page 93, along with a photograph made on the same day in the red line of un-ionized hydrogen at the Sacramento Peak Observatory in New Mexico. The two zones of activity that coincide with the sunspot zones extend from east to west in both hemispheres and show up clearly as regions of above-average brightness. The dark areas and patches in the un-ionized-hydrogen photograph are comparatively cool clouds of gas seen in projection against the solar disk.

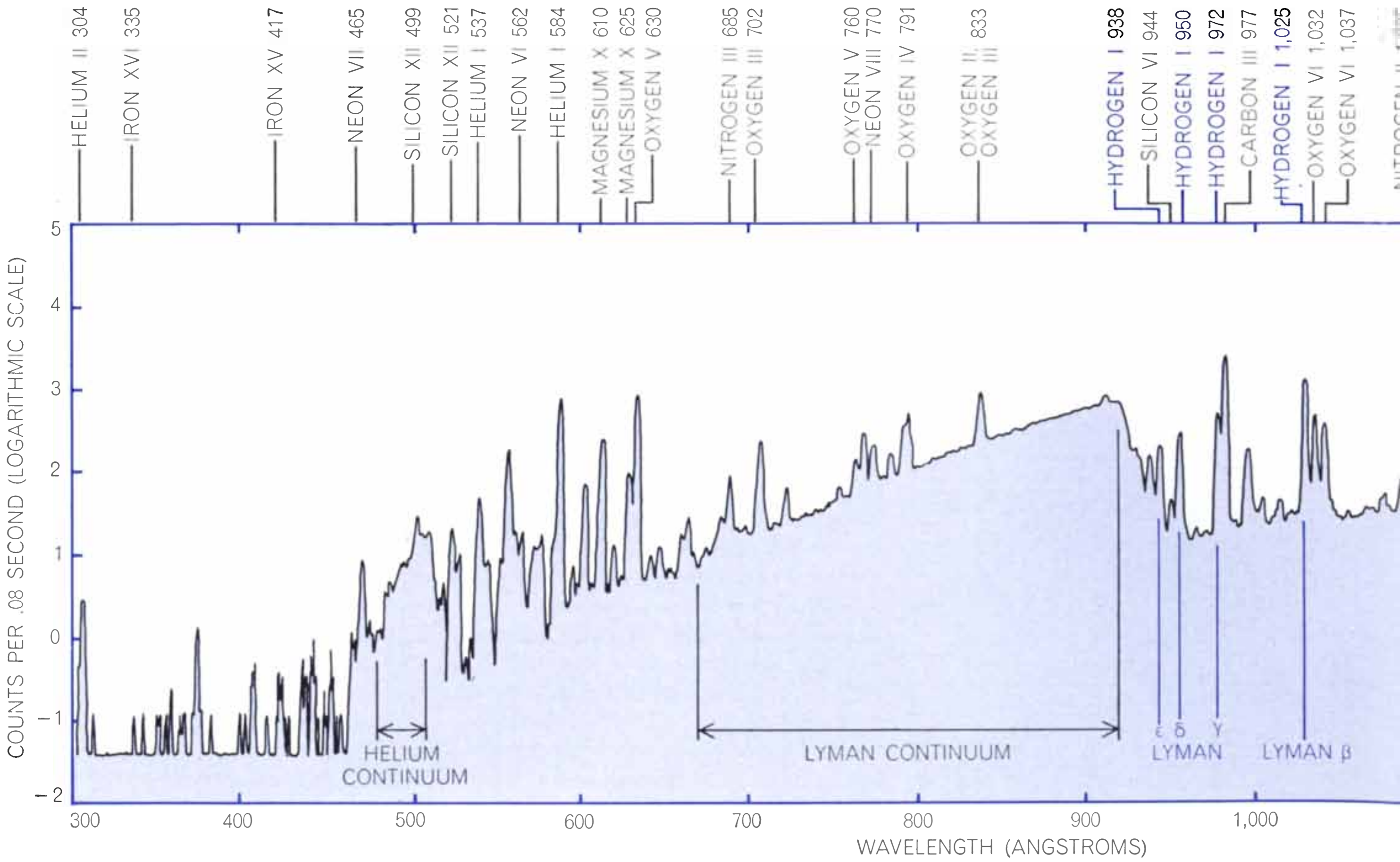
The sequence of ultraviolet images in this illustration is arranged in order of increasing height and therefore of increasing temperature in the solar atmosphere. The first ultraviolet image (Image 1) is formed by the radiation of ionized hydrogen in the chromosphere, where the temperature is about 10,000 degrees K. It shows both the bright active regions and the dark clouds and is generally similar in appearance to the photograph made in un-ionized hydrogen. Image 2 is formed by the radiation of doubly ionized nitrogen at a temperature of about 100,000 degrees K. In addition to the bright centers of activity the image shows a higher contrast between bright and dark regions; a bright ring of emission from the chromosphere-corona transition region is beginning to show along the edge, which takes on a jagged appearance. Image 3 is made in the radiation of oxygen atoms from which five electrons have been removed (O VI) at a height where the temperature is about 325,000 degrees K. The solar edge has now become very bright and above the limb some radiation can be detected from the lower corona, where the temperature is a million degrees.

Image 4 is made in the radiation of magnesium atoms from which nine electrons have been removed (Mg X); this ion is formed at an average temperature of 1.4 million degrees. An image of Mg-X emission is therefore a picture of the corona, which can be seen well beyond the limb as well as in front of the disk. In fact, the size of the image is sharply bounded at the east and west by the instrument's limited field of view. Note the increasing contrast between the brightness of the active regions and that of the surrounding "quiet" corona. Finally, Image 5 is formed by silicon ions (Si XII) radiating at an average temperature of 2.25 million degrees and therefore shows the hotter parts of the corona in the active regions. Both poles, for example, appear to be relatively dark and hence must be considerably cooler than the regions closer to the equator.

The ultraviolet spectroheliograph extends the study of solar atmospheric structure from the low chromosphere to the corona. This new technique will be particularly valuable when it is applied to the investigation of solar flares, which have traditionally been studied with the aid of solar images made in the red line of hydrogen. The series of images on the cover of this issue of *Scientific American* shows the development of a bright flare at the sun's limb as it appeared in the radiation of O VI on November 29, 1967. Each contour interval represents a factor of 1.6 in brightness; thus the flare center was seen to increase in brightness by a factor of about six. Since successive images were made at five-minute intervals, much important activity was probably missed.

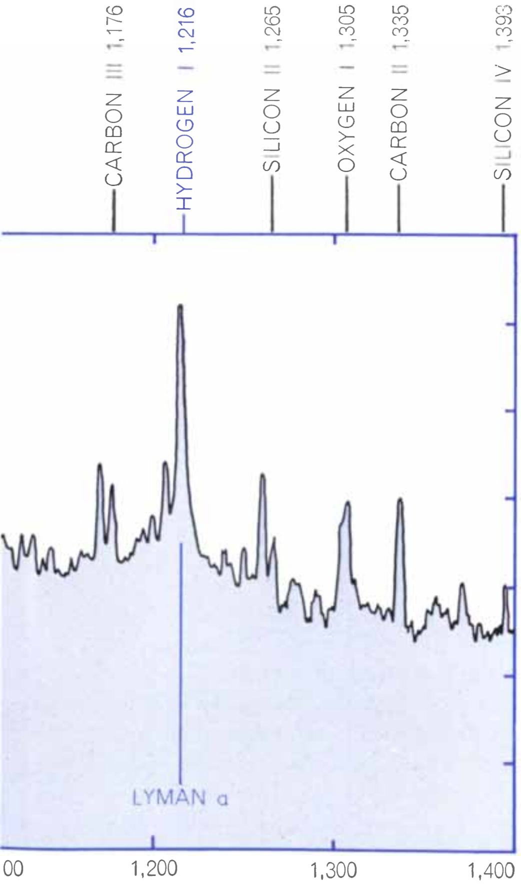
Late in 1969 an improved spectroheliograph, with twice the angular resolution of the present instrument and a time resolution of 30 seconds, will be placed in orbit. In 1972 a still more advanced version of the Harvard instrument will be carried aloft on the Apollo Telescope Mount (ATM), a manned observatory that will also carry other types of solar instrument now being prepared by groups at the High Altitude Observatory in Boulder, Colo., the Naval Research Laboratory, the Goddard Space Flight Center of NASA and American Science & Engineering, Inc.

In the case of the sun extraordinary advances in knowledge have been made with very small space telescopes, no more than an inch or two in diameter and with a pointing accuracy limited to about one minute of arc. For stars, how-



SOLAR SPECTRUM at ultraviolet wavelengths between 300 and 1,400 angstroms was recorded from the center of the solar disk by the Harvard spectroheliograph on OSO-IV. The height of the curve at each wavelength is the number of counts per counting interval of 80 milliseconds. Some of the more prominent emission lines are labeled, including the first five members of the Lyman series of neutral hydrogen (*color*). The great bulge of emission extending from 650 to 910 angstroms is the Lyman series continuum created when ever, the requirements are much more stringent: telescope apertures must be 10 inches or more and errors in pointing no greater than a few seconds of arc. This requirement for larger and more complex instrumentation has kept stellar ultraviolet astronomy from developing as rapidly as its solar counterpart, but the recent successful launching (on December 7, 1968) of the Orbiting Astronomical Observatory (OAO-II) has given enormous impetus to stellar space astronomy.

Considering the huge variety of objects in the stellar and galactic universe, it is almost a certainty that the observations of ultraviolet radiation from the new space platforms will transform our ideas in radical and unpredictable ways. Although the element of unexpected discovery is an important one, there are also well-defined scientific reasons for expecting that stellar ultraviolet astronomy will help us to answer basic questions about the universe. Current ideas about the evolution of stars, derived from observations with large optical telescopes on the earth, suggest that a star begins its life as a condensation of interstellar dust and gas. If such a protostar happens to have the right mass (between a hundredth and 100 times the mass of the sun), it will contract further and become a star. During the life of the star thermonuclear processes convert hydrogen into heavy elements with the release of energy; the more massive the star, the more rapid its rate of energy production and hence the shorter its life expectancy. A star with a mass roughly the same as the mass of the sun may last for about 50 billion years. On the other hand, a very hot, massive star—one weighing 10 times as much as the sun—will radiate 10,000 times more energy per unit time, or 1,000 times more energy per unit mass; as a result its life expectancy will be only a thousandth that of the sun, or about 50 million years. Assuming that the age of our galaxy is about 10,000 million years, there has been ample time for several generations of hot, massive stars to have lived and died.



ionized hydrogen atoms (protons) in the lower chromosphere capture free electrons. A similar but weaker continuum is radiated by ionized helium around 500 angstroms.

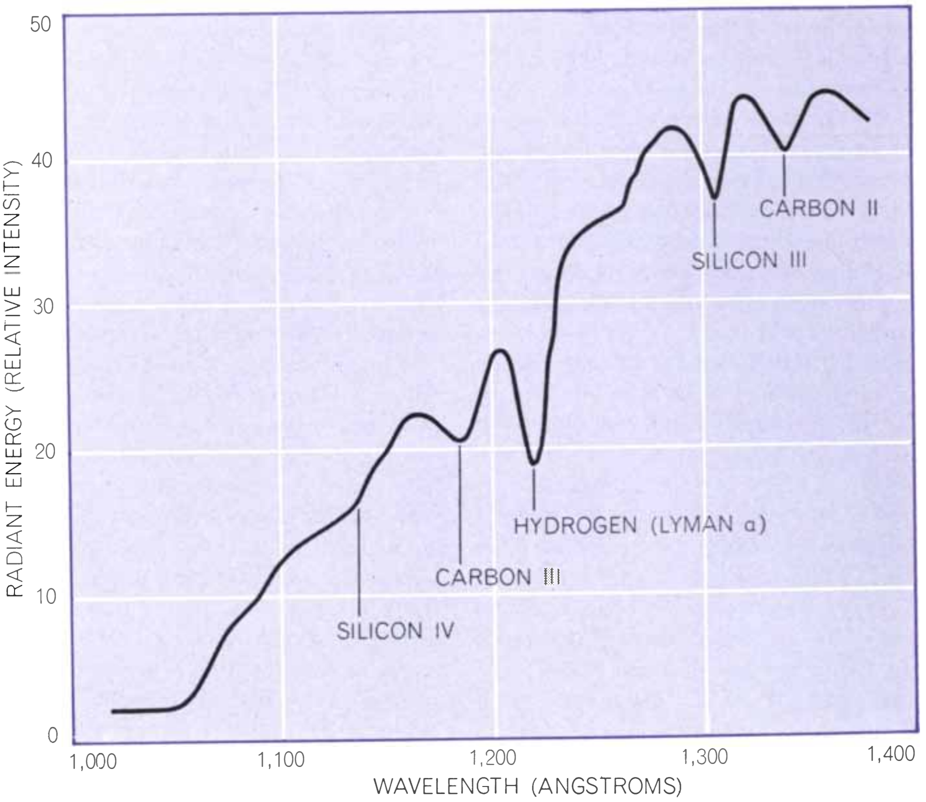
Thus the hottest stars, having short lives, are more likely to show the effects of aging than the cooler ones. Unfortunately for earth-based astronomy only a small part of the radiation of hot stars is in the form of visible light. For example, two stars that appear equally bright to the unaided eye will emit totally different ratios of ultraviolet and infrared radiation if one star is at 6,000 degrees K., like the sun, and the other is at 25,000 degrees [*see top illustration on next page*]. Theoretically the hotter star should emit more than 80 percent of its energy in the ultraviolet region of the spectrum, whereas the cooler star should emit more than 80 percent of its energy in the visible, infrared and radio regions. One must add that calculating such energy-distribution curves in the absence of observational data is a bit like inferring the shape of an iceberg from the part seen above water. Real stars are likely to radiate in quite a different way, and until their ultraviolet energy fluxes are observed we cannot really determine how fast they are evolving.

When a star is born, it consists almost entirely of hydrogen, the fuel that feeds its thermonuclear furnace. When the fuel gives out, the star may explode into a supernova and eject a large fraction of its mass into interstellar space. The ejected material increases the proportion of heavy elements in interstellar space; therefore the hot, massive stars formed a few tens of millions of years ago should contain a higher percentage of heavy elements in relation to hydrogen than the slowly evolving sun, which we believe was formed a few billion years ago. A crucial test of this theory would be the accurate measurement of the chemical composition both of hot stars and of the interstellar gas. Chemical abundances are derived from measurements of the intensity of emission lines in stellar spectra, or alternatively from the blackness of absorption lines. Some of these lines are formed in the atmosphere of the star, as in the Fraunhofer spectrum of the sun; others are impressed on the spectrum when the starlight passes through the interstellar gas on its long journey to the earth. Although some information on chemical composition can be obtained from measurements of the visible spectrum, the most important lines lie in the ultraviolet. It is there that decisive tests of theories of stellar evolution will be conducted.

In reality the intensity of an interstellar absorption line in the spectrum of a distant hot star tells us only the quantity of a given element that is in a particular state of ionization. It is not always possible to observe spectral lines from all the states of ionization in which a given element may exist in the interstellar medium; therefore the fraction of ions in a given state must be calculated from theory. Since atoms in the interstellar gas are ionized by the action of ultraviolet radiation from nearby hot stars, it becomes essential to measure the flux of this radiation. A valuable by-product of the calculation is the evaluation of the density of free electrons in space. Knowledge of this density not only helps us to determine the properties of the interstellar gas but also is vital for estimating the distance of pulsars.

Another question that ultraviolet astronomy may help to answer is whether or not (and if so, to what extent) other stars have coronas and activity cycles like the sun's. The sunspot cycle and its attendant quasi-periodic variation of solar activity is one of the most spectacular features of the sun's behavior. The number of sunspots visible on the disk of the sun at any one time varies on the average in a cycle of about 11 years. The same regions of the sun where sunspots occur are also the scenes of violent storms and eruptions, the most spectacular of which are the solar flares. The activity is much more widespread and violent near the maximum of the sunspot cycle than near the minimum. The activity zones are not confined to the visible surface but extend up through the

SPECTRAL ENERGY DISTRIBUTION of two stars shows the vastly greater flux of ultraviolet radiation emitted by one whose surface temperature is 25,000 degrees K. compared with one whose temperature is that of the sun. The curves, based on theoretical calculations, are plotted so that the areas under the curves represent the total output of radiant energy. The sharp drop in the hotter star's emission at about 900 angstroms indicates where atomic hydrogen in the stellar atmosphere absorbs ultraviolet radiation emitted from the surface.



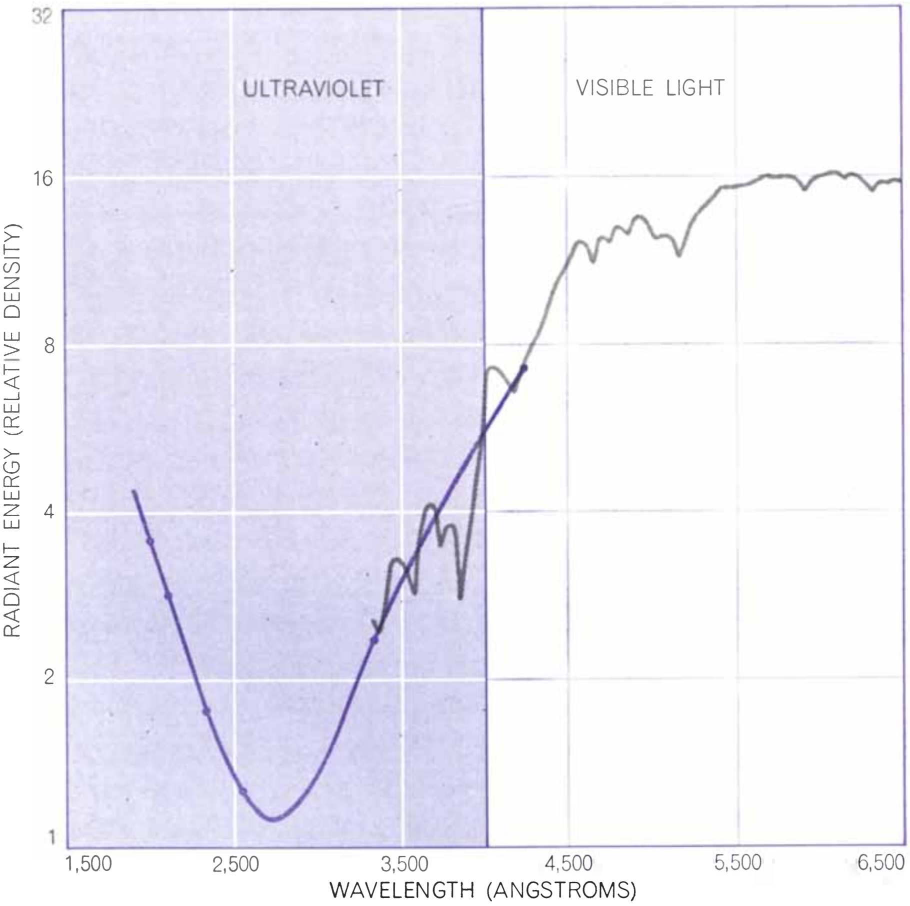
ULTRAVIOLET SPECTRUM of star Alpha Virginis was made by University of Wisconsin telescopes on OAO-II. Four absorption lines are produced by carbon and silicon ions in the star's atmosphere. Hydrogen in interstellar space produced the Lyman-alpha absorption.

corona [*see illustration on page 93*]. The material in sunspots is highly magnetized. It is in fact fairly well established that sunspots and their associated activity are caused by the highly localized buildup of the magnetic fields, which have a strength of several thousand gauss in sunspots compared with an average intensity of very few gauss in quiet regions.

Exactly what causes such periodic accumulations of magnetic fields in the sun is not known, but it seems reasonable to suppose that a large fraction of other stars must have similar cycles of activity and that they too are surrounded by extended high-temperature coronas. Since stars are too far away to show a visible disk, it is unlikely that "star spots" can be directly observed in the foreseeable future. Just as the existence of a high-temperature solar corona can be inferred from the observation of emission lines of highly ionized atoms in its ultraviolet spectrum, however, so might similar observations of stellar spectra reveal the existence of stellar coronas. On the average the intensity of these emission lines should go up and down in unison with the star-spot cycle in a period not unlike that of the solar 11-year cycle, although the length of the period might be quite different. It would be extremely valuable if sunlike activity were discovered in other stars, not only because it might explain the variability of our own star but also because it should help to answer basic questions about the physics of a hot magnetized plasma—a gas of ions and electrons.

Much of what has been said here about the importance of observing the ultraviolet radiation of stars applies equally to external galaxies. In addition to the comparatively normal stars and galaxies, a great variety of objects (whose importance in studies of stellar and galactic evolution far outweighs their small number) can be expected to emit spectacular amounts of ultraviolet radiation. They include several types of star known to have an extended atmosphere; several classes of variable stars, including magnetic stars and novas; the planetary nebulas that surround stars on their way to extinction; remnants of supernovas, such as the Crab nebula, and several types of abnormal external galaxy, including the quasi-stellar objects (quasars) and the Seyfert galaxies (rather normal-looking galaxies with bright, starlike nuclei).

Since 1965 sounding rockets have been equipped with pointing systems capable of aiming a telescope at an individual star so precisely that its spectrum can be photographed or recorded by photoelectric cell during an exposure lasting several minutes. Early measurements made in this way by Donald C. Morton, Edward B. Jenkins and their collaborators at Princeton University with ultraviolet spectrographs flown in Aerobee sounding rockets were a harbinger of the scientific discoveries and surprises that were expected from the more extensive and systematic observations to be made with telescopes aboard large orbiting observatories. The Princeton astronomers found, for example, that the average density of hydrogen between the sun and the stars in Orion's Belt is about .1 atom per cubic centimeter, which is about a tenth the density of un-ionized hydrogen derived from measurements of the hydrogen radio line at a wavelength of 21 centimeters. They have also discovered that certain hot supergiant stars in Orion and Puppis are ejecting a shell of gas containing significant amounts of mass at a speed of about 1,500 kilometers per second. Such definite indications of mass-loss are important in considerations of stellar evolution.



SPECTRUM OF NUCLEUS OF ANDROMEDA GALAXY is a composite of measurements made by ground-based telescopes (*gray curve*) and by a University of Wisconsin telescope on OAO-II (*colored curve*). The rise in the ultraviolet flux was unexpected and implies that the nucleus contains more hot stars than had been suggested by studies from the ground.

The first satellite measurements of stellar ultraviolet radiation were made by the Russians with automatic equipment carried by *Cosmos 51.* Shortly thereafter, in 1966, stellar spectrograms were made with small hand-held instruments by astronauts in *Gemini X* and *Gemini XI.* The first comprehensive results from satellite ultraviolet astronomy began to arrive in December, 1968, with the launching of OAO-II. (The first OAO vehicle was launched in 1966, but a power failure occurred before any measurements could be made.) The second OAO has now been in continuous operation for more than five months, and its performance has been essentially perfect.

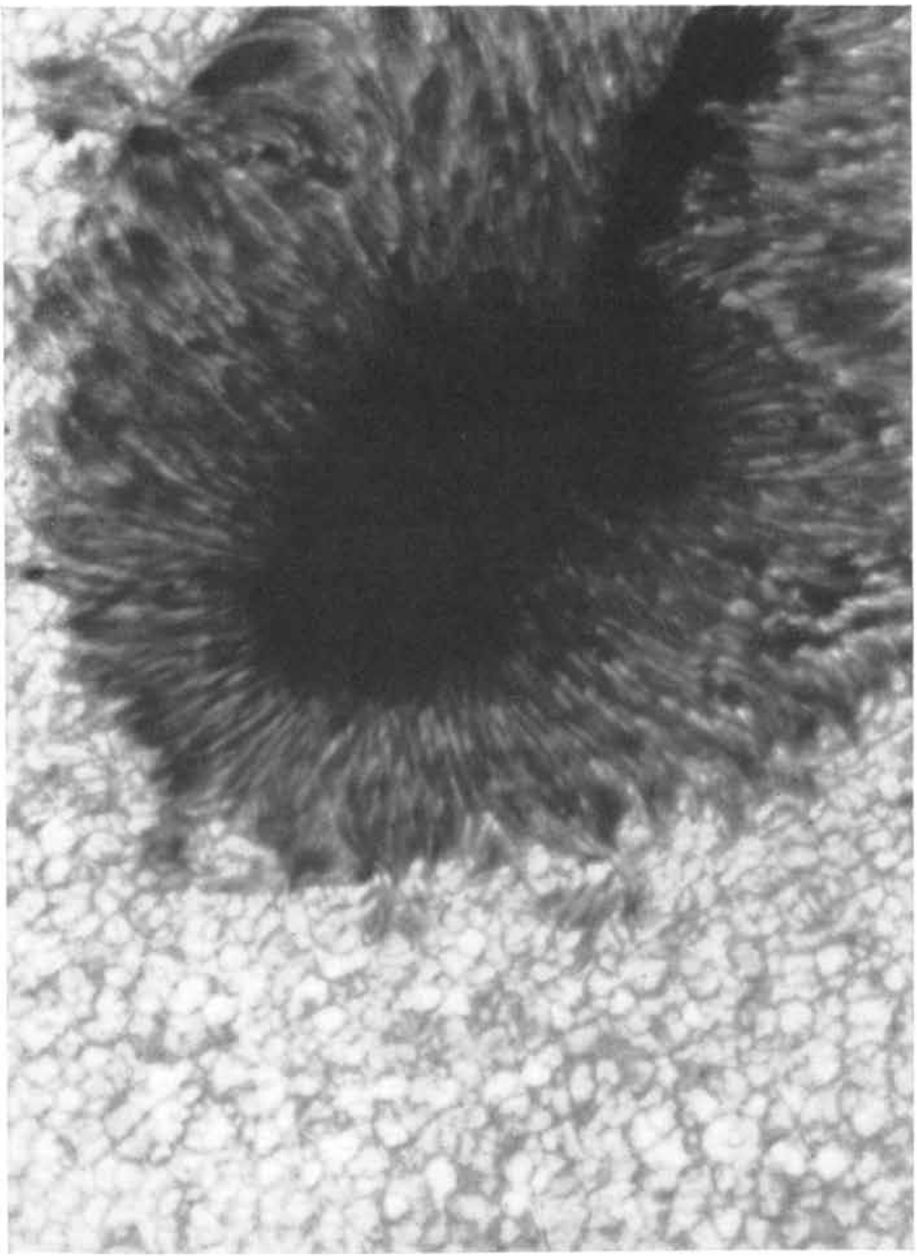
OAO-II is the first of three astronomical observatories that were planned in 1958–1959 as one of the earliest programs initiated by NASA during the first year of its existence. It carries 11 small telescopes, four of them supplied by the Smithsonian Astrophysical Observatory and seven by the University of Wisconsin. The objective of the Smithsonian Observatory's Project Celescope is to map the entire sky in four different ultraviolet wavelength bands centered on 1,400, 1,500, 2,300 and 2,700 angstroms. The mapping is being carried out with four telescopes of 12-inch aperture, each equipped with an appropriate combination of optical filter and television camera tube to isolate and record one of the four wavelength bands. It is estimated that some 50,000 stars will be recorded by the Smithsonian telescopes. Three of the Wisconsin telescopes have a diameter of 16 inches, and four smaller telescopes have a diameter of eight inches. These telescopes are being used to make accurate photoelectric measurements of the ultraviolet spectra of stars, of nebulas within our own galaxy, of other galaxies and of planets.

The OAO-II experimenters were rewarded with unexpected and provocative findings within the first few weeks of the satellite's launching. According to Robert J. Davis, principal experimenter of Project Celescope, about 1 percent of the stars observed so far are between six and 40 times brighter in the ultraviolet than had been expected. The stars in the Pleiades, believed to be a collection of very young stars and therefore bright in the ultraviolet, turn out to be brighter than anticipated by a factor of from three to six.

Arthur D. Code, director of the Washburn Observatory of the University of Wisconsin, has generously given me two series of spectral scans made by the Wisconsin instruments aboard OAO-II. The first is a scan in the region between 1,000 and 1,400 angstroms of the spectrum of the bright star Alpha Virginis (Spica) [*see bottom illustration on opposite page*]. Although the resolution of the spectrum is comparatively low, a number of lines can be clearly identified. Note particularly the Lyman-alpha line of hydrogen, which is absorbed out of the spectrum when the starlight passes through a column of interstellar hydrogen gas between Spica and the solar system. The average density of the un-ionized hydrogen in this path is about one atom per cubic centimeter, or about 10 times greater than the density found by Morton and Jenkins in the directions of the constellations Orion and Puppis.

The second ultraviolet scan supplied by Code [*see illustration above*] is a plot of the absolute energy radiated by the nucleus of the great spiral galaxy in Andromeda in the region between 2,000 and 4,000 angstroms. The more detailed curve depicting wavelengths longer than 3,300 angstroms was obtained from ground-based observations, which had suggested that most of the visible radiation from the nucleus, or central region, of the Andromeda galaxy is to be attributed to cool and old giant stars with a surface temperature of between 4,000 and 5,000 degrees K. If the galaxy's nucleus were populated solely by such stars, the energy curve would be expected to continue its downward trend from the visible to the ultraviolet. The reversal of the trend reveals the presence of a certain number of much bluer and therefore hotter stars. These hotter stars are presumably younger than the majority of the stars in the galactic nucleus.

The ultraviolet radiation from the Andromeda galaxy is also of much importance in accounting for the ionization of the interstellar gas in the galaxy's nucleus and in attacking problems of cosmology. The external galaxies comprising the observable universe are receding from our own local group of galaxies, of which the Andromeda galaxy is a member, with speeds up to a substantial fraction of the velocity of light. Furthermore, the velocity of recession, as measured by the shift of the spectrum toward the red, is proportional to the distance, as would be expected if all the galaxies had originated in a gigantic explosion 10 to 20 billion years ago.



SUNSPOT, photographed from a balloon at an altitude of 80,000 feet, has a cool center that shows up dark in white light. The granulated appearance of the surrounding disk is caused by convection currents in the sun's atmosphere. This photograph is one of several thousand made by Project Stratoscope, directed by Martin Schwarzschild of Princeton University.

The effect of a red shift, of course, is to displace a galaxy's ultraviolet radiation toward the visible region. Code comments that if all galaxies radiate as strongly in the ultraviolet as the Andromeda galaxy does, the most distant ones, whose red shifts are greater than 60 percent of the rest wavelength, would appear bluer at visible wavelengths than nearby galaxies with very small red shifts. This means that if all galaxies are about equally old, the more distant ones would appear younger than those in our local group by an amount equal to the difference in the travel time of the light from the galaxies, which could amount to a billion years or more. Measurement of the colors of distant galaxies should therefore tell us whether these objects have spectral characteristics different from those of the Andromeda galaxy, and thus whether they were formed at an earlier time or whether they are significantly different in their physical makeup. In either case the implications for cosmology are profound.

Neither the Smithsonian instruments nor the Wisconsin ones have sufficient resolution to record the number of spectral lines required, for example, to study the relative abundances of the elements. Investigations of line spectra will be conducted on the two subsequent missions of the OAO program. Both OAO-III and OAO-IV will carry 36-inch telescopes and their associated spectrographs. OAO-III is being instrumented by the Goddard Space Flight Center and OAO-IV by Princeton. The Princeton project is aimed at measuring the relative abundances of the elements in the interstellar medium.

The emergence of so many new opportunities for exploring the universe has also focused attention on the greatly increased importance of the traditional methods of observation with optical and radio telescopes on the earth. The cost of space telescopes is such that they must not be used to make observations that can be accomplished more efficiently and cheaply on the ground. Although certain kinds of problem can best be attacked by one technique or the other, the most rapid advances will surely be made by coordinated observing programs that exploit both ground and space telescopes to observe the same object over the entire range of its spectrum, from gamma rays to long radio waves. It is this prospect and not merely the establishing of an observatory in space that makes the future of astronomy so appealing not only to astronomers but also to a growing number of physicists.

(Announcing)

"TRAINING THE HARDCORE"

The first complete, practical guide for the hiring and training of the hardcore unemployed. In twelve volumes. Published by the Urban Research Corporation, under the sponsorship of The Urban Coalition and The National Alliance of Businessmen.

"Training the Hardcore" is based upon studies of the experience of 200 companies, including actual formats of training programs from such companies as Ford, Bell Telephone, North American Rockwell, Armour, Milwaukee Road, Polaroid, KLH, Philadelphia Gas Works, and others.

THE CONTENTS

Volume

1 Introduction
2 Case Studies
3 Developing a Training Capacity
4 Recruitment
5 Training Courses
6 Orientation for the Hardcore
7 Support Services
8 Orientation for Other Employees
9 The Roles of Top Management
10 Community Relations
11 Union Relations
12 Government Support and Funding

Policies which have led to failures as well as those which have led to successes are both presented in detail. Particular emphasis is also given to interviews and conversations with "hardcore unemployed" persons themselves; this on the theory that the persons in whose names training programs are instituted know the most about what sort of program will be meaningful to them and what sort will not be.

The twelve volumes contain more than 1,400 pages, delivered in loose-leaf binders, which are updated monthly as additional information becomes available.

Published in cooperation with General Learning Corporation. $485.00.

If you would like to inspect Volume 1, which contains a thorough summary of the subsequent volumes, kindly forward the coupon. There is, of course, no charge for this preliminary inspection, but should you decide against buying the complete set, we ask that Volume 1 be returned to us.

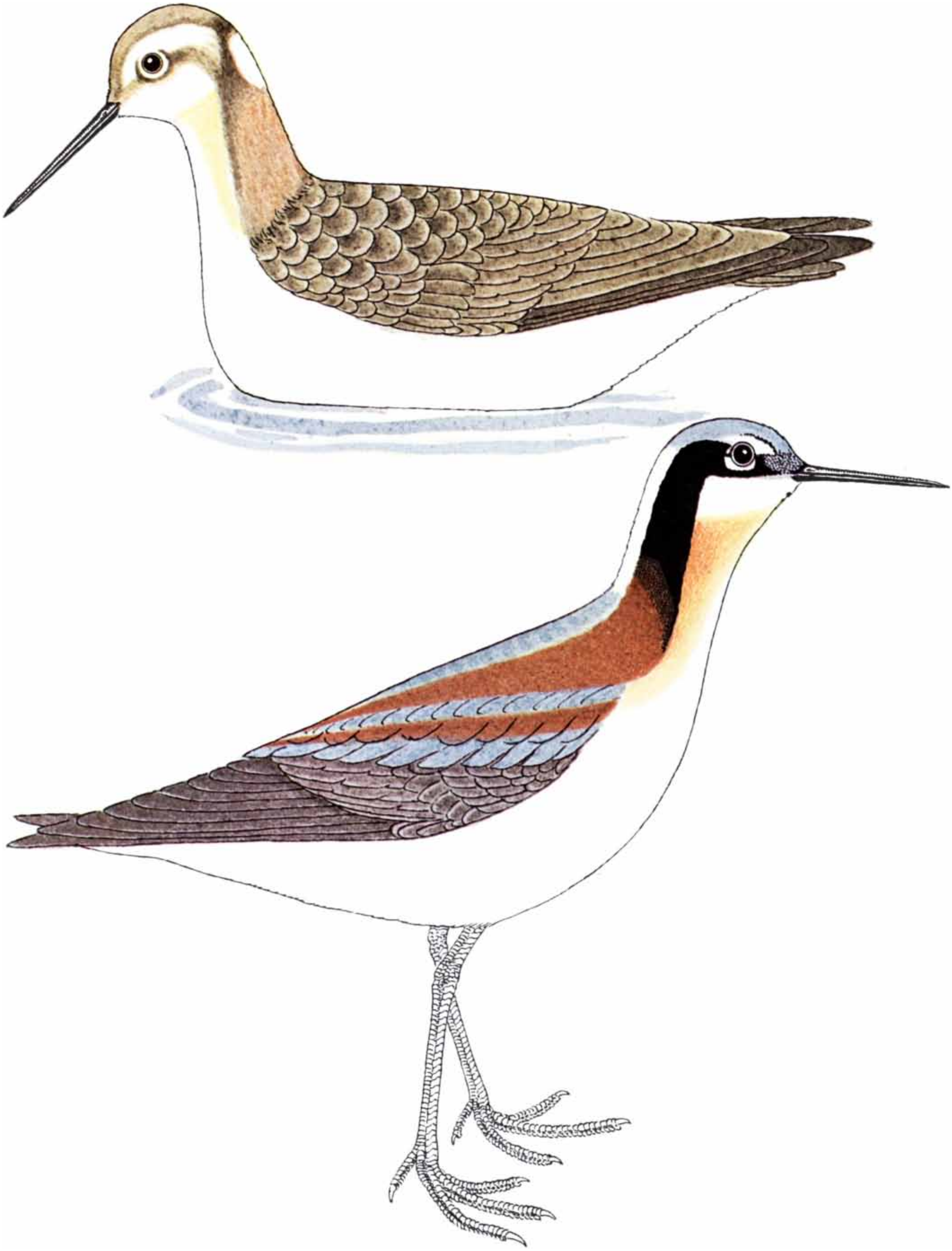
Thank you.

Please send me, for purposes of acquainting myself further with "Training the Hardcore," one copy of Volume 1.

Name____________________
Position____________________
Company____________________
Address____________________

Urban Research Corporation
5464 South Shore Drive, Chicago 60615
(312) 955-3050

© 1969 SCIENTIFIC AMERICAN, INC



BREEDING-SEASON PLUMAGE of the female phalarope (*bottom*) is more colorful than the plumage of the male (*top*). In this somewhat stylized rendering the birds are drawn to actual scale; the female is larger as well as more colorful than the male. These birds are Wilson's phalarope (*Steganopus tricolor*), a species that breeds in Canada and over much of the U.S. west of the Great Lakes.

THE PHALAROPE

The reproductive behavior of this aquatic shorebird is most unusual. The female phalarope aggressively courts the male; the male phalarope warms the eggs and rears the young birds

by E. Otto Höhn

The difference or similarity in coloration between a male bird and a female of the same species usually reflects the behavior of the species in courtship and breeding. In species where the male bird is more colorful than the female, the male tends to play a dominant role in courtship and a minor one in rearing the young. If the breeding plumage of the two sexes is similar, the female usually participates actively in the events of courtship and the male has a large share in the care of the eggs and fledglings. Then there are a few species in which the female is more colorful than the male, and here the female takes the initiative in courtship and the male alone rears the young. In this last group is the small, long-necked shorebird known as the phalarope.

In the process of selecting a mate the brightly colored and (compared with the male) large female phalarope is the aggressor. Soon after the eggs are laid the female bird typically leaves the breeding ground. Then for 21 days the small, dull-colored male phalarope incubates the eggs and for 10 more days looks after the downy young. Shortly before the incubation period the male bird sheds feathers from its abdomen and the bare skin thickens and becomes engorged with blood. This "brood patch" enables the male to warm the eggs and the young birds.

In addition to its unusual reproductive behavior the phalarope has other singular characteristics. For one thing, in shallow water, where other shorebirds wade, the phalarope swims. Floating high in the water, it looks as if it were made of cork. The phalarope is an excellent swimmer, and associated with this ability are lobed membranes that extend from the sides of its toes. These flexible paddles are much like the ones found on the highly aquatic coot and grebe but are considerably smaller. The structure of the phalarope's foot is responsible for the bird's name: the genus *Phalaropus* is from the Greek for "coot-footed."

As it swims about the phalarope daintily picks up bits of food from the surface of the water or searches for food below the surface by dipping its head into the water, in the manner of surface-feeding ducks. At times the phalarope engages in an unusual feeding maneuver: it rotates like a top, dabbing at the water with its beak as it turns. Specimens of Wilson's phalarope (a species named after the 18th-century American naturalist Alexander Wilson) have been observed spinning as fast as 60 revolutions per minute, spearing the water with their beak once in each revolution. These rotations may serve more than one purpose, depending on environmental circumstances. In shallow water the spinning movement could stir up food particles lodged on the bottom. In deep, cold pools on Greenland, which are frequented by phalaropes, the spinning is believed to activate mosquito larvae, making them more conspicuous to the birds. In some instances the rotations may simply be due to an abundance of food. Turning to the side for a bit of food, the bird may see another morsel farther to the rear, and after snapping up the first it continues the turning movement to reach the second.

There are three species of phalarope;

ARCTIC PHALAROPES are the two other phalarope species. At top is the northern phalarope (*Lobipes lobatus*) in summer breeding plumage; at bottom is the red phalarope (*Phalaropus fulicarius*), which is also seen in breeding plumage. The birds are females.

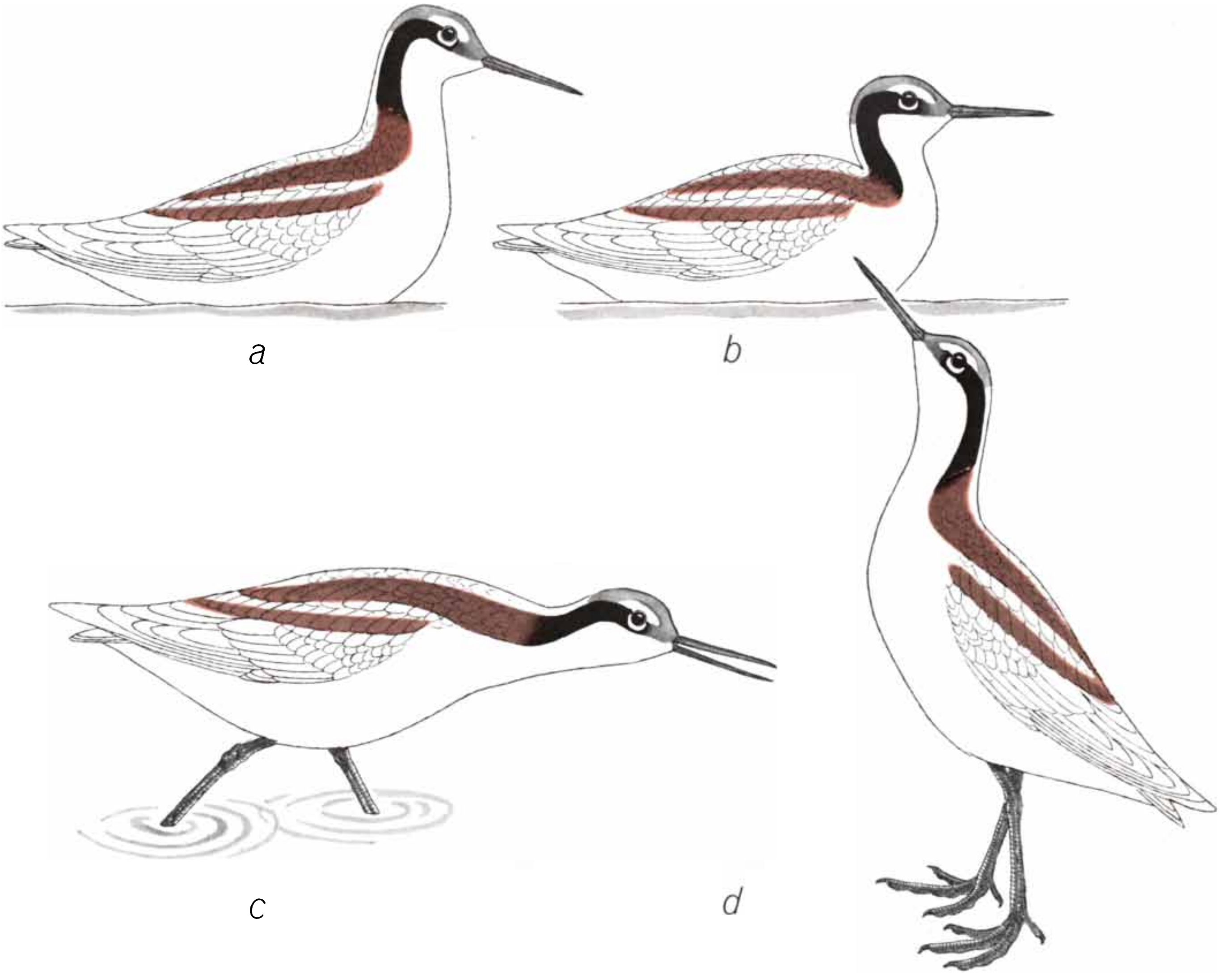
they belong to the order Charadriiformes, which includes plovers and sandpipers. One species, the red phalarope (*Phalaropus fulicarius*), breeds on the tundra of the arctic zone in both the Eastern and the Western Hemisphere. The nest of the red phalarope is never more than a few miles inland, and the birds spend the nonbreeding season—the winter—at sea, often many miles from land. They feed on crustaceans and other small animals of the marine plankton, and they rest and sleep afloat. Thus apart from the summer breeding season they are true seabirds. The wintering areas of this species are off the east and west coasts of South America and the west coast of Africa [*see top illustration on page 108*]. On its migrations the red phalarope apparently flies over the sea; the birds are rarely seen on land south of their breeding grounds.

A second species, the northern phalarope (*Lobipes lobatus*), nests in the subarctic region of the Eastern and Western hemispheres. This species tends to build its nests farther inland than the red phalarope. It too winters at sea. Its wintering areas, which are nearer the Equator than the red phalarope's, lie off the coasts of Peru, Argentina, northwest Africa, southern Arabia and Indonesia.

Wilson's phalarope (*Steganopus tricolor*) is the most familiar of the three species. It is confined to the Americas. The breeding grounds of Wilson's phalarope extend from the Great Lakes westward across the prairie provinces of Canada into British Columbia and farther south over a large area of the western U.S. The main wintering area of the bird is inland on the pampas of Argentina.

All the phalaropes have a dull fall and winter plumage, gray above and white below. It is the same in both sexes. In the breeding plumage of spring and summer other colors appear, although more faintly in the male bird [*see illustration on page 104*]. The only other bird species in which the female is more colorful than the male are the tropical and subtropical painted snipes (*Rostratula*) and the quaillike hemipodes (*Turnices*) of the Eastern Hemisphere. The female is somewhat larger than the male in all three phalarope species. The average weight of Wilson's phalaropes in a group I collected during the breeding season was 2.4 ounces for the females but only 1.7 ounces for the males.

I have observed the reproductive behavior of Wilson's phalarope in southern Canada and the behavior of the other phalarope species along Hudson Bay and in Alaska. The process of pair for-



POSTURES OF THE PHALAROPE are associated with reproduction. During the period of pair formation the female Wilson's phalarope stays near a particular male, normally swimming with her head up (*a*). A threat posture is assumed (*b*) when another female phalarope approaches. While walking, the female bird assumes a different threat posture to ward off intruders (*c*). The head-up pose (*d*) is taken by a pair of phalaropes before their first mating. It probably signifies that no threat is intended.

mation is best known in Wilson's phalarope. In a group of the birds one sees a female keeping close to a particular male, following him as he swims here and there. When another female approaches the pair, the first female lowers her head and pulls it in until her long neck is scarcely visible. This posture is a threat gesture: it sets the stage for a forward lunge. With her head down the threatening female swims purposefully toward the intruder. When she is about a foot away from her, she takes wing and with her neck extended and her legs dangling flies at her. Usually the threatened female retreats, and after a brief chase the threatening bird returns to the vicinity of the male. If one female threatens another while the birds are walking on land, the threatening bird extends her head—sometimes with her bill partly open—toward the enemy. Actual fights, however, are rare. I have seen only one. The two birds hovered in the air beak to beak, each trying to stab the other as they fluttered with their legs dangling. The aerial fight lasted for only a few seconds; the intruding female gave up and flew away.

It appears, then, that the female phalarope chooses a particular male and wards off other females by threat and, if necessary, attack. Thus the male bird becomes accustomed to the company of a particular female and later directs his sexual advances toward her. Male phalaropes show very little aggressive behavior at any time.

Before two phalaropes mate they sometimes face each other with their heads raised, the bill of each bird aimed at a point above the other's head. For a few seconds the birds hold this pose. Since the phalarope threatens and fights with its beak, the posture apparently signifies that no threat is intended. The female then indicates readiness for mating by taking a crouching position. The male rises up like a miniature helicopter, hovers over the female's back and mating follows. The raised-beak display seems to be a prelude only to the first mating between two individuals. When the birds are more familiar with each other, the display is omitted; the male simply hovers aloft and lands on the female's back.

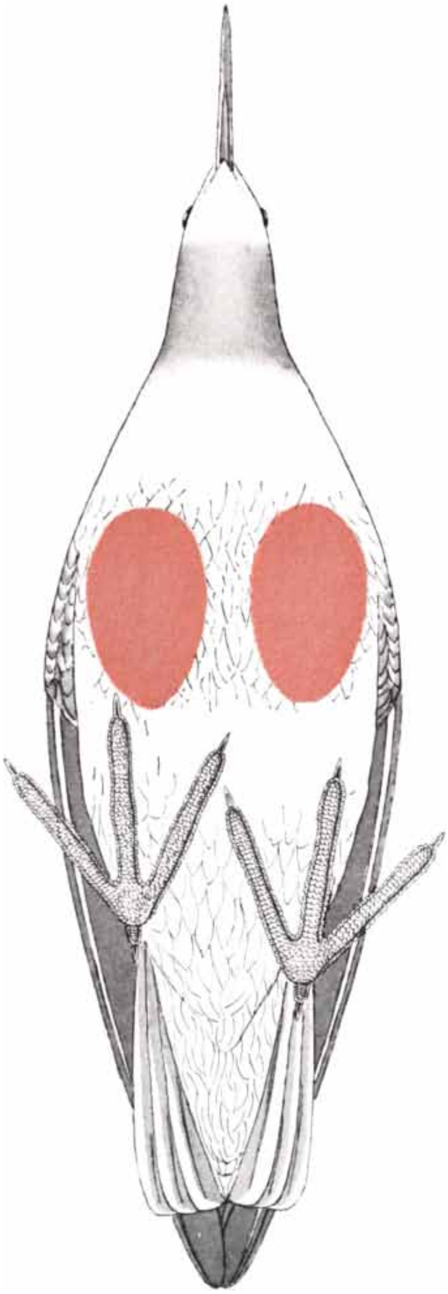
What is the biological significance of the reversed roles of the sexes in the phalarope? At the University of Alberta my colleagues and I have been studying the bird to learn the physiological factors that underlie its behavior. The greater aggressiveness of the female phalarope compared with the male suggests that in the earlier phases of the breeding cycle the female may secrete more androgen (male sex hormone) than the male does. There are many studies of animals that link aggressiveness to male sex hormones. It has been shown that among birds the administration of androgen increases the aggressive behavior of doves, female herring gulls and castrated young male chickens.

That androgens play a role in the production of the phalarope's breeding plumage has been shown by J. E. Johns of Montana State University. In his experiments with Wilson's phalaropes and northern phalaropes in their drab winter plumage he plucked feathers from the shoulder of the birds. He then found that if the plucked phalaropes were given the male sex hormone testosterone, they grew new feathers like those of the normal female in breeding plumage, regardless of whether the experimental bird was a male or a female. Phalaropes that received no hormone after plucking grew feathers of the dull winter plumage, similar to the feathers that had been plucked. Other experimental phalaropes were given a female sex hormone, namely estradiol, which in combination with the pituitary hormone prolactin induces formation of the brood patch. Whether administered alone or together, these hormones did not give rise to female nuptial plumage.

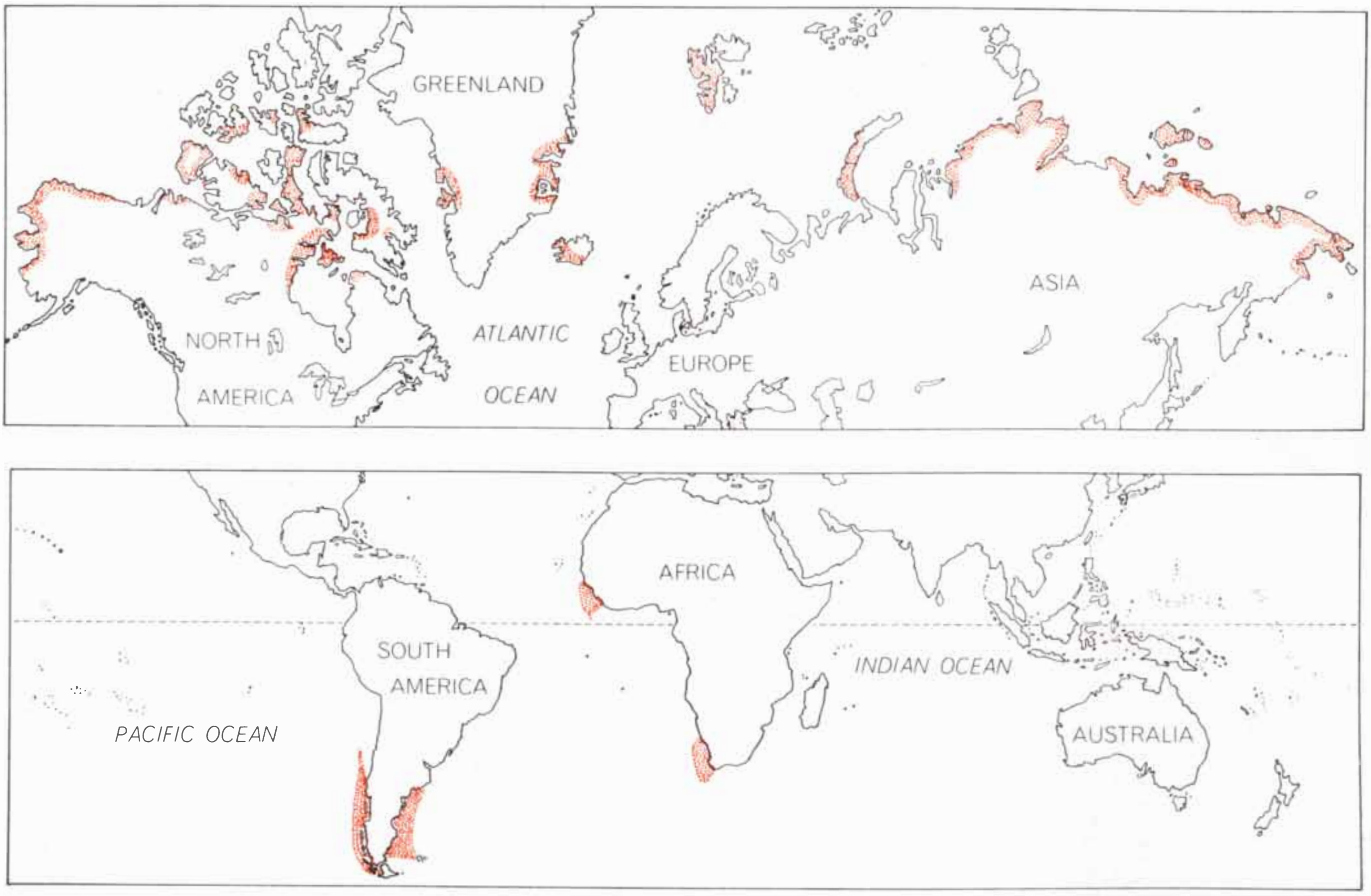
It remained to be shown that the female phalarope does in fact produce large amounts of androgen. My colleagues and I collected ovaries and testes from a group of phalaropes and, for comparison, from other bird species in which the male is the dominant sex: mallard ducks, domestic fowl and red-winged blackbirds. We isolated the individual hormones (by means of chromatography) from extracts of the birds' gonads and measured the amount of each hormone. In birds where the male is dominant the amount of testosterone in the testis was five to 10 times greater than the amount in the ovary. In phalaropes, however, there was about as much testosterone in the ovary as in the testis, and in some instances there was more testosterone in the ovary.

Evidently the brilliant plumage of the female phalarope is due to the high production of androgen by her ovaries. The same factor is probably responsible for the aggressive behavior of the female. We ruled out a possible involvement of hormones of the adrenal cortex, because the amount of androgen in this gland was slightly greater in the male than in the female both in phalaropes and in the control birds (mallard ducks).

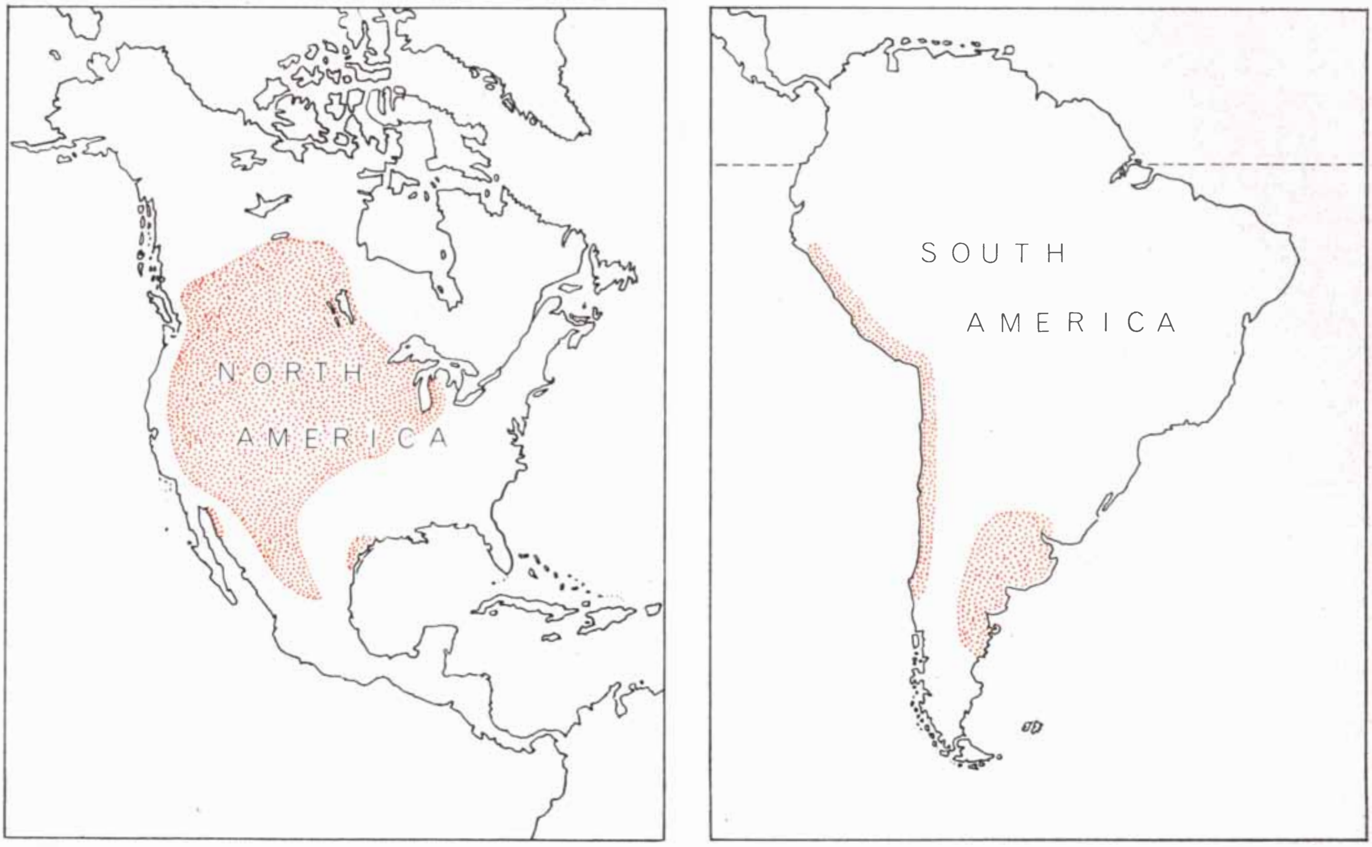
How it is that the male phalarope and not the female develops a brood patch and incubates the eggs has also found an endocrine explanation. It has been known for some time that an estrogen, or female hormone, together with prolactin will cause finches of both sexes to develop a brood patch. (The brood patch normally appears only on the female finch.) In phalaropes we found that only a combination of androgen and prolactin



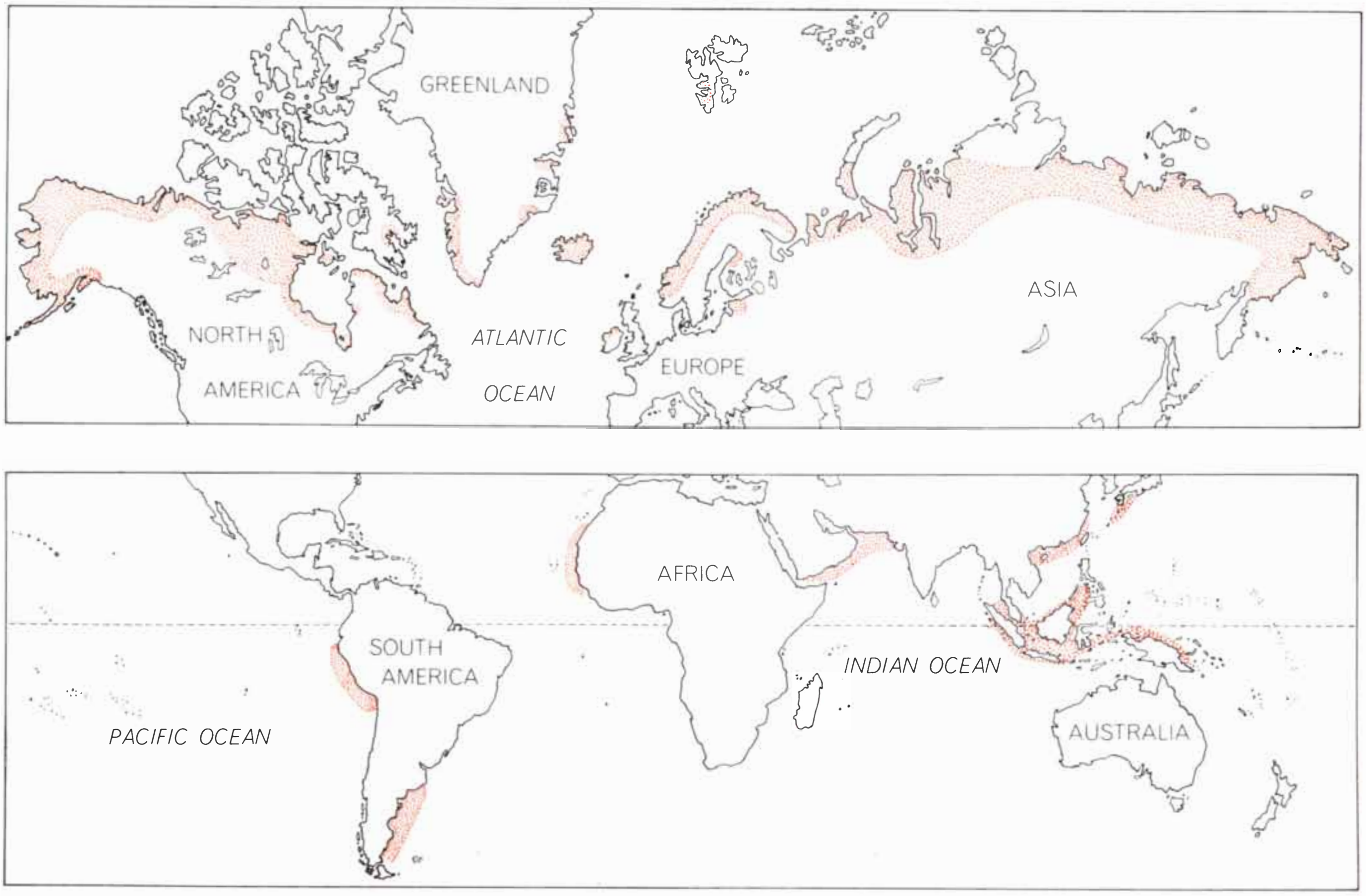
BROOD PATCHES appear on the abdomen of the male phalarope before the period of incubation. The exposed skin, which becomes engorged with blood after the feathers are shed, is brought in contact with the eggs and with the young birds. The female phalarope does not produce brood patches.



DISTRIBUTION OF RED PHALAROPE indicates the bird's arctic breeding grounds (*top*) and the areas where it spends the arctic winter (*bottom*). The nests of this species are never more than a few miles inland and their migratory flights are over the sea.



DISTRIBUTION OF WILSON'S PHALAROPE (which is named after the 18th-century American naturalist Alexander Wilson) shows the bird to be in North America during the breeding season (*left*) and in South America during the nonbreeding season (*right*).



DISTRIBUTION OF NORTHERN PHALAROPE indicates that this species tends to nest south of the red phalarope (*top*) and to winter nearer the Equator (*bottom*). Like the red phalarope, the northern phalarope feeds on small animals of the marine plankton.

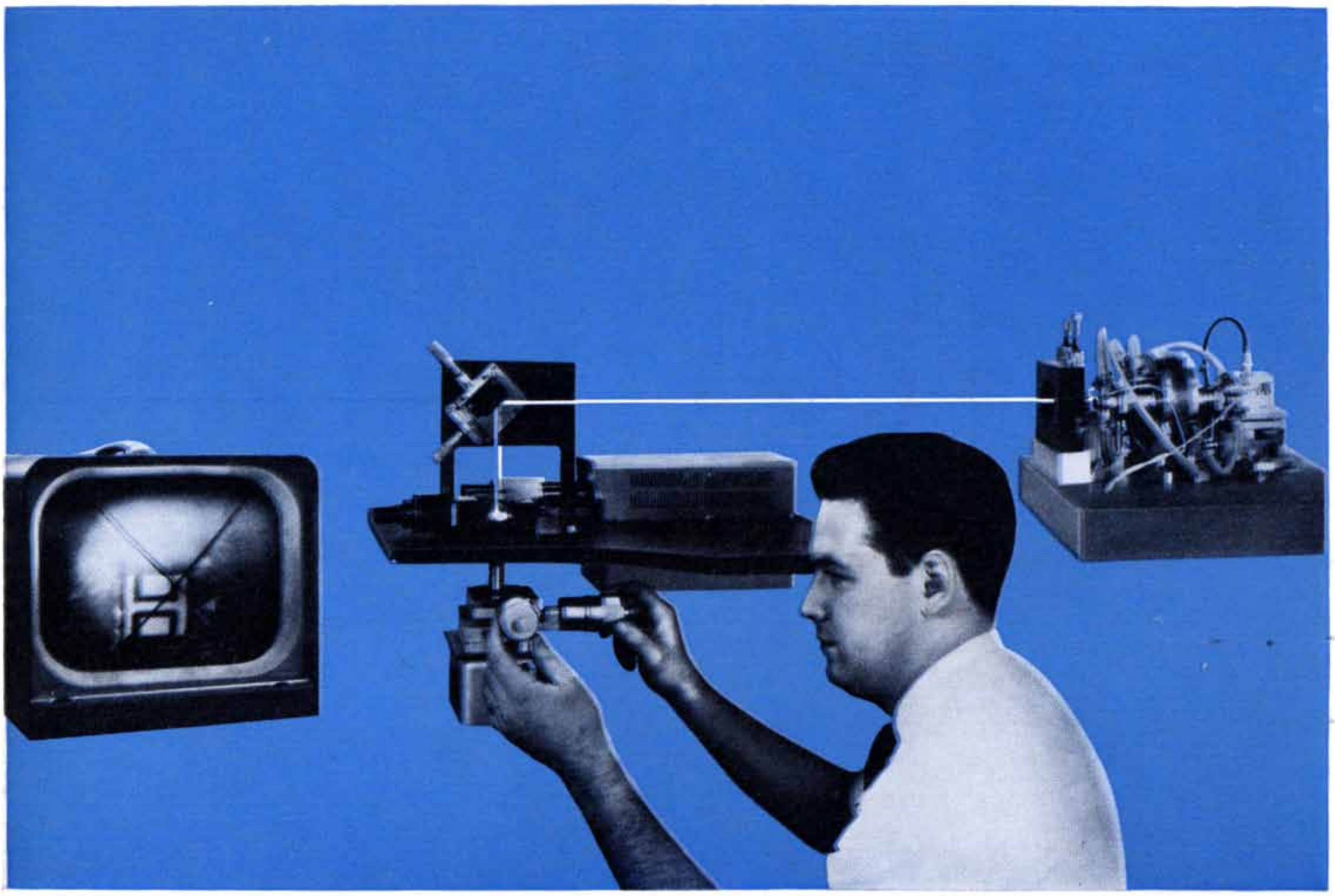
could induce the formation of brood patches, both in females, which in nature do not develop these patches, and in males at a time of the year when they do not naturally form them.

Our earlier investigations had shown that the female phalarope is an efficient producer of androgen. The failure to develop brood patches could hardly be due to a lack of androgen, but it might result from insufficient production of prolactin. Prolactin is readily detected with a test employing the crop sac of the pigeon, which is a dilated section of the pigeon's gullet. In a nonbreeding pigeon the lining of the crop sac is thin and smooth, but in a breeding pigeon, just before the young hatch, the tissue of the lining proliferates into a wrinkled mass known as the crop gland. (The surface of the crop gland, sometimes called "pigeon's milk," is shed by the pigeon and eaten by its young.) This proliferation results from the increased secretion of prolactin from the pigeon's pituitary gland. The same effect can be achieved by injecting a sufficient amount of prolactin; therefore whether or not a given sample contains this amount of the hormone can be determined by injecting it into the skin over the crop sac and observing whether or not the crop gland develops.

With this technique we tested an extract prepared from the pituitaries of 39 breeding male phalaropes. Injected into a test pigeon over a four-day period, the extract gave rise to a response similar to one produced in another pigeon by 10 micrograms of prolactin. A similar extract from 40 female phalaropes did not affect the crop sac at all. Apparently the pituitary of the female phalarope secretes much less prolactin than the pituitary of the male. This result was confirmed by further experiments. We employed test pigeons that had been made more sensitive to prolactin by the injection of a subthreshold dose of the hormone; such a dose does not affect the crop sac, but it reduces the amount of prolactin required for a response. Tests with pigeons thus prepared indicated that during the breeding period the pituitary of the male phalarope contains about three and a half times more prolactin than the pituitary of the female. Evidently the female phalarope fails to develop brood patches because of an insufficiency of prolactin, one of the two hormones required for the formation of the patches. This lack of prolactin is almost certainly the reason female phalaropes also have no urge to incubate their eggs.

One may speculate on how the partial reversal of the sexual roles of the male and female phalarope evolved. Suppose that in an ancestral phalarope both sexes had an androgen-dependent colorful breeding plumage, and that both sexes shared in rearing their young, as many living shorebirds do. Then suppose that a hereditary inability to secrete enough prolactin to form brood patches and to induce the urge to incubate eggs arose, perhaps as a mutation, in certain females. Pairs of birds in which the female was so affected could still reproduce if the male entirely took over incubation and the rearing of the young. Freed from the urge to incubate, these somewhat nonmaternal females would leave the breeding grounds earlier than normal females. Then they would be less exposed to predation, since they would not have to remain in the nest to incubate the eggs and rear the helpless young. As a result more of these females would survive to reproduce again.

The exposure of the male bird to predation while incubating the eggs and rearing the young would tend to favor individual males with an inconspicuous breeding plumage and thus with a



Micromachining with the laser

Bell Laboratories engineers M. I. Cohen and B. A. Unger have developed experimental techniques for using lasers in certain delicate thin-film integrated circuit work: machining circuit patterns, making "gap" capacitors, trimming tantalum thin-film resistors and monolithic quartz resonators, and cutting masks for circuit fabrication.

Our experimental system (above) combines a solid-state YAG (yttrium aluminum garnet) laser, manual positioning of the circuit, and television observation. The optical part of the system was developed by Western Electric's Engineering Research Center, located at Princeton, New Jersey.

The high spectral purity of the continuous-wave YAG laser, invented at Bell Laboratories, lets us focus the light to a very small spot for precision cuts less than 5 microns (1/5 mil) wide and resistor trimming accurate to better than 0.1 percent. And, through Q-switching, the YAG laser produces high peak power at high repetition rates—over 1,000 pps—giving us the cutting speed necessary for practical circuit work.

Laser beams pass through any transparent atmosphere or material and can be accurately concentrated onto tiny areas. With the proper wavelength, we can machine components inside a transparent encapsulation without damaging it. Also, since we can regulate cutting depth, we can "micromachine" thin films without harming underlying materials.

To make capacitors, for example, Cohen and Unger use a laser to cut (vaporize) a narrow gap between conductors. In gold conductors on sapphire or alumina substrates, they have cut gaps from 5 microns to 600 microns wide with good control.

Similarly, Bell Labs engineers have adjusted thin-film quartz crystal resonators to frequencies as precise as one part in 10^8. The laser vaporizes part of the thin-film electrode, raising the resonator frequency to the desired value.

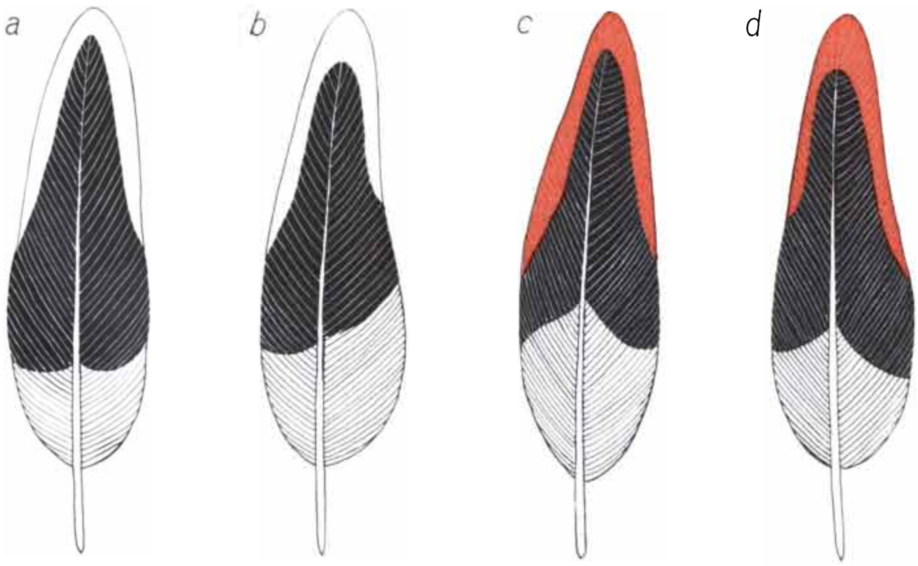
By removing hairline shorts, we have also repaired expensive integrated circuits that could not be reclaimed by standard techniques.

Pioneered at Bell Laboratories and Western Electric, laser micromachining is already in pilot and volume production use at Western Electric and other major integrated circuit manufacturers.

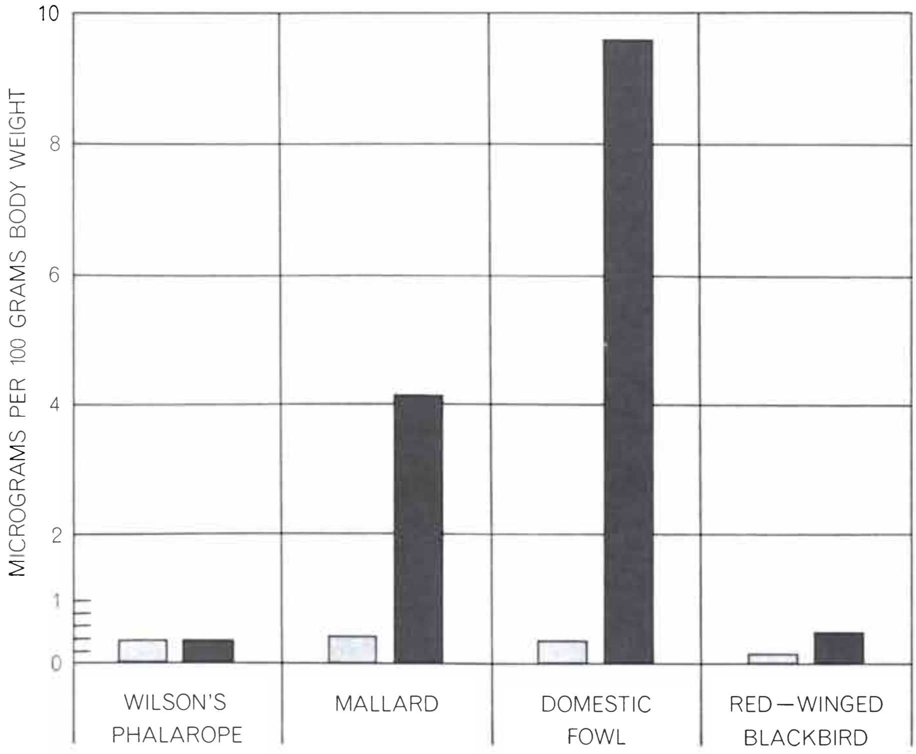
From the Research and Development Unit of the Bell System—

Bell Labs

low secretion of androgen at the time the plumage appeared. Brightly colored feathers would not, however, be a disadvantage to a female bird that had been liberated from the duties of the nest. On the contrary, a more colorful female would probably be more attractive to males and perhaps more dominant in encounters with females lacking a mate. Selection, then, would favor a female bird with an enhanced androgen secretion at the time breeding plumage is grown. The continuation of this increased androgen production into the early breeding season would account for the aggressiveness with which the brightly colored female phalarope courts the rather pallid male.



PLUMAGE IS TRANSFORMED when phalaropes receive testosterone (male sex hormone). The shoulder feather of a northern phalarope in dull winter plumage looks much the same whether the feather was grown in the wild (*a*) or in captivity (*b*). When a bird in this plumage is plucked and given testosterone, the new shoulder feather that is grown (*c*) resembles the shoulder feather of the female of the species in breeding plumage (*d*). The bright-colored female plumage is produced whether the phalarope is a male or a female.



DISTRIBUTION OF TESTOSTERONE in male (*dark gray*) and female (*light gray*) birds is based on measurements of the amount of the hormone in their gonads. In the species compared with the phalarope the male is more colorful and more aggressive than the female.

Analytic Instruments in Process Control

Self-regulating manufacturing processes call for systems that sensitively monitor the performance of the process. An example of such a system is a gas chromatograph linked to a computer

by F. W. Karasek

The manufacture of ketchup on a large scale provides a homely example of a practice that is becoming increasingly common in industrial technology: the use of instruments to control processes more or less automatically. Commercial ketchup manufacturing once required a human operator to keep an eye on the batch so that cooking could be stopped when the mixture of tomatoes, spices and water had reached the right consistency. Now samples of ketchup from the vat can be continuously passed through a refractometer, which determines the consistency of the mixture optically and turns off the heat when enough water has been evaporated.

This theme of advancing capabilities in the closely related functions of measurement and control runs through the entire history of technology. In a number of industries, particularly the manufacture of chemicals and the refining of petroleum, the techniques of combining measurement and control have reached a highly automatic stage with the introduction of digital computers that receive information from a number of measuring instruments and accordingly control the quantities of substances being fed into the manufacturing operation and the steps taken to process the substances. It is more than likely that the trend toward highly automatic control of manufacturing processes will accelerate in a number of industries.

Instrumental control of processes in the chemical and petroleum industries has been going on for some 50 years, but for much of that time the technique involved only single-function instruments that measured temperature, pressure or flow and actuated simple control mechanisms. By about 1945 it was beginning to be realized that this type of control was crude compared with what could be done by means of sensors that would continuously analyze a process stream and provide detailed information on its composition. The result was the development of a number of powerful process instruments.

Two additional technologies, closely related to process control, were required before process control could be achieved in an optimum way. One of them is computer technology, particularly as it relates to small digital computers devoted to a single task, such as the control of a manufacturing operation. The other is control technology, which involves studying the controllability of processes in such ways as understanding the reactions that are taking place, determining how flows respond dynamically to changes in valve positions and working out the mathematics of processes. These two technologies have evolved rapidly during the past 10 years, with the result that it is now possible to conceive of systems that will control a process at its optimum, whether that optimum is defined in terms of economy, purity, yield or a combination of such objectives.

A designer faced with the problem of creating an automatic analytical instrument for a process stream naturally turns to principles that have been found useful in the laboratory. Usually the instrument that results embodies the principles of the laboratory instrument but looks quite different. The reason is that the production instrument must operate more or less continuously over long periods, often in a harsh environment; hence the instrument must be much more rugged than the laboratory model. Of the many laboratory principles that have been employed in process-control instruments, the three in widest use today are refractometry, absorption spectroscopy and gas chromatography.

Refraction is of course the bending of a beam of light as it passes from one medium into another. In refractometry one essentially measures the difference between the velocity of light in one medium and its velocity in the other. The measurement is usually obtained by applying Snell's law (named for the 17th-century Dutch mathematician and physicist Willebrord Snell), which expresses the refractive index of a material as the ratio of the sine of the angle of incidence of the incident light beam to the sine of the angle of the refracted beam. Refractive index, like molecular weight, density and boiling point, is a characteristic property of a substance. Since every substance has its own specific refractive index, variations in the index of a mixture containing different proportions of two compounds reflect changes in the composition of the mixture. For example, if one were assessing a mixture of water (refractive index 1.33) and benzyl alcohol (1.54), a reading of 1.43 would show that the proportions were almost equal.

Refractive index as a measure of chemical composition can be applied with assurance only to binary, or two-compound, mixtures; the measurement becomes ambiguous when more components are present. (An exception is the "pseudo-binary" mixture, where the components separate into two distinct groups, with each group behaving as one component.) Notwithstanding this limitation, process refractometers have proved to be surprisingly useful. This can be attributed largely to their simplicity, sensitivity and well-developed design. About 1,000 of these instruments have been installed over the past 10 years, and the trend is advancing, particularly in the food and chemical industries.

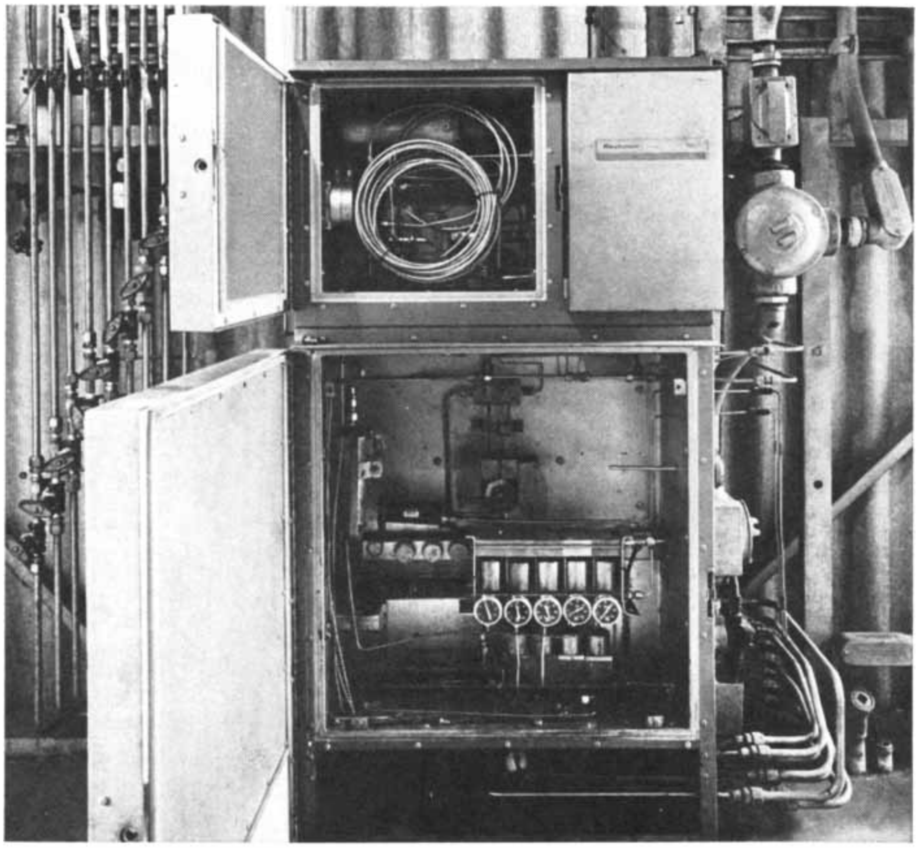
Measurement of refractive index is a

simple matter and can be carried to a high degree of precision: it is possible to determine a change of five units in the seventh decimal place of the refractive index. The first process refractometers, which appeared in 1952, operated on the differential-cell principle, wherein the difference in index between the sample in one part of the cell and a reference material in another part is measured. The early applications involved controlling the fractionation of complex hydrocarbon mixtures by choosing the sampling point so that the mixture could be regarded as pseudo-binary.

By 1957 about 20 process refractometers were operating successfully, even though they had two main weaknesses: they produced only a limited, single-point measurement, and they gave ambiguous information when the pseudo-binary relation changed. Today most of these instruments have been replaced by the more powerful chromatograph, which completely analyzes the process mixture, making possible the use of computers and control techniques that were beyond the horizon at the time the earlier refractometers were installed. (I shall return to chromatographic systems later.)

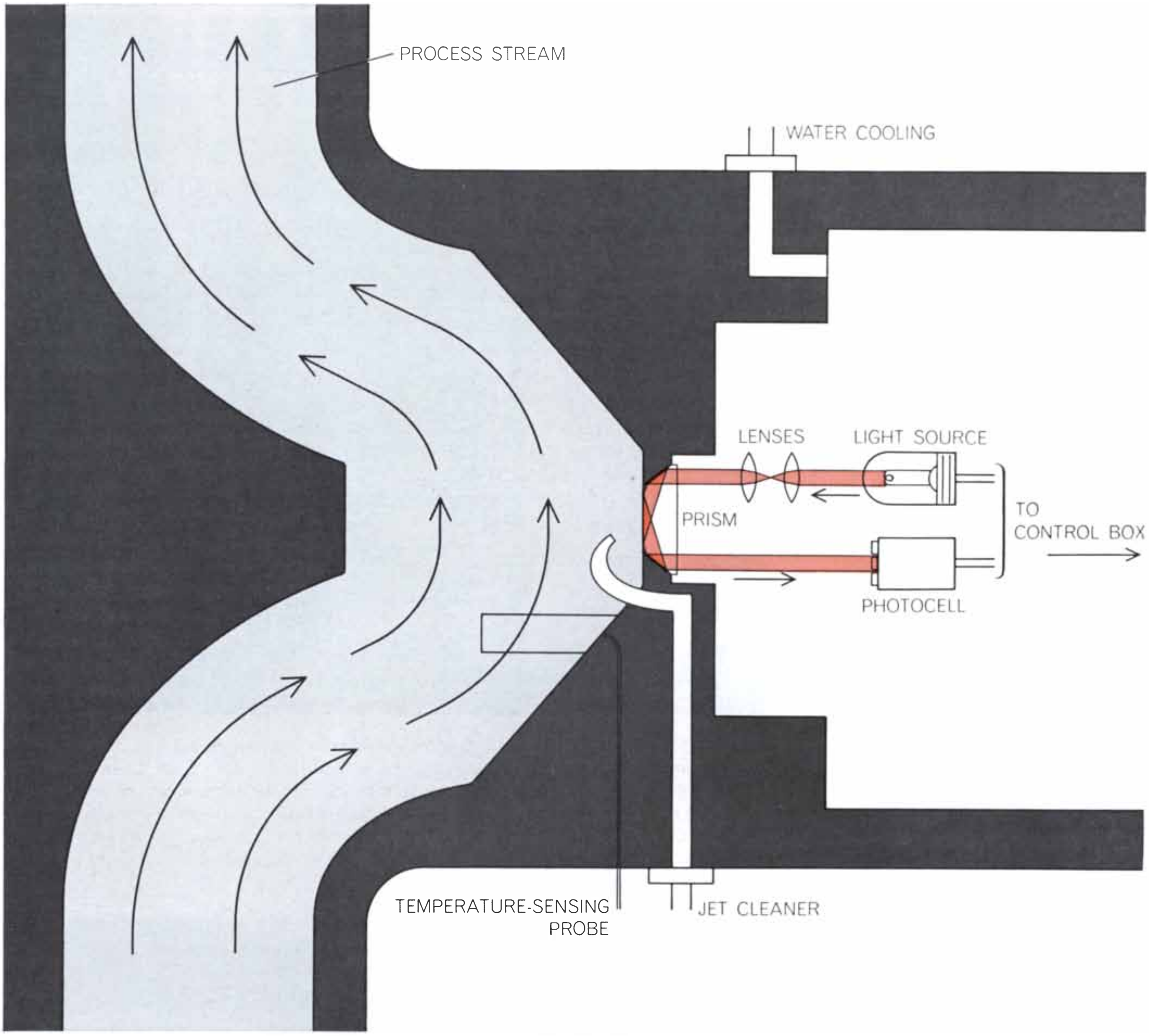
A new dimension was added to refractometry in 1962 with the development of the critical-angle refractometer [*see illustration on next page*]. This instrument measures refractive index by finding the angle at which a light beam striking the interface of a glass surface and the sample is totally reflected. (It is called the critical angle because rays arriving at larger angles penetrate the liquid, whereas at lesser angles they are reflected.) In a process system a circular beam of light is directed at the face of a prism that is in contact with a solution; there the beam is divided into a bright portion representing totally reflected light and a dark portion representing the absorption of light by the solution. The relative areas covered by the bright and dark portions of the beam depend on the critical-angle value of the solution. As the composition of the solution changes, so do the proportions of the bright and dark areas. The position of the dividing line between the areas is maintained by a servomechanism and becomes a measure of the critical angle.

The significant feature of this method is that only a surface reflection is needed; light does not have to be transmitted through the solution, as with other refractometers. As a result opaque solutions can be measured as readily as transparent ones, and the measurements can be made directly in the process



PROCESS CHROMATOGRAPH samples hydrocarbons at the refinery of the Mobil Oil Corporation in Torrance, Calif. Samples of the process stream enter from the slender pipes at left and are carried by a gas through the chromatographic column, which is the coiled tubing in the top compartment; there the components of the stream are separated. Data from eight such chromatographs analyzing 27 process streams go to a digital computer that directs the chromatography and sends the analyses to the control room of the refinery.

CONTROL ROOM at the Torrance refinery includes the printer in the foreground, which provides a computer print-out of the chromatographic analyses of the 27 process streams.



REFRACTION PRINCIPLE is employed in the process instrument known as the critical-angle refractometer. It continuously samples a solution and measures changes in the proportions of two substances in the solution by recording changes in the refractive index. The instrument measures the critical angle, which is the angle at which a beam of light striking the interface between the prism and the sample is totally reflected. Light is refracted from the interface and directed to a detector consisting of two photocells.

stream. There is no need to use a separate sampling system for the instrument.

The characteristics of the critical-angle refractometer make it particularly suited to the food industry. It is the refractometer employed in the ketchup process. The instrument also works well for solutions of sugar and water. Many of the carbonated beverages sold in the U.S. are made in processes that employ a refractometer to control the preparation of the sugar-water base to within .05 percent of the desired concentration. Another system of this kind, by including a small analogue computer, brings a syrup concentration to the proper point and automatically adds seed crystals to begin crystallization. The result is high-quality crystalline sugar, produced in about a third of the time needed for conventional processes.

The principle of absorption spectroscopy is applied in the process instrument known as the infrared analyzer. The infrared region of greatest practical interest lies between the electromagnetic wavelengths of two and 15 microns, which is where the strong absorption spectra characteristic of most organic and inorganic substances are found. A particular substance has unique spectra, because the frequency of the absorption bands is determined by the positions and configurational relations of the vibrating atoms in the molecules making up the substance. Since the amount of energy absorbed by a substance at any one of its characteristic frequencies is a measure of the substance's concentration in a mixture, infrared spectroscopy provides a very sensitive and powerful technique for the analysis of mixtures.

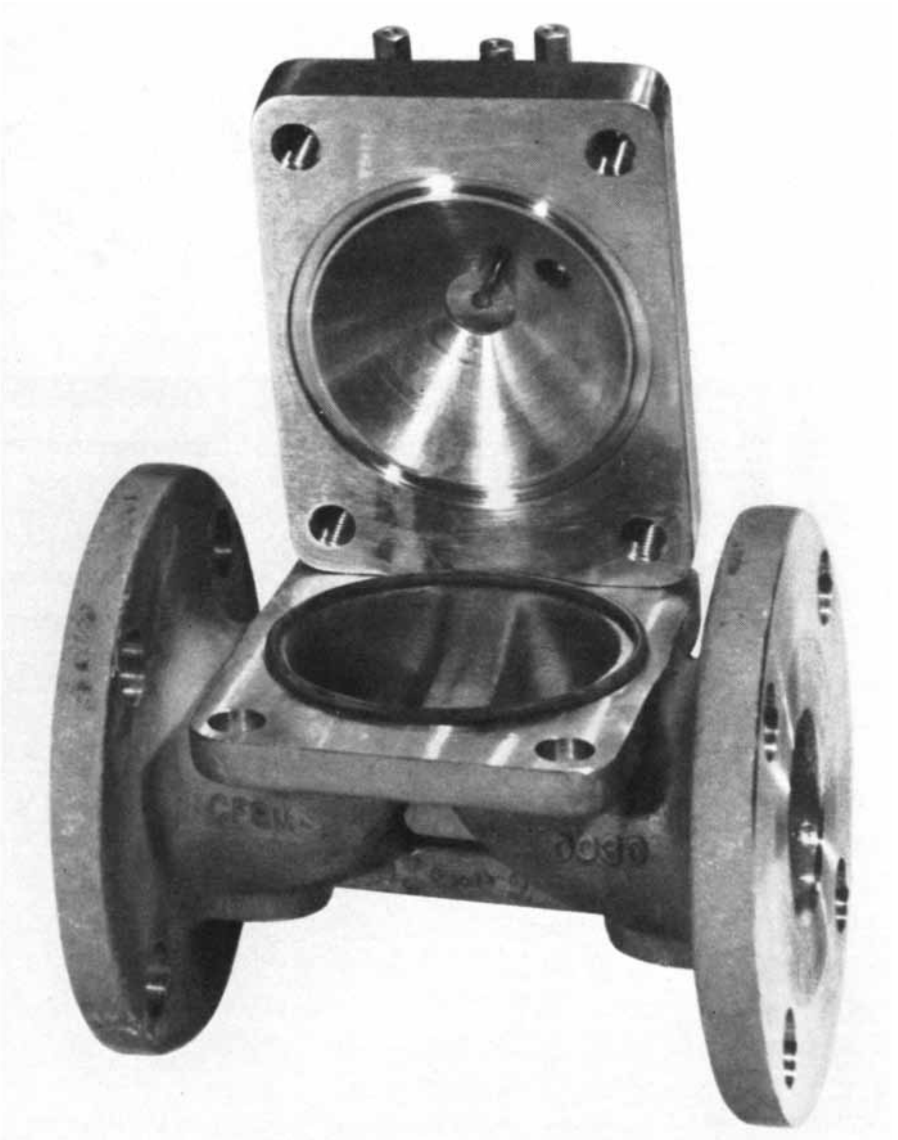
An infrared spectrometer used in the laboratory has an optical system capable of passing a selected wavelength of infrared radiation through a substance; with such a system the substance can be examined at each of its characteristic wavelengths. The designers of the first infrared analyzers for industrial processes did not incorporate such a system because it is too complicated and delicate for such processes; they turned instead to the positive-filter photometer, which remains the basis for the infrared analyzers in process streams today. A typical instrument consists basically of a sample cell, a comparison cell and a detector, together with a motor, a filter for each cell and a source of infrared radiation for each cell [*see illustration on next page*].

Suppose it is desired to analyze a stream of air for its content of carbon monoxide. A quantity of carbon monoxide is sealed in the detector, so that the detector will absorb infrared radiation at the frequencies characteristic of carbon monoxide. Pure air is put in the comparison cell, and air from the process stream passes through the sample cell. A chopper in front of the sources of infrared radiation sends pulses of infrared radiation, containing all infrared wavelengths, alternately through the comparison cell and the sample cell. If the sample cell contains any carbon monoxide, a fraction of the radiation passing through that cell will be absorbed before it reaches the detector, whereas none of the radiation passing through the comparison cell will be absorbed because that cell contains no carbon monoxide. When there is carbon monoxide in the sample cell, the difference in radiation reaching the detector from the two cells causes a membrane in the detector to move back and forth in uneven intervals, generating an alternating electric current that is recorded as a signal measuring the amount of carbon monoxide in the stream of air being sampled.

If the stream also contained ether, which partly overlaps the absorption spectra of carbon monoxide, the instrument could be rendered relatively insensitive to the ether by putting this substance in both filters. The same technique could be used for all other infrared absorbers in the mixture being sampled, provided that the mixture is not unduly complicated. By this positive-filter method, in short, components that would interfere with the desired measurement are filtered out.

The infrared analyzer, like the refractometer, is capable only of measuring the concentration of a single component in a sample stream. Unlike the refractometer, however, it can isolate that component from a mixture of many others. Among the compounds measured by infrared analyzers in industrial processes are (in addition to carbon monoxide) carbon disulfide, carbon dioxide, nitrous oxide, nitric oxide, sulfur dioxide, hydrocarbons, acetone, water vapor, ammonia and Freon. Considering only the industrial processes where the measurement of carbon dioxide is important, one finds such applications as the control of combustion, the manufacture of cement and the production of ethylene oxide, phthalic anhydride and ammonia.

Gas chromatography is a procedure whereby a volatile mixture is separated into its components by a moving inert gas passing over a sorbent. Although the sorbent can be a solid that has a large surface area, it usually consists of a nonvolatile liquid coated on an inert solid support. A typical separation column is a tube of small diameter, perhaps 1/8 inch, packed with small particles of the solid support that are coated with up to 20 percent by weight of a sorbent liquid. The support for the liquid can be simply the walls of the tube; in this case the tube has an inner diameter of .02 inch or less. In either configuration the gas, passing through the tube, carries the sample mixture. The components of the mixture separate themselves in characteristic ratios between the gas phase and the liquid phase [*see illustration on page 117*]. The separation takes place because components that are more soluble in the liquid move more slowly down the column than the less soluble components. A detector at the exit sees the components emerge in a series of peaks; the time of emergence of a peak identifies the component, and



CRITICAL-ANGLE REFRACTOMETER is opened to show the chamber where the solution flows past the prism, which is at center of upper section. The curving tube is part of the washing system that keeps the prism clean; the dark circle is the housing for a temperature-sensing probe that electronically compensates for temperature changes in refractive index.

the area of the peak reveals the concentration of the component in the mixture. As a method of separating the individual components of a complex mixture gas chromatography has no equal.

A process chromatograph resembles its laboratory counterpart more closely than most plant instruments do. The differences result from requirements imposed by the plant environment and automatic operation. The analyzer must obtain its own sample, analyze it and process the information from the detector into a useful quantity. Depending on the end use, this quantity is then delivered to either a recorder, a controller or a computer. The process chromatograph must continue to repeat the same cycle, perhaps as often as once a minute when high speeds are necessary for direct control, for many months at a time and with a high degree of reliability.

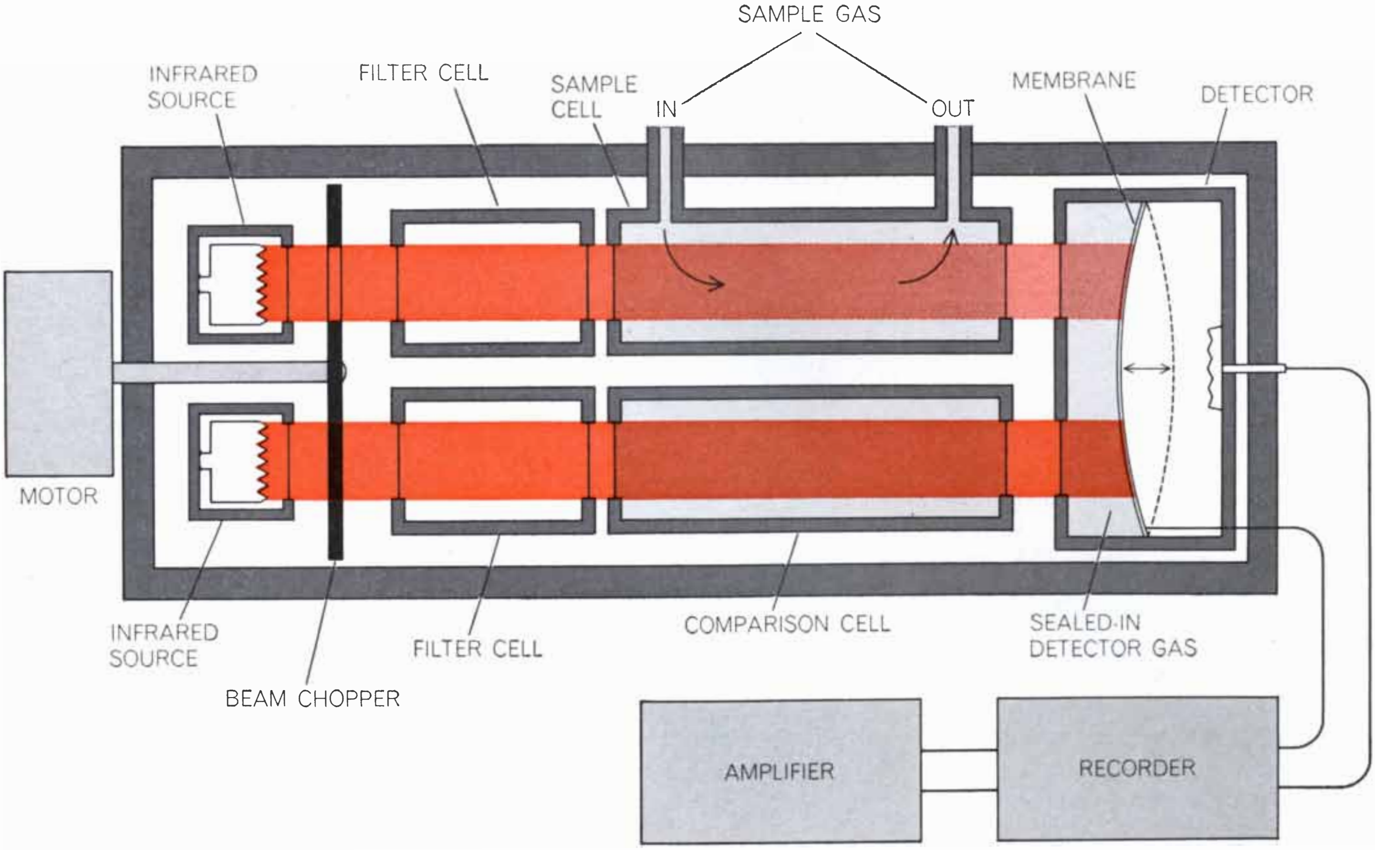
Process chromatographs appeared on the industrial scene about 10 years ago. They offered for the first time an analytical method that could give a complete compositional analysis and yet was simple in principle and execution. The development of the chromatographic analyzer and its use in processes has increased rapidly, to the point where about 8,000 are now in service. The usual installation consists of the analyzer, which is close to the process stream, and a programmer that can be anywhere. The programmer directs the operation of the analyzer.

The applications of these instruments range from the simple monitoring of one component in a mixture to providing a computerized output analyzing as many as 100 components in a stream. The instruments have found particularly wide use in the petroleum industry, where products of high value are made in large amounts and are both well suited for the chromatograph's analytical ability and amenable to automatic control.

Although one can see a progression in analytical capability from the refractometer to the infrared analyzer to the chromatographic analyzer, the fact is that these instruments have changed little from their original forms. Their performance has been refined and their reliability has been increased, but that is all. It seems likely that the next stage will be the development of new concepts based on the analytical principles employed in these instruments.

Chromatography has the greatest potential of the three methods. The future of process chromatography is suggested by the computer systems being installed in plant-control laboratories where many repetitive analyses are needed. In these systems a man injects a sample into a laboratory chromatograph and then directs the computer to carry out the analysis. The computer thereupon operates the instrument, reduces the data and types out a complete analysis. A single computer can share its time among as many as 30 instruments. It is a comparatively small step from these systems to similar ones involving process chromatographs. A few such processing systems have begun to appear; in one at the Mobil Oil Corporation's refinery in Torrance, Calif., eight chromatographs analyze 27 process streams and one computer controls the entire operation.

So far, however, all the systems in-

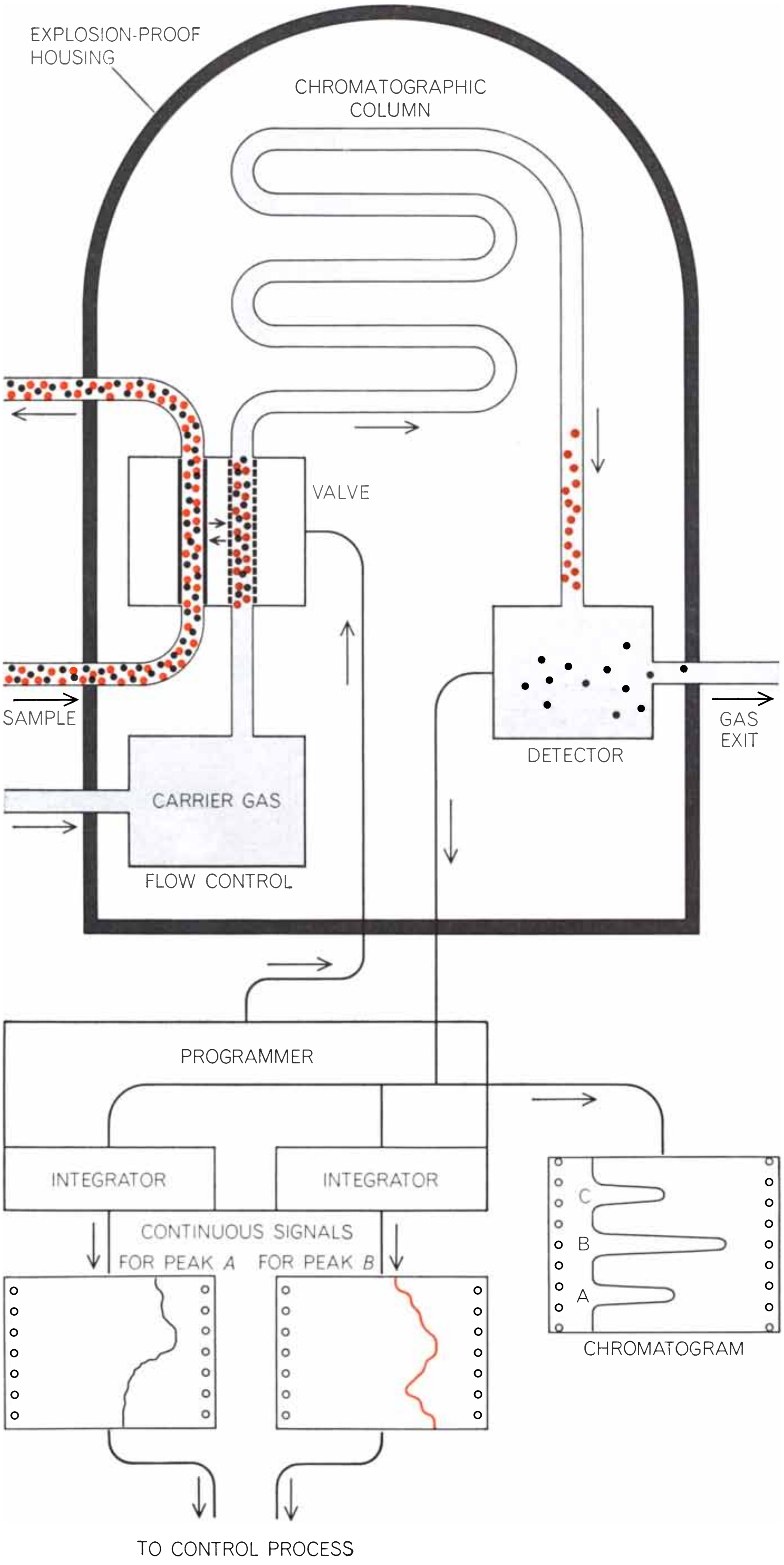


INFRARED ANALYZER measures the concentration of a single component in a process stream. Carbon dioxide, for example, would be measured by sealing a quantity of that gas in the detector. A beam chopper directs infrared radiation alternately through the comparison cell, which contains pure air, and the sample cell, which may or may not contain carbon dioxide. The carbon dioxide in the detector absorbs part of the radiation in a characteristic way. Absorption causes the gas to expand, and the membrane moves accordingly, generating an electrical signal. When carbon dioxide enters the sample cell, it absorbs part of the radiation before the beam from that cell reaches the detector, and uneven pulses result. They can be translated into the percentage of carbon dioxide in the sample stream. Gases with absorption spectra overlapping carbon dioxide can be put in the filter cells, nullifying the interference.

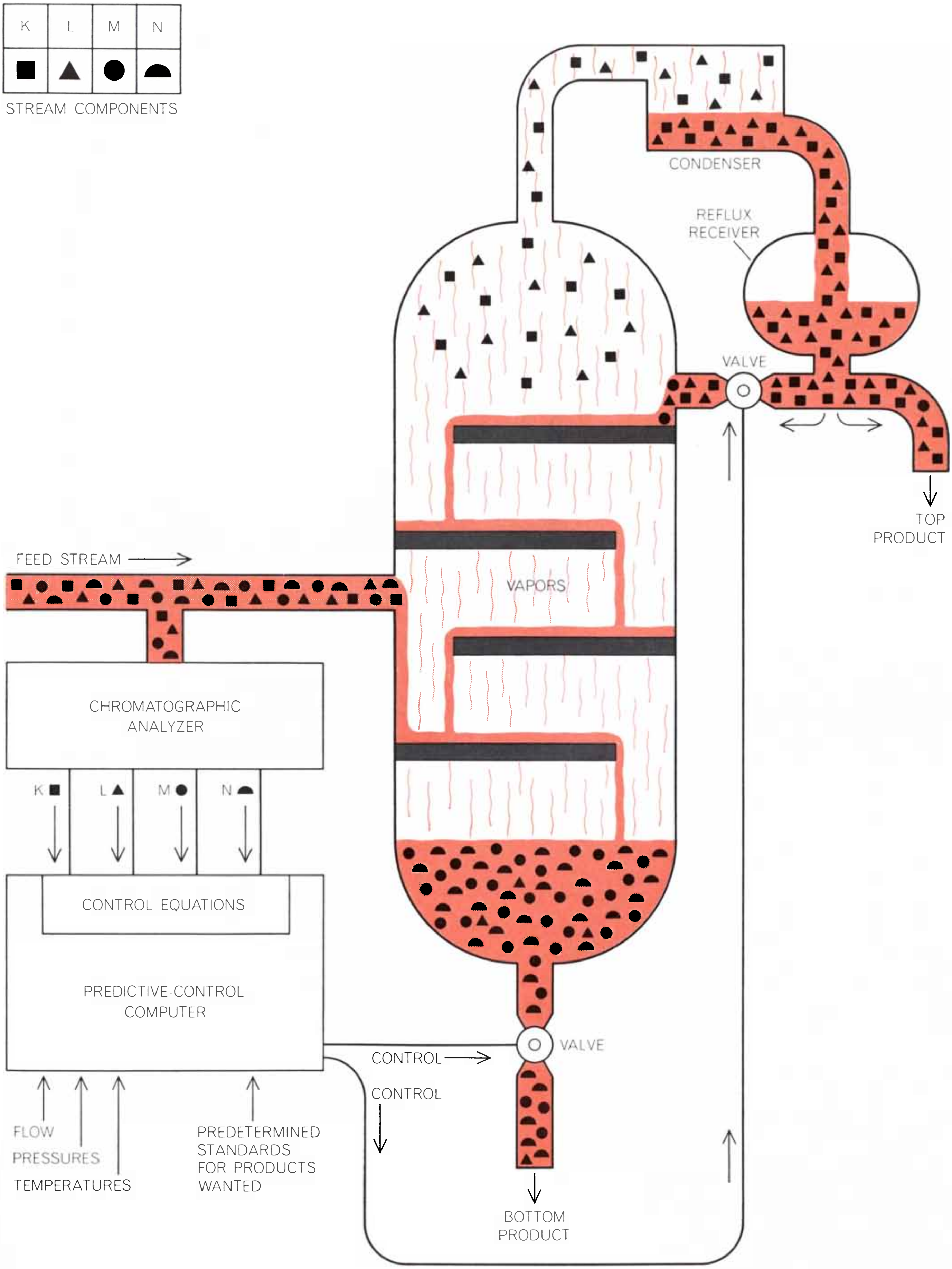
volving chromatographs and computers merely hitch the instruments to the computer in the most compatible fashion possible. In effect the computer is mechanizing the functions of the laboratory workers and coordinating the instruments. If one were to make a fresh start, a better approach would be to design a system that incorporated all the latest knowledge of chromatography and took fully into account the capacities of the computer for the high-speed handling of data and the making of decisions.

A computer-chromatograph system designed along these lines is nearly ready for testing in the laboratories of Applied Automation Inc. in Bartlesville, Okla. The system has five independent temperature zones (to allow for the capacity of the chromatograph to give different separations at different temperatures) that can hold as many as eight different chromatographic columns. A given sample will go concurrently through all the columns, receiving a different separation in each one. The peaks coming through each detector are sensed, identified and stored by the computer until it is ready to assemble all the information in the form of a complete analysis. The system also has the capacity to identify components that might otherwise be unidentified because of limitations in a less sophisticated system. The components can be mathematically separated by the computer because of the abundance of information available from the different columns. For example, components *A* and *B* might be unresolved in the first peak emerging from a chromatographic column and components *C* and *D* in the second peak, whereas in a different column the unresolved components are *A* and *C* in the first peak and *B* and *D* in the second. The computer can set the components as unknowns in four equations, which it solves by the standard mathematics of simultaneous equations.

One further step is needed to achieve the ultimate in analytical power for a process instrument. The remaining weakness of chromatography is the identification of component peaks. The identification of a component eluting from a chromatograph is made by comparing its time of elution with the elution time of the pure component under the same conditions. A number of disturbances can change the conditions and therefore the time. Moreover, a new compound appearing in the stream may have the same elution time as a compound already there. Identification of components also becomes progressively more difficult as



PROCESS CHROMATOGRAPH employs helium to carry intermittent samples of a process stream through a chromatographic column, where the components are separated according to their affinity for a sorbent liquid in the column. They are recorded as peaks on a chromatogram. Time of emergence of a peak identifies the component; area under peak reflects the concentration of the component. *C* represents a component not shown in the illustration.



PREDICTIVE-CONTROL TECHNIQUE combines a chromatograph and a computer for the automatic control of a distillation or fractionation process such as the separation of ethylene from a mixture of hydrocarbons. In predictive control the computer takes into account conditions that change, such as temperature and flow, and adjusts the operation of the fractionation column accordingly so that its products meet predetermined standards. Among the components of the feed stream the one identified as *K* is the most volatile, so that it emerges from the top of the column as a vapor; *N*, the least volatile component, leaves the bottom as a liquid.

the number of compounds in a mixture increases.

The need, then, is for an unambiguous sensor to identify the compounds in the flow emerging from a chromatographic column. Such sensing is achieved routinely in laboratories by directing a small portion of chromatographic effluent to a mass spectrometer, where the mass spectral patterns uniquely caused by each component provide an excellent means of identification. Recently mass spectrometers that are small, sensitive and stable enough to be considered as process instruments have appeared, and it seems likely that the technique of mass spectrometry will soon be available for process chromatography. Indeed, a mass spectrometer and a gas chromatograph have already been linked in an instrument for the space vehicle that will land on Mars; the instrument will analyze the Martian soil.

Clearly the trend in process control is toward a broader scope for automatic operations. A simple but truly automatic process is found increasingly in distillation, which involves the separation of components on the basis of their volatility. When a liquid is fed into a distillation column and heat is applied, the most volatile component will evaporate, so that the vapor leaving the top of the column will be rich in that component; correspondingly the liquid leaving the bottom of the column will be rich in the least volatile component.

Suppose the material being fed into a distillation column contains four components, which in terms of decreasing volatility can be called *K*, *L*, *M* and *N*. In a series of steps involving vaporization and condensation the most volatile component (*K*) will emerge from the top of the column and the least volatile one (*N*) will go out the bottom. The separation is never quite perfect, however, so that some of the components of medium volatility (*L* and *M*) will be in both products. The composition of the end products is determined by the interrelated control of such variables as flow, pressure and temperature.

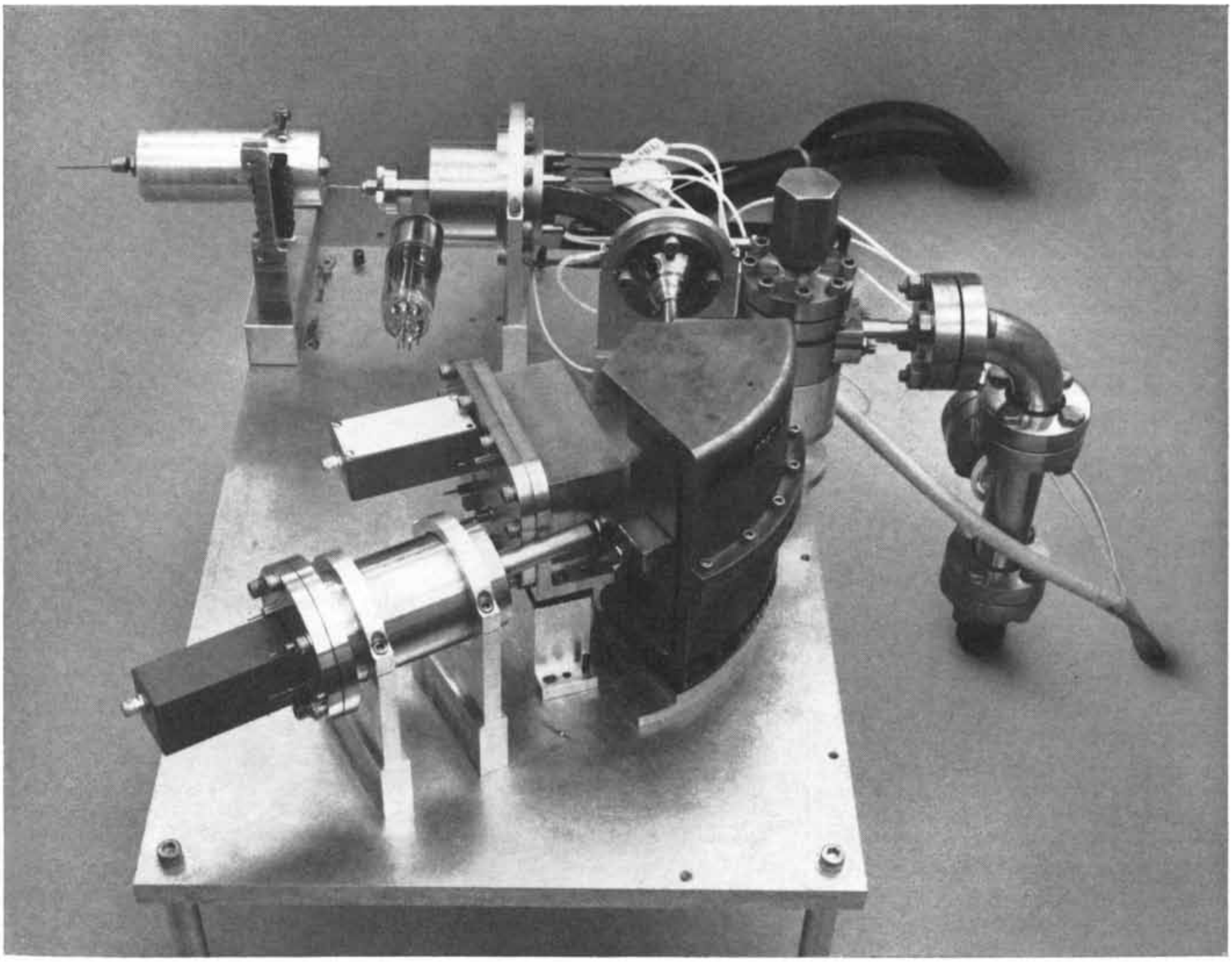
The recent development that has allowed this process to be made highly automatic in a number of petroleum and petrochemical plants is called predictive control. It can be illustrated by an analogy involving a home heating system. Most houses now have a heat-control system based on a thermostat, which operates by ordinary feedback: the desired temperature is set, and when the temperature of the house goes below that level, the thermostat generates an error signal to turn on the heat. An error signal is always necessary. Predictive control, in contrast, would sense all disturbances—the opening of a door, a drop in the outdoor temperature—and control the furnace accordingly, so that the house temperature would be constant and no error in temperature would be needed to actuate the control.

Suppose one wants to establish a predictive-control system for a distillation column where 17 variables that influence the distillation process must be taken into account. Equations can be written that relate the desired amount of impurity in the top and bottom products of the distillation column to the 17 variables. Basic controls are employed to eliminate the variability of as many of the 17 factors as possible, until one arrives at equations that express two flow controls as functions of the remaining variables. The two flow controls are the rate of flow of the bottom product (*B*) and the reflux rate (*R*), which concerns the material that is diverted from the top product and returned to the column for further distillation. These are the manipulated controls.

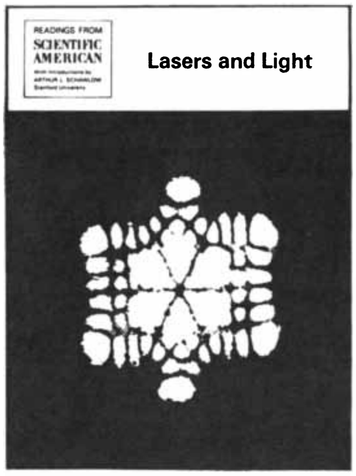
The predictive equations can be put into a small digital computer. The computer is also provided with measurements of feed composition (made by a high-speed process chromatograph), flow rate and temperature and with specifications of the column pressure to be maintained and the concentration of impurities to be allowed. The computer will calculate minute by minute the values of *B* and *R* and will adjust them continuously to maintain the specified purity of products.

Predictive control has arrived just in time to be of great value to the petrochemical industry as it confronts a rising demand for ethylene, which is the unit building block in the manufacture of polyethylene. The ethylene process is the most difficult one in the petrochemical industry to control in an optimum manner. The production of ethylene begins with a hydrocarbon feed that is thermally decomposed to a mixture rich in ethylene. The rest of the process involves separating the ethylene from the other components by means of a series of fractionation columns, which all interact through common heating and refrigeration systems. Predictive control systems are able to optimize this complicated operation.

Predictive control as applied to distillation is a step toward another prospect that is just beginning to emerge: direct digital control. Thus far the evolution of direct digital control has involved replacing large numbers of conventional controllers in a plant with a computer, which has the principal function of process regulation. As the evolution of direct digital control proceeds one can foresee an era when computers will control not just processes but entire plants. The day of the computer-operated factory appears to be close at hand.



CHROMATOGRAPH AND MASS SPECTROMETER are combined to yield precise identification of components. The instrument, which is indicative of the trend in process-control devices, was developed by the Jet Propulsion Laboratory for the analysis of soil on Mars. A sample of soil is treated by heat, and the resulting gases are injected into the instrument for analysis of components.



1969, 376 pages
394 illustrations (22 in color)
clothbound $10.00
paperbound $5.95

Lasers and Light

Readings from **SCIENTIFIC AMERICAN**

With Introductions by ARTHUR L. SCHAWLOW, Stanford University

The thirty-two readings gathered here provide a look at some of the exciting advances made in the study of light during the past decade and exemplify the richness and diversity that characterize modern optics. Its range and authoritativeness make this collection one that will delight and inform both the general reader who is interested in the science and technology of light and the professional scientist.

Included are many of the articles on laser science and technology that appeared in SCIENTIFIC AMERICAN through 1968 and all of the articles from the September 1968 issue devoted to the topic of light. Each article is reproduced as it originally appeared, with full text, illustrations, and color. Many of the readings were written by men who have played important roles in the development of the laser, and their papers convey a sense of the excitement laser light has generated. Arthur L. Schawlow, one of the proposers of the optical maser, has provided an introduction for each section of the book.

From your bookseller, or from Dept. J
W. H. Freeman and Company
660 Market Street, San Francisco, California 94104
Warner House, Folkestone, Kent, England

© 1969 SCIENTIFIC AMERICAN, INC

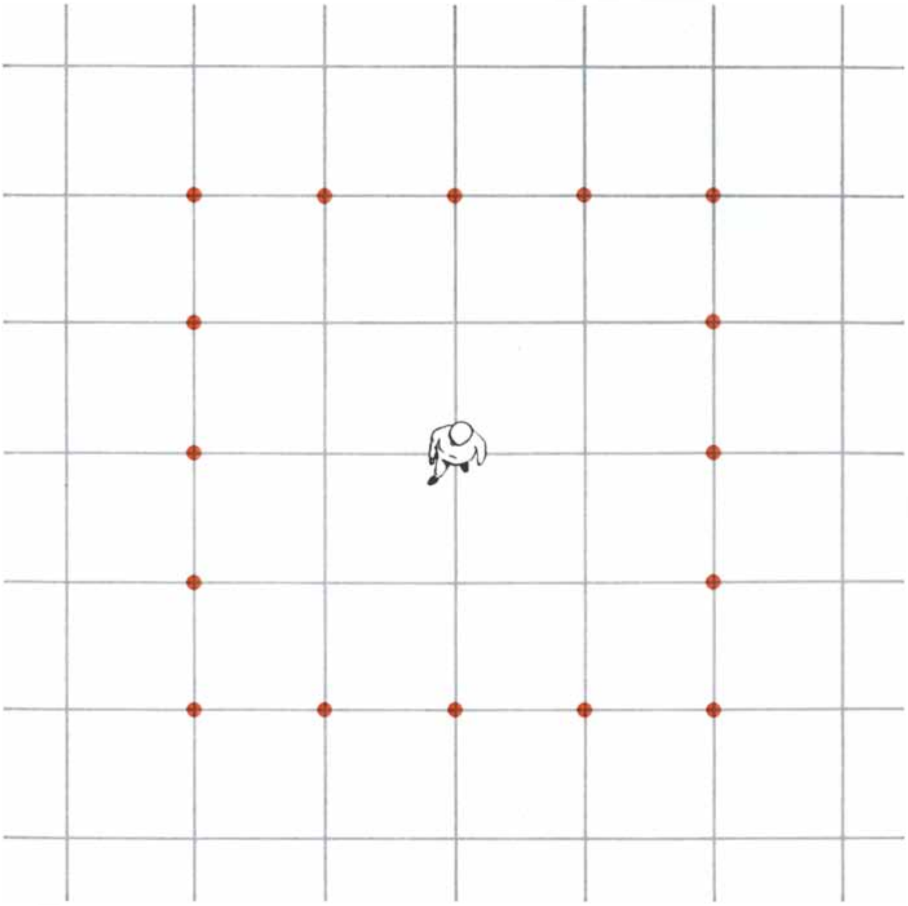
MATHEMATICAL GAMES

Random walks, by semidrunk bugs and others, on the square and on the cube

by Martin Gardner

Last month I discussed the random walk of discrete steps along a line, with or without absorbing barriers, and pointed out its curious equivalence to various two-person betting games. This month the random walk moves onto the plane and into space.

One type of two-dimensional random walk that has been much studied is one that goes from vertex to adjacent vertex on the infinite checkerboard lattice shown below. Every step is one unit and the walk is "symmetric" in the sense that each of the four possible directions is picked with a probability of 1/4. The walk can be made finite by surrounding the walker with absorbing barriers, shown as colored spots; when he steps on one of these spots, he is "absorbed" and the walk ends. (The surrounding barriers need not form a neat square. They may form a boundary of any shape.) As in the analogous finite walk on the line, it is not hard to calculate the probability that a walk starting at any vertex inside the boundary will end at a specified barrier. One can also determine the expected number of steps (the average in the long run of repeated walks) that will be taken before a walk ends. The formulas involved in such calculations have many unexpected scientific applications, such as determining the voltages at interior parts of electrical networks.



Random walk on a square lattice

When the walker is not trapped inside barriers but can escape to wander over a square lattice that covers the infinite plane, the situation grows more complicated and gives rise to many problems that are as yet unsolved. Some of the established theorems are deep and paradoxical. Consider a random walk on an infinite lattice with no barriers. If the walk continues an arbitrarily long time, the proportion of visits the walker makes to any specified corner approaches zero as a limit. On the other hand, if the walk continues long enough, the walker is certain to touch every vertex, including a return visit to his starting spot. As John G. Kemeny points out in "Random Walks," an excellent nontechnical article in *Enrichment Mathematics for High School* (National Council of Teachers of Mathematics, 1963), this introduces a profound distinction between logical and practical possibility. It is *logically* possible that such a walker can travel forever without reaching a given corner. To the mathematician, however, it has a *practical* probability of zero even though the expected number of steps for reaching any specified corner is infinite. The distinction is often encountered where infinite sets are concerned. If a penny is flipped forever, for example, it is logically possible that heads and tails will forever alternate, although the practical probability that this will happen is zero.

Kemeny expresses it this way: If you stand at an intersection on the infinite lattice while a friend, starting at any other spot, wanders randomly over the lattice, he will be practically certain to meet you if you are able to wait an arbitrarily long time. The statement can be even stronger. After the first meeting the probability is again 1 that if your friend continues wandering, he will eventually return to you. In other words, it is practically certain that such a walker, given enough time, will visit every intersection an infinity of times!

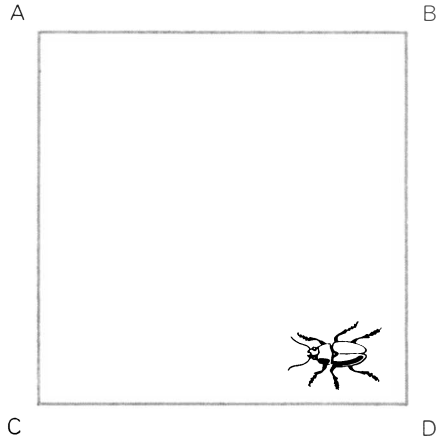
Suppose two walkers move haphazardly over an infinite square lattice. Are they certain to meet? (If they begin an odd number of steps apart and step in unison, they can never meet at a corner, but they can bump into each other at the middle of a line segment.) Once more the answer is that they will meet infinitely often if they walk long enough. If three men step in unison and wander over the infinite lattice, and if each pair of their starting spots is separated by an even number of steps, all three are certain to meet at *some corner*. The probability of their meeting at a specified corner, however, drops to less than 1.

For four or more walkers the probability that all four will meet *somewhere* also becomes less than 1.

The biggest surprise comes when we generalize to a space lattice. If such a lattice (it need not be cubical) is finite, a random walker is practically certain to reach any intersection in a finite time. As Kemeny puts it, if you are inside a large building with a complex network of corridors and stairways, you can be sure of reaching an exit in a finite time by walking randomly through the building. If the lattice is infinite, however, this is not the case. George Polya proved in 1921 that the probability is less than 1 that a random walker will reach any assigned corner on such a lattice even if he walks forever. In 1940 W. H. McCrea and F. J. W. Whipple showed that the probability of a walker's returning to his starting spot, after wandering an infinite time on an infinite cubical lattice, is about .35.

Turning from planar lattices to the plane itself, allowing the walker to step a unit distance in any randomly selected direction, the situation becomes more complicated in some ways and simpler in others. For example, the expected (average) distance of the walker from his starting spot, after n equal steps, is simply the length of the step times the square root of n. This was proved by Albert Einstein in a paper on molecular statistics published in 1905, the same year he published his celebrated first paper on relativity. (It was independently proved by Marian Smoluchowski.) Without knowing that Brownian motion—the random walks of minute particles suspended in a liquid—had been observed 78 years earlier by the Scottish botanist Robert Brown, Einstein predicted such motions and analyzed them mathematically. (Readers will find a simple proof of the square-root theorem in George Gamow's *One Two Three... Infinity.*)

Discrete random walks in space obey the same square-root formula. As on the plane, the steps need not be uniform in length. The expected distance from the origin, after n steps, is the average length of a step times the square root of n. It is here that random walks become invaluable in the study of diffusion phenomena: the random movements of molecules in a liquid or gas, the diffusion of heat through metals, the spread of rumors, the spread of a disease and so on. The recent drift of the "Hong Kong flu" eastward across the U.S. was a blend of millions of random walks by microbes. There are applications in almost every science. The first major use of the Monte Carlo method—a way of using computers to simulate difficult probability problems—was in calculating the random walks of neutrons through various substances. In such diffusion phenomena, as well as in Brownian motion, the square-root formula must be modified by many other factors such as temperature, the viscosity of the embedding medium and so on. Moreover, such motions are usually continuous, not discrete; they are called Markov "processes," as distinct from Markov "chains." The square-root formula provides only a first approximation for estimating expected distances. (For recent work in this field, beginning with Norbert Wiener's brilliant first paper on Brownian motion in 1920, see "Brownian Motion and Potential Theory," by Reuben Hersh and Richard J. Griego; Scientific American, March.)



START OF STEP	FINISH OF STEP			
	A	B	C	D
A	0	1/2	1/2	0
B	1/3	0	0	2/3
C	1/3	0	0	2/3
D	0	1/2	1/2	0

"Ergodic walk" on a square (left) *and the matrix of its transition probabilities* (right)

The outward drift of a random walker from his starting spot, on the plane or in space, is not at a constant rate. If the walk itself is at a steady pace, the square root of the number of steps increases at a steadily decreasing rate. The longer the walk, the slower the drift. Gamow, in the book cited above, gives a dramatic illustration. A light quantum near the sun's center takes about 50 centuries to perform a "drunkard's walk" to the surface. Once free of the sun it instantly sobers up and, if it is headed in the right direction, reaches the earth in about eight minutes.

Here is a simple question to be answered next month. Two men start at the same spot on the plane. One makes a random walk of 70 unit steps, then stops. The other stops after a random walk of 30 unit steps. What is the expected distance between them at the finish?

We turn now to a type of random walk that is different from any we have considered so far. Assume that a bug starts at corner A of the square shown at the left in the illustration above and crawls randomly along its edges. Instead of equalizing the "transition probabilities" from corner to corner, as in earlier examples, assume that at corners B and C the probability of the bug's heading toward D is twice the probability of its returning to A. At A and D the bug chooses between two paths with a probability of 1/2 for each, whereas at B and C it chooses the path to D with a probability of 2/3 and the path to A with a probability of 1/3. The network is finite but, since there are no absorbing barriers, the walk never ends. Such a walk is usually called an "ergodic walk." We should like to calculate what fraction of visits, in the long run, the bug makes to each corner.

One way to do it is to draw the "stochastic matrix" at the right in the illustration, which shows the transition probabilities from any corner to any other. Zeros on the matrix indicate transitions that cannot occur. Since every state of this ergodic Markov chain must lead to another, the sum of the probabilities on any horizontal row, called a "probability vector," must be 1.

The probability that the bug will visit a given corner is the same as the sum of the probabilities that it will move to that corner from an adjacent corner during its endless walk. For instance, the probability that it is at D is the probability that it will go to D from B added to the probability that it will go to D from C. (These are long-run probabilities, not the probabilities of going to D when the bug *starts* at B or C.) Let d be the probability that at any moment when the bug is at a corner it will be at corner D. Let a, b, c be the probabilities that it is at corners A, B, C. From the D column of the matrix we can see that the probability that, in the long run, the bug will be going from B to D is $b(2/3)$ and from C to D is $c(2/3)$. The long-run probability of being at D is the sum of these two probabilities, and therefore we can write the following equation: $d = b(2/3) + c(2/3)$. This simplifies to

$$d = 2b/3 + 2c/3 .$$

The other three columns give us similar formulas for a, b and c:

$$a = b/3 + c/3$$
$$b = a/2 + d/2$$
$$c = a/2 + d/2$$

When the bug is not on an edge, it is certain to be at a corner, and so we have a fifth equation:

$$a + b + c + d = 1 .$$

A glance at the four preceding equations shows that $b = c$ and $d = 2a$, making it easy to solve the five simultaneous equations: $a = 1/6$, $b = 1/4$, $c = 1/4$, $d = 1/3$. The bug will spend 1/6 of its corner time at A, 1/4 at B, 1/4 at C, 1/3 at D. It will make twice as many visits to D as to A.

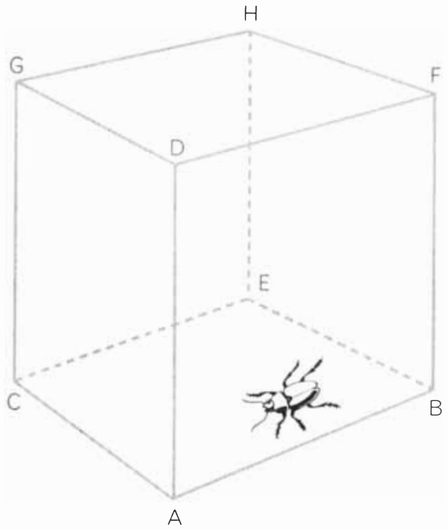
Readers might like to try the same technique on the cube analogue of this problem, given by Kemeny in the article mentioned above. On the cube shown at the left in the illustration below the bug is twice as likely to step toward H as toward A. The stochastic matrix of transition probabilities is shown to the right of the cube. The eight simultaneous equations derived from the eight columns, together with the equality $a + b + c + d + e + f + g + h = 1$, have a unique solution. The bug, performing its perpetual ergodic walk, will make 3/54 of its corner visits to A, 5/54 to each of B, C, D, 8/54 to each of E, F, G and 12/54 to H. It will visit H four times as often as A.

If an ergodic walk of this type is symmetric, in the sense that at each vertex the possible next steps are chosen with equal probability, the fraction of visits spent at any two given corners is proportional to the number of different ways to walk to those two corners. For example, a cat performing a symmetric random ergodic walk along the edges of the Great Pyramid of Egypt will visit the apex four times for every three times it visits a base corner because there are four ways to walk to the top corner and only three ways to reach each base corner. It is easy to draw the matrix and write the equations showing that the cat will make 1/4 of its visits to the apex and 3/16 to each base corner.

For another easy problem to be answered next month, assume that on the cube shown below, with the same matrix, a fly starts a random walk at A. At the same time a spider begins a random walk at H. Both move at the same speed. What is the probability that they will meet at the middle of an edge after each has gone the minimum distance of 1½ edges?

Many other pleasant problems arise in connection with ergodic random walks along the edges of a cube and other regular solids. If a drunk bug starts at a corner of a cube and walks to the most distant corner, at each corner choosing one of the three paths with equal probability, its average walk will have a length of 10 edges. If the bug is only semidrunk, never going back along an edge just traversed but selecting the other two paths with equal probability, its average walk to the most distant corner is 6. In *both* cases the bug's average walk back to its starting corner is 8, the number of corners on a cube.

This is no coincidence. Thomas H. O'Beirne of Glasgow (whose term "semidrunk" I have borrowed) has shown in an unpublished proof that, on any regular network with every vertex topologically the same as every other, a random walk back to a starting vertex has an expected number of steps equal to the total number of vertices. It is true regardless of whether the walker chooses all paths at each vertex with equal probability or only those that exclude the path just traveled. A drunk or semidrunk bug stepping from corner to corner on a square will have an average walk back to the starting corner of four steps. The edges of all Platonic and Archimedean solids form regular space networks of



		FINISH OF STEP							
		A	B	C	D	E	F	G	H
START OF STEP	A	0	1/3	1/3	1/3	0	0	0	0
	B	1/5	0	0	0	2/5	2/5	0	0
	C	1/5	0	0	0	2/5	0	2/5	0
	D	1/5	0	0	0	0	2/5	2/5	0
	E	0	1/4	1/4	0	0	0	0	1/2
	F	0	1/4	0	1/4	0	0	0	1/2
	G	0	0	1/4	1/4	0	0	0	1/2
	H	0	0	0	0	1/3	1/3	1/3	0

Ergodic random walk on a cube (left) *and the matrix of its transition probabilities* (right)

The Innovation Group!

What **is** the innovation group?

It is a set of men very important to this country. You may be part of it . . . or your boss may . . . or the man in the next office . . . or a key man who reports to you.

These are the men who manage technological change: Men who manage research or design. Who steer a company through a technological changeover. Men who set federal policy for a technological age. That's some of them.

Such men are a lonely crowd, mostly. For each one of them, all by himself, has had to learn an art that has no teachers or precedents or tradition. It's not quite a matter of knowing science or engineering, their art. Not the same as conventional business management. Their job—which may be your job—is something new in the world, born in the last couple of decades. It's the art of technical management.

We in Technology Communication, Inc. have set out to bring these isolated and important men together.

Our goal: An organization that combines the intellectual standards of a professional association with the vigor and the professionalism which is commanded by a company operating under the discipline of the profit motive.

That is not an easy combination to achieve. We believe we know some fresh and exciting ways to do it. As the first step, we announce formation of . . .

The Innovation Group

What is the **Innovation** group?

It is the readers of a remarkable new journal called INNOVATION. The journal is published monthly. It carries absolutely no advertising. It aims to be the handsomest and most readable business or professional periodical in existence. It is—its cover declares—"about the art of managing advancing technology."

Issue number one of INNOVATION will appear in May. In it, if you were a member of the Group, you would read about a profound and surprising social change underway in America—a change which helps explain why your peculiar job is twice as difficult and twice as important as it used to be. As a Group member, you'd read how one of the country's most technical industries is in danger of destroying itself right now because its executives don't yet know how to control technical change. You'd learn from one of the world's most successful technical managers how the insights of systems theory and of ecology can give you a new understanding of what that peculiar job of yours really is. You'd hear about the agony that your technical kind of management creates for financial men—and about the ways the smartest of them are finding to ease the pain. You'd gain *understanding* of a new restlessness that has started to leap from student activists to just those smart young men you yourself need to hire and to manage; you'd better know how they think.

With a couple of special exceptions, you or anyone else who wants to read the charter issue of INNOVATION must be a member of . . .

The Innovation Group

What is **The Innovation Group**?

It is some thousands of imaginative men who are participating in a project of discovery. Together, we in the Innovation Group are setting out to discover the principles by which technology *can* be managed. Sharing our experiences. Thinking out the dangers and the opportunities of technological change—and figuring out how to deal with them.

We will be putting our minds together in many ways. Face to face. On paper. Over the telephone. By video and audio tape. Participants in the Group will share their experience in the pages of their journal, at conference-workshops, on tours, in telephone hook-ups . . . in books, for that matter.

What we are going to do is to chart the skills and the standards of a newly meaningful profession—the management of technology.

The whole stirring enterprise is proceeding with the intellectual guidance of a very distinguished and very active Advisory Board. The members of the board are: Warren G. Bennis, vice president of the State University of N.Y. at Buffalo; Emilio Daddario, chairman of the House subcommittee on science, research, and development; Eugene G. Fubini, consultant; C. Lester Hogan, president of Fairchild Camera and Instrument Co., Inc.; J. Herbert Hollomon, president of the University of Oklahoma; Warren Kraemer, president of Warren Kraemer Associates, management consultants; Donald G. Marquis, director of the research project on research and technology at MIT's Sloan School; Emmanuel G. Mesthene, director of the Harvard program on technology and society; Jack Morton, vice president of Bell Telephone Laboratories; E. C. Williams, chief scientist of the British Ministry of Power.

If *you* would like to join men like this in a project of this sort, we can now invite you to join . . .

The Innovation Group

The Innovation Group

Technology Communication, Inc.
265 Madison Ave., New York, N.Y. 10016

If you manage technology . . . if you lead an organization dependent on technological advance . . . if you are professionally concerned with the effects of technological change . . . perhaps you belong with us.

☐ Tell me more about benefits of membership in The Innovation Group.

☐ Send the facts to attached list of technical managers with my compliments.

NAME TITLE

COMPANY DIVISION

STREET ADDRESS CITY STATE ZIP

UNUSUAL BARGAINS BY MAIL

SHOPPING MART

A selection of products available by mail for readers of Scientific American
All merchandise sold on money-back guarantee. Order by No. Send check or M.O.

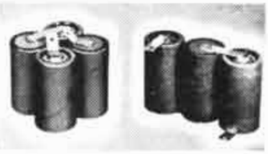
GIANT FREE CATALOG 148 PAGES

CONTINUOUS WAVE HE-NE LASER

Top quality instrument—budget price. Portable, versatile, easily operated. Just plug in reg. outlet-no warmup req. Ideal for research, lab experiments, holography, optical testing & alignment, surveying, etc. 1 milliwatt minimum output in TEMOO mode. Long tube life guaranteed 1,000 hrs. (3m typical). Produces continuous intense beam of coherent, monochromatic light at 6328 angstroms. 1.5 milliradian max. divergence; 1.0 mm beam diam. 1" diam. x 14" laser tube. Instant starting cold cathode. Safe, solid state, low ripple regulated power supply.
Stock No. 79,002S. $220.00 Ppd.

4¼ ASTRONOMICAL TELESCOPE

See craters on moon, rings of Saturn, double stars. Up to 270X. New vibration-free equatorial mount. f/10 mirror corrected to better than ¼ wave length. Gives theoretical limited resolution. Rack and pinion focusing. Aluminum tube, 6x finder telescope. 2 standard eyepieces and Barlow lens gives power up to 270x. Free with scope: Star Chart, 272-p. "Handbook of the Heavens"; "How to Use Your Telescope."
Stock No. 85,105S...$84.50 fob.
3" ASTRONOMICAL TELESCOPE
Stock No. 85,050S..$29.95 Ppd.



NEW SURPLUS Ni-Cd CELLS

Save more than 75%! Long life-accept 300 charge and discharge cycles. 1.25 Volts per cell—750 milliampere hours capacity. Excel. charge retention. Hermetically sealed. Indefinite storage life. Multiple cells welded in series—easily cut. Combine to form btry. ⅞" dia. x 1⅝" high. Spec. price for 100 up.

Order #	Cells	DCV	Ppd.
60,634S	4	5.00	$4.80

Trickle Charger (1-10 cells)
70,812S..........$11.50

GIANT FREE CATALOG

148 Pages! More than 4,000 UNUSUAL BARGAINS! Completely new. Enormous variety of telescopes, microscopes, binoculars, magnets, magnifiers, photo components, lenses, prisms, optical instruments, parts and accessories. Write for Free Catalog "S".

NAME
ADDRESS
CITY
STATE ZIP

ANALOG COMPUTER KIT

Demonstrates basic analog computing principles. Can be used for multiplication, division, powers, roots, log operations, trig problems, physics formulae, electricity and magnetism problems. Operates on two flashlight batteries. Electric meter and 3 potentiometers mounted on die-cut box. Answer indicated on dial. 20" long, 9" wide, 2" deep.
Stock No. 70,341S..$14.95 Ppd.

SPILHAUS CLOCK NOTES TICK OF THE UNIVERSE

A unique space-age time piece—gives 19 different readings of earth and sky at a glance. Beautiful fruitwood case houses 3 illuminated synchronized sky-blue dials. Requires only one geographical setting to day and month for accurate forms and positions of stars, a timetable of rising and setting of sun, moon, and stars and stages of the tides. Left dial gives local time, right dial world time. 16" x 11½" x 4½" deep. Shipping wt. 13 lbs. Uses house current.
No. 1202S.........$250.00 FOB

ORDER BY STOCK NUMBER • SEND CHECK OR MONEY ORDER • MONEY-BACK GUARANTEE

EDMUND SCIENTIFIC CO. 300 EDSCORP BUILDING
BARRINGTON, NEW JERSEY 08007

Most SLRs under $200 stop short of at least one of the new Miranda Sensomat's basic features: optics independently judged superior to those on cameras selling for twice as much; the unique Miranda 3-year guarantee; all shutter speeds to 1/1000 second; a lens mount that accepts over 1,500 lenses and accessories; and a through-the-lens zone-metering system that isn't fooled by subjects poised against bright backgrounds.

Sensomat also has several smaller extras, the biggest of which is the removable eye-level prism finder. Just slip it off, shade the ground glass with the accessory you already own—your hand—and you're ready for that low-angle shot. (Of course, you can also add accessory waist-level and critical-focusing finder hoods when you need them.)

Yet for all its versatility, the Sensomat costs less than $190* with the brand-new Miranda 50mm f/1.8 automatic lens (less than $240* with the new f/1.4). See it at your dealer.

You can't buy more camera for $190*

Miranda Sensomat®

Allied Impex Corp., 168 Glen Cove Rd., Carle Place, L.I., N.Y. 11514. Chicago; Dallas; Glendale, Calif. In Canada, Kingsway Film Equipment Ltd., Ontario

*YOUR DEALER WILL DETERMINE EXACT RETAIL PRICE IN YOUR AREA. ® REGISTERED BY ALLIED IMPEX CORP., EXCLUSIVE U.S. IMPORTER.

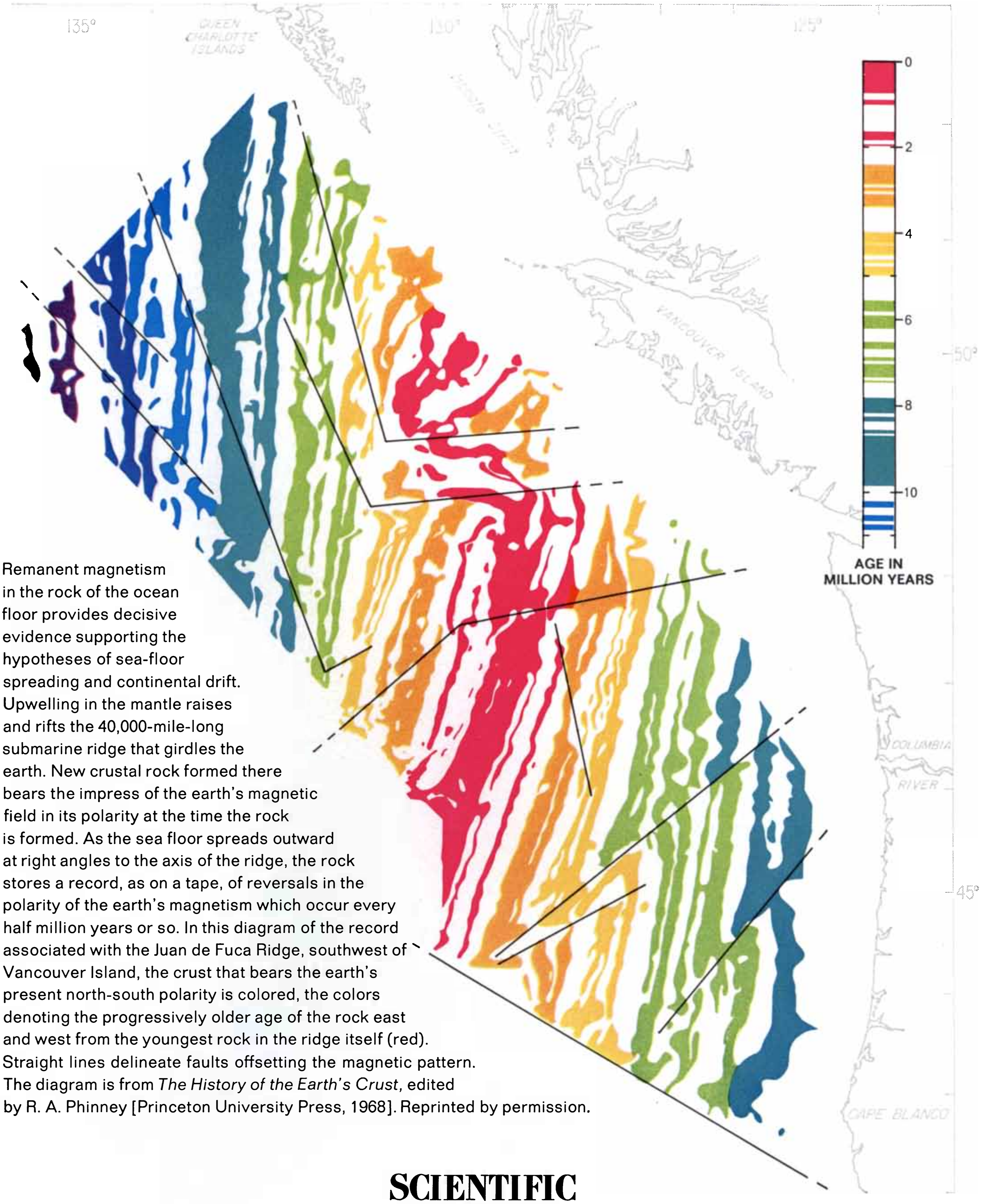
the same kind. On a tetrahedron a drunk or semidrunk bug will traverse four edges in an average walk back to a starting corner; on a dodecahedron it will traverse 20 edges, and so on. Readers interested in how to set up equations for calculating average walks on such networks will find the method explained, with reference to the dodecahedron, in solutions to problems E1752 and E1897 in *The American Mathematical Monthly*, February, 1966, page 200, and October, 1967, pages 1008–1010. For the rhombic dodecahedron, which is not regular, see O'Beirne's article "A Nonsense Result in the Traffic Statistics of Drunk Flies," in the *Bulletin of the Institute of Mathematics and Its Applications*, August, 1966, pages 116–119.

Steps need not be to adjacent spots on a network. Consider the symmetric random ergodic walk of a rook over a chessboard, assuming that on each move the rook chooses between all possible moves with equal probability. Since a rook can reach any cell from 14 other cells, the transition probability for every move is 1/14. The rook will therefore spend the same time on each cell.

The situation with respect to other chess pieces is different because their transition probabilities vary. A knight, for example, can reach a corner cell only from two other cells, whereas it can reach any of the 16 central cells from eight other cells. Since the proportion here is 2/8 or 1/4, it follows that during an endless random walk over the chessboard a knight will visit any designated corner square one-fourth as often as it will visit a given square among the central 16. For a proof see "Generalized Symmetric Random Walks," by Eugene Albert, in *Scripta Mathematica*, August, 1964, pages 185–187.

Last month's betting problem was intended as a joke. If player *A* begins with a certain number of dollars and if, on each draw of a card from a packet with an equal number of red and black cards, *A*'s opponent *B* is allowed to bet half of *A*'s current capital, the game obviously is the same as the one explained earlier in the same column, in which *A* always bets half of his *own* capital. *B*, who now calls the bets, will win precisely what *A* would have lost in the earlier game. Therefore the formula given last month for the first game applies here. We were told that the loser starts with $100 and that a deck of 52 cards is used as a randomizer. The winner is sure to win exactly $100 - [100(\frac{3}{4})^{26}]$ dollars, leaving the loser with less than a dime.

© 1969 SCIENTIFIC AMERICAN, INC

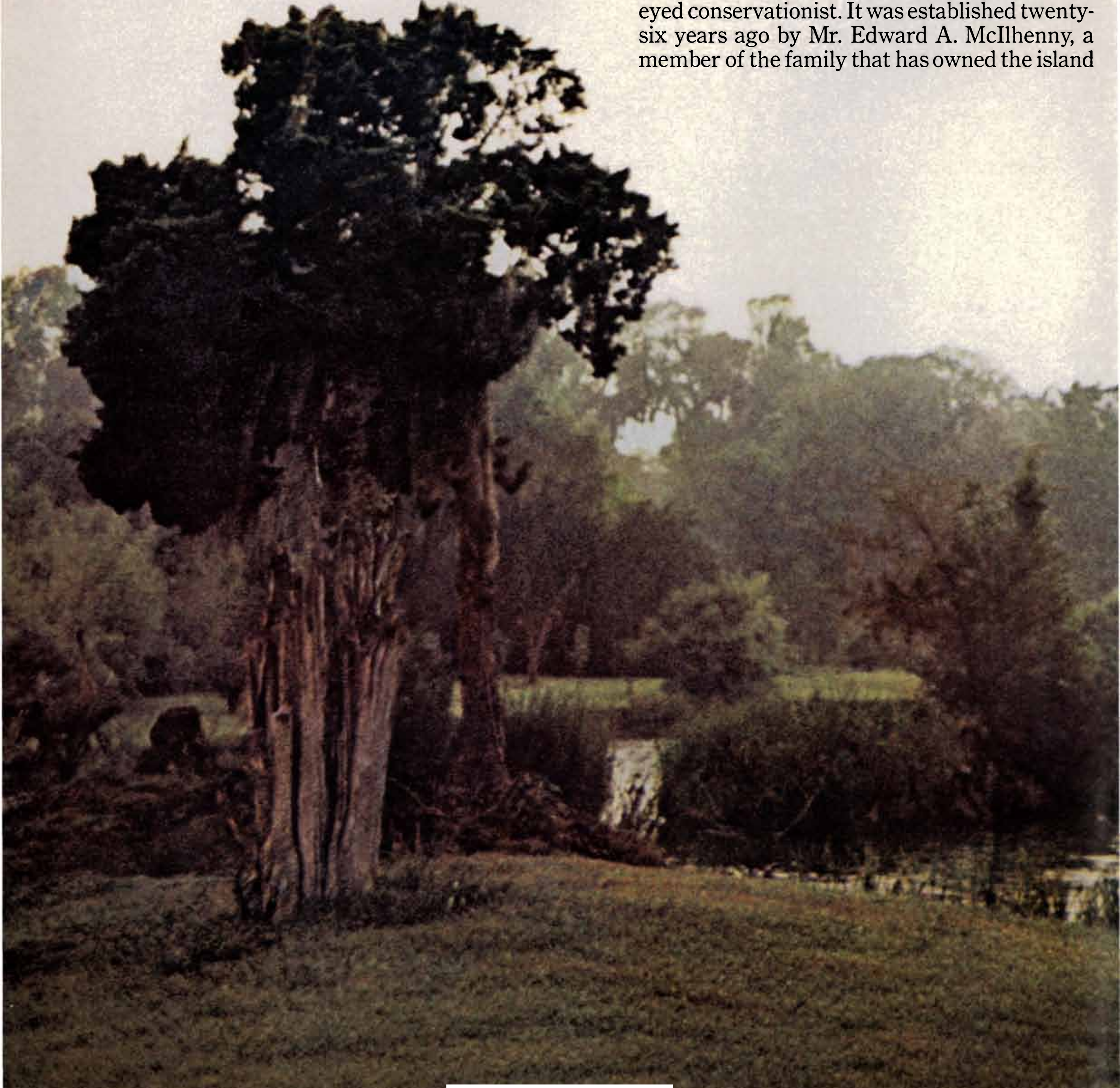


Remanent magnetism in the rock of the ocean floor provides decisive evidence supporting the hypotheses of sea-floor spreading and continental drift. Upwelling in the mantle raises and rifts the 40,000-mile-long submarine ridge that girdles the earth. New crustal rock formed there bears the impress of the earth's magnetic field in its polarity at the time the rock is formed. As the sea floor spreads outward at right angles to the axis of the ridge, the rock stores a record, as on a tape, of reversals in the polarity of the earth's magnetism which occur every half million years or so. In this diagram of the record associated with the Juan de Fuca Ridge, southwest of Vancouver Island, the crust that bears the earth's present north-south polarity is colored, the colors denoting the progressively older age of the rock east and west from the youngest rock in the ridge itself (red). Straight lines delineate faults offsetting the magnetic pattern. The diagram is from *The History of the Earth's Crust,* edited by R. A. Phinney [Princeton University Press, 1968]. Reprinted by permission.

SCIENTIFIC AMERICAN

Will devote the entire September 1969 issue to

THE OCEAN



is an oil field.

for 152 years. It had one purpose. To save the snowy egret from extinction.

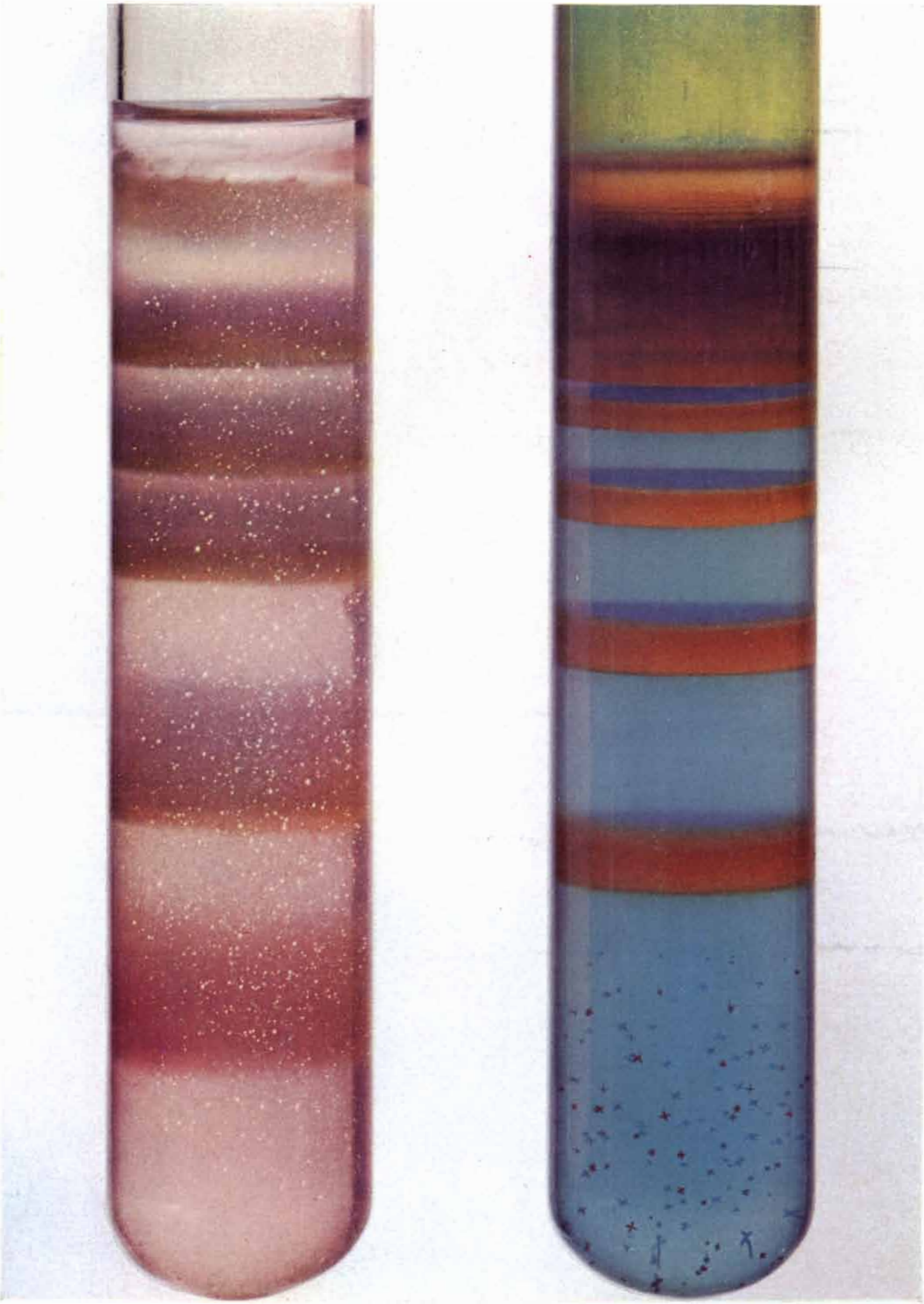
Known as Bird City, the sanctuary started with only seven egrets. Now, over 100,000 nest around its man-made lake every year. To see these alabaster birds sharing their Eden with herons, ducks, coots, swans, cormorants, turtles, deer and alligators is almost a primeval experience. It seems to put the clock back to the beginning.

And, wherever you wander on this peaceful island, you have to look hard to spot the oil wells. Many are hidden by grandfatherly oak trees bearded with Spanish moss. Others are screened by banks of azalea and rhododendron. To Jersey's affiliate, Humble Oil & Refining Company, this respect for environment is only right and proper.

The oil industry provides Louisiana with one-third of its total revenue. But even this contribution would be a poor excuse for defiling beauty or disturbing wildlife.

Amen say the egrets.

Standard Oil Company
(New Jersey)



Liesegang bands: colloidal gold (left) *and copper chromate* (right), *with many gold and copper-chromate crystals also visible*

THE AMATEUR SCIENTIST

Salts react in a gel to make the colorful Liesegang bands

Conducted by C. L. Stong

In 1896 the German chemist R. E. Liesegang, experimenting with photographic materials, discovered a puzzling effect. He had spread in a glass dish a thin layer of gelatin gel that contained potassium chromate. Then he had placed a crystal of silver nitrate in the middle of the gel. Within a few days a pattern of concentric rings grew outward from the silver nitrate, the clear spaces between the rings increasing logarithmically from the center. Liesegang found that the colored rings were crystals of silver chromate embedded periodically in the gel. It was not surprising that crystals should form by the reaction of silver nitrate with potassium chromate, but why did the salts react to form periodic rings?

Other investigators have since repeated the experiment using a number of salts and gels that produce rhythmic banding. Some of the reactions produce secondary logarithmic bands; others form bands that are periodic but not logarithmic. Collectively the patterns display all the colors of the rainbow.

Similar patterns are found in nature, some in minerals such as limonite, variscite and chalcedony and others in animals, including the wings of certain colorful butterflies. Not everyone agrees that any of these natural patterns are formed by the Liesegang reaction, nor has the reaction itself been explained to everyone's satisfaction. The puzzle continues to intrigue experimenters, including Roger Sassen of 80–21 231st Street, Jamaica, N.Y. 11427. Sassen, who is a graduate student in geology at Lehigh University, writes:

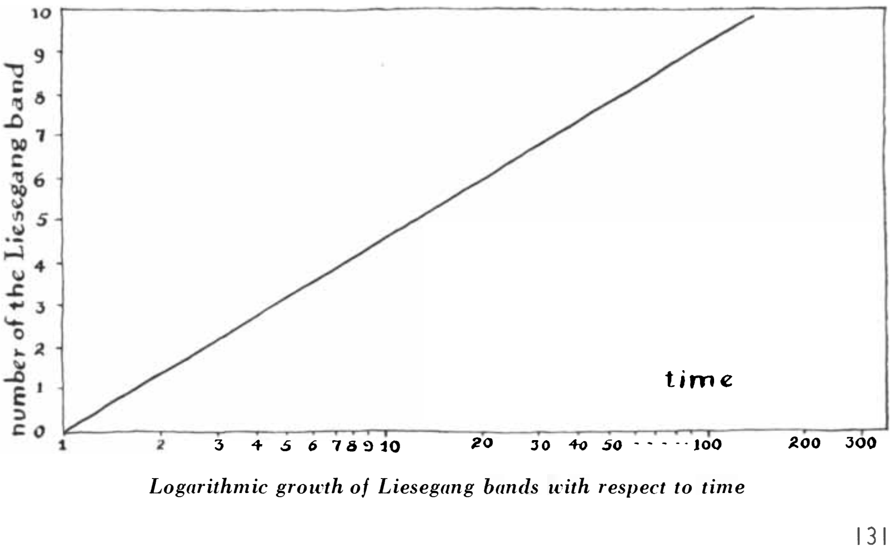
"The Liesegang experiment requires only a few pieces of fairly inexpensive apparatus: a balance that is capable of weighing chemicals to .1 gram, a graduated cylinder of 10-milliliter capacity, another one of 50-milliliter capacity, and a one-milliliter pipette graduated in divisions of .1 milliliter or, preferably, .01 milliliter. In addition you should have a few flat dishes and a dozen test tubes with a capacity of 30 to 50 milliliters for use as reaction vessels.

"The success of the experiments depends in part on preparing solutions of specified concentration. I shall specify concentration in terms of molarity. By convention a one-molar solution (1M) of any chemical contains 6.023×10^{23} molecules of that chemical per liter of solution. This number of molecules, known as Avogadro's number, is exactly equal to the molecular weight of the chemical expressed in grams. For example, a molecule of acetic acid (CH_3COOH) contains two carbon atoms (each of atomic weight 12), two oxygen atoms (each of atomic weight 16) and four hydrogen atoms (each of atomic weight 1). Adding the atomic weights $(24 + 32 + 4)$ gives 60 as the molecular weight of acetic acid. A quantity of 60 grams of the acid contains 6.023×10^{23} molecules; diluted to a volume of 1,000 milliliters with water, it has a concentration of 1M. A .1M solution would be prepared by diluting with water six grams of the acid to a volume of 1,000 milliliters, and so on. The atomic weights of all chemical elements and the molecular weights of many compounds are listed in handbooks of chemistry.

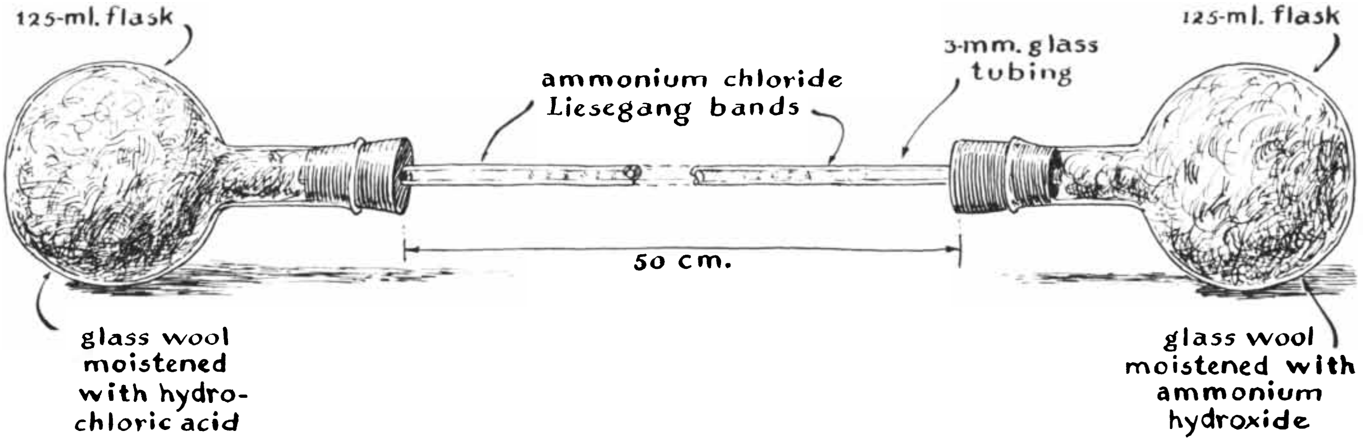
"Liesegang made his experiment with gelatin gel, but a number of other gels can also be used. Among them are agar, gelled blood plasma, cellulose-derived thickeners and silica gel. Most of my experiments have been done with silica gel, prepared by adding acetic acid to a solution of sodium silicate that is commonly called water glass.

"Sodium silicate solution is available from druggists. It must be diluted to a density of 1.06 grams per milliliter. The density of the commercial material is usually unknown, but it can be determined by weighing a specimen of the solution. Weigh a clean, dry container. Transfer to the container exactly 100 milliliters of the solution and weigh again. Subtract the weight of the empty container from the weight of the filled container to determine the net weight of the solution. Typically the net weight of 100 milliliters of a commercial solution of sodium silicate will be about 130 grams.

"Assume that 130 grams is the weight in this example. The weight indicates that the solution contains 30 grams of sodium silicate, because 100 milliliters of water weighs approximately 100 grams. The solution to be used for making gel must contain only six grams of sodium silicate per 100 milliliters of solution. The necessary dilution is found by dividing 6 into 30, thereby obtaining a ratio of 5. Sodium silicate at a density



Logarithmic growth of Liesegang bands with respect to time



Roger Sassen's apparatus for developing bands in a gas

of 1.06 can be made in this case by adding one part of the commercial solution to four parts of water.

"For convenience I usually make up several liters of the diluted stock solution at a time. Use distilled water for all solutions. Diluted sodium silicate solution is converted to gel immediately before use in each experiment. Gel is prepared by adding to the diluted solution an equal volume of acetic acid solution at concentrations ranging from .5M to 1M, depending on the requirements of the experiment. The gel forms in less than an hour.

"When the gel is examined under a microscope, it resembles a water-soaked sponge. The speed of chemical reactions in silica gel is governed by the slow pace at which fluids diffuse through the porous mass. For this reason crystals grow slowly in the gel, but they show a remarkable perfection of form.

"As an introduction to the gel technique I suggest that you grow crystals of silver acetate, which are long, colorless and needle-like. Mix 10 milliliters of dilute sodium silicate of 1.06 density with 10 milliliters of 1M acetic acid and transfer it to a 30-milliliter test tube. After the gel has formed fill the space above it with a .5M solution of silver nitrate. The crystals appear within a matter of days as the silver nitrate diffuses into the gel.

"To reproduce a version of Liesegang's experiment dissolve three grams of plain unsweetened, unflavored gelatin in 96.6 grams of water heated to 140 degrees Fahrenheit. Stir .4 gram of potassium chromate into the mixture. Pour the mixture into a shallow dish and let it cool. While it is cooling dissolve 1/4 gram of silver nitrate in one milliliter of water. When the gel has formed, gently put a single drop of the silver nitrate solution on the gel near the center. Rings of silver chromate should begin to form within minutes.

"I prefer to let the reactions proceed in test tubes, so that bands rather than rings will appear. The tubes can be closed with stoppers to prevent evaporation so that the reactions can continue without attention. The test tube technique is particularly convenient for reactions that continue for several weeks, as in the case of copper acetate and potassium chromate, which combine to form interestingly spaced bands of brown copper chromate.

"Prepare the gel for this reaction by mixing 20 milliliters of sodium silicate with 20 milliliters of a solution consisting of two milliliters of 1M potassium chromate and .6 milliliter of concentrated acetic acid. Transfer the mixture to a 50-milliliter test tube. After the gel has formed fill the remainder of the tube with .25M copper acetate solution, and thereafter replace it with fresh copper acetate solution weekly. When a number of bands have appeared, it is possible to verify the equation that relates the distances of successive Liesegang precipitations to a constant, $Y_n/Y_{n-1} = K$, in which Y is the distance from the interface between the liquid and the gel

Rings of limonite in basalt

to the center of a selected band, *n* is the number of that band, counting from the interface, and *K* is the constant.

"Liesegang bands can be grown within a week by using glass capillaries instead of test tubes. The capillaries should have uniform diameter, as do those used for determining the melting point of chemicals. The sealed ends of melting-point tubes must be broken off. To grow copper chromate bands by this technique I hold the capillary almost horizontally and dip one end in a freshly prepared solution of sodium silicate and an equal volume of .5M acetic acid that is also .025M in terms of potassium chromate. The tube fills by capillary attraction. When the gel has formed, the tube is placed for development in a stoppered test tube containing .25M copper acetate solution.

"The fully developed capillary is removed from the test tube for study. The distances between the bands can be recorded with a pencil by taping the capillary to a sheet of paper and measuring with a metric ruler and machinist's dividers. The bands form so rapidly in capillaries that in some cases it is possible to verify an equation stating that the distance of a band from the interface, divided by the square root of the time required for the growth of the band, is equal to a constant, $Y_n/t_n^{1/2} = K$, in which *Y* is the band of interest and *t* is the time.

"A particularly attractive experiment involves the precipitation of metallic gold and the development of a banded pattern by exposing the gel to sunlight. To make this experiment prepare a sodium silicate solution that includes one milliliter of a 1 percent solution (by weight) of yellow gold chloride. Transfer the mixture to a test tube and make it gel by adding an equal volume of 1.5M sulfuric acid. Gelling takes a week or so. Fill the space above the gel with a solution made by dissolving as much oxalic acid as possible in water, creating a saturated solution. Within a matter of days thousands of minute, sparkling crystals of gold will form in the gel. If all has gone well, Liesegang bands will appear after the gel has been exposed to sunlight.

"The nature of the gel influences the form of the crystallized products, as can be demonstrated by growing crystals of lead iodide. In a test tube prepare a silica gel with 1M acetic acid solution that is .050M with respect to lead acetate. Place over the gel a saturated solution of potassium iodide. Translucent crystals of yellow lead iodide will appear in the form of feathery dendrites and free-growing hexagonal plates. When gelatin gel is substituted for silica gel, the same reaction yields Liesegang banding.

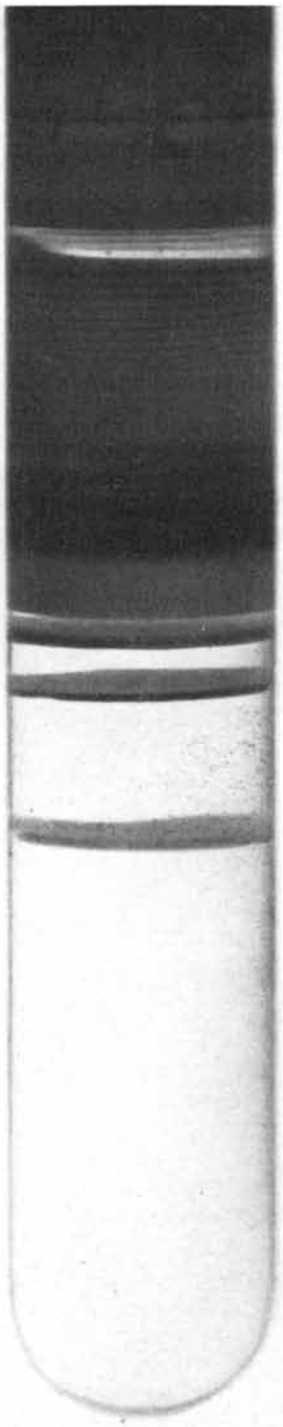
"Spiral patterns can also be grown. An example involves a green precipitate of cobalt hydroxide. Mix five milliliters of .1M cobaltous nitrate solution with 30 milliliters of hot, 10 percent gelatin solution in a 50-milliliter test tube. After the gel cools and sets fill the space above the gel with a .2M solution of ammonium hydroxide. A spiral or somewhat spiral pattern may form, but it does not always appear. If it does not, try again.

"Liesegang suggested that the rings and bands of agates arose when the Liesegang phenomenon operated in natural silica gels that were later altered to microcrystalline quartz, but geologists now generally agree that agate patterns can be attributed to the successive deposit of layers of silica gel and impurities. Banded patterns are also observed in some stalactites, oölites, a few sediments and, of course, tree rings. When the number of each band in the pattern of these materials is plotted against the logarithm of the distance of the band from its origin, the resulting graph is usually a curved line. On the other hand, comparable graphs of nearly all known Liesegang patterns are straight lines [*see illustration on page 131*]. Structures of the mineral limonite that grow in the form of concentric rings during the weathering of sedimentary rocks yield graphs that approximate a straight line. The rings in most specimens, however, are distorted and incomplete.

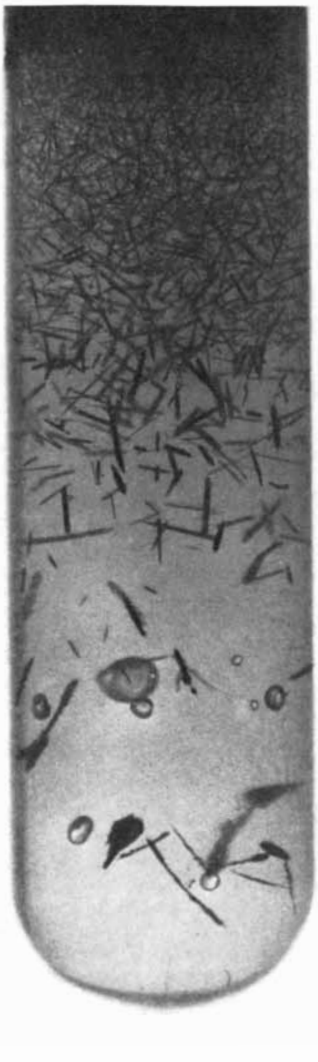
"During the 1920's synthetic agates were prepared in the laboratory with the aim of discovering conditions that could have led to the formation of Liesegang rings in a geological environment. Silica gel containing potassium ferrocyanide was placed in collodion bags and immersed in copper sulfate solution. After a few weeks copper ions diffused into the balls of gel and reacted to form three-dimensional banded precipitates. The synthetic agates were then dried slowly under pressure until they acquired a hardness of about 5 on Mohs's scale, meaning that they were so hard they could barely be scratched with a knife. Although the experiment was suggestive, it failed to prove that agates are so formed.

"A two-dimensional version of this experiment is easy to do. Grow concentric rings of copper chromate in a thin layer of silica gel sandwiched between microscope slides. Pour a silica gel mixture made with .5M acetic acid, which is .025M with respect to potassium chromate, onto one plate and put the other plate on top, taking care not to trap air bubbles in the gel. A thicker layer, in which the rings are easier to see, can be made by spacing the glass plates apart with a few particles of crushed glass.

"Immerse the sandwich in a .25M solution of copper acetate. Within a week, as the copper ions diffuse inward, concentric rings will precipitate. They closely resemble the rings of limonite found occasionally in the cracks of rocks, suggesting that such formations may result from the penetration of ground-



Bands of mercury vanadate



Silver crystals in silica gel

water solutions into cracked rock [*see bottom illustration on page 132*].

"One can demonstrate that the presence of gel in cracked rock is not essential to the growth of such patterns. Sandwich between a pair of glass plates a solution of .01M potassium iodide and let a solution of .5M silver nitrate diffuse into the thin film. Delicate but irregular rings will usually appear. The experiment does not always work, so that you may have to resort to trial and error.

"Other examples of periodic structures that form without gel can be observed by letting a drop of saturated potassium dichromate solution evaporate from a glass surface. Concentric rings of orange crystals usually develop. The phenomenon is called periodic crystallization. Fine spirals of potassium dichromate crystals can be grown by allowing a thin film of solution to evaporate from a warmed glass slide. Crystallization usually begins at the edge and spirals inward, but if a dust particle happens to lodge near the center of the film, crystallization may start at the particle and spiral outward.

"Periodic crystal structures can also be prepared by cooling thin films of molten substances. An example is a thin film of molten sulfur that cools slowly on the sides of a Pyrex test tube. Molten organic substances, such as benzil and acetanilide, will also crystallize as concentric rings.

"Another interesting example of the Liesegang phenomenon, with air replacing gel as the medium, occurs when the fumes of hydrochloric acid and ammonium hydroxide diffuse into a long glass tube from opposite ends. Within an hour a finely banded precipitate of white ammonium chloride smoke will form near the middle of the tube. Construct the apparatus by cutting a piece of three-millimeter glass tubing to a length of 50 to 75 centimeters, fire-polish the ends, wash the tube with detergent solution, rinse it with distilled water and let it dry. (Do not use tubing larger than three millimeters in diameter.)

"Attach to the ends of the tube bulbs that have a capacity of about 100 milliliters and are filled with loosely packed glass wool. The bulbs can be improvised from straight drying tubes or any equivalent scheme such as a pair of 125-milliliter flasks closed with perforated stoppers [*see top illustration on page 132*]. The apparatus must be airtight: a small leak can spoil the reaction. The diffusion tube should be level and protected from abrupt changes in temperature. Moisten the glass wool in one bulb with two milliliters of 10M hydrochloric acid and the wool in the other bulb with the same volume of 1.5M ammonium hydroxide. These solutions can be made from stock reagents; concentrated hydrochloric acid is usually 12M and concentrated ammonium hydroxide 15M.

"Most Liesegang patterns appear as rings or bands, but some of them are symmetrical figures. A straight line drawn through the axis of symmetry divides these patterns into mirror images that resemble the markings of certain living creatures. This resemblance, coupled with the fact that living body cells contain gel, has suggested to some biologists that Liesegang phenomena might be of biological significance.

"Charles Darwin described patterns of color in many organisms. Liesegang was tempted to explain the patterns as examples of periodic precipitation. 'As no ornaments are more beautiful,' he said, 'than the ocelli on the feathers of various birds, on the hairy coats of some mammals, on the scales of reptiles and fishes, on the skin of amphibians, on the wings of many Lepidoptera and other insects, they deserve to be specially noticed. An ocellus consists of a spot within a ring of another color, like the pupil within the iris, but the central spot is often surrounded by additional concentric zones.' Darwin described a South African moth 'in which a magnificent ocellus occupies nearly the whole surface of each hinder wing; it consists of a black center . . . surrounded by successive, ocher-yellow, black, ocher-yellow, pink, white, pink, brown and whitish zones. Although we do not know the steps by which these wonderfully beautiful and complex ornaments have been developed, the process has probably been a simple one.'

"Fascinating patterns of similar design and color can be grown in gel. For example, in one of my experiments the effect was demonstrated by letting a solution that contained both silver nitrate and mercurous nitrate diffuse into a layer of gel between glass plates. The gel, made with .5M acetic acid solution, was .025M with respect to potassium chromate. Incidentally, colonies of certain microorganisms grow in structures that consist of spirals or concentric bands, and Liesegang bands of growth-inhibiting substances have been used to grow cultures of bacteria in the form of concentric rings.

"Not many Liesegang patterns have been made with organic reactants, although several experimenters have induced the periodic precipitation of compounds by reacting inorganic substances with organic compounds. These reactions go quickest if the slowly diffusing organic compound is placed in the gel. I was able to grow bands by placing nickel nitrate solution over silica gel prepared with .5M acetic acid solution that contained a trace of dimethylglyoxime. The possibility that a vast number of organic reactions might proceed in gel media and that they may have biochemical significance suggests that the search for new organic examples of the Liesegang phenomenon could be an exciting and challenging hobby.

"Most of the chemicals used in these experiments are toxic. Some can cause severe burns. Handle them accordingly. Use a rubber squeeze bulb rather than your mouth for sucking solutions into the pipette. Work in a ventilated space and close to a source of running water so that in case of accident the chemicals can be washed quickly from the skin. Store the reacting chemicals out of the reach of children and pets."

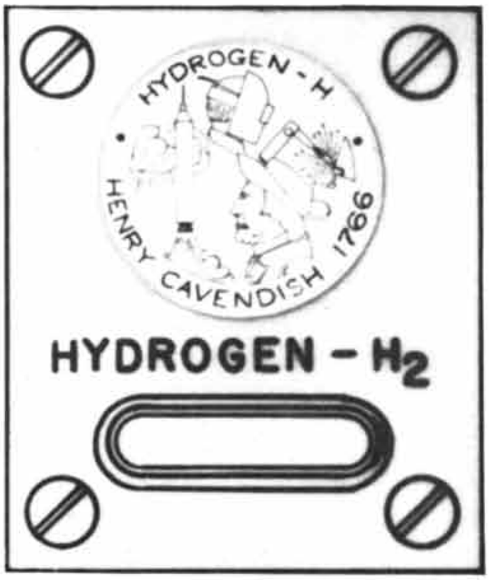
INTRODUCING A UNIQUE, VALUABLE, & EDUCATIONAL POSSESSION . . .

Now You Can Own A Sample Of Every Element Now Known

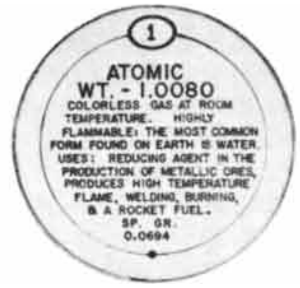
The Mentor Medal-Art Co., Inc. Announces

"DISCOVERIES OF THE ELEMENTS"

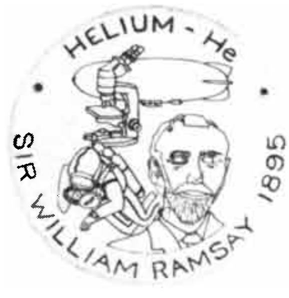
Sterling Silver & Bronze Medal Series.



Actual Size Of Medal & Sample In A Clear Plastic Case.
-Patent Pending-



HYDROGEN-Reverse



HELIUM-Obverse

THE MENTOR MEDAL-ART CO., INC. proudly announces a fantastic new medal series in SOLID STERLING SILVER or SOLID BRONZE.

This commemorative series will portray the discoverer of each of the 103 elements now known to man, along with the principle uses of the element on the obverse. The reverse will list the chemical and physical properties and principle uses.

Included will be an ACTUAL SAMPLE OF THE ELEMENT, whether it is solid, liquid or gas, whenever possible. Otherwise a stable form such as an oxide, or a facsimile, as in the case of the radioactive elements will be used; both will be housed and protected in a handsome, optically clear PLASTIC CASE.

THIS IS A STRICTLY LIMITED EDITION SERIES . . . Only 2,000 consecutively edge numbered solid sterling Silver medals and 10,000 Bronze medals split between edge numbered and unnumbered, will be struck, or the number reached at the time of the subscription deadline, July 1, 1969. After the minting, the dies will be destroyed thus protecting the subscribers from over-production.

The medals will be one and one half inches in diameter in high relief with a "proof-like" surface. SUBSCRIPTIONS WILL BE FILLED ON A FIRST COME BASIS. The medals will be issued at the rate of two per month beginning with HYDROGEN and HELIUM in late August or September of this year.

The cost of the medals, sample element, and plastic case is only $12.00 each for the numbered Silver, $6.50 each for the numbered Bronze, and $5.00 each for the unnumbered Bronze.

DON'T MISS OUT ON THIS ONCE IN A LIFETIME OPPORTUNITY. FILL IN AND SEND THE ATTACHED ORDER BLANK ALONG WITH A CHECK OR MONEY-ORDER FOR THE DESIRED SET OR SETS.

APPLICATION FORM

Mentor Medal Art Co., Inc., Dept. E2,
Box 160, Willoughby, Ohio 44094

NAME
ADDRESS
CITY
STATE ZIP
OCCUPATION
AGE
SIGNATURE
DATE

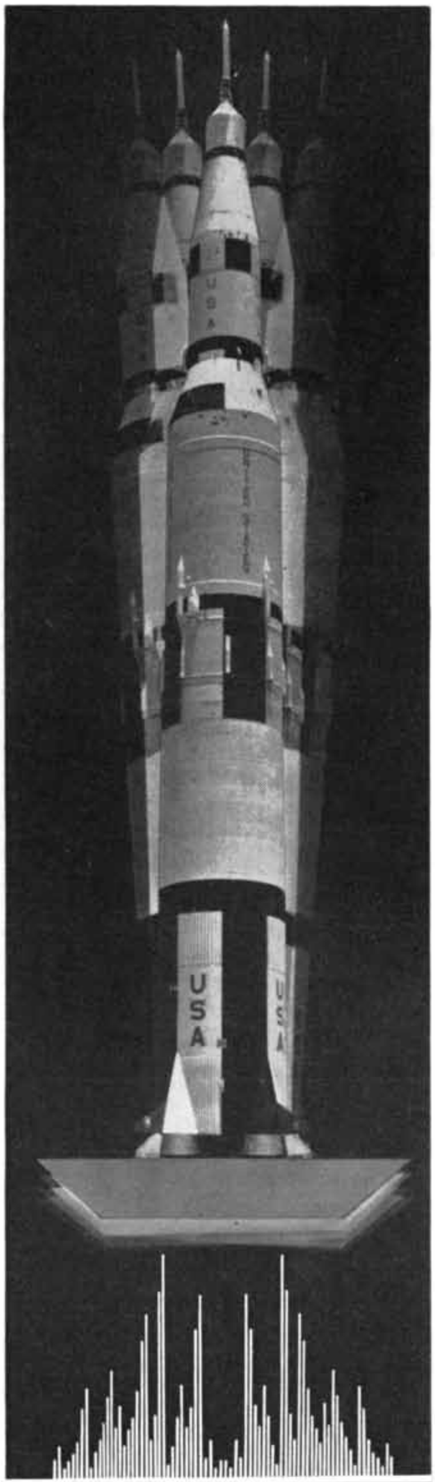
"DISCOVERIES OF THE ELEMENTS"
-Limit One Set of Each Type Per Subscriber-
I have enclosed a check or money-order for the first two issues in:

☐ Numbered silver @ 12.00 each or 24.00
☐ Numbered bronze @ 6.50 each or 13.00
☐ bronze @ 5.00 each or 10.00
Subtotal
Ohio res. please add 4% sales tax
Total

I understand that this subscription entitles me to two Medal/sample issues monthly for which I agree to pay upon being invoiced on a monthly pre-payment basis.

© 1969 SCIENTIFIC AMERICAN, INC

Frequency measurements mean a safe launch for Saturn. A pinpoint re-entry for Apollo. A man's life in Boston.



Shaking the life out of a Saturn V.

Imagine a vibrating platform like the one you stand on for a foot massage. Except this one's 50 feet by 50 feet. And on it stands a rocket over five stories tall. Yet this vibrating platform can shake the life out of a Saturn V launch vehicle. Some massage!

What the engineers would like to know are the critical vibration frequencies that could destroy a rocket at launch or in flight. Hewlett-Packard has just the instrument to make sense out of all the shakes and rattles.

With this same instrument, electronics engineers can perform more detailed circuit analyses, doctors can look closer at EKG and EEG signals, architects can probe building structures, physicists can dig deeper into nuclear magnetic resonance, and astronomers can minutely scan the noisy universe.

In the case of Saturn, the HP 5450A Fourier Analyzer can perform in seconds a complete vibration analysis, resolving all the component frequencies to display on a scope the power spectrum indicating any dangerous resonance. At the push of a button, it will convert the power spectral density function to the correlation function—using Fourier transforms.

The HP Fourier Analyzer can perform as a Power Spectrum Analyzer, a Correlator, an Averager, a Digital Filter, or you can select from a keyboard any combination of these functions. Using digital techniques and Fast Fourier Transform Algorithm you get the speed and versatility to perform Time Series Analysis up to 25 kHz. Since a straightforward keyboard is used for programming you don't have to be a digital computer expert. However, as a side benefit you do have the full computing power of a standard HP Model 2115A Computer whenever you call for it from the keyboard. System price is approximately $50,000, depending on options.

The full 5450A story, far more intriguing than just massaging Saturn V, is yours for the asking. Simply write for the 5450A Data Sheet and mention your application.

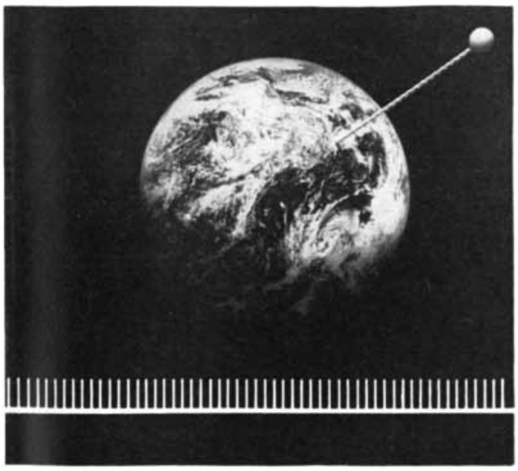
Apollo, journey into time and space — and back.

Apollo re-entry and splashdown precision is no happy accident; it stems from astronaut practice and skill—and, in part, HP dedication to microseconds of time.

Microseconds might seem irrelevant in week-long, million-mile space flights. But such precision timing is critical to mission success. It helps tracking-radars detect minute increments of Doppler shift caused by spacecraft movement—even at lunar distances—so as to determine the vehicle's speed precisely and to plot its position within feet. NASA's worldwide computer network, 15 land stations and 3 ships, depends on meticulously synchronized data exchange necessary for correlating the trajectory calculations. In the event of engine malfunction, behind the Moon or in deep space for example, such information would be vital in getting the astronauts safely back to Earth.

NASA's timing system was completely updated before Apollo 8 and played a vital role in the lunar orbit and subsequent Apollo experiments. Tracking stations in California, Spain and Australia were each equipped with a new primary frequency standard—the HP Precision Frequency Source (E02-5061A), keyed to an HP cesium clock. These primary references and their slaved rubidium and quartz standards were synchronized with a portable HP atomic clock that was flown on a pre-launch round-the-world trip.

If minute details—in time—govern your time- or frequency-dependent studies, write for HP Application Note 52, "Frequency and Time Standards," a 100-page discussion of the practical aspects of their applications.



Hearts shouldn't lose their tempo undetected.

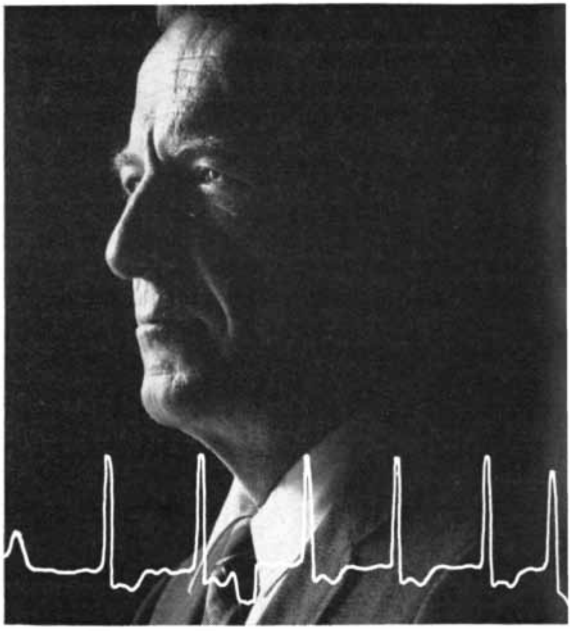
The chances of complete recovery for patients who have suffered a common type of heart attack called myocardial infarction are generally good. But complications after this trauma can greatly complicate the physician's job, to say nothing of threatening his patient's life. Some 40% of all fatalities in the hospital due to myocardial infarction are caused by arrhythmias—electrical disturbances within the heart causing irregularities in the beat such as speed-ups, pauses or double-beats.

Doctors know arrhythmias can often be successfully prevented by medication, but the usual coronary instrument alerts the doctor only when a serious arrhythmia is in process. Again, doctors know that major arrhythmias—the potentially fatal ones—are almost always preceded by less dangerous ones. The problem is to detect these small irregularities.

Since monitoring for medicine is part of Hewlett-Packard's business, we've been searching for a means to help doctors detect smaller arrhythmias, to give them a head start on treating major ones. The new 7822A Arrhythmia Monitor is such a means. When inserted into a patient monitoring system, it measures a pattern of heart signals that the physician establishes as "normal" for a particular patient. Characteristics of this "normal" pattern are stored in the instrument's memory, and each subsequent beat is compared to it instantaneously.

Beats which are abnormal as compared with this stored "normal" are classed as ectopic or out-of-place. The number of these beats occurring in a given time period is numbered and recorded so that the doctor can assess the success of drug treatment. Any excessive ectopic activity will trigger alarms to indicate a more serious condition.

This "extra ounce" of prevention can make monitoring even more effective toward achieving a complete recovery from myocardial infarction. If you would like a definitive description of this new instrument, drop us a note asking for publication No. 7822A. We'll also send you our pamphlet, "How Patients are Helped by Intensive Care Monitoring."



If measurement of frequency isn't of personal concern to you, let us know how we can help you—in measurement, analysis or computation. Hewlett-Packard, 1503 Page Mill Road, Palo Alto, Calif. 94304; Europe: 1217 Meyrin-Geneva, Switzerland.

HEWLETT hp PACKARD

Measurement, Analysis and Computation

00904

BOOKS

How man and his domesticated plants crossed the oceans, and other matters

by Philip Morrison

PEOPLES AND CULTURES OF THE PACIFIC: AN ANTHROPOLOGICAL READER, edited by Andrew P. Vayda. The Natural History Press ($7.95). SEA ROUTES TO POLYNESIA: AMERICAN INDIANS AND EARLY ASIATICS IN THE PACIFIC, by Thor Heyerdahl. Rand McNally & Company ($5.95). AGRICULTURAL ORIGINS AND DISPERSALS: THE DOMESTICATION OF ANIMALS AND FOODSTUFFS, by Carl O. Sauer. The M.I.T. Press ($2.95). Over almost an entire hemisphere the broad Pacific curves, a fluid barrier and a way of travel for man. Shoaling from its depths are found more than 20,000 islands, and perhaps one man in every 1,000 lives there, formed into a wildly distinct and diverse set of cultures. How and when did man get there? Whence did he come? Was the ocean a barrier or a road for man before the Iberian voyagers? These are the main topics of the early history of man in the Pacific, a central theme for these three books. *Peoples and Cultures of the Pacific* is a thick anthology, systematically presenting papers from many sources, mostly written within the past decade, that amounts to a course of study in the whole field. The geography is well laid out, the language studies, history, folklore and politics are all touched on, rather in the sense of a wide sampling than of a brief survey. It is the scale of the ocean itself that emerges most strikingly from the story of Pacific man. First, there is a preponderance of evidence to the effect that accidental one-way voyages, in which the canoes of islanders, out for a day's fishing or a three-day voyage between the islands of an archipelago, were swept thousands of miles away by a sudden typhoon, occasionally to float the survivors ashore on a distant island, to which they added their genes and their small inventory of goods and ideas. Sometimes they were the first men on the new shore. This strange mode of migration has clearly had more impact than the alternative possibilities, deeply embedded in folklore, of intrepid or perhaps desperate planned exile by sea, with the hope of a landfall. There are many features of the Polynesian culture-dispersal pattern that support this chancy model, first argued a decade ago by Andrew Sharp.

The sea giveth and the sea taketh away. If man has been sown like seed by the willful typhoon, he has also been destroyed by its blasts. Here we have the moving account of the fate of the 500 people of Ulithi Atoll before Typhoon Ophelia in the fall of 1960. The winds reached a velocity of 100 or 120 knots; the sea surged four feet deep over the coral. Most of the canoes, vital for fishermen, were destroyed and most of the crops were ruined. In the old days there would have been few survivors. The history of Polynesia is the history of sudden decimation by an unpredictable storm; now the shock was more personal and cultural. The people survived with outside help, but their taboos, the old deference of women, the restrained silence of children, with many of the old ways already threatened and in slow flux under the external influence of loran and Coast Guard stations, were swept away like the trees. Nothing like that had happened in the 1907 typhoon; it had killed many of the people but the social forms stood firm. The eyewitness account by William A. Lessa makes a remarkable chapter.

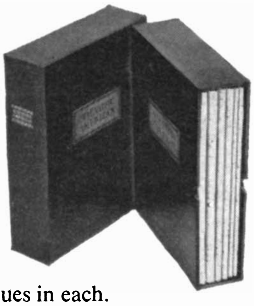
Thor Heyerdahl made his voyage on the balsa raft in defense of a theory, the theory of the Pacific as a highway by which Peruvian men and ideas came to Polynesia. Here we have a set of nine of his papers, mostly written in the 1960's, expounding and elaborating on this theme. His work on the curious multiple centerboards of the Incas that make balsa rafts such as *Kon-Tiki* capable of tacking and sailing into the wind, and his understanding of long sea voyages under plausible early circumstances, are fully convincing. He argues well that the straight lines of the map are not good measures of sea routes on the curved Pacific, where a voyage from China to Peru is as long whether taken via New Zealand or Seattle or Fiji. The sea surface moves in persistent currents, and that motion is dominant compared with the slow sea speed of a raft or a sailing canoe. Thus Heyerdahl argues that, whereas there are no good sea routes out of America across the Atlantic, there are plausible routes, one equatorial and one southerly, out of America toward Asia, and one "natural passage" from Indonesia north of Hawaii to Mexico, found by Spanish shipmasters in 1565. Much less compelling is the evidence that by such routes Polynesia was settled, or that cotton, gourds, sweet potatoes and yams were taken to Asia.

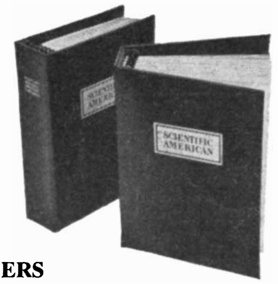
The last of these three books is a familiar and brilliant set of lectures by the geographer who more than anyone else has analyzed the nature of agricultural invention. He has seen that there is a real difference between propagation by seed and by cutting. He argues that agriculture begins with a sedentary tropical fishing people, hand-tending their watery fields until their root crops, propagated vegetatively, have become man-made plants no longer setting seed at all, such as taro, bananas and yams. The cereals are an entirely different technology. In Professor Sauer's view, Aleutian voyages of ancient times brought such an art of farming to the Americas. We face once more postglacial Pacific crossings. To these well-known and stimulating lectures this second edition in paperback adds a few smaller themes: one is the pre-Columbian use of maize, even in Italy. These notions mingle with those of Heyerdahl, but it is not argued that the main spread of maize in Europe and Africa came much before Columbus. It is possible on this view that the millenniums that brought maize across the Pacific and the Indian Ocean to Milan ended by chance at just about the same time the Atlantic round trips began to short-circuit the voyage. That will require evidence stronger than the comparison made in a single phrase by an observant cleric in 1493, writing from Spain to his old friend in Rome that the Indians' maize was like the millet of

Milan. How much alike were the maize and the millet? Petrus Martyr describes maize rather fully for an already familiar plant. Maize still seems the prize of Columbus, even after that story.

The Stanford Two-Mile Accelerator, R. B. Neal, General Editor. W. A. Benjamin, Inc. ($35). The longest straight line in the world—two miles end to end and straight to about a fiftieth of an inch—runs in an evacuated two-foot aluminum pipe in a 10-foot-square concrete tunnel buried in the foothills just west of the Stanford University campus. The line is set by a laser beam, with an ingenious series of Fresnel lenses to fix its position. The stout pipe supports 1,000 equal lengths of brazed copper and stainless steel disk-and-cylinder parts that join to form a flanged high-vacuum channel for electrons and fields. Into this wave guide nearly 250 multi-megawatt pulsed klystrons pump their radio-frequency power from above the ground, at a wavelength of about 10 centimeters, up to 360 pulses per second. The carefully phased voltages that appear in the wave guide pull electrons along the tube, a few million electron volts gained per foot, until two miles downstream (although only about a yard along according to the fast-moving electron's Lorentz-contracted and equally valid point of view) they possess 20 billion electron volts of energy and flow with a peak current of about a fortieth of an ampere. Every few months the machine is realigned; sections may drift a millimeter or so during that time as a result of the earth's settling. The San Andreas rift zone lies about half a mile from the injector end of the machine, and one can expect the line to curve by a couple of millimeters a year. The beam is only a few millimeters across, so that it has plenty of room in its three-quarter-inch pipe. Small deflections are sensed and the beam is centered by the operator, who can set currents in small coils to bend the path slightly. The earth's field is of course degaussed out, and even the effect of a great earthquake with a nearby epicenter is expected to be met by the lateral adjustment of a foot each way that the main pipe design provides for.

This fat book, with 90 experts contributing its 27 chapters, is a full account of the machine and all its systems, from architecture and administration to cafeteria and cooling water. The machine reached the 20-billion-electron-volt design energy at about half design current in early 1967, 10 years after the proposal was submitted. The cost was with-





in the budget of $114 million, the time was as promised, and operation is steady, looking toward 80 percent productive time. The limit on current is so far set by an unexpected and subtle interaction between sections that is mediated by the beam current itself, inducing transverse radio-frequency fields that cumulatively break up the beam at somewhat low currents, although they are still 10 times higher than the world record for billion-volt electrons.

The general reader, no less than the specialist, can find a great deal of understandable and interesting material in this knowing, crisp, detailed account of a project in big science. For example, what do you do with this deadly beam once you have studied the processes it starts up in a room-sized spark chamber or a bridge-sized spectrometer? If the fast beam is stopped in high-atomic-weight material, say copper or iron, the electron and photon cascade stops short in a few inches. The power density is enough to melt anything you might have, no matter how hard you try to cool it. Therefore the beam is stopped gradually instead, in hydrogenous stuff—flowing water—where the stopping electrons can spread out many feet in the material and cooling is practical. Much of the water is decomposed, but there is a catalytic recombiner to keep the hydrogen from leaking too much into the unventilated, rather radioactive room shielded under many yards of concrete and rock fill, a shielding beefed up with five dozen surplus 105-millimeter cannon barrels.

The report is here and there too bland. The long and interesting fight over the main 220,000-volt power lines, which spoiled the view of a nearby well-to-do suburb, is merely alluded to. The solution reached was not the expensive underground line the neighbors sought but an inexpensive and quite unobtrusive design of tubular steel poles replacing the big latticework towers. The history of this compromise ought to have been outlined with more candor. How much money was saved? What does Woodside think now? The story of the great linear accelerator is a genuine success story; it need not fade into a company handout.

MAGNETIC COMPASSES AND MAGNETOMETERS, by Alfred Hine. University of Toronto Press ($30). The fact that a pivoted magnetic needle aligns itself with a generally northward terrestrial field has been known for a couple of millenniums since its discovery in China. In wooden ships, as in a forest, the needle works well enough. It remains useful in these days of iron ships, although the casual experience of the city dweller would suggest that the compass points not north but to the nearest steam radiator. Captain Bligh added to his administrative shortcomings the fault of storing his iron pistols in the ship's binnacle.

In this detailed and devoted volume by a late scientific officer at the Admiralty Compass Observatory is laid out the theory and practice of giving the magnetic needle "a fair chance" to provide an indication of direction. The first problem is of course dynamical: ships and aircraft turn, pitch and roll. The simplest therapy is gimbals. The needle naturally oscillates; the design of a compass card with a long period—weighted by a peripheral ring—and air damping was the work of Lord Kelvin. Liquid damping came later because of many practical difficulties, entering the Royal Navy in dreadnought days and the merchant marine only during World War II. The field the compass measures can only be the local field; a modern ship contributes to this field by the permanent magnetization of its many steel parts, and by an induced field that varies with the ship's heading. Bligh's pistols are as nothing compared with a big steel deck! It was Denis Poisson and then Sir George Airy who showed a century ago how to correct the field by homeopathic means; permanent magnets compensate for the ship's fixed perturbations, and adjustable soft-iron rods and spheres—often seen at the binnacle—for the rest. Theory and practice are all here, although the theory is given in the old-fashioned form, without vectors, matrices or explicit multipole expansion.

Between the two world wars came the gyrocompass. This expensive and elegant device swept the Royal Navy, until its officers "scarcely appreciated the advantages of the magnetic compass." In World War II the advance of electronics meant cheaper servomechanisms to repeat the magnetic compass readings for many to use, and to allow placing the compass at a suitable spot, often high on the mast or at a wing tip. The magnetic devices have also made a strong comeback in hybrid form: the gyromagnetic compass. Here the gyroscope holds a platform steady against the faster motions, while the magnetic needle responds to an average north-seeking force, without the hours of settling needed for the usual gyrocompass to feel the weak Coriolis force. All these devices are described in some detail, with much of the evolution of each kind. The treatise does not omit the needleless magnetic sensors, such as the saturable inductor, and even the nuclear-magnetic-resonance devices now used for field-strength measurements. The heart of the tale is the north-seeking needle, seen here, as it rarely is, in its genuine context, not in the artificial simplicity we too often impose on the familiar. A well-swung compass is true to magnetic meridian to a tenth of a degree on shipboard; it is better than the North Star.

THE EPITOME OF ANDREAS VESALIUS, translated from the Latin with a preface and an introduction by L. R. Lind, with anatomical notes by C. W. Asling. The M.I.T. Press ($3.45). TABULAE ANATOMICAE, by Bartholomeus Eustachius, with Latin notes by Jo. Maria Lancisius. Editrice Parnaso, Modena ($22). Vesalius' 1543 publication of the *De humani corporis fabrica,* in seven books, with its woodcut anatomical illustrations of extraordinary beauty and unprecedented accuracy, drawn from the cadaver, is often regarded as reflecting the beginning of modern science. The brilliant young Flemish professor at Padua also published, only two weeks after his great *Fabrica,* this compact manual, a digest probably intended for medical students without time or money for the bigger work. We now have the *Epitome* in paperback, a new reprint of a fine translation done by a Kansas historian 20 years ago. The Latin text is here too, complete in facsimile, photographically reproduced with all the figures, at about half the original size. Here is a chance for everyone to own the famed Vesalius skeleton shown, like Hamlet, holding a skull in bony silent reflection. There are about a dozen other full figures of the body in various stages of dissection, drawn by the engraver of the bigger work. The book is a bargain.

The second work is altogether grander. It is a full-sized photo-offset reproduction of a beautiful anatomical atlas published in Rome in 1714, with the text entirely in Latin. Nearly everyone who flies knows of the author by the ducts from middle ear to throat named after him, those bypasses whose slow functioning causes the pains of depressurization. His own unhappy story is much less well known. This book presents his work in 50-odd full-page copper engravings, distinctly more precise and detailed than the work of Vesalius, although by comparison austere and unornamented. Eustachius was a contemporary of Vesalius', and produced his own atlas of plates probably within a decade after the *Fabrica.* His techniques—injection of vessels, chemical

UNABRIDGED
MASTER CATALOG of
BOOKS in PRINT

aids to dissection—were remarkably modern. He managed to publish only the first eight plates, in a monograph on the kidney. Neither he nor his literary executor ever saw his full-scale anatomy published. More than 150 years later the plates were found in the possession of a descendant of Eustachius' most loyal disciple. Sponsored by Pope Clement XI, they came out in this volume with a set of notes written by the Pope's physician. About half of the plates show the full figure; the rest are elegantly drawn presentations of respiratory or reproductive systems, of the parts of the eye and so on. The central nervous system appears here in a form as complete and correct as the state of 18th-century knowledge. It has been considered by judicious historians for a generation that, had Eustachius been able to publish his work, anatomy would have advanced along its path of growth earlier by more than a century. The figures are compelling in their fineness of line and richness of understanding. Although the marvelous landscape background of the *Fabrica* is lacking, these plates are handsomely and naturally composed. The volume is an inexpensive attraction for schools, libraries, collectors and physicians.

WHOLE EARTH CATALOG: ACCESS TO TOOLS. Portola Institute ($5). On the heavy paper cover of this large newsprint brochure there appears, true to the name, the whole earth photographed in color from a synchronous satellite. The 60 inside pages list and discuss, in a kind of reasoned and personal set of reviews, with copious facsimile citations and illustrations, many books, services and materials, all of which are easily gained by mail. The *Whole Earth Catalog* is more a pointer than a seller, as it engagingly says, although the institute will accept orders for many items listed.

The document is good reading: it assembles the input and output channels of a strangely attractive contemporary subculture, that of the dissenter and reformer who seeks to construct a philosophical, personal and economic refuge from this curious industrial society. The philosophy is the appealing one of "understanding whole systems"—from Buckminster Fuller and D'Arcy Wentworth Thompson to the famous children's book of Kees Boeke (the universe seen in 40 order-of-magnitude jumps). The economics is consistent: it includes domes for houses (both their established study and their improvised construction—"the crystallographic theory and junkyard practice"), mushrooms, beekeeping, teepees, village technology and campcraft. There is, for example, an attachment for converting your chain saw to a one-man lumber mill, a knowing account of the best kerosene lamps ("Coleman lamps are terrible—they hiss and clank and blind you, just like civilization") and information about automobile repair data, consumer guides, extrication of Government documents, discount houses and the better mail-order suppliers of camping gear, radio parts, optical components—even art reproductions. The best educational reform and popular science are devotedly included, with a catholic spread of taste. Most of the reviews are from readers and users themselves; the catalogue solicits more, for which a small payment is offered. The personal side is conservationist, rural and iconoclastic; the books of Yoga and Tantra are here, with the journal of the Modern Utopians. Just as in the wider society, there is a strong component of the irrational and the indulgent, more from the supplement, which contains the letters of readers, than from the more disciplined rebelliousness of the editors. There is a good stock of a few Anglo-Saxon words, and some touching visual exhibits of drug-induced mental deterioration. Yet even this ugly side of the subculture is more benign, one can argue, than the cancers of war and cruelty, which proliferate in the publications of our straight world and which are completely absent here.

VERTEBRATE PALEONTOLOGY, by Alfred Sherwood Romer. The University of Chicago Press ($12.50). NOTES AND COMMENTS ON VERTEBRATE PALEONTOLOGY, by Alfred Sherwood Romer. The University of Chicago Press ($6.95). *Vertebrate Paleontology* has been out in its third edition for a couple of years now. It is a fitting continuator of the work, for a generation the standard text on this entire subject. It is a "hard" book—nearly all the figures are of bony skeletons, close to the data, with no soft reconstructions—but it is not a dry one; its lightness of manner and its wisdom of thought and phrase carry the reader through many a page of detail. It hardly requires comment at this late date. Its new little companion volume is a lively first-person gloss on the problems and progress of the field for the past two decades. Together they constitute a modern classic that will endure until the coming of a more quantitative paleontology.

A few of Professor Romer's opinions are worth mentioning here. "There came a generation of students whose interests appeared to lie not in...the study of animals...but in the curious field of the study of the *names* of animals." The strict adherence to formal priority that arose in the international zoological bodies during the 1920's has unsettled much familiar usage. Professor Romer urges a "deliberate refusal to submit," very much in the tone of the times! One foresees a museum sit-in. The drifting of continents is another timely topic. Romer writes: "I was brought up in an atmosphere in which it was maintained that continents had always been in the exact position in which they are found today.... I have found myself weakening...and gradually falling away from the 'true faith.'... I can say that Paleozoic faunal distributions could be much more readily understood if the Atlantic Ocean had not been in existence in Paleozoic days." Finally, he holds to his clever interpretation of the big "sail" on the back of the finny dinosaur *Dimetrodon*. The beast did not participate in regattas; its sail was in fact a solar heater and a cooling radiator. The evidence cited is that the fin area varies directly with the volume of the animal, so that the bigger species have huge sails indeed.

THE NEW COSMOS, by Albrecht Unsöld. Translated by William H. McCrea. Springer-Verlag New York Inc. ($6.50). There are several up-to-date and handsomely illustrated introductory astronomy texts by able authors. All of them have a single audience in mind: the college students who take elementary astronomy as a first or only science. There are also many works that are specialized in topic and employ the astronomer's argot in more or less sprightly fashion. Here is a paperback introduction to modern astronomy with a difference. Its intended audience is a student of science, or a scientist in another field, to whom quantitative argument in the full language of analysis is a tool and not a barrier. Concise, vigorously written and ably translated from the German, its main ideas allowed "to stand out plainly in their scientific and historical settings," it is the way into this science for such a reader. Its depth and authority remind one of the classic American text of a generation ago: Russell, Dugan and Stewart. Unsöld's book is more advanced in mathematical level, although it is only a third as long, and is thus rather too concise as an overall introduction. It spans everything from the meteorite flux as a function of particle mass to the number density of galaxies as a function of their luminosity, all in under 400 pages.

INDEX OF ADVERTISERS

JUNE 1969

BIBLIOGRAPHY

Readers interested in further reading on the subjects covered by articles in this issue may find the lists below helpful.

SCIENCE POLICY IN THE U.S.S.R.

Soviet Research and Development: Its Organization, Personnel, and Funds. Alexander G. Korol. The M.I.T. Press, 1965.

Science and the Soviet Economy. R. W. Davies. University of Birmingham, 1967.

Science Policy in the U.S.S.R. Eugène Zaleski, Joseph P. Kozlowski, Helgard Wienart, R. Amann, M. J. Berry and R. W. Davies. Organisation for Economic Co-operation and Development, 1969.

NUCLEAR TRACKS IN SOLIDS

Solid-State Track Detectors: Applications to Nuclear Science and Geophysics. R. L. Fleischer, P. B. Price and R. M. Walker in *Annual Review of Nuclear Science*, Vol. 15, pages 1–28; 1965.

Tracks of Charged Particles in Solids. R. L. Fleischer, P. B. Price and R. M. Walker in *Science*, Vol. 149, No. 3682, pages 383–393; July 23, 1965.

Fossil Alpha-Particle Recoil Tracks: A New Method of Age Determination. W. H. Huang and R. M. Walker in *Science*, Vol. 155, No. 3765, pages 1103–1106; March 3, 1967.

WOUND HEALING

Wound Healing. Leslie B. Arey in *Physiological Reviews*, Vol. 16, No. 3, pages 327–406; July, 1936.

Some Biochemical Aspects of Fibrogenesis and Wound Healing. David S. Jackson in *The New England Journal of Medicine*, Vol. 259, No. 17, pages 814–820; October 23, 1958.

Cell Fine Structure and Biosynthesis of Intercellular Macromolecules. Keith R. Porter in *Biophysical Journal*, Vol. 4, No. 1, Part 2, pages 167–196; January, 1964.

The Connective Tissue Fiber Forming Cell. Russell Ross in *Treatise on Collagen, Vol. II, Part A: Biology of Collagen*, edited by Bernard S. Gould. Academic Press, 1968.

The Fibroblast and Wound Repair. Russell Ross in *Biological Reviews*, Vol. 43, No. 1, pages 51–96; February, 1968.

TRUTH AND PROOF

The Semantic Conception of Truth. Alfred Tarski in *Semantics and the Philosophy of Language*, edited by Leonard Linsky. University of Illinois Press, 1952.

The Concept of Truth in Formalized Languages. Alfred Tarski in *Logic, Semantics, Metamathematics. Papers from 1923 to 1938*, translated by J. H. Woodger, Clarendon Press, 1956.

L'Antinomia del mentitore nel pensiero contemporaneo. Da Peirce a Tarski. Francesca Rivetti Barbò. Societa editrice Vita e Pensiero, 1961.

Introduction to Logic and to the Methodology of Deductive Sciences. Alfred Tarski. Oxford University Press, 1965.

On Formally Undecidable Propositions of the Principia Mathematica and Related Systems, I and On Undecidable Propositions of Formal Mathematical Systems. Kurt Gödel in *The Undecidable*, edited by Martin Davis. Raven Press, 1965.

HORMONES IN SOCIAL AMOEBAE AND MAMMALS

The Cellular Slime Molds. John Tyler Bonner. Princeton University Press, 1967.

Cyclic AMP. G. A. Robison, R. W. Butcher and E. W. Sutherland in *Annual Review of Biochemistry*, Vol. 37, pages 149–174; 1968.

Cyclic AMP: A Naturally Occurring Acrasin in the Cellular Slime Molds. T. M. Konijn, D. S. Barkley, Y. Y. Chang and J. T. Bonner in *The American Naturalist*, Vol. 102, No. 925, pages 225–233; May–June, 1968.

ULTRAVIOLET ASTRONOMY

The Sun. Leo Goldberg and Edward R. Dyer, Jr., in *Science in Space*, edited by Lloyd V. Berkner and Hugh Odishaw. McGraw-Hill Book Company, 1961.

Ultraviolet and X Rays from the Sun. Leo Goldberg in *Annual Review of Astronomy and Astrophysics*, edited by Leo Goldberg, David Layzer and John G. Phillips, Vol. 5, pages 279–324; 1967.

Astronomical Research with the Large Space Telescope. Lyman Spitzer, Jr., in *Science*, Vol. 161, No. 3838, pages 225–229; July 19, 1968.

Ultraviolet Solar Images from Space. Leo Goldberg, Robert W. Noyes, William H. Parkinson, Edmond M. Reeves and George L. Withbroe in *Science*, Vol. 162, No. 3849, pages 95–99; October 4, 1968.

THE PHALAROPE

Testosterone-induced Incubation Patches of Phalarope Birds. J. E. Johns and E. W. Pfeiffer in *Science*, Vol. 140, No. 3572, pages 1225–1226; June 14, 1963.

Gonadal Hormones in Wilson's Phalarope (Steganopus tricolor) and Other Birds in Relation to Plumage and Sex Behavior. E. O. Höhn and S. C. Cheng in *General and Comparative Endocrinology*, Vol. 8, No. 1, pages 1–11; February, 1967.

Prolactin and Nesting Behavior in Phalaropes. Charles S. Nicoll, E. W. Pfeiffer and H. R. Fevold in *General and Comparative Endocrinology*, Vol. 8, No. 1, pages 61–65; February, 1967.

ANALYTIC INSTRUMENTS IN PROCESS CONTROL

Process Control with Infra-Red Gas-Analysers. A. E. Martin, A. M. Reid and J. Smart in *Control*, Vol. 2, No. 18, pages 108–110; December, 1959.

Refractometers. L. E. Maley in *Journal of Chemical Education*, Vol. 45, No. 6, pages A467–A485; June, 1968.

Process Control by Gas Chromatography. I. G. McWilliams in *Advances in Chromatography: Vol. VII*, edited by J. Calvin Giddings and Roy A. Keller. Marcel Dekker, Inc., 1968.

MATHEMATICAL GAMES

Finite Markov Chains. John G. Kemeny, Hazleton Mirkil, J. Laurie Snell and Gerald L. Thompson in *Finite Mathematical Structures*. Prentice-Hall Incorporated, 1959.

Random Walks: Part Three of Mathematical Conversations. E. B. Dynkin and V. A. Uspenskii. Translated and adapted from the first Russian edition by Norman D. Whaland, Jr., and Olga A. Titelbaum. D. C. Heath and Company, 1963.

THE AMATEUR SCIENTIST

Crystallization: Theory and Practice. Andrew VanHook. Reinhold Publishing Corporation, 1961.

A good hose has to come on like Gangbusters. It has to withstand an awful lot of pressure. For if it doesn't, whatever depends on it will fall flat on its face.

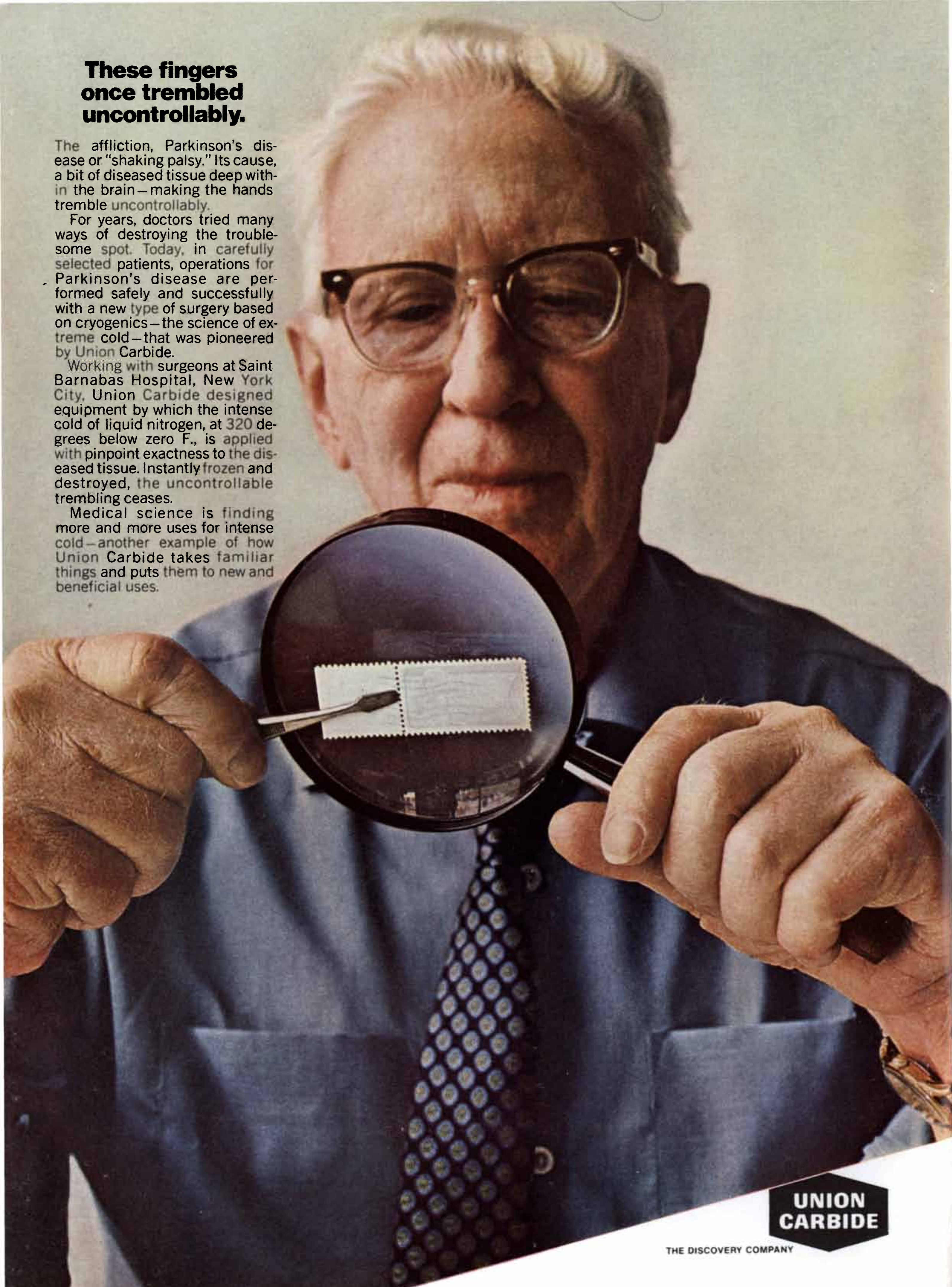
Industrial hose comes from the same tough neighborhood where tires come from. In tires, the strongest, safest ones made are reinforced with rayon cord. Little wonder that when "the boys" wanted extra strength for their hoses, they reinforced them with rayon fiber.

Avisco® rayon is the yarn that's tough enough to take it. Besides greater strength, we can build in better adhesion and extra flexibility. Being man-made we can keep a pretty close watch on every inch of yarn, every inch of the way.

Maybe you have a job to do where rayon can help. Give us a call and let us show you what we're made of. Avisco fibers and yarns. Empire State Building, New York, N.Y. 10001

fmc® CORPORATION **AVISCO FIBERS**

Is there anything we can't make them do?



These fingers once trembled uncontrollably.

The affliction, Parkinson's disease or "shaking palsy." Its cause, a bit of diseased tissue deep within the brain – making the hands tremble uncontrollably.

For years, doctors tried many ways of destroying the troublesome spot. Today, in carefully selected patients, operations for Parkinson's disease are performed safely and successfully with a new type of surgery based on cryogenics – the science of extreme cold – that was pioneered by Union Carbide.

Working with surgeons at Saint Barnabas Hospital, New York City, Union Carbide designed equipment by which the intense cold of liquid nitrogen, at 320 degrees below zero F., is applied with pinpoint exactness to the diseased tissue. Instantly frozen and destroyed, the uncontrollable trembling ceases.

Medical science is finding more and more uses for intense cold – another example of how Union Carbide takes familiar things and puts them to new and beneficial uses.

UNION CARBIDE

THE DISCOVERY COMPANY