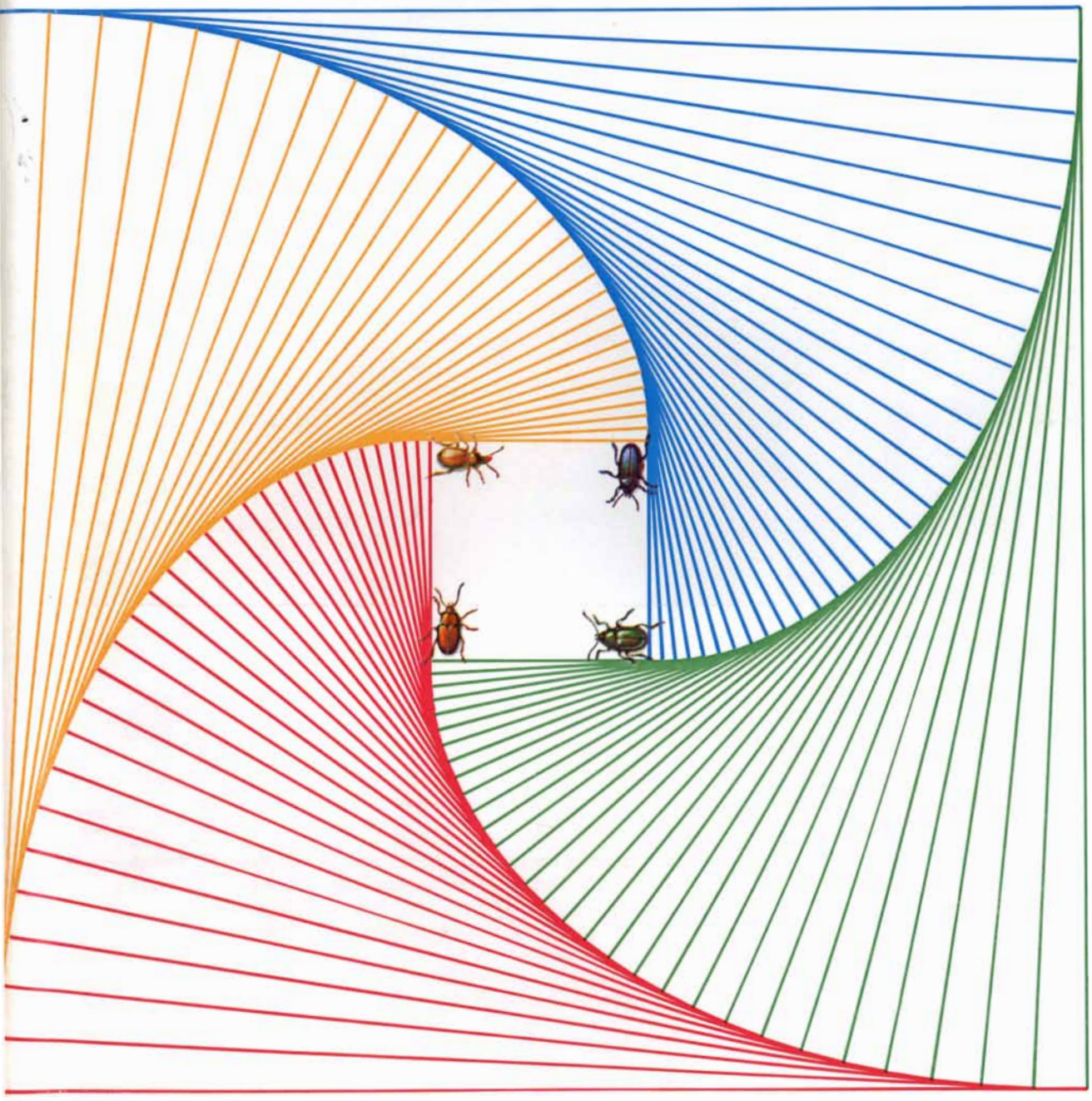
SCIENTIFIC AMERICAN



FOUR-BUG PROBLEM

SIXTY CENTS

July 1965

THIS BANKER HAS A PhD IN CHEMISTRY

Read how outstanding technical ability and complete banking facilities help First National City Bank find imaginative, resourceful solutions to corporate financial problems.

WHEN a First National City banker has a problem, there are a lot of places he can turn within the Bank to help solve it.

First National City Bank's Petroleum Department has banking experts, engineering experts, an economics expert and a PhD in chemistry.

The Bank's Transportation Department makes a career out of the airline, barge, bus and transit, railroad, shipping and trucking industries—including the latest financing and leasing methods.

The Bank's Public Utilities Department specializes in electricity, natural gas transmission and distribution, atomic energy, telephones and telegraphs, and water.

The Bank's Aerospace and Electronics Department is staffed by six professional bankers who spend *all* their time working with companies in today's fast-moving scientific industries.

Specialized knowledge is in constant flow throughout the Bank on a daily basis. In addition, First National City Bank maintains the most complete financial library in all New York.

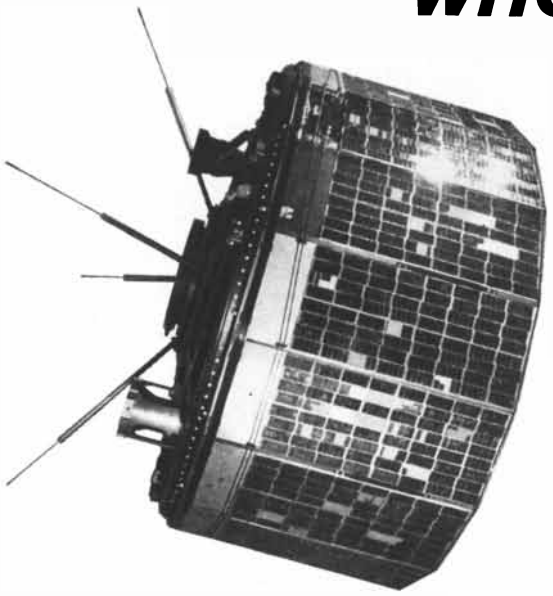
Why all this emphasis on making technical expertise part of every First National City banker's portfolio?

Because the needs of our customers grow more complex every day. The better the banker, the better the chance of meeting these needs—and men who know more make better bankers.

FIRST NATIONAL CITY BANK

399 PARK AVENUE, NEW YORK, N. Y. 10022 • MEMBER FEDERAL DEPOSIT INSURANCE CORPORATION

WHO HAS...



the world's largest airline weather center, with daily reports from weather satellites...

jet schedules you can count on...

and maybe the juiciest steaks in the sky...

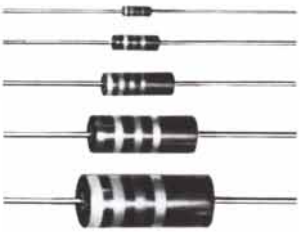
Who but

UNITED AIR LINES

Welcome aboard the Extra Care Airline

This label of "quality" A-B QUALITY covers everything made by Allen-Bradley

EXCLUSIVE HOT MOLDED RESISTORS



World renowned for their conservative ratings and stable characteristics. Due to uniformity of production, long term performance can be accurately predicted. With billions of these resistors in service, there is no known instance of catastrophic failure. Rated 1/8, 1/4, 1/2, 1, and 2 watts at 70°C. Available in all standard EIA and MIL-R-11 resistance values and tolerances, plus values above and below standard limits.

SUBMINIATURE LOW PASS FILTERS

Especially designed for use in cable connectors, FO type filters provide a maximum reduction of RFI in a minimum of space—attenuation is greater than 50 DB over the frequency range from 100 MHz to 10 GHz. With these filters mounted through a ground plane in the connector, there's complete shielding to prevent rf coupling between input and output. Individual filter replacement is possible.

ADJUSTABLE FIXED RESISTORS Type R

Type R built to withstand environmental extremes. Resistance element and terminals hot molded into integral unit with insulated mounting base. Has stepless adjustment and is noninductive. Watertight, can be encapsulated. Rated ¼ watt at 70°C. Values from 100 ohms to 2.5 megohms. Tol. ±10 and 20%. Type N similar in construction but for less critical environments. Rated ⅓ watt at 50°C.

HOT MOLDED POTENTIOMETERS Type J and Type K

Type J potentiometers have solid, hot molded resistance element. Smooth, quiet control. Available in single, dual, and triple units, also with vernier adjustment. Rated 2.25 watts at 70°C. Values to 5 megohms. Type K have similar construction rated 1 watt at 125°C, 2 watts at 100°C, and 3 watts at 70°C.

HOT MOLDED POTENTIOMETERS Type G and Type L

Type G potentiometers are miniature controls with solid molded resistance element. Only ½" in diameter. Quiet, stepless operation. Rated ½ watt at 70°C. Values to 5 megohms. Type L are similar to Type G but rated ½ watt at 100°C. Can be used at 150°C with "no load."

HOT MOLDED POTENTIOMETERS Type H 5 Watt

Provides higher voltage and wattage ratings for industrial and commercial electronic equipment. Quiet, stepless control. Life exceeds 100,000 complete cycles of operation on accelerated tests with less than 10% resistance change. Rated 5 watts at 40°C and 3 watts at 70°C, with a maximum of 750 volts.

FEED-THRU AND STAND-OFF CAPACITORS

Discoidal design eliminates all parallel resonance effects at 1000 Mcps and less. Insulation resistance exceeds 100,000 megohms. Standard values: 470 mmf ±20% and 1000 mmf GMV. Special values from 6.8 mmf to 1500 mmf. Rated to 500 v DC maximum.

FERRITES AND CERAMIC MAGNETS

Ferrite cores are available in a wide range of sizes and shapes, such as flared yokes; U, E, and cup cores, and torroids.

Ceramic permanent magnets are available in nonoriented and oriented materials. Also types especially suited for use in the design of high efficiency d.c. motors.

Please send for Publication 6024:
Allen-Bradley Co., 1204 S. Third Street, Milwaukee, Wisconsin 53204. In Canada: Allen-Bradley Canada Ltd., Galt, Ontario. Export Office: 630 Third Ave., New York, N.Y., U.S.A. 10017.

ALLEN-BRADLEY

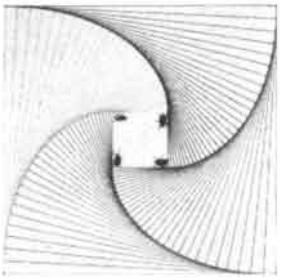
QUALITY ELECTRONIC COMPONENTS

Established 1845 **SCIENTIFIC AMERICAN** July 1965 Volume 213 Number 1

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: $7 PER YEAR.



THE COVER

The Op-art pattern on the cover, a set of nested and rotating squares, is an example of the close relation of such patterns to many problems encountered in recreational mathematics (see "Mathematical Games," page 100). This painting illustrates the "four-bug problem." The bugs start at the corners of a square. Simultaneously they begin to move clockwise, each crawling at the same speed and each heading always for its neighbor. At any moment the bugs mark the corners of a square that becomes smaller and rotates as they crawl. Each bug's path (shown here in the appropriate color) to their meeting point in the center is a spiral the length of which is exactly equal to a side of the original square.

THE ILLUSTRATIONS

Cover by Joan Starwood and Thomas Prentiss

Page	Source	Page	Source
19	John Lutnes, Kitt Peak National Observatory	64	C. M. Fowler, Los Alamos Scientific Laboratory
20	Air Force Office of Scientific Research	65–73	Dan Todd
21	Stanford University	74	A. E. Vatter, University of Colorado Medical Center
22–25	Allen Beechel	75	Roderic B. Park and John Biggins, University of California at Berkeley
26–33	Alex Semenoick	76–82	Joan Starwood
34	Todd Walker	85	Kathleen M. Kenyon, St. Hugh's College, Oxford
36–37	Bunji Tagawa	86	Matson Photo Service
38	Todd Walker	87	Kathleen M. Kenyon, St. Hugh's College, Oxford
39–40	Bunji Tagawa	88–89	Eric Mose
41	Bunji Tagawa (*top*), Todd Walker (*bottom*)	90	Kathleen M. Kenyon, St. Hugh's College, Oxford
42	Todd Walker	92–96	Thomas Prentiss
44	Joan Starwood	97	James Egleson
51	Eric Mose	98	Thomas Prentiss
52	Richard J. Wurtman and Julius Axelrod	99	James Egleson
53	Eric Mose (*top*); David Wolfe, Harvard Medical School (*bottom*)	100–103	Irving Geis
54–56	Eric Mose	104	Irving Geis (*top*), Charles B. Slackman (*bottom*)
57	Douglas E. Kelly, University of Washington	106–110	Roger Hayward
58–60	Eric Mose		

Another RD and E capability report from ALLIS-CHALMERS

Fuel cell A-OK after 1,400 hours . . . and still running!

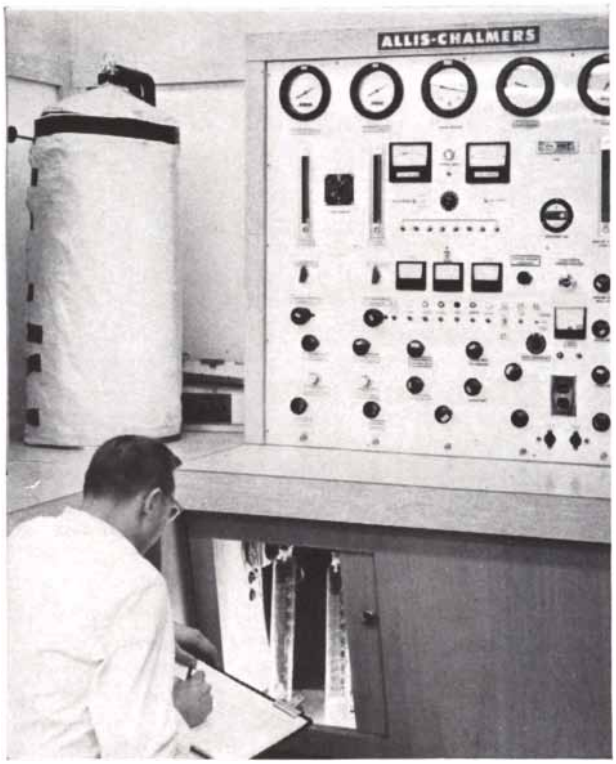
An Allis-Chalmers 28-volt fuel-cell system has successfully operated for more than 1,400 hours — producing both electrical power and drinking water from hydrogen and oxygen. And it is continuing to rack up running time during performance testing in Milwaukee. Allis-Chalmers built this advanced system for NASA.

What's so important about 1,400 hours? A spacecraft could make nine round trips from earth to moon in 1,400 hours. Allis-Chalmers scientists and engineers feel this system's performance proves such fuel cells are suitable power sources for prolonged space missions — such as space stations in orbit 30 to 45 days.

The Allis-Chalmers system converts chemical energy of hydrogen and oxygen directly into 28-volt electricity. During its initial 1,400-hour performance, it efficiently supplied power outputs ranging from 800 to 2,300 watts. Total energy delivered so far is 1,430 kilowatt-hours. Plenty of power for aerospace applications!

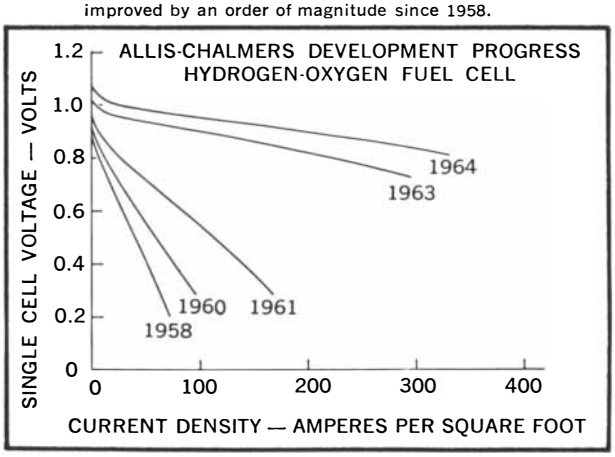
The system is producing 2½ to 3 gallons of water per day, enough for two astronauts. Purity of the water is assured because a distillation process is used to remove moisture from the Allis-Chalmers system. A built-in capillary membrane (we don't need pumps) takes out the water in vapor form. Simply condense it, collect it, and drink it as needed.

May we discuss your aerospace and defense fuel-cell application? Write: Space and Defense Sciences Department, ALLIS-CHALMERS, Box 512, Milwaukee, Wisconsin 53201.



A-C research scientist logs the production of drinking water during 1,400-hour testing of the 28-volt fuel-cell system.

Chart shows Allis-Chalmers fuel-cell technology has improved by an order of magnitude since 1958.



FUEL-CELL design for Aerospace and Hydrospace applications is but one of the opportunities open today for qualified scientists and engineers at Allis-Chalmers. For information concerning employment write to: Manager of Professional Placement, Allis-Chalmers, Milwaukee, Wisconsin 53201.

An Equal Opportunity Employer

ALLIS-CHALMERS

6513

LETTERS

Sirs:

The article "Intense Magnetic Fields" by Henry H. Kolm and Arthur J. Freeman [SCIENTIFIC AMERICAN, April] contains a rather misleading account of the pulsed-field work of Kapitza and his successors. Kolm and Freeman state: "As early as 1924 the Soviet physicist Peter L. Kapitza, then working at the University of Cambridge, succeeded in generating fields as high as 500,000 gauss for several thousandths of a second. He did this by short-circuiting a large alternating-current generator across a small, jelly-roll-shaped coil for a single half-cycle of current." The authors have apparently confused two actual Kapitza experiments.

In 1924 Kapitza used battery discharges into ordinary copper-tape coils of the jelly-roll type. In an article devoted mainly to 40- to 80-kilogauss coils he speaks of a 500-kilogauss pulse through a coil of .1-centimeter bore, but he does not actually describe the coil or its condition after the pulse.

In 1927 Kapitza made the brilliant technological advances that laid the basis for his great work on high-field phenomena in solids. He developed a specially contoured solenoid of cadmium copper with a machined helical inner layer and several wound helical outer layers. He drew pulsed power from a generator and reached 320 kilogauss *without* coil damage. Kolm and Freeman state subsequently: "The first advance since Kapitza came in 1956, when Simon Foner and one of us (Kolm) generated pulsed fields of 750,000 gauss...." This statement disagrees with the literature. (For a review see H. P. Furth, *Science*, Vol. 132, page 387, 1960).

The principal point to be made is that the level of repeatable pulsed magnetic fields has progressively increased since Kapitza's experiments and the technology has been contributed by many people. F. Coensgen's 1954 experiment in the Sherwood project with a repeatable 450 kilogauss in 100 cubic centimeters' volume represented one such advance.

The last decade's progress in ultrahigh pulsed fields has been characterized by two basic technological innovations: the use of high input currents (ranging up to megamperes) and the use of massive single-layer and single-turn coils. These two innovations were made in 1955 by Furth, Levine and Waniek (*Scientific American*, February, 1958). Their single-turn technique, used with hardened steel, still represents the best published method of approaching the megagauss level in permanent coils.

STIRLING A. COLGATE

President
New Mexico Institute
of Mining and Technology
Socorro, N.M.

Sirs:

We regret that it was not possible to do full justice to every detail of Kapitza's work in the four sentences we were able to devote to him, and that our synopsis may have been misleading. Mr. Colgate is correct in pointing out that, although Kapitza's main innovation was the use of kinetic energy from a rotating generator, his highest field was actually achieved by the older method of using a chemical storage battery. Kapitza makes no mention to the effect that his 500,000-gauss coil failed, and we consider it proper to grant him the record.

Our decision to cite as "the first advance since Kapitza" the 1956 magnet of Foner and Kolm (at the risk of seeming immodest) is quite thoroughly justified on two counts. In the first place, the repeatable field of 750,000 gauss we achieved was in fact the first improvement over Kapitza's record. In the second place, the improvement was based on a novel approach made possible by the use of capacitor-storage, one that had not been recognized by earlier workers who used capacitors. We are referring to the technique of using an underdamped or oscillatory discharge, in which the resistive energy losses are small compared with the energy transferred into the magnetic field. The reason this innovation is of such fundamental importance is that in an oscillatory discharge the current is determined by circuit inductance rather than coil resistance, which makes it possible for the first time to use coil materials that are much stronger than copper but whose higher resistance would have completely precluded their use in any of the earlier magnets. It was precisely this innovation that permitted the use of strong, massive coils, whether of single-turn or multiple-turn construction. An oscillatory discharge requires the switching of enormous input currents, a problem we solved by developing a triggered spark gap for this purpose.

Contrary to Mr. Colgate's statement, the innovation of using high input currents was not made by Furth and Waniek. At the time our first 750,000-gauss fields were achieved Furth and Waniek, working at the Harvard Cyclotron Laboratory, were using a pulse transformer to lower the input current so that it could be handled by conventional ignitrons. As a result resistance was so critical in this system that their copper coils were cooled to liquid-nitrogen temperature to achieve lower resistance. They did not exceed the copper limit, or about 350,000 gauss, until they abandoned their method in favor of ours and adopted the use of monolithic coils.

It is unfortunate that Furth, Levine and Waniek failed to make reference to our prior work in any of their subsequent publications, a circumstance that accounts for Mr. Colgate's erroneous impression.

HENRY H. KOLM

ARTHUR J. FREEMAN

National Magnet Laboratory
Massachusetts Institute of Technology
Cambridge, Mass.

Scientific American, July, 1965; Vol. 213, No. 1. 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 Martin M. Davidson, Advertising Manager, SCIENTIFIC AMERICAN, 415 Madison Avenue, New York, N.Y. 10017.

Subscription correspondence should be addressed to Jerome L. Feldman, Circulation 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.

Microfilm correspondence and orders should be addressed to Department SA, University Microfilms, Ann Arbor, Mich. 48107.

Subscription rates: one year, $7; two years, $13; three years, $18. These rates apply throughout the world. Subscribers in the United Kingdom may remit to Midland Bank Limited, 69 Pall Mall, London SW 1, England, for the account of Scientific American, Inc.: one year, two pounds 11 shillings; two years, four pounds 14 shillings; three years, six pounds 10 shillings.

Change of address: please notify us four weeks in advance of change. If available, kindly furnish an address imprint from a recent issue. Be sure to give both old and new addresses, including ZIP-code numbers, if any.

It's a diesel ...it's a gas.

There is nothing prettier than a Mercedes-Benz.

Unless it's two of them.

Consider the beauties above.

Identical except for the wonderful power plants under the bonnets.

At left, the 190 D. The "D" stands for diesel and for extra distance.

At right, the similarly elegant 190, with gasoline engine.

Now, if you drive 20,000 to 30,000 miles per year the 190 D may be for you. Here's real class with smart, business-like economy.

Proved, too. One of the country's top petroleum marketers recently put this car through a 7,000 mile fuel consumption test in a big city.

The 190 D traveled up to 39 miles per gallon in regular city traffic and up to 43.9 miles per gallon over the big city expressways!

And remember—diesel fuel, available practically everywhere, costs about 40% less than gasoline.

But if the idea of a diesel is a little too advanced to digest right now, you might think of the 190.

Here is the same classic beauty. The same quiet, enduring design. Simply runs on gasoline (22 m.p.g., too). That's the difference.

Automatic transmission, if you want, both models.

MERCEDES-BENZ

GOING TO EUROPE? FOR TRAVEL GUIDE, WRITE HANS von BROCKHUSEN, MERCEDES-BENZ OF NORTH AMERICA, INC., 158 LINWOOD PLAZA, FORT LEE, N. J., 07024
MERCEDES-BENZ OF NORTH AMERICA, INC. A SUBSIDIARY OF DAIMLER-BENZ A. G., GERMANY

1

3

2

1. High resolution and simple use of the DY-2801A Quartz Thermometer are shown in this actual readout of the temperature difference between the ice-point and triple-point of water.

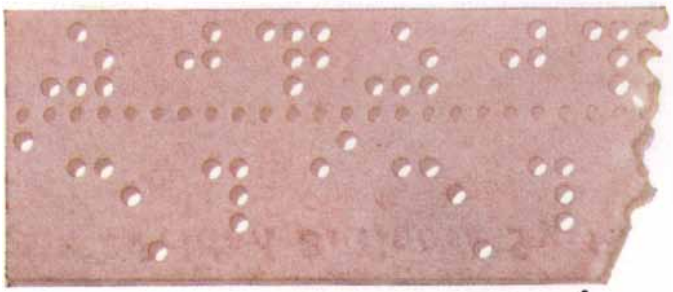
2. One probe of the DY-2801A is pictured actual size; longer and shorter probes also are available. The probe is only .375 inch in diameter.

3. Because the DY-2801A derives digital temperature information from frequency rather than voltage variations, accurate measurements at great distances are easier, less error-prone and free from typical voltage-measuring problems such as noise, cable resistance and inaccurate amplification. It is therefore useful in oceanography, temperature measurement beneath the surface of the earth and industrial processing. The instrument's high resolution, repeatability and stability make it particularly valuable in the laboratory.

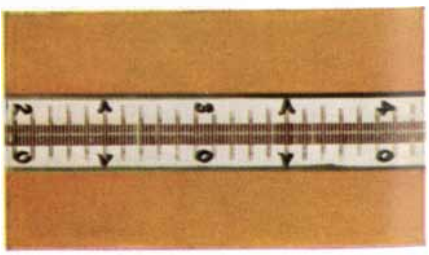
4. Permanent digital recordings of temperature measurements are easily available.

5. The traditional mercury lab thermometer is crude in relation to more advanced platinum resistance thermometers or thermistor devices. New techniques in the Hewlett-Packard quartz thermometer offer high accuracy and resolution, combined with ease of use.

6. The sensor crystal (shown about 5 1/2 times actual size) resonates at 28.208 megacycles at 0° C. Frequency change is linear at 1000 cycles per second per degree centigrade. The Hewlett-Packard thermometer, whose probes contain this crystal, then measures and displays frequency in terms of degrees C or F.



4



5

6

First Major Advance in Thermometry Since 1887—from Hewlett-Packard

Hewlett-Packard's new quartz crystal thermometer achieves a new order of accuracy and resolution, combined with unprecedented ease of use. It is the first major advance in thermometry since the development of temperature-measuring methods which depend on voltage or resistance change.

The traditional wide-range mercury lab thermometer provides resolution no greater than 0.01 degree. Electronic thermometry devices which derive temperature information from variations in voltage or resistance require complicated bridge-balancing procedures and cannot be used at great distances from the point to be measured.

A "reverse concept" in the Hewlett-Packard crystal laboratory has made possible a thermometer which responds to temperature variations with frequency change rather than with voltage change. This technique overcomes many of the inherent disadvantages of previous instruments. In the design and manufacture of quartz crystal oscillators used in frequency and time standards, Hewlett-Packard went to great lengths to isolate crystal resonators from temperature changes which alter their frequency of oscillation. Why not, then, devise a temperature-sensitive crystal whose frequency-versus-temperature characteristic is linear? Analytical procedures led to a practical solution with which such a crystal was produced.

The result was a new thermometer which measures temperature with 0.0001° resolution and displays it in direct digital readout, with positioned decimal and plus or minus indication. Measurements are made by means of a probe which encloses the quartz crystal. There is no need for bridge balancing, conversion charts, temperature references or computer processing. Sensors may be located up to 1000 feet from the readout unit without noise pickup, cable resistance problems or amplifiers. Longer cables can be used with appropriate amplification.

Chemists are using the new Hewlett-Packard Quartz Thermometer in calorimetry, measuring heat changes caused by chemical reaction, and in measuring molecular weights of unknowns by dissolving them in a known benzine compound and noting the change in its freezing point. For the petroleum industry, the device is useful for temperature measurements far below the surface of the earth, as well as for monitoring various processes through long cable systems. A special version for oceanographic studies will measure temperature at depths as great as 35,000 feet.

By combining research into crystal technology with its experience and know-how in electronic measuring techniques, Hewlett-Packard has introduced new and exciting capabilities in thermometry.

The hp Quartz Thermometer measures temperature and temperature difference directly, providing digital readout of degrees C (−40 to +230) or F (−40 to +450). Resolution is as high as 0.0001°C. With standard equipment, measurements may be printed on paper tape, recorded in computer-compatible format on punched cards, punched paper tape or magnetic tape, or produced in strip-chart form. Probes can be exposed to pressures of up to 3000 psi, 10,000 G's shock, 1000 G's vibration to 1000 cps, without change in calibration. When necessary, calibration is simple; only one temperature point need be checked. Two models are available, DY-2801A with 2 sensor probes for 2-point monitoring or difference readings to 0.0001°, $3250; DY-2800A with one probe, 0.1° resolution (optionally 0.01°), $2250. For more information write Hewlett-Packard/Dymec Division, Palo Alto, California 94306, Tel. (415) 326-1755; Europe: Hewlett-Packard S. A., 54 Route des Acacias, Geneva; Canada: Hewlett-Packard (Canada) Ltd., 8270 Mayrand Street, Montreal.

Data subject to change without notice. Prices f. o. b. factory.

HEWLETT hp PACKARD

An extra measure of quality

50 AND 100 YEARS AGO

SCIENTIFIC AMERICAN

JULY, 1915: "The following description, given by a British officer and published in the daily papers, shows the effects of the German poison gases upon the men who survived them and were carried into the hospital: 'When we got to the hospital, we had no difficulty finding out in which ward the men were, as the noise of the poor devils trying to get breath was sufficient to direct us. We were met by a doctor belonging to our division, who took us into the ward. There were about 20 of the worst cases in the ward on mattresses, all more or less in a sitting position propped up against the walls. Their faces, arms and hands were of a shiny gray-black color, with mouths open and lead-glazed eyes, all swaying backward and forward trying to get their breath."

"The announcement is authoritatively made that Henry Ford has practically perfected an automobile tractor, upon which he has been working for a long time, and that within a few years he expects to develop an immense new plant at Dearborn, Mich., where these machines will be built. Nothing in the way of details nor of the character of this tractor has yet been given out; but the inventor says that it will be of world-wide importance in reducing the labor of tilling the soil and doing the heavy work on the farm, with a corresponding effect in reducing the cost of food."

"It is reasonable to assume that the number of stars in space having a temperature so low that their radiations do not affect our eyes or photographic plates is extremely large. That these invisible stars are far more numerous than the luminous stars is suggested by Mr. F. A. Lindemann, who attempts in the *Monthly Notices of the Royal Astronomical Society* a rough calculation of their relative number, based on the assumption that new stars (novæ) are due to collisions. He concludes that there are about 4,000 times as many dark stars as bright ones."

"Out of Work: A Study of Employment by Frances A. Kellor thoroughly acquaints the reader with the present machinery of employment, both private and governmental, and discloses its imperfections and its inadequacy. It exposes also the fallacies of complacency, which assume that a man can always get work of some kind; that snow shoveling and farm labor provide means of existence for the willing worker; and that unemployment is, in short, a moral rather than an industrial problem. There is no deliberate emotional appeal, but the bare statement of undeniable facts is sufficient to touch the heart and quicken the intelligence toward organized effort in the direction of bettering conditions. Much of the existing evil might be obviated by the adoption of methods indicated by the author, whose work delves into the primary causes of industrial distress and efficiently grapples with conditions the evil of which can scarcely be exaggerated, and which the more favored individual but vaguely appreciates. The volume comprises a practical study, not an academic discussion, and bears internal evidence of first-hand knowledge, organized reasoning and logical conclusions."

Scientific American.

JULY, 1865: "The Pacific Railroad is now one of the most urgent needs of the nation. Capital has been to a great extent thrown out of its recent channels and has become stagnated by the sudden cessation of the war. Its most natural and useful employment would be found in opening and developing the vast mineral resources of the mountain regions of the West. So much would such a result redound to the general good that it is to be hoped that Congress will lend all possible aid to the work. One hundred and sixteen millions of dollars, the estimated cost of the central route from Council Bluffs via South Pass to Benicia (a distance of 2,032 miles), would soon be repaid by the road to the nation in more ways than one. Many considerations, involving both the national defense and the general welfare, seem to make the completion of at least one through trunk line a duty as well as a privilege of the general Government. There is but little doubt that the people at large clearly recognize the common interest in the matter and are unanimously minded to

Report from

BELL LABORATORIES

Strip of postformable aluminum-polyethylene laminate developed by Bell Laboratories (left) shows best stiffness-to-weight ratio compared with equal-weight, -length and -width strips of glass-mat-reinforced polyester (center) and cold-rolled steel. Samples of the laminate after deep-drawing are also shown. Note the uniform thickness possible with this process.

Aluminum-polyethylene structural laminates formed without adhesive

A laminate of polyethylene sandwiched between lightweight metals such as aluminum would have highly desirable strength-to-weight properties. But conventional methods of making such a laminate by using intermediate adhesive agents are incapable of producing a structural bond.

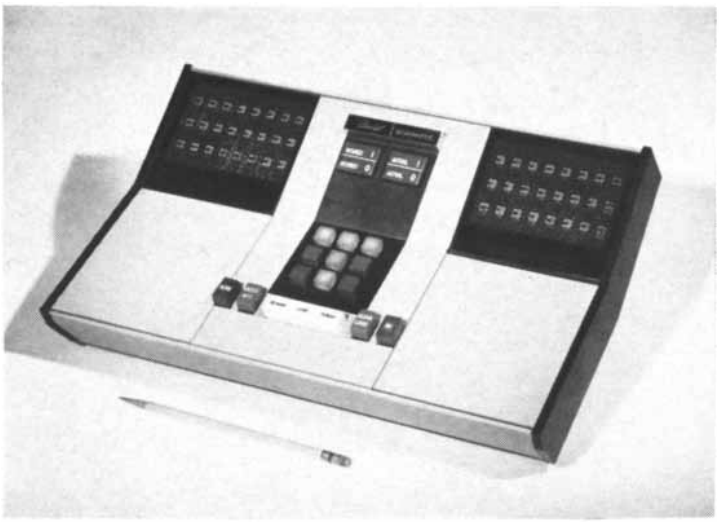
A. T. Spencer (left) and K. H. Pohl, originators of the structural laminate, demonstrate the results of postforming it with a heated die.

Recently, however, K. H. Pohl and A. T. Spencer of Bell Laboratories developed a structural-quality aluminum-polyethylene laminate *without* using adhesives. Such a laminate was concurrently predicted by Bell Laboratories' work on adhesion, which indicated that thermoplastics generally considered "unbondable" without surface modification can form strong joints with materials on which they spread spontaneously.

To produce the aluminum-polyethylene laminate, sheet polyethylene is molded between properly prepared sheets of aluminum. No preparation of the polyethylene is required, other than cleaning. The aluminum, however, is roughened, degreased and etched to produce a suitable surface. In a heated press the polyethylene melts, spreads spontaneously on the aluminum and, on cooling, produces a strong, durable bond.

In addition to its high flexural strength, the laminate has other characteristics which make it desirable for use in electronic equipment and elsewhere. It can be bonded, riveted, bolted, punched and sheared, and unlike reinforced plastics, it can be welded. It can also be bent readily or otherwise postformed at temperatures which soften the polyethylene core and thus allow independent motion of the aluminum skins. A Bell Laboratories-developed process for deep-drawing the laminate, in which the edges remain cool and only the portion to be deformed is heated, produced the samples seen in the illustrations.

Bell Telephone Laboratories
Research and Development Unit of the Bell System



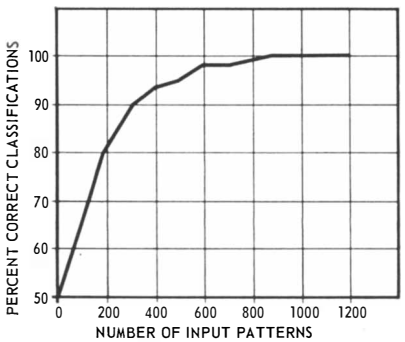
Trainable networks for pattern recognition

Bendix Research Laboratories has recently developed a new type of logic network which can be trained to process a wide variety of complex data patterns. This approach eliminates the need for designing and building special equipment for each new type of pattern recognition problem. Computer simulation of the trainable network shows that it can successfully learn complex classification problems. To demonstrate hardware feasibility, an experimental model has been built with readily available integrated circuits.

The trainable network concept is applicable to both spatially oriented and non-spatial patterns. Examples of the former are recognition of alphanumeric characters and processing of aerial photographs; examples of the latter are speech recognition from audio waveforms and analysis of sonar return data.

In the current approach, a logic network of general configuration is first designed. The input-output relationship of this network is initially random. To solve a given type of problem, the network is gradually "trained" to produce a desired input-output relationship by applying a "reward" signal when it produces the correct output for a given input pattern, and a "punish" signal when it produces an incorrect output. These reward and punish signals modify internal parameters of the network so that it eventually converges to an optimum state for the problem at hand. Re-training of the network for a different problem is accomplished by the same procedure.

A unique type of logic called nonlinear threshold logic has been developed for implementing the trainable network. In



Typical learning curve

conventional linear threshold logic, each input variable is multiplied by a stored adjustable analog value, or "weight," and the weighted inputs are summed to generate a single analog quantity known as a weighted sum. Comparison of the weighted sum with a stored threshold value determines the output decision (a logic ONE or ZERO). Such a circuit can be trained by incremental adjustment of the weights and threshold values, but the logic capacity is somewhat limited. The nonlinear threshold logic entails the formation of certain cross-products of the input variables prior to the weighting and summing operation. Training of the circuit includes the selection of cross-product terms in addition to the incremental adjustment of weights and thresholds. Inclusion of these terms results in significant advantages in logic capacity and flexibility.

The photograph shows an experimental pattern recognizer which is capable of being trained to recognize binary input patterns arranged in a 3 by 3 matrix. The model shown is constructed entirely of silicon integrated circuits.

A typical learning curve obtained from computer simulation of a trainable network is shown in the illustration. Input data used in this process was provided as binary patterns arranged in a 3 by 3 matrix to form various distorted shapes of the letters X, C, T, and J. The problem was to separate the X and C group of patterns from the J and T group of patterns. Extension of this concept to more complex and higher resolution patterns has been shown to be feasible, and is undergoing continued study.

Bendix Research embraces a wide range of technology including space and extreme environment technology, solid-state and thin-film research, mass spectrometry, photoelectronics and electro-optical systems, electron beam and tube technology, measurement science, applied mechanics, energy conversion systems, dynamic controls, fluid-state technology, systems analysis and computation, navigation and guidance, microwave technology and communications, photogrammetric instruments and techniques, data processing and control systems. Motivation: to develop new techniques and hardware for The Bendix Corporation to produce new and better products and complete, integrated, advanced systems for aerospace, defense, industrial, aviation, and automotive applications. Inquiries are invited. We also invite engineers and scientists to discuss career position opportunities with us. An equal opportunity employer. Write Director, Bendix Research Laboratories Division, Southfield, Michigan.

Research Laboratories Division

THE Bendix CORPORATION

WHERE IDEAS UNLOCK THE FUTURE

have the Pacific Railroad built in the quickest possible way. They perceive the propriety of a common contribution of expense where the benefit is common, without quibbling about its exact distribution, which is impossible. Especially so in a work like this, so vast that private enterprise cannot grapple with it and would accomplish it, if at all, only by piecemeal, nibbling its way along as it felt itself supported or impelled by the oncoming tide of population. It is doubtful whether indeed private enterprise would ever get through with it. The vast tracts of unarable land on these routes seem to make the work impossible except as a public one."

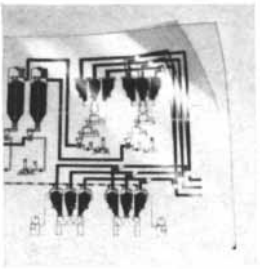
"Harold, Payne, Atzerodt and Mrs. Suratt, having been found guilty of the crime of assassinating President Lincoln, were executed on Friday, the 7th inst., in the city of Washington. Dr. Mudd, Arnold and O'Laughlin have been sentenced to imprisonment and hard labor for life, and Spangler to six years' imprisonment and hard labor in the Albany penitentiary."

"A flying machine of novel form is now in process of construction at Hoboken for the United States Government. It was commenced during the war and was intended for use in aerial reconnaissance of the enemy's position. The war is over, but the machine is going on till its success or failure is an established fact. The idea of the invention is an old one, but this is the first time that an attempt has been made to put it into practice. The Government was induced to embark in the enterprise upon the strength of certain experiments made by the late distinguished General (and Professor) Mitchell. He had long been interested in the subject of aerial navigation and believed that the principle of screw propulsion could be made to work in air as well as in water. The Government toy—as some persons will probably call it—is a cigar-shaped canoe built of copper with iron ribs. An engine is placed in the center with sufficient power to work a horizontal screw fan with 20-foot blades. There are four fans connected to the engine—one below, one above the canoe and one at each end. The upper and lower fans are worked together to produce an ascent; and the terminal fans are made to revolve together or separately in the same direction or in opposite directions for the purpose of propelling the craft horizontally. The weight of the whole, fully equipped and manned, is about six tuns."

Temperature flags! Nothing is left to chance with new Celanar Polyester Film

There are no unwanted "surprises" with new Celanar polyester film by Celanese. Because we leave nothing to chance. Such as the possibility of undetected harmful environmental changes suffered in transit. Celanar film is shipped with *temperature recording flags* to alert you to possible damage before you put it on the processing line.

But that's just one of the meaningful service advantages that are causing so many manufacturers to switch to new Celanar film. For magnetic tape. Packaging. Engineering reproduction. Metalizing. Stationery and office supplies. And electrical applications.



The *basic* reason why Celanar film has no peer is, of course, *cleanliness*. New Celanar film is produced in a sealed-off "White Room" clean enough for surgery—the most modern in the industry. And the cleaner the polyester film, the better it processes.

It's nice, though, to have such other unique advantages accrue to you as *protection against dust contamination* by use of non-fibrous plastic cores. And that, for critical applications, Celanar film may be *shipped with Impact Recorders* to protect you against accepting film jolted and damaged during shipment. Or that its splice-free roll lengths are tailored to your specifications.

This is the kind of meaningful service you would expect from Celanese Plastics—whose operating philosophy is that the customer, not the supplier, is always right. Celanese Plastics Company, 744 Broad Street, Newark 2, N. J.

Celanese® Celanar®

Celanar® Polyester Film

Temperature recording flags, to warn you against possible damage in shipment, is just one of *six* meaningful service advantages you get when you switch to new Celanar polyester film.

Celanese

THE AUTHORS

DAEL WOLFLE ("The Support of Science in the U.S.") is executive officer of the American Association for the Advancement of Science and publisher of the association's magazine, *Science*. He was graduated from the University of Washington in 1927 and received a doctor's degree from Ohio State University in 1931. After a number of years of teaching psychology at Ohio State, the University of Mississippi and the University of Chicago and as a Federal official during World War II, he was appointed executive secretary of the American Psychological Association in 1946. From 1950 to 1954 he was director of the Commission on Human Resources and Advanced Training, an organization formed by several learned societies. Wolfle took his present post in 1954. He has served as a consultant to most of the major grant-making agencies of the Federal Government.

BRIAN ROBINSON ("Hydroxyl Radicals in Space") is senior research scientist at the Radiophysics Laboratory of the Commonwealth Scientific and Industrial Research Organisation in Sydney. Born in Melbourne, he was graduated from the University of Sydney in 1952 and obtained a master's degree there a year later. After working briefly with the radio astronomy group at the C.S.I.R.O. he won a Royal Society scholarship to the University of Cambridge, where he received a Ph.D. in 1958. From 1958 to 1961 he was with the Netherlands Foundation for Radio Astronomy, leaving it to take up his present post. He writes that his research interests are "the study of neutral hydrogen in external galaxies and in intergalactic space; development of high-sensitivity receivers for radio astronomy and detection of new spectral lines at radio wavelengths."

NOEL DE NEVERS ("The Secondary Recovery of Petroleum") is assistant professor of chemical engineering at the University of Utah. After being graduated from Stanford University he spent a year as a Fulbright fellow at the Technische Hochschule in Karlsruhe, Germany, and then obtained a Ph.D. in chemical engineering at the University of Michigan. Thereafter he joined the California Research Corporation, a subsidiary of the Standard Oil Company of California, where he did work on the processes he describes in his article. He is continuing that work at the University of Utah.

RICHARD J. WURTMAN and JULIUS AXELROD ("The Pineal Gland") are respectively a fellow in endocrinology at the Massachusetts General Hospital and chief of the section on pharmacology in the Laboratory of Clinical Science at the National Institute of Mental Health. Wurtman majored in philosophy as an undergraduate at the University of Pennsylvania and was graduated in 1960 from the Harvard Medical School. There, he writes, he "became interested in the new field of neuroendocrinology, especially in how environmental factors such as light and stress influence glandular function and how circulating hormones affect nervous function and behavior." Axelrod was graduated from the City College of the City of New York in 1933. He received a Ph.D. at George Washington University 22 years later, having in the interim spent 13 years in a commercial laboratory before joining the National Heart Institute in 1950 to work on "the fate of drugs and the enzymes involved in their metabolism." At the National Institute of Mental Health, which he joined in 1955, he has developed techniques for studying how hormones regulate the action of neurotransmitters.

FRANCIS BITTER ("Ultrastrong Magnetic Fields") is professor of geophysics at the Massachusetts Institute of Technology, with which he has been associated since 1934. He founded there the first magnet laboratory, a precursor of the National Magnet Laboratory now operated by M.I.T. under the sponsorship of the Air Force Office of Scientific Research. Bitter spent three undergraduate years at the University of Chicago and one at Columbia University, obtaining a Ph.D. at the latter institution in 1928. He was with the Westinghouse Research Laboratories for four years before going to M.I.T. As a young man he made the acquaintance of such physicists as Albert Einstein, Max Planck and P. A. M. Dirac—an experience he recalls as "tremendous." Of his present situation he writes: "I think it is a great privilege to be associated with a good university nowadays. The combination of teaching, research and public service is very stimulating."

EUGENE I. RABINOWITCH and GOVINDJEE ("The Role of Chlorophyll in Photosynthesis") are respective-

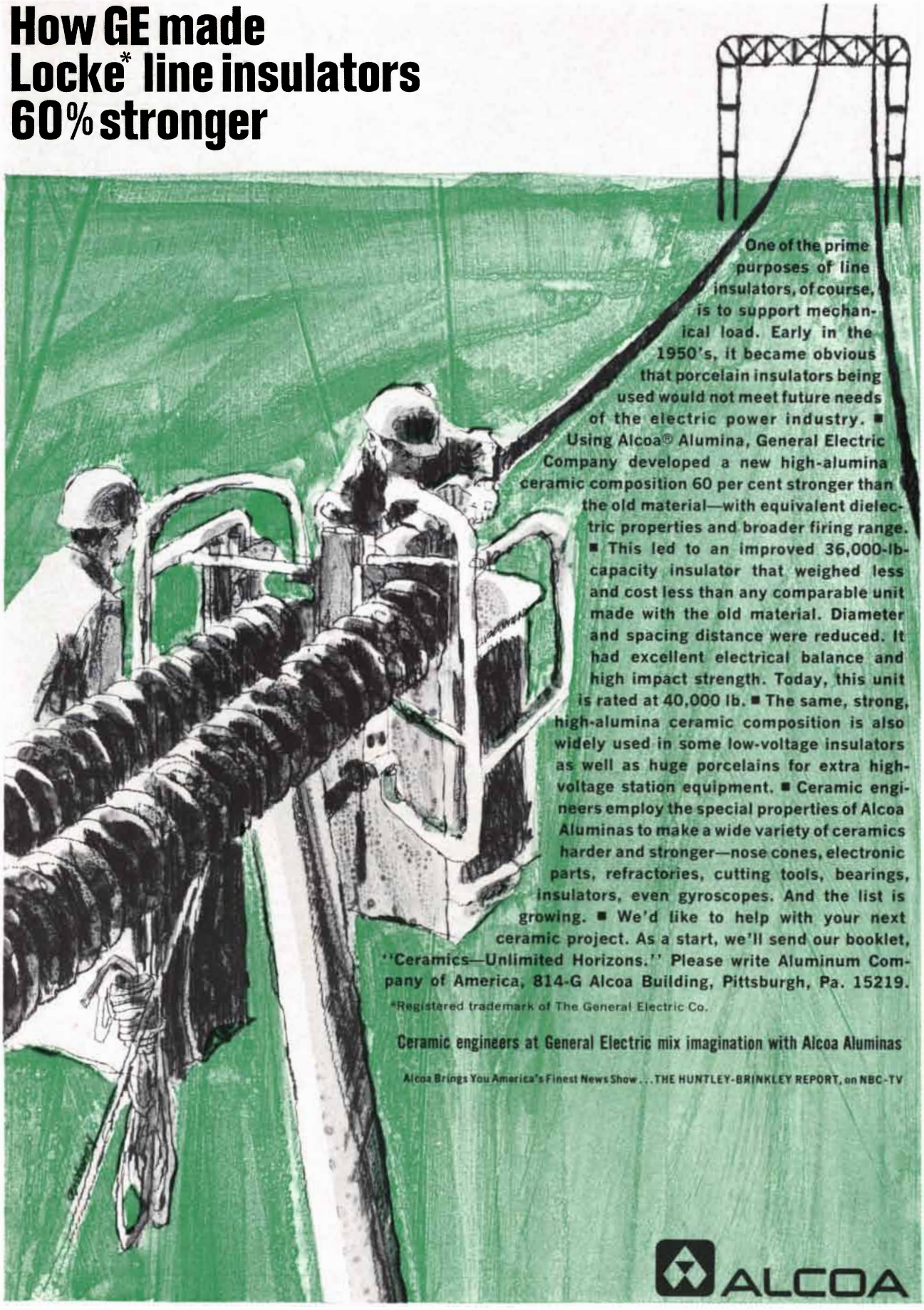
ly professor of botany and biophysics and assistant professor of botany at the University of Illinois. Rabinowitch was born in Russia, received a Ph.D. in inorganic chemistry at the University of Berlin in 1926 and worked in Germany, Denmark and Britain before coming to the U.S. in 1938. During World War II he worked on the Manhattan project. He joined the staff of the University of Illinois in 1947. Rabinowitch is the author of many books and papers, editor of the *Bulletin of the Atomic Scientists* and a translator of Russian poetry. In addition a book of poems that he has written in Russian is being published in Paris. Govindjee, who has no other name, is a graduate of the University of Allahabad in India; he obtained a Ph.D. in biophysics at the University of Illinois in 1960. His father dropped the family name, Asthana, in an effort to wipe out caste distinctions; one can often judge the caste or subcaste of Indians by their family names. Rabinowitch says he has "often suggested to Govindjee that he should invent a first name to make life easier for abstracters and indexers, but he seems to enjoy the distinction."

KATHLEEN M. KENYON ("Ancient Jerusalem") is principal of St. Hugh's College of the University of Oxford and director of the British School of Archaeology in Jerusalem. She has participated in many archaeological expeditions in Great Britain, the Middle East and Africa since 1929 and has written about several of them.

BERNARD GREENBERG ("Flies and Disease") is associate professor of zoology at the College of Pharmacy of the University of Illinois. He was graduated from Brooklyn College in 1944 and obtained master's and doctor's degrees in entomology and microbiology at the University of Kansas. His association with the University of Illinois dates from 1954; he has interspersed it with periods as a visiting scientist at the Higher Institute of Health in Rome and the Institute of Health and Tropical Diseases in Mexico City. In his travels he indulges an interest in antiquities, notably the Etruscan and Roman cultures in Rome and the pre-Columbian culture in Mexico.

EDWIN G. BORING, who in this issue reviews *Mental Retardation: A Review of Research,* edited by Harvey A. Stevens and Rick Heber, is Edgar Pierce Professor of Psychology emeritus at Harvard University.

How GE made Locke* line insulators 60% stronger



One of the prime purposes of line insulators, of course, is to support mechanical load. Early in the 1950's, it became obvious that porcelain insulators being used would not meet future needs of the electric power industry. ■ Using Alcoa® Alumina, General Electric Company developed a new high-alumina ceramic composition 60 per cent stronger than the old material—with equivalent dielectric properties and broader firing range. ■ This led to an improved 36,000-lb-capacity insulator that weighed less and cost less than any comparable unit made with the old material. Diameter and spacing distance were reduced. It had excellent electrical balance and high impact strength. Today, this unit is rated at 40,000 lb. ■ The same, strong, high-alumina ceramic composition is also widely used in some low-voltage insulators as well as huge porcelains for extra high-voltage station equipment. ■ Ceramic engineers employ the special properties of Alcoa Aluminas to make a wide variety of ceramics harder and stronger—nose cones, electronic parts, refractories, cutting tools, bearings, insulators, even gyroscopes. And the list is growing. ■ We'd like to help with your next ceramic project. As a start, we'll send our booklet, "Ceramics—Unlimited Horizons." Please write Aluminum Company of America, 814-G Alcoa Building, Pittsburgh, Pa. 15219.

*Registered trademark of The General Electric Co.

Ceramic engineers at General Electric mix imagination with Alcoa Aluminas

Alcoa Brings You America's Finest News Show . . . THE HUNTLEY-BRINKLEY REPORT, on NBC-TV

ALCOA

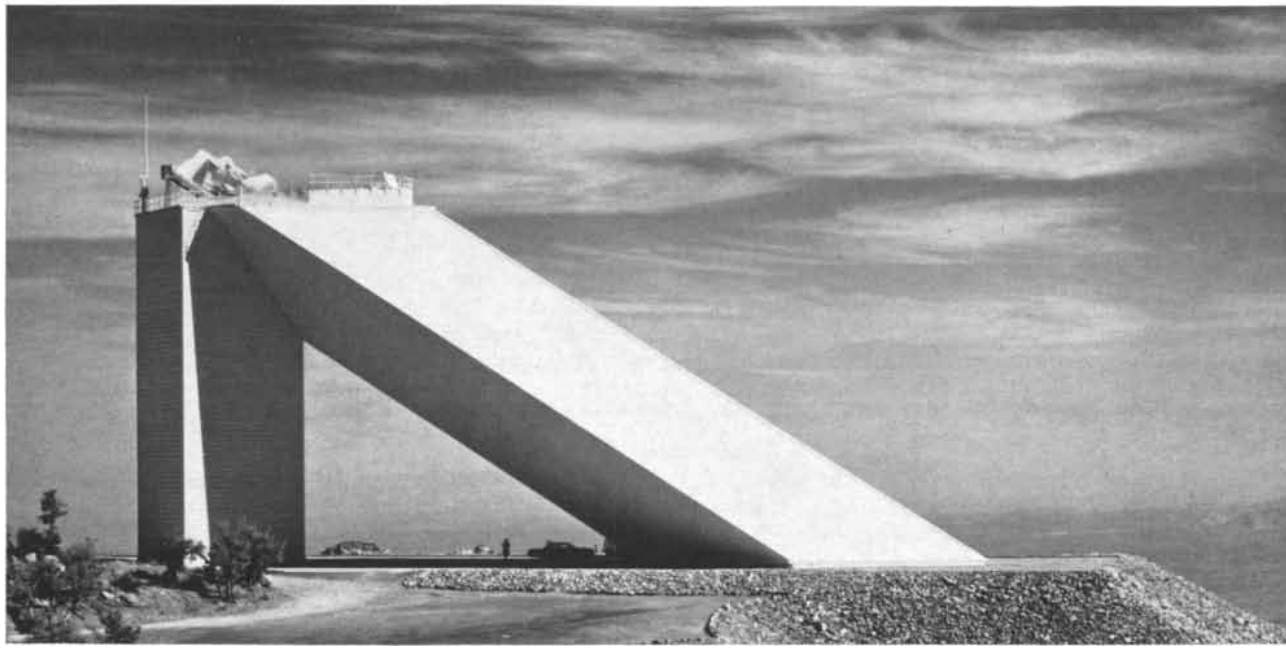
Established 1845 **SCIENTIFIC AMERICAN** July 1965 Volume 213 Number 1

The Support of Science in the U.S.

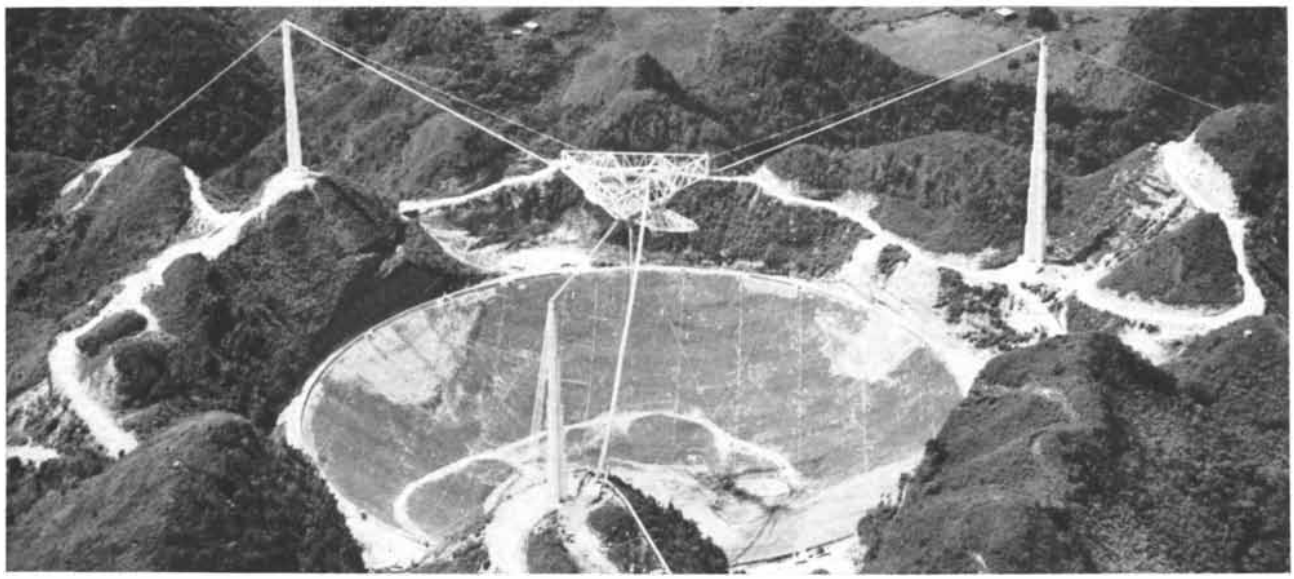
The sharp and sustained increase in funds has improved research and has benefited the investigator. Nevertheless, serious questions are being raised about the financing of research in universities

by Dael Wolfle

This year in the U.S. nearly $21 billion—3.2 percent of the gross national product—will be spent for research and development. Some two-thirds of the funds will be supplied by the Federal Government. "Research and development" includes basic research, applied research and engineering, design and even the development of prototypes; it is a broad category, but it does encompass all forms of scientific research. Not long ago the support of science was primarily the business of the colleges and universities and some voluntary agencies; before World War II the Federal Government's contribution was largely in agricultural research and the work of such agencies as the U.S. Geological Survey and the Naval Observatory. It was not until 1942 that the country's expenditures on science reached $1 billion. A steady growth in the support of science continued through the war and afterward; beginning in 1953 there was a sharp and sustained rise of huge proportions [*see top illustration on page 22*]. Since 1953 the country has increased its expenditures for science at an average rate of 13 percent a year. The most striking rise has



SOLAR TELESCOPE, the world's largest, dominates the Kitt Peak National Observatory in Arizona, which is an example of "big science." The observatory southwest of Tucson was built and is maintained by the National Science Foundation, a Federal agency. It is operated, under contract to the N.S.F., by the 10-university Association of Universities for Research in Astronomy, Inc.



IONOSPHERIC OBSERVATORY near Arecibo, Puerto Rico, is a combined radar and radio telescope. Its reflector, the largest in the world, is a 1,000-foot bowl in the mountains, lined with steel mesh; the radar "feed" is suspended about 500 feet above the bowl. The observatory, designed by Cornell University workers, was built with $9 million from the Advanced Research Projects Agency of the Department of Defense and is operated by Cornell under contract to the Air Force Office of Scientific Research.

been in the contribution of the Federal Government, which has grown at a rate of nearly 20 percent a year. Although spending for development is leveling off, appropriations for academic research will continue to increase at about the present rate for some years.

The funds spent for scientific work during the past two decades have provided research opportunities on a scale previously unimagined. All fields of science have benefited from the better equipment, special facilities, greater freedom from constraints and larger number of workers made possible by the increased budgets. The award of Nobel prizes is one measure of the growing strength of basic research in this country; in the 1930's Nobel prizes were awarded to nine American scientists, in the 1940's to 13 of them and in the 1950's to 27. Meanwhile the economy of the country has gained enormously from the upsurge in technological research and development. In 1953 research and development accounted for 11 percent of all industrial investment; in 1962 research and development absorbed about 25 percent.

The subject is nonetheless surrounded by disquiet. In Congress and in the Executive branch, in the universities and learned societies and foundations questions are being raised about the manner in which science is financed. Most of the questions deal not with the adequacy of the national effort but with the effects of the massive Federal contribution on the course of science and in particular on the conduct of basic research in the universities.

Evidence of this concern is found in a rapidly growing list of policy studies and program analyses. The National Academy of Sciences is midway in a series of reports dealing with various aspects of the scientific enterprise. The U.S. Chamber of Commerce has advocated the establishment of a national review body that would decide on major new programs. Two committees of Congress—the House Select Committee on Government Research and the House Subcommittee on Science, Research and Development—have reviewed many aspects of the Federal program, and their reports have become valuable sources of detailed information. Moreover, Congress has begun to insist that executive agencies prepare special reports on certain areas of investigation such as oceanography so that the Federal effort can be examined as a whole instead of in its budgetary and departmental fragments. The White House Office of Science and Technology has appointed a blue-ribbon committee of industrial, scientific and educational leaders to review the policies and programs of the National Institutes of Health. The Bureau of the Budget has taken the lead in reexamining the administrative practices of the Federal agencies that support basic research. The National Science Foundation has reorganized and strengthened its staff sections responsible for studies of scientific policy, planning and resources. "Science policy" has become the topic of a number of university seminars and analyses.

All this ferment of analysis and reexamination makes it clear that major changes in policies governing the support of science are under way or in the offing. These analyses have also served to provide reassurance that many of the past policies and practices are sound and should be continued. The magnificent achievements of recent decades are evidence that the support system has been a fundamentally healthy one.

Support for research and development comes from many sources; some contribute only a few dollars, others billions. Some 300 firms provide 80 percent of the industrial money that goes into research and development; another 13,000 firms provide the remainder. Some 200 private foundations grant significant amounts to science and medicine. Universities and many colleges provide research talent, laboratories and financial help. A number of private research institutions finance their own investigations. State and local governments conduct a variety of research programs. Four agencies are responsible for 95 percent of the Federal funds: the Department of Defense, the National Aeronautics and Space Administration, the Atomic Energy Commission and the Department of Health, Education, and Welfare. In addition to these giants there are another four agencies that account for 4 percent of the Federal total: the Department of Agriculture, the Na-

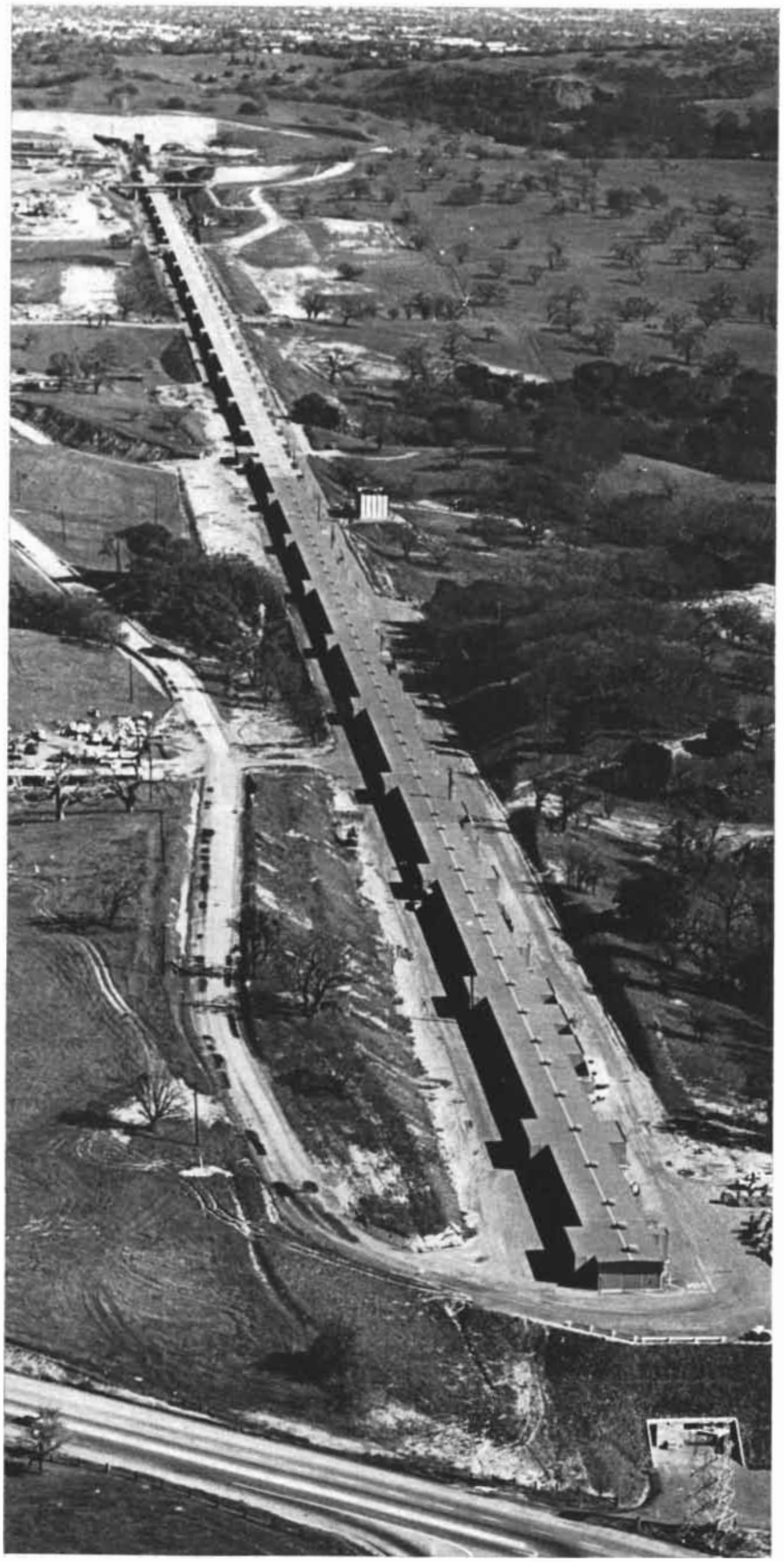
tional Science Foundation, the Department of the Interior and the Federal Aviation Agency. The remaining 1 percent of Federal research and development funds is spent by 21 other agencies.

In the 12 years from 1953 to 1965 every major source of research and development funds increased its support substantially. Federal funds are five times what they were in 1953. Industrial support has tripled, and the universities have done almost as well. The other nonprofit institutions are contributing six times their 1953 amount.

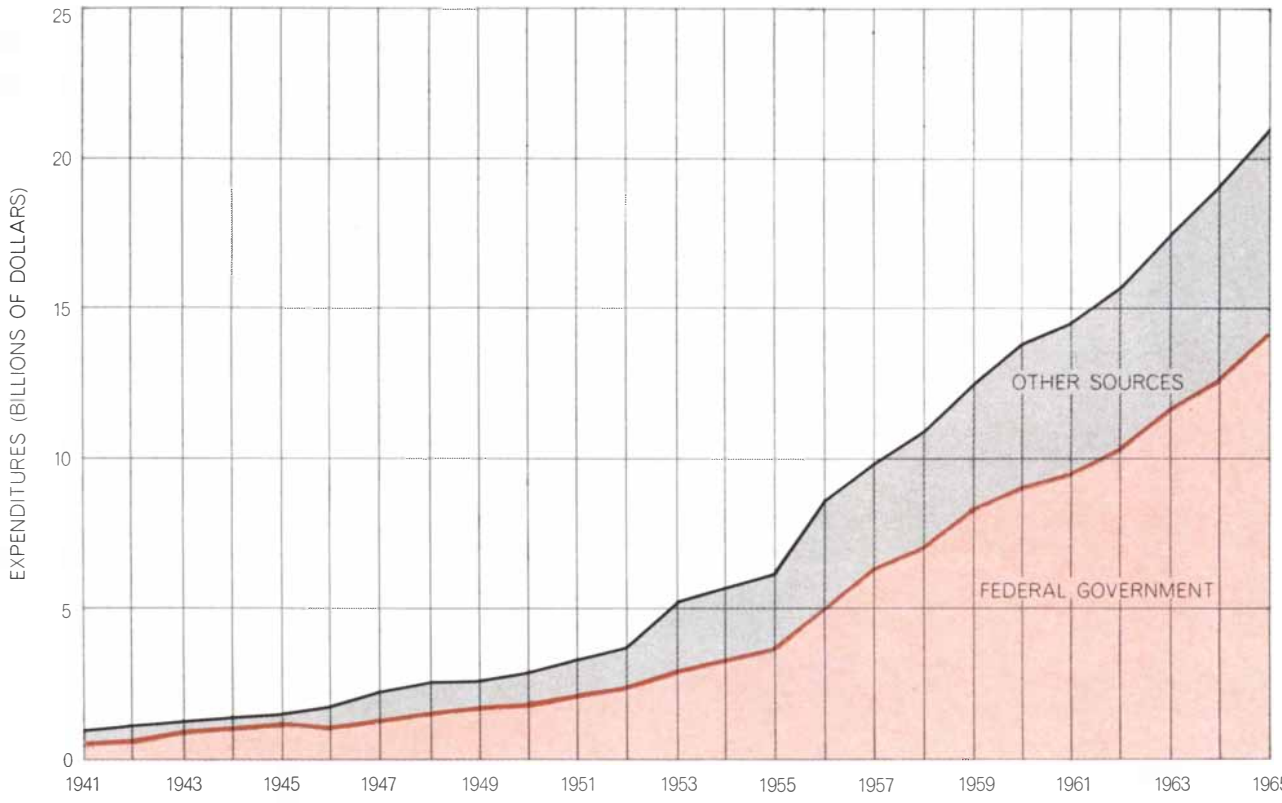
Just as the amounts of money supplied by these four sectors vary greatly, so do the amounts they use [*see illustration on page 23*]. The Federal Government supplies two-thirds of the funds, but Federal laboratories carry out less than 15 percent of the work. Industry contributes a third of the funds but conducts three-fourths of all the work (mostly with Federal funds). The colleges and universities provide about a tenth of the funds, and the other nonprofit institutions about a fortieth. (The universities' contribution is underrepresented in the financial reports, perhaps by several hundred million dollars a year; they provide substantial additional support, in the form of laboratory facilities and faculty time, that is not budgeted explicitly for research.)

From 1953 until 1960 about 8 percent of the nation's research and development budget was devoted to basic research. The percentage has been rising since 1960, reaching almost 12 percent in 1965. As for the Federal Government's funds, in 1953 less than 7 percent went for basic research. The figure has been rising since 1960, to about 11 percent in 1965. The universities are relatively much more prominent in basic research than in the total research and development effort, being responsible for almost half of all basic research. In contrast the industrial laboratories, which dominate in development activity, conduct only about a fourth of the basic research.

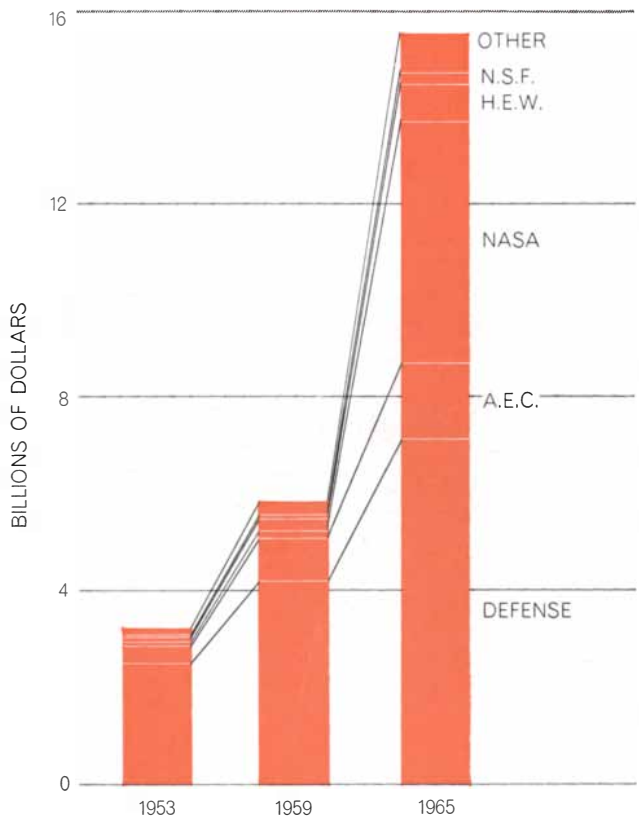
Development activity is directly associated with indentifiable industrial, economic, military or other practical objectives. Its cost and the cost of any associated research are therefore justified and budgeted in terms of its expected contribution to the attainment of specific objectives. In the case of basic research the situation is quite different. The ultimate beneficiaries of basic research are many, but they are hard to identify in advance. As a result the costs of basic research tend to be

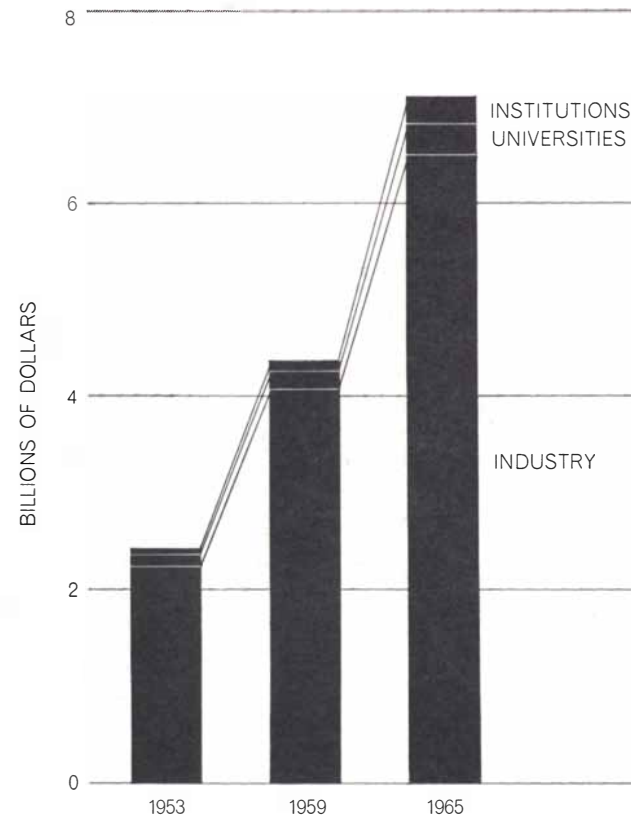


TWO-MILE ELECTRON ACCELERATOR is being built for $114 million by Stanford University under a contract with the Atomic Energy Commission. It required specific approval by Congress. Electrons will be accelerated through a two-mile tunnel by an electric field generated by klystron tubes housed in the long structure visible in the photograph.



U.S. SPENDING FOR SCIENCE has soared since 1953, as shown by this chart of total expenditures for research and development by the Federal Government and other sources. 1953–1963 figures are from the National Science Foundation; others are estimates.





SOURCES of Federal (*left*) research and development funds include National Science Foundation, Department of Health, Education and Welfare, National Aeronautics and Space Administration, Atomic Energy Commission and Department of Defense. Figures are for fiscal years. Among the non-Federal sources (*right*) "Institutions" include nonprofit institutes, foundations and so on.

shared widely. Some basic research of notable quality is done in industrial laboratories, but most of it is conducted in universities with support from public funds. In some cases this public support involves Congress directly in decisions on priorities. Modern basic research sometimes calls for large-scale facilities such as particle accelerators, oceanographic research vessels and astronomical observatories. Such "big science" enterprises are so expensive that they must be considered individually at top Government levels, where the cost and promise of each can be compared with those of other claimants for available funds.

On the other hand, "little science," typically the work of a university faculty member and his assistants and advanced students, will continue to be budgeted on an *a priori* basis and to be supported by means of a large number of "project" grants. Little science, the principal subject of the remainder of this article, is an area of central concern to science as a whole, not least because it involves the education of future scientists. It is the kind of science that is most characteristic of academic research and hence is most often involved in Government-university relations. It is also the area in which those relations are most likely to change.

Sustained scientific work of high quality requires the effective union of three elements: a self-renewing population of able scientists; appropriate research facilities with the necessary supporting structure for institutional management; a source of money. In a few well-endowed research institutions all three elements are happily present in an almost totally self-contained and self-supporting organization. Such unity, however, is rare. More commonly under present conditions there is a scientific staff, a university with multiple obligations, and an external source of funds. All three sides of this triangle are interested in science, but their interests differ in detail; tensions arise and compromises become essential. The scientist must serve three masters: the internal logic and the opportunities of his own discipline, the policies and requirements of his institution, and the customs and wishes of his financial supporter. The university must meet the demands of science, of its many other endeavors and of the agencies that provide support. The Government agencies have an equally complex problem: in supporting a large number of individual scientific projects they must also consider the general welfare of the universities and be mindful of the wishes of Congress and the public it represents.

One useful change in the interrelations of scientists, universities and Federal agencies would be the simplification and standardization of what has grown to be a maze of rules and regulations governing fiscal and administrative details and reports. The complexity of grant administration was summarized last year by the House Select Committee on Government Research: "One of the ironies of the research grant is that while it is sometimes itself a simple one-page (if not a one-paragraph) document, it is accompanied by a bulky manual of instructions, explanations, and amendments. For example, although the N.I.H. [National Institutes of Health] grant form is a one-page instrument, it incorporates by reference the N.I.H. grant manual, which runs to more than 100 pages."

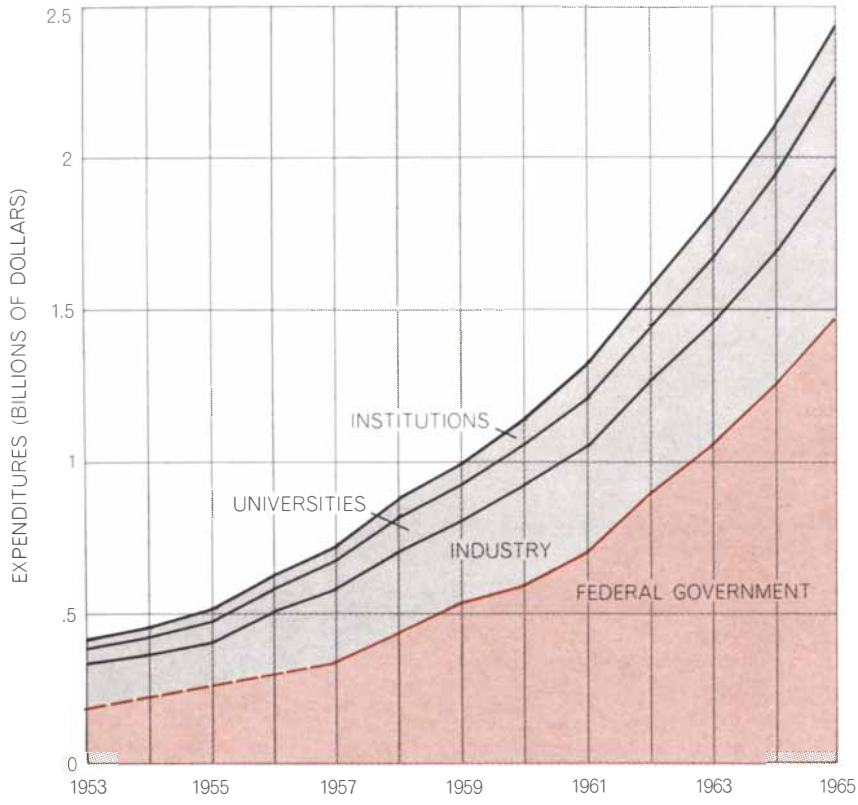
The National Institutes of Health manual of course explains only N.I.H. procedures and requirements; other agencies have adopted different rules and procedures. Congress has sometimes added to the confusion by setting arbitrary limits on the amounts that some agencies can pay to reimburse an institution for the indirect costs of conducting research. This "overhead" rate varies, moreover, depending on the agency that grants the funds. Sometimes overhead can be paid on some budgetary items but not on others, or at one rate on some items and at another rate on other items. The multiplication of administrative red tape slows decisions, harasses both agency and university personnel and puts the emphasis on form rather than substance. Fortunately these difficulties are widely recognized, and simplification and standardization would bring such obvious advantages that they will surely come about.

Standardization of procedures will be welcome, but more fundamental changes are required. Project grants are nominally made to a university or other institution, but in reality they are awarded to an individual. The scientist and Government official frequently deal directly with each other on both substantive and budgetary matters, largely excluding the university administration from any important role in reaching decisions about the research done in the university. Not all of the consequences have been happy ones.

When a faculty member looks outside his university for the major sources of support for his work, his interest and loyalty are likely to go where the dollars are. When the continuation of his work depends on his maintaining good relations and an effective record with private foundations and Washington agencies, and when his professional reputation depends primarily on his research productivity, he is likely to devote more and more of his time to writing project proposals and reports and to supervising the increased number of research assistants that liberal grants enable him to hire. Correspondingly less of his interest and loyalty go to

SUPPLIED BY \ USED BY	FEDERAL GOVERNMENT	INDUSTRY	COLLEGES AND UNIVERSITIES	OTHER INSTITUTIONS	TOTAL
FEDERAL GOVERNMENT	2,900	8,900	1,600	400	13,800
INDUSTRY		6,280	80	140	6,500
COLLEGES AND UNIVERSITIES			300		300
OTHER INSTITUTIONS			90	160	250
TOTAL	2,900	15,180	2,070	700	20,850

RESEARCH AND DEVELOPMENT FUNDS were supplied and expended in 1965 as shown in this matrix. Figures, in millions of dollars, are estimates based on National Science Foundation tables for prior years. Industry conducts most of the research and development.



FUNDS FOR BASIC RESEARCH from all sources have increased since 1953 (1964 and 1965 estimated). The Government supplies most of these funds, in largest part to universities.

the university that happens to be his home for the present, and less of his time is devoted to teaching and to doing actual laboratory work with his own hands.

There are many contentions that the increase in research has been bought at the expense of a depreciation of teaching. The research programs at most colleges and universities are not large enough to have an adverse effect on teaching. In the universities with large research budgets, however, complaints are heard that there is a schism between the teachers and the researchers; that the ablest graduate students are research assistants, whereas the less able ones become teaching assistants; that the big-time research operator has become the admired model in the eyes of graduate students; that in return for the explosive growth of research we are building up a deficit in the training of future scientists and in the general education of other students in science. There is a substantial body of opinion to the effect that whereas education at the graduate level has improved as a result of the availability of better equipment and larger and more competent staffs, undergraduate teaching has suffered.

The emphasis on research supported by outside funds on an individual-project basis has also tended to strengthen the divisive forces and weaken the integrative forces that are always at work on a university campus. By and large faculty scientists like the change to off-campus support; it means that each researcher is judged by colleagues in his own field of specialization. Physicists judge physicists, biochemists judge biochemists and geologists judge geologists. A man can take pride in the fact that specialists from other institutions have judged his work and found it worthy of support.

Bringing new funds to the campus enhances the scientist's prestige and gives him some freedom from local control. He can buy equipment or hire a secretary, travel to a national meeting to discuss work with other people in his field and even invite a man from another institution to pay him a visit—with expenses paid—to consult on research plans. And he can do all this without having to ask his dean or president for permission, because the grant is his. (That is, he can pay for these extras if he has had the foresight to provide for them in his project proposal. If not, it may take weeks for a busy office in Washington to let him know whether or not he can transfer $100 from one budget category to another!)

The result of all this is that the project-grant system undoubtedly weakens the scientist's ties with his own university. It means that many decisions about the research conducted on a campus are made in Washington instead of at the campus level and are made piecemeal rather than with full account taken of all the other programs and responsibilities of the university. A university is not solely a group of individualistic faculty members. It is a community of scholars and of students who wish to learn from them. It includes a central administration responsible for the development of the entire university, not simply the uncoordinated expansion of individual units or empires. Professor X would rather entrust his research proposal to the judgment of his professional colleagues on a Washington reviewing panel than to what he may consider the uninformed or biased decisions of his own dean and president. President Y, however, would prefer to have a larger measure of control at the university level, because he remembers that the university is responsible for teaching as well as research, for history and philosophy as well as physics and biochemistry, for the library as well as the observatory—and he wants funds that can be used in the best interests of the university.

Not only may the institutions in which research is carried out be changed by the methods of support; science itself may also be affected. One cannot help worrying about what subtle distortions in the course of scientific progress may result from the fact that nearly all of the Federal support now comes from "mission-oriented" agencies. The National Institutes of Health are interested in certain diseases, the Atomic Energy Commission in nuclear energy, the Department of Defense in weapons systems and countermeasures. Each supports basic research, but each selects projects in terms of its own mission. Of all the Federal grant-making agencies, only the National Science Foundation is free from this necessity. To be sure, many researchers have secured support from the mission-oriented agencies for exactly what they as scientists most wanted to do. The fact remains that, of all the money spent for basic research in the U.S., only about one dollar in five comes from a source that does not have specific missions in mind. It is still a matter of opinion whether or not this fact is threatening

the future health of basic science, but there is a wide-spread feeling that the National Science Foundation should assume a greatly increased share of the responsibility for supporting basic research.

Certainly agencies with special missions will continue to support basic research; funding decisions will often be controlled by immediate objectives; projects will continue to be supported largely on the basis of their individual merits and those of the scientists involved. Yet basic improvements in the system are possible. Now that massive Federal support is accepted as an obligation, the most necessary change is to shift a substantial amount of the decision-making responsibility closer to the point of research. The fact is that decisions that should be made by the executive agencies are now being made by Congress. Decisions that should be made by the universities are being made by the agencies.

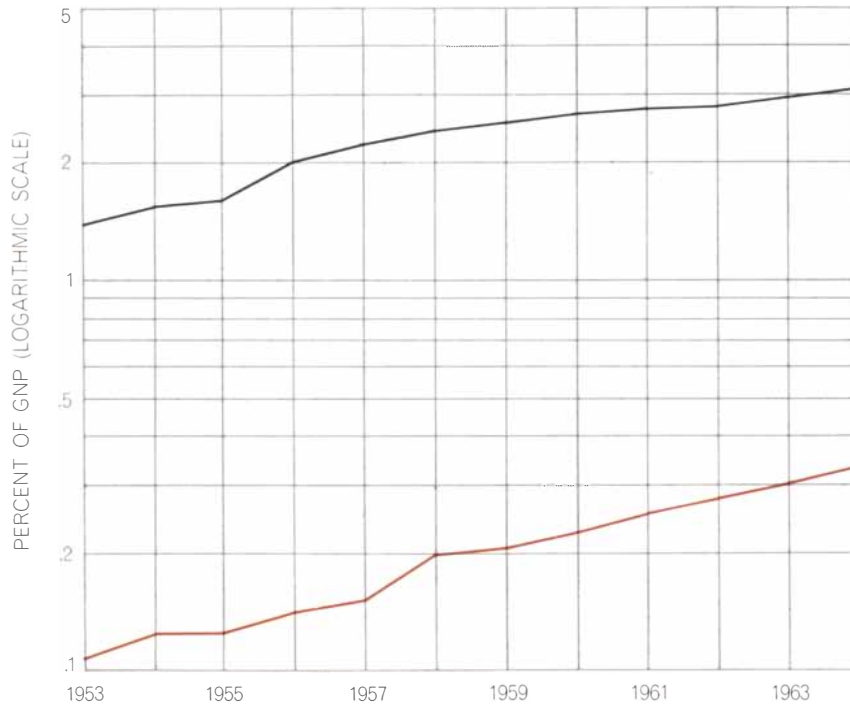
In Great Britain, Parliament avoids political and governmental control of science and education by making block grants to the University Grants Committee, which in turn allots funds to the British universities. For a number of reasons this mode of operation is not feasible in the U.S. Don K. Price of the Harvard School of Public Administration has pointed out that Congress takes a very different attitude toward the relation between ends and means than Parliament does. Parliament is content to decide on the ends, authorize the necessary funds and leave the details of the means to administrative agencies and the civil service. Congress, on the other hand, pays much attention to the means by which national objectives are to be attained. It reviews the budgets of Federal agencies in great detail, sometimes instructing an agency that no more than (and occasionally no less than) a stipulated amount is to be spent on a particular kind of activity. Congress is not likely to surrender its control of means as well as ends but it might well give the agencies a freer hand with the details and subcategories of their research budgets.

There will have to be, in turn, a substantial shifting of responsibility for research decisions from the supporting agencies to the universities. Some of the Federal agencies are now supplementing the project grant with newer forms of support that will help the universities to regain this responsibility: "general-research support" grants; "program" grants that support an established group of research colleagues not for a specific project but for work in an area in which they have demonstrated their competence; "institutional" grants that can be used in whatever way the university officials believe will best advance science on the campus; grants to help with the construction or equipping of laboratories, and the new "science development" grants to help selected institutions that are already quite good take a major step up the quality ladder.

These newer forms of grants will help to shift responsibility back to the campus, but the universities also have some work to do. A university's functions include both teaching and research; it has to maintain a reasonable balance between the two and also decide on the kind and amount of research that make sense in the context of its total program. The university president will sometimes say wearily that he knows these are his responsibilities but that his hands are tied—that there is no way to stop the very competent Professor X when he wants to start a new project because half a dozen other universities are eager to have the professor, willing to take him on his own terms and confident that plentiful grants will follow him to his new home. If the president lets his hands remain tied, the project grant will continue to be the dominant form of research support. It now seems likely, however, that universities that develop strong institutional controls and excel in the management of research funds can expect to receive a larger amount of support in more flexible forms.

Both the Government and the universities need to reconsider their interrelation. The makers of science policy must recognize that the nation is as dependent on the universities as the latter have come to be on the Government. The universities are institutions with major responsibility for the nation's future and not just for its present eminence in science; institutions with a broad role in the nation's intellectual life and not merely laboratories qualified to solve current problems.

The universities have always adjusted their policies and programs to changing social and economic requirements, and they will have to continue to do so. The universities cannot, however, merely respond to outside forces. They must also be independent innovators and stubborn conservators of old values. The weight of history urges that control of the universities by any one benefactor must be prevented if they are to preserve their independence, play their full roles as critics, conservators and innovators, and retain control over their own destinies.



SCIENTIFIC ACTIVITY accounts for a small but growing part of the country's total output. The curves show expenditures for research and development (*black*) and basic research (*color*) as percentages of the gross national product. The increase is somewhat greater in the case of basic research. As in other charts, figures for 1964 are estimates.

HYDROXYL RADICALS IN SPACE

Although these molecules are far less abundant in the interstellar medium than atoms of hydrogen are, radio astronomers find that they reveal many fresh facts about the dynamics of the galactic center

by Brian J. Robinson

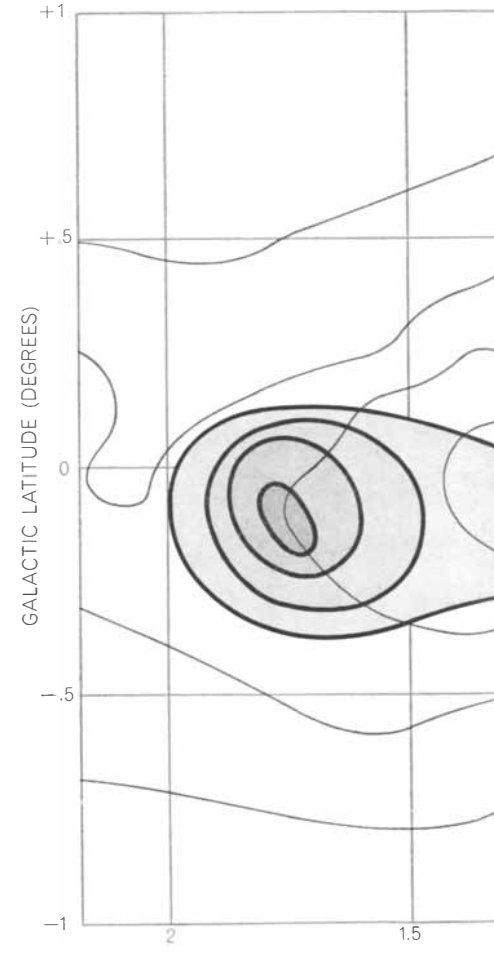
The discovery in 1951 of radio waves emitted by hydrogen in space added an exciting new dimension to astronomy that has proved to be richly informative. It made possible explorations of how hydrogen—the primary building material of all matter in the universe—is distributed in interstellar space. The greatest triumph in the use of the radio signals broadcast by hydrogen has been in mapping the spiral arms of our galaxy. The Doppler shift of these signals, which can be measured far more accurately than the Doppler shift of light, has revealed the motion of hydrogen clouds and has done much to elucidate the dynamics of galaxies in general. Studies of the proportion of hydrogen in distant galaxies have supplied valuable clues as to how they have evolved. The 21-centimeter radio waves emitted by hydrogen have become a potent astronomical tool that has yielded a wealth of new information over the past decade, already reported in several articles in *Scientific American* [see particularly "Radio Waves from Interstellar Hydrogen," by Harold I. Ewen, December, 1953, and "Hydrogen in Galaxies," by Morton S. Roberts, June, 1963].

Now a new signal from interstellar space has joined the signal of hydrogen. Its source is the oxygen-hydrogen molecule OH, known in chemistry as the hydroxyl radical. The hydroxyl signal has been detected only in its muted form—by absorption rather than by emission. In 1963 radio astronomers discovered absorptions by concentrations of the hydroxyl molecule in our galaxy. Unlike the 21-centimeter monotone of hydrogen, the hydroxyl signal consists of four different frequencies. The exploration of hydroxyl clouds in the interstellar medium has become an inquiry of great interest to astrophysicists, because it gives promise of providing information about the formation of molecules in space and about the dynamics of the central region of our galaxy.

Soon after Ewen and E. M. Purcell of Harvard University discovered the hydrogen emission at the 21-centimeter wavelength in 1951, radio astronomers began to search for radio emissions from other atoms in interstellar space. They realized that detection of any such signals would be difficult, because radio emissions from atoms are rare and extremely weak when they do occur. To begin with, there is the rarity of matter in space: the amount of hydrogen averages less than one atom per cubic centimeter, and all other atoms are very much sparser. Secondly, atomic processes that could give rise to detectable radio emissions are considerably less common than those that produce emissions of light. The energy of a quantum of radiation depends on its frequency, and a quantum at radio frequencies is a million times weaker than a quantum at the much higher frequencies of visible light. Consequently radio emissions can arise only from atomic processes that involve comparatively slight changes in the atoms' energy state. Furthermore, because the probability of emission drops sharply with decrease in frequency (or increase in wavelength) of the radiation, any radio signal from a collection of atoms is very weak.

The hydrogen atom's emission of 21-centimeter radio waves arises from an interaction of the spin of the atom's electron and the spin of the nucleus. The two spins may be in opposite directions or parallel. When they are parallel, the electron's spin can flip over so that it is opposite to that of the nucleus, resulting in a state of lower energy. The transition releases a quantum of energy at the frequency of 1,420.4057 megacycles per second (or the wavelength of 21 centimeters).

Few other atoms likely to be reasonably abundant in space have a nuclear spin that would make such interactions possible. The nucleus of nitrogen has a



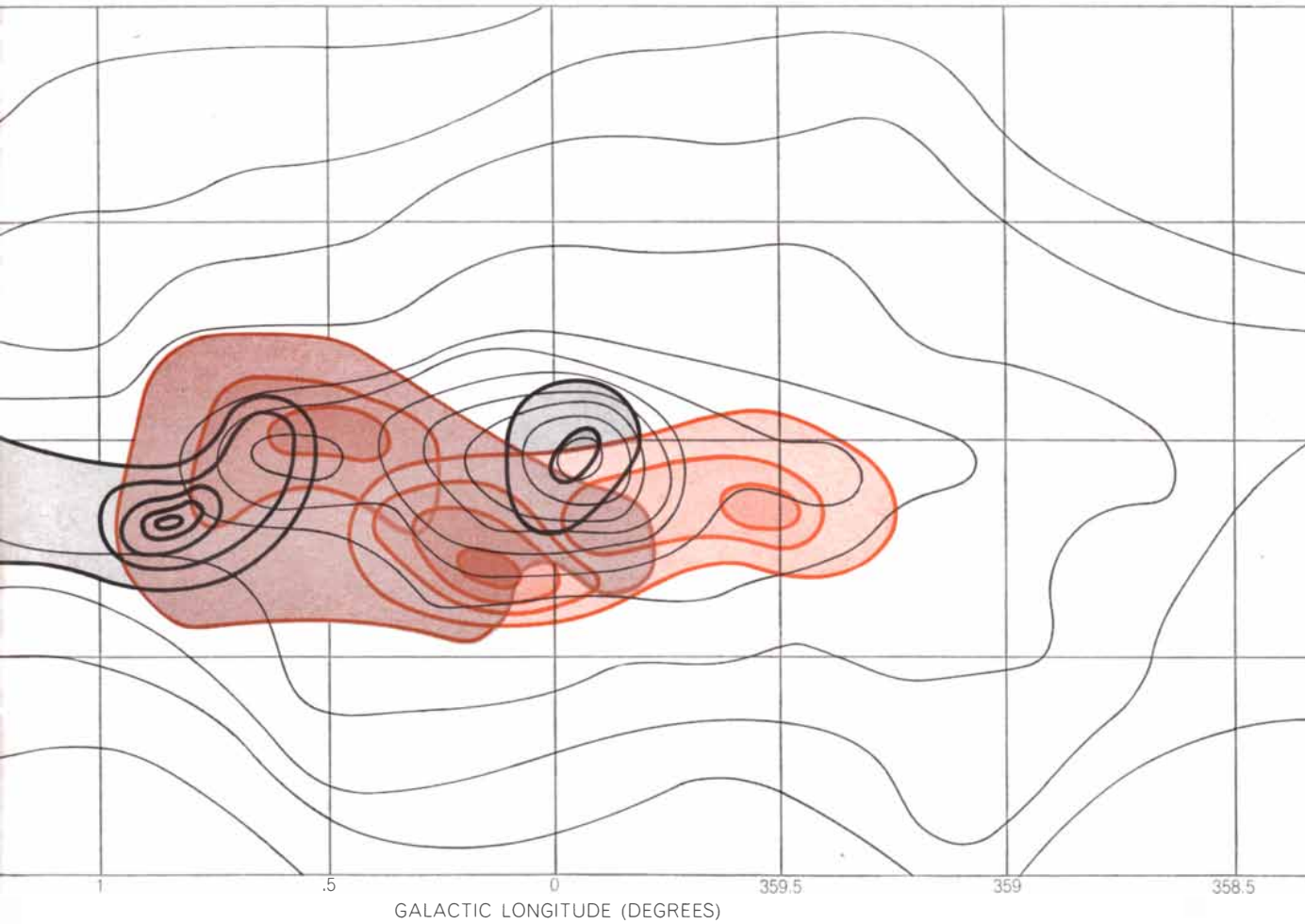
THREE GAS CLOUDS near the center of our galaxy are composed of atoms of oxygen and hydrogen linked together in diatomic molecules of the hydroxyl radical (OH).

spin, but its interaction with the electron spin is too weak to produce radio signals detectable on the earth. On the basis of what was known about atomic radio emissions from laboratory experiments, radio astronomers chose deuterium (heavy hydrogen, or H^2) as their first target of search for a second radio emitter in space. The interaction of the nuclear and electronic spins in deuterium is considerably weaker than in ordinary hydrogen. Moreover, the deuterium isotope must be many thousands of times less abundant in space. But deuterium at least offered a definite objective to aim for, because the frequency of the radio waves it emits (or will absorb) was known with high precision: it is 327.3843 megacycles.

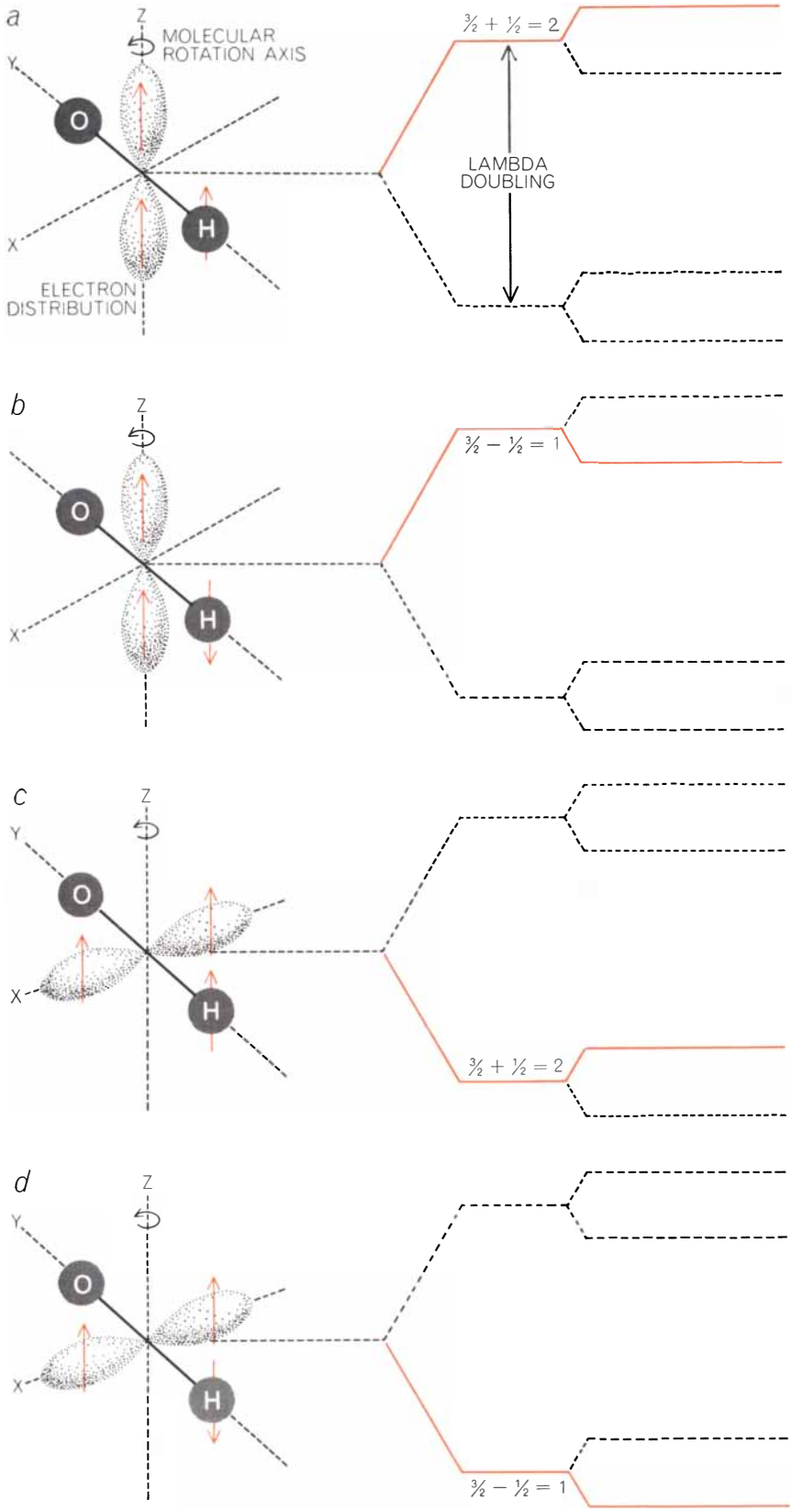
For several years radio astronomers in the U.S.S.R., Australia and Britain made painstaking efforts to discover the 327-megacycle line in the spectrum of radio waves from space. They turned their radio telescopes to strong radio sources, particularly the powerful emitter in the direction of the constellation Cassiopeia called Cassiopeia *A*, and examined the spectrum of each source closely to try to detect absorption of its radiation by deuterium in interstellar space. In this search the investigators developed refined techniques and produced the most accurate measurements that had ever been made in radio astronomy. No sign of the hunted line could be found. The 250-foot radio telescope at Jodrell Bank in Britain, pushing the observations to a level that would have detected absorption of as little as one part in 15,000 of the radiation from Cassiopeia *A*, failed to record any absorption by deuterium. This indicated that the proportion of deuterium atoms to hydrogen atoms was less than one in 4,000. In 1962 Sander Weinreb of the Massachusetts Institute of Technology undertook a new search with a more sensitive technique. In extensive observations of Cassiopeia *A* (totaling more than 1,800 hours) with the 85-foot radio telescope at the National Radio Astronomy Observatory in Green Bank, W.Va., he too failed to detect any deuterium absorption. His technique established that if any deuterium existed in interstellar space, its ratio to ordinary hydrogen there must be less than one atom in 13,000—less than half the ratio of the deuterium isotope in hydrogen on the earth.

In truth, deuterium was not a particularly good candidate for this kind of exploration. It had been chosen for concentrated attention because it was the only plausible candidate whose radio frequency was accurately known. In the meantime, however, accurate measurements of radio emissions from the hydroxyl radical were obtained in the laboratory, and in 1963 Weinreb applied his new techniques to a search of the heavens for that molecule.

In theory radio absorption by the hydroxyl radical and by certain other molecules should be much stronger than that by deuterium. In a two-atom molecule (such as the hydroxyl radical) with



These are a few of the many clouds of such molecules detected here and elsewhere within our galaxy as a result of their absorption of radio frequencies between 1,612 and 1,720 megacycles. The emitter in this region is a complex of radio sources known as Sagittarius *A*; the narrow contour lines show how its radio brightness increases toward the galactic center. Doppler shifts show that two of the clouds (*dark and light color*) are moving outward; the third is receding. OH molecules were first detected in space in October, 1963.



FOUR ENERGY LEVELS of the hydroxyl molecule are shown schematically. The motion of the unpaired electron interacts with the rotation of the diatomic molecule. Two states are possible: *a*, where the electron distribution is along the axis of molecular rotation, and *c*, where it is in the plane of rotation. A transition between the two absorbs or emits a quantum of energy at microwave frequencies. Still another emission or absorption can occur when one nucleus of the molecule (*the hydrogen nucleus in this illustration*) has a magnetic moment opposed to, rather than aligned with, the molecule's internal magnetic moment (*compare "a" with "b" and "c" with "d"*). Numbers show the quantum arithmetic of each level.

an odd number of electrons several interactions giving rise to radio emissions are possible. The rotation of the two nuclei interacts with the orbital motion of the odd electron, and the electron's motion can assume either of two orientations with respect to the rotation axis of the molecule; this produces a "lambda-doubling" of the molecule's rotational states [*see illustration at left*]. The transition from one orientation to the other involves an energy change that corresponds to the frequency of radio waves in the microwave range, well below the frequencies at which transitions in atmospheric molecules absorb waves from space. In the hydroxyl molecule each of these two states in turn has two possible configurations that depend on the direction of the hydrogen nucleus's spin. Thus the hydroxyl molecule has four possible energy transitions and may emit (or absorb) radio waves at four different frequencies. When the motion of the unpaired electron in the molecule is changed by the absorption of radio energy, it is the radio wave's electric field that produces this change, in contrast to the case of the single hydrogen atom, where the reversal of the electron's spin involves only the comparatively weak energy of the magnetic field. Consequently radio absorption by the molecule is a great deal stronger—about 10,000 times stronger—than that by a hydrogen atom. Radio-absorbing molecules in space therefore should be detectable even at concentrations well below one molecule per cubic meter.

It seemed likely that the hydroxyl radical and a few other two-atom molecules (CH, SiH and NO) might be found in detectable concentrations in interstellar space. But the failure to find deuterium, and the lack of precise information about the radio frequencies of the molecules in question, discouraged any thorough search for them. Nonetheless, in 1958 an exploratory attempt was made by Alan H. Barrett and A. Edward Lilley at the Naval Research Laboratory; it was unsuccessful. Then, in 1959, Charles H. Townes and his group at Columbia University did succeed in establishing the frequencies of two radio lines for the hydroxyl molecule; they measured these frequencies in the laboratory and found them to be 1,665.46 and 1,667.34 megacycles, with an estimated possible error of 30 kilocycles. Not until 1963, however, were any serious attempts made to use this information to probe for the hydroxyl radical in space. In the fall of that year Weinreb and his co-workers,

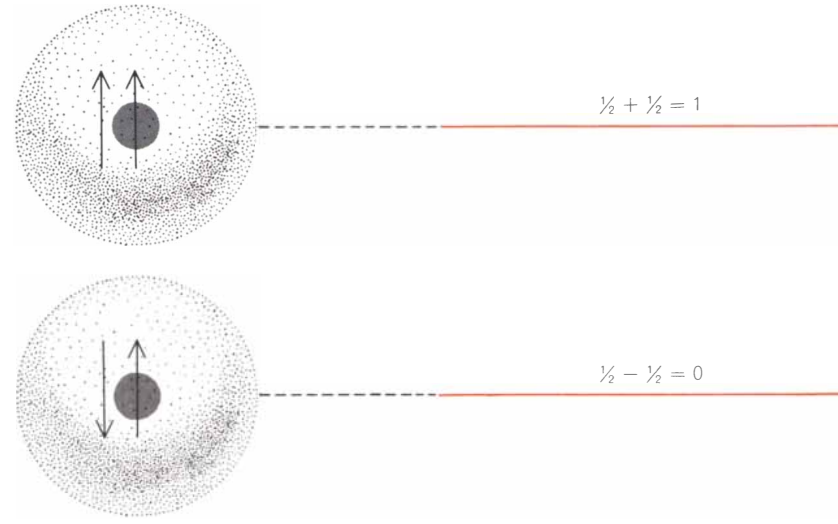
applying his new techniques, met success on their first try.

Working with Barrett, M. Littleton Meeks and J. C. Henry, Weinreb attached his ultra-stable receiver to M.I.T.'s 84-foot radio telescope on Millstone Hill in Massachusetts, pointed the dish toward Cassiopeia *A* and sifted its emissions in a band of frequencies around 1,667 megacycles. The first evening's observations, on October 15, 1963, showed a significant absorption at the 1,667-megacycle line of the hydroxyl radical. Soon afterward the receiver also disclosed absorption at the other known OH line, 1,665 megacycles. The relative intensities of absorption by the two lines were in the ratio of nine to five, which agreed well with the values for the hydroxyl radical that had been established by theoretical calculations and laboratory experiments. There could be little doubt, therefore, that the absorption lines signaled the presence of hydroxyl molecules in interstellar space between Cassiopeia *A* and the earth.

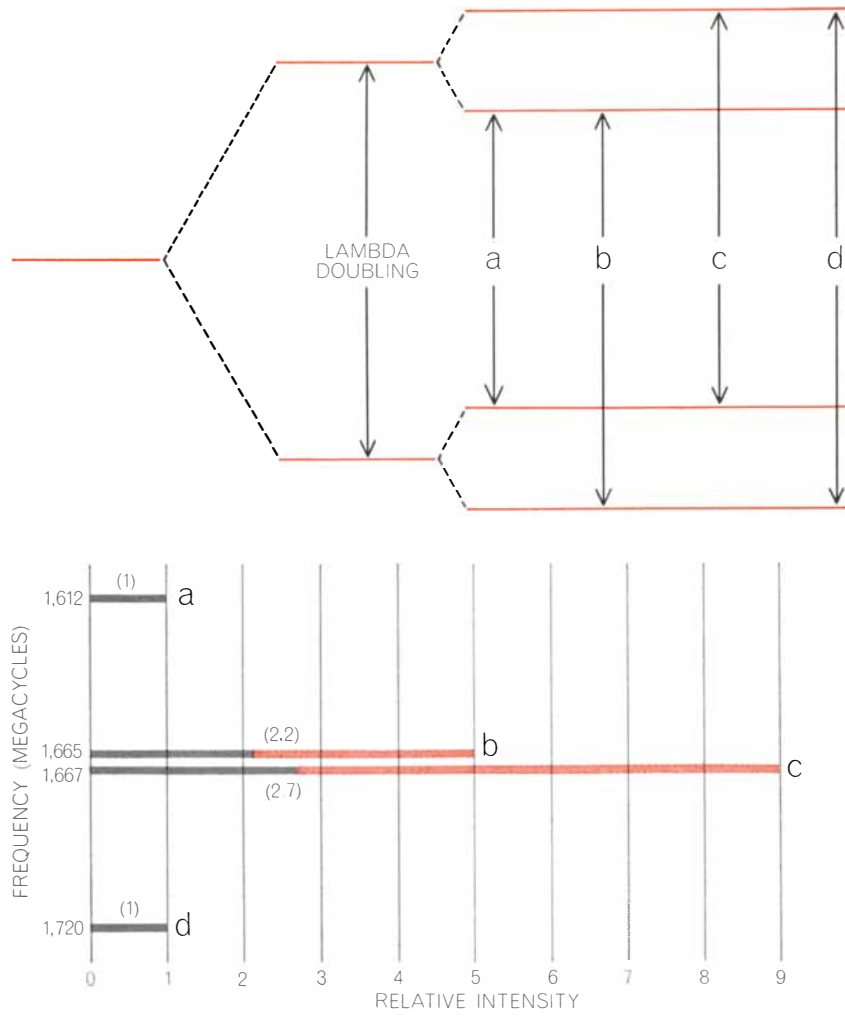
The amount of absorption was small and indicated that the average OH density in that direction was about one molecule per 10 cubic meters. From the shapes of the absorption spectra it could be deduced that in the various gas clouds observed the proportion of hydroxyl molecules to hydrogen varied slightly. The observations also yielded precise frequencies for the two hydroxyl lines: 1,665.402 and 1,667.357 megacycles, with a margin of uncertainty of only seven kilocycles plus or minus.

The announcement of the discovery at M.I.T. prompted several other radio observatories to undertake immediate surveys of the same kind. In Australia, John G. Bolton, Frank F. Gardner, Karel J. van Damme and I quickly improvised a suitable receiver and turned the 210-foot radio telescope of the Parkes Observatory to Sagittarius *A*, a radio source at the center of our galaxy lying behind several spiral arms rich in gas clouds. On November 20, a month after the M.I.T. discovery, we detected hydroxyl absorption at the two frequencies in the Sagittarius direction. Three weeks later Ewen and Nannielou H. Dieter of Harvard, using the U.S. Air Force's 84-foot radio telescope on Sagamore Hill in Massachusetts, confirmed the presence of OH in both the Cassiopeia and the Sagittarius directions. In the same week Harold F. Weaver and David Williams of the University of California at Berkeley also observed absorption by hydroxyl molecules in the direction of the center of the galaxy.

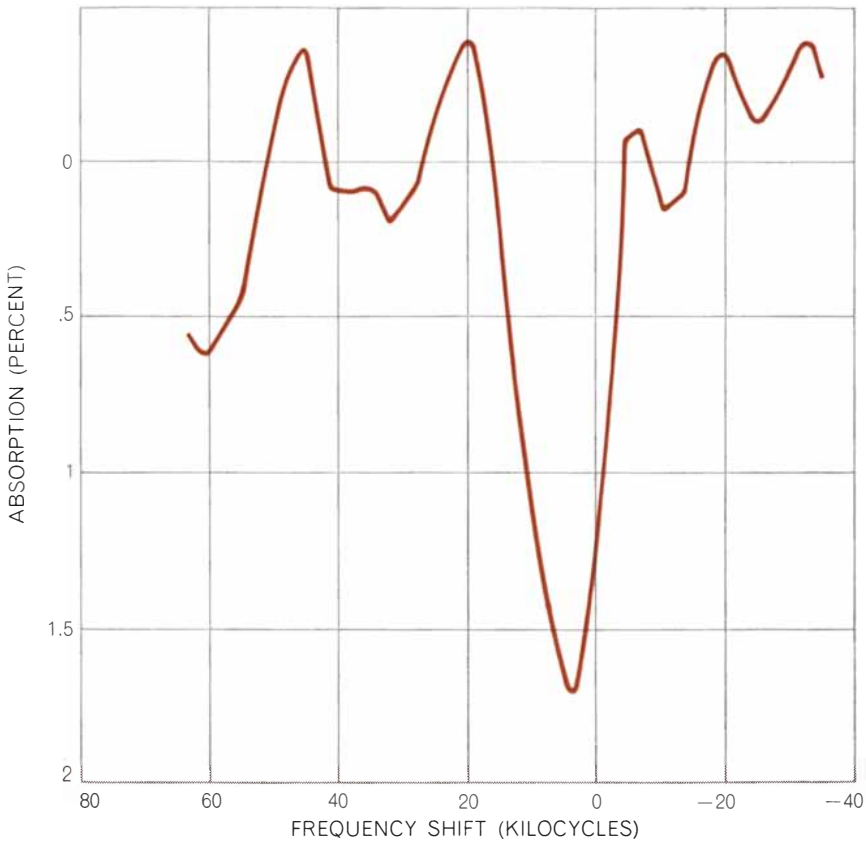
What could the hydroxyl radical, as



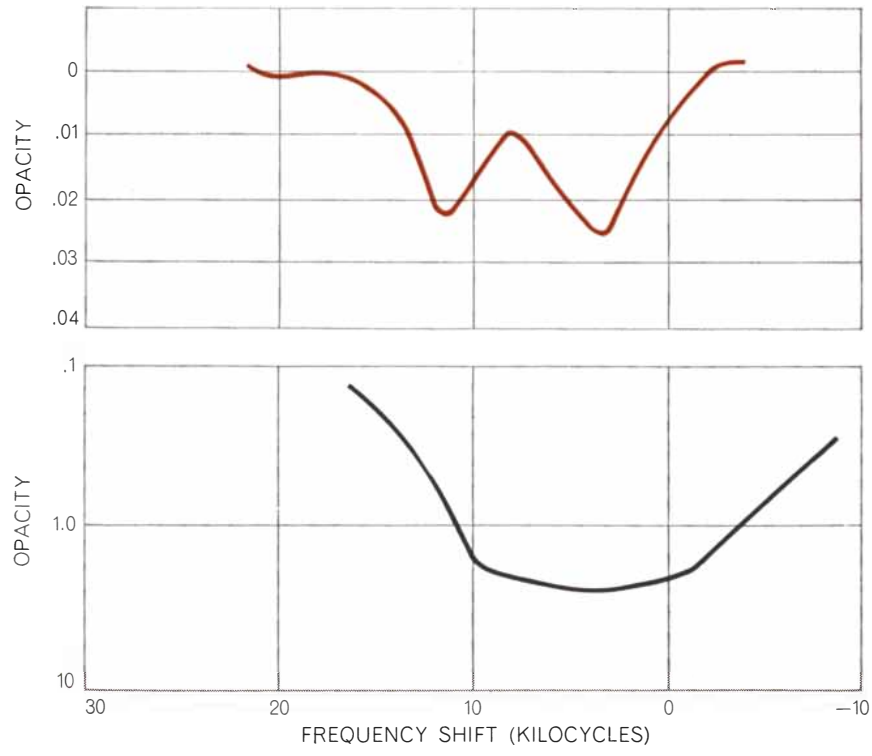
HYDROGEN SPIN-FLIP is the parallel phenomenon on the atomic level responsible for absorption and emission of quanta at the 21-centimeter wavelength. When nuclear and electron spins are parallel (*top*), the atom contains more energy than when the spins are opposed.



ABSORPTION FREQUENCIES related to the four energy levels of the hydroxyl molecule that stem from "lambda-doubling" are not equally common in occurrence. In theory whole-step transitions ("*b*" *and* "*c*" *at top*) are far more probable than semistep (*a*) or superstep (*d*) ones. Observation shows this is so but not to the extent predicted. Black lines (*bottom*) give the values observed at the galactic center; colored lines, those predicted for whole steps. Observations also show that these values are not alike in all regions of the galaxy.



FIRST EVIDENCE that hydroxyl molecules existed in space was this absorption dip recorded by Massachusetts Institute of Technology investigators in the radio spectrum of Cassiopeia *A* near 1,667.357 megacycles (*zero on horizontal scale*). Smaller fluctuations represent "noise." The absorption indicates an OH density of only one molecule per 10 cubic meters.



FINER RESOLUTION shows that the OH absorption of Cassiopeia *A* emissions near 1,667 megacycles actually consists of two separate dips (*top*), corresponding to a pair of clouds moving with different velocities. The higher random velocity of hydrogen atoms prevents detection of this motion; 21-centimeter readings produce only a single profile (*bottom*).

a new marker in the sky, tell us about our galaxy and events in interstellar space? Not very much, it was thought at first. No emissions from the molecule could be detected, so observations of its presence were limited to narrow sections of space where it absorbed radiation from strong radio sources behind it. The molecule appeared to be so sparsely distributed in space that it absorbed only 1 to 2 percent of the radiation from these sources. As a tool for astrophysical studies it seemed to hold little promise.

Soon, however, it began to turn up various clues. Barrett and his colleagues at M.I.T. found that OH absorption could serve as a means of resolving different gas clouds in space and estimating their relative speeds and temperatures. Hydrogen atoms, being light, have comparatively rapid thermal motions in a gas cloud. This has the effect of broadening their radio-absorption line—that is, widening the span of frequencies they will absorb. The hydroxyl molecule, on the other hand, has a thermal velocity only one-fourth that of the hydrogen atom, because it is 17 times heavier. An OH absorption line therefore should be sharper than the hydrogen line. Barrett and his colleagues examined one hydroxyl line with fine tuning and found that they could resolve it into two separate absorptions [*see bottom illustration at left*]. They interpreted this to mean that the two dips represented two separate clouds of gas, each moving at a different velocity in the line of sight from the earth and having a different internal temperature—respectively 90 degrees and 120 degrees Kelvin (degrees centigrade above absolute zero).

At the Radiophysics Laboratory of the Commonwealth Scientific and Industrial Research Organisation in Sydney we were struck by the curious shape of the absorption spectrum in our first observations of OH at the center of the galaxy. The general shape of the spectrum was so unusual that we took our receiver apart to see if an aberration of the instrument was responsible for it. After the equipment was modified and reassembled the hydroxyl-radical absorption spectrum in the direction of Sagittarius *A* still showed the same curious shape.

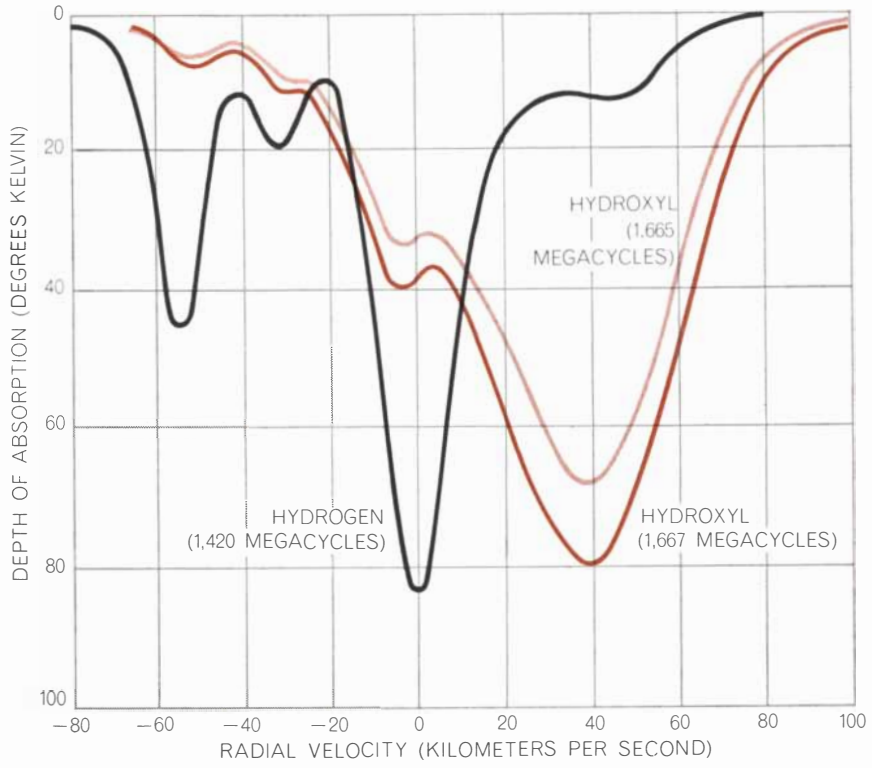
A spectrum of absorption by gas clouds in space represents, of course, a mixture of absorptions by various clouds in the line of sight. They may be moving at different velocities toward or away from the earth, and these dif-

ferences will produce slightly different Doppler shifts of a given absorption line. Therefore the profile of the absorption spectrum will show a series of separate dips [*see illustrations at right*]. The absorption in the Sagittarius *A* direction by hydrogen shows well-marked dips presumably representing gas in spiral arms moving toward the solar system (that is, outward from the center of the galaxy) at speeds of 30 and 53 kilometers per second. The same "features" (absorption dips) showed up in the hydroxyl absorption profiles. But the hydroxyl spectrum also contained a broad, dominating feature that had no marked counterpart in the hydrogen-absorption picture. It showed a strongly absorbing cloud or clouds of hydroxyl radicals moving *toward* the center of the galaxy at a velocity of 40 kilometers per second! This was completely unexpected, because the hydrogen observations indicated that gas clouds generally move *away* from the center.

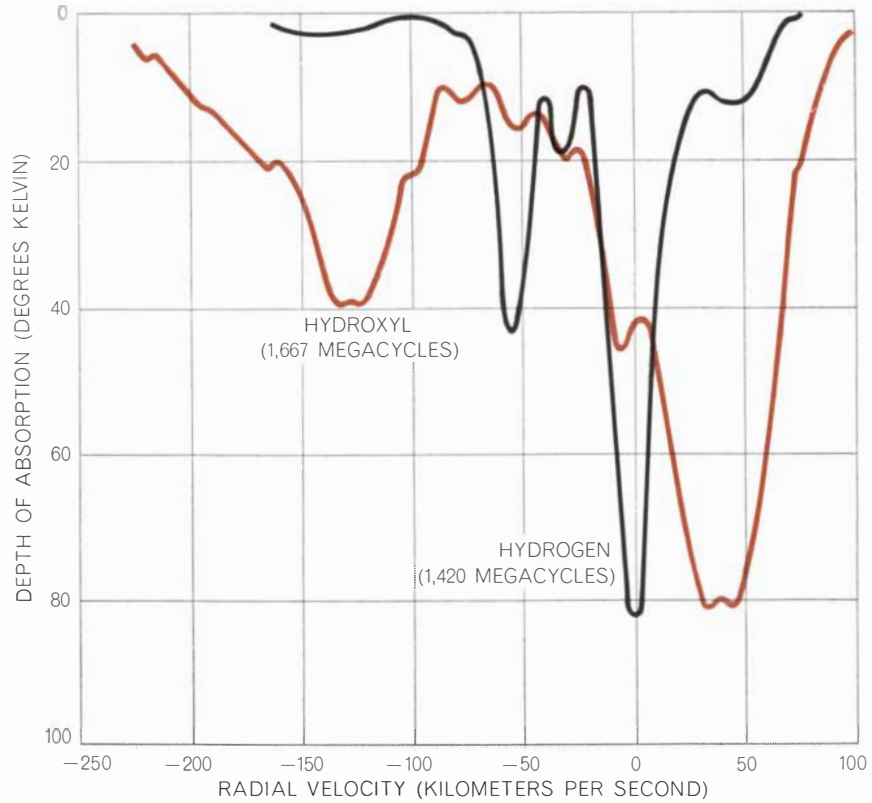
Further observations revealed another surprise. The inward-moving OH did not blanket the entire complex of sources of Sagittarius *A* (which extend for several degrees) but covered only the small, intense source near the very center of the galaxy; in that direction, however, the hydroxyl concentration was so high that it absorbed 60 percent of the radiation at its absorbing lines! In contrast, the hydroxyl clouds in the direction of Cassiopeia *A* absorb only 1.6 percent of the radiation at these frequencies. Surveys of other radio sources in our galaxy have since shown, in fact, that the proportion of hydroxyl molecules to hydrogen in different gas clouds varies enormously: by a factor of more than 1,000. The concentration of hydroxyl molecules increases rapidly toward the center of the galaxy, where an appreciable fraction of the oxygen atoms must have formed OH radicals.

Shortly after our observations in Australia, Lilley, Samuel J. Goldstein, Ellen J. Gundermann and Arno A. Penzias surveyed the hydroxyl absorption of Sagittarius *A* over a wider frequency range with the Harvard 60-foot dish and a maser receiver. They found another strong absorption by OH, but this one was moving *away* from the galactic center at 120 kilometers per second. There was no marked absorption by hydrogen associated with it, which again indicated a high proportion of hydroxyl radicals near the center of the galaxy. A month later we confirmed the Harvard observations with our telescope in Australia.

Needless to say, the finding that OH clouds can serve as a probe for the

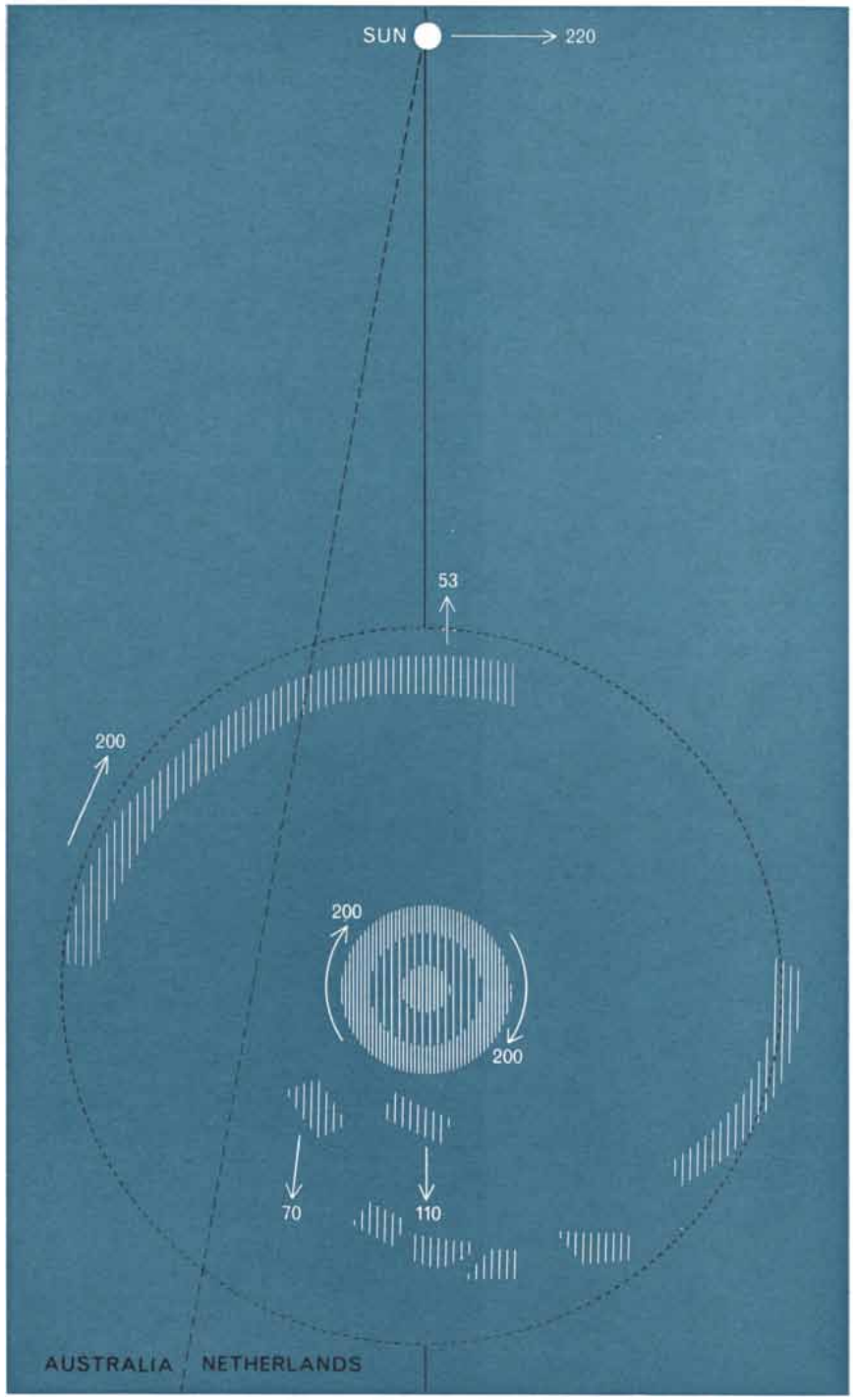


CONTRASTING MOTIONS of hydrogen atoms and hydroxyl molecules in the gas clouds near the galactic center were revealed by Australian studies of Sagittarius *A* in 1964. The two strongest OH lines (*dark and light color*) show positive values on the horizontal scale, or inward movement; the negative value for hydrogen (*left*) means outward movement.



FURTHER EXTENSION of measurements by Harvard University workers revealed an OH accumulation near the galactic center with an outward velocity of 120 kilometers per second. Colored line shows OH absorption at 1,667 megacycles as recorded by Australian radio telescope. Black line shows hydrogen 21-centimeter band at a much reduced scale.

center of our galaxy has aroused keen interest among astronomers. The nuclei of galaxies have recently come in for attention because of remarkable events that have been discovered in some of them [see "Quasi-stellar Radio Sources," by Jesse L. Greenstein, SCIENTIFIC AMERICAN, December, 1963, and "Exploding Galaxies," by Allan R. Sandage, SCIENTIFIC AMERICAN, November, 1964]. The events taking place at the center of our galaxy are less catastrophic but no less interesting. We cannot see the center because of the heavy banks of dust clouds that lie in the way. The hydrogen line opened the center to radio probes, but it leaves many of the observations unclear because they contain a mixture of emissions by the gas through the full depth of the galaxy and absorptions by the hydrogen on our side of the center. In the hydroxyl radical we have a sharp tool that tells its story only by absorption, without confusing the picture with emissions. In combination with the hydrogen probe it may soon help to unravel some of the complex phenomena taking place at the galaxy's center.



MOVEMENT OF HYDROGEN near the galactic center is of two kinds. Within 2,000 light-years of the center (*radius of inner band*) the gas seems to be rotating rapidly. From that region to a distance 10,000 light-years away from the center (*radius of broken circle*) the hydrogen is streaming radially outward at varying speeds. Arrows show direction of motion; figures are velocities in kilometers per second. The view is from above the galactic north pole; the broken diagonal separates Australian from Netherlands astronomical findings.

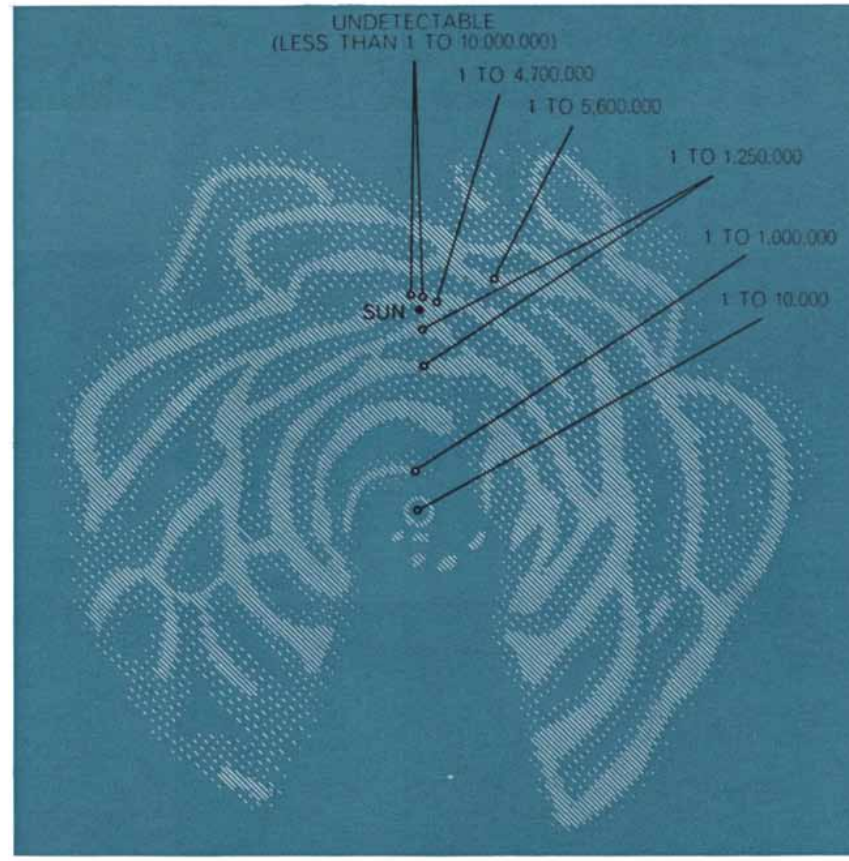
In general the hydrogen clouds in our galaxy revolve in nearly circular orbits around the center under the control of gravitational forces. Near the center, however, the gas is flowing radially outward at a high rate, and at the nucleus itself it seems to be whirling at extremely high speed. What appears to be blowing the hydrogen away from the center? Is it the force of a gigantic explosion that took place sometime in the past? Is it a driving force of the hydromagnetic type? Perhaps close comparison of the distribution and motions of hydrogen and the hydroxyl radical will help to provide an answer.

Our group in Sydney has surveyed the whole central region of the galaxy with the Parkes 210-foot radio telescope. It is an ideal location for such observations, because in Australia the galactic center passes almost directly overhead. The survey shows that, whereas hydrogen covers the entire center, the hydroxyl radical is concentrated in small clouds. Some are traveling away from the center, some toward it, each at its own velocity [*see illustration on pages 26 and 27*]. There are startling differences between the motions of the hydrogen and the hydroxyl clouds, with the movements of the latter corresponding only in the most general way to those of the hydrogen in which they are embedded.

Equally puzzling is the contrast between the high abundance of hydroxyl radicals near the center and the scarcity of these molecules out in the spiral arms of the galaxy. How, indeed, is the OH molecule formed at all? The probability of hydrogen and oxygen atoms uniting by direct collision in space is extremely low. The most likely process is the formation of molecules on grains of dust on which the atoms collect. These "dust" grains are believed to be composed mainly of frozen water; presumably the grains evaporate and

release H_2O molecules, which are then dissociated into H and OH. The central part of most spiral galaxies, probably including our own, contains much less dust than the spiral arms. It may be, then, that the hydroxyl molecules in the center are remains of dust clouds that have been evaporated by warming collisions, whereas in the spiral arms the OH is still trapped in dust grains. Near the center of the galaxy there is comparatively little ultraviolet radiation to dissociate the hydroxyl molecule; in the spiral arms, on the other hand, any OH that is released from the dust grains may have only a short lifetime because it is exposed to copious ultraviolet radiation from hot blue stars.

As soon as the OH absorptions at 1,665 and 1,667 megacycles were discovered, efforts were of course made to detect absorption at the other two radio lines expected in the hydroxyl molecule. These minor lines were so weak that no attempt had been made to measure their frequencies in the laboratory. But in April, 1964, on the basis of new theoretical calculations of the frequencies, both lines were detected in the Sagittarius direction with the Parkes telescope: one at 1,612.2 megacycles, the other at 1,720.6 megacycles. These new discoveries brought to light a new anomaly. Theoretical calculations and later laboratory measurements by H. E. Radford showed that the relative intensities of absorption at the four OH lines should be in the ratios 1:5:9:1 (for the lines at 1,612, 1,665, 1,667 and 1,720 megacycles respectively). The actual absorptions recorded in the telescope receiver, however, put the ratios at 1:2.2:2.7:1 for the strong Sagittarius A absorption. These ratios are incompatible with simple self-absorption effects and imply unusual physical conditions at the center. Moreover, it turned out that the ratios vary in different parts of the galaxy. This is an enigma that obviously calls for further investigation; it may lead to better understanding of the conditions in interstellar space and the formation of hydroxyl molecules at the center of the galaxy.



VARYING CONCENTRATIONS of hydroxyl molecules throughout the interstellar medium are plotted on a generalized sketch of the distribution of hydrogen atoms in the spiral arms of the galaxy. The ratio of hydroxyl to hydrogen increases a thousandfold between the region near the solar system and the galactic center. There one hydroxyl molecule is found for every 10,000 hydrogen atoms, but surveys in other directions show less than one per four million hydrogen atoms or are unable to detect any hydroxyl absorption.

What are the prospects for discovering other substances in space by their radio lines? One likely candidate is the diatomic molecule CH, which has a lambda-doubling in its lowest energy state and should absorb radio energy in the region around 3,400 megacycles. Preliminary surveys in that neighborhood have failed to find any absorption, but the molecule may become detectable when a more precise estimate of the frequency has been obtained. Another molecule being hunted is the hydroxyl radical containing not common oxygen 16 but the rare isotope oxygen 18. The four radio lines of $O^{18}H$ should be slightly displaced from those of $O^{16}H$, with the strongest line at 1,639.3 megacycles. The amount of the heavy version of the hydroxyl molecule in space must be very small indeed, but the chances of detecting the strongest line seem good.

What about the hydrogen molecule itself: H_2? We have theoretical grounds for believing that there must be a great deal of hydrogen in the molecular form in interstellar space. One of these grounds is that our galaxy, considering the gravitational motions within it, seems to have about 50 percent more mass than can be accounted for by its known content of stars and atomic hydrogen; the missing mass may be in the form of molecular hydrogen in space. Unfortunately we cannot hope to detect molecular hydrogen in interstellar space, because in its lowest energy state the molecule has no transitions at either optical or radio frequencies. There is, however, a chance that the molecule might be detected close to hot stars, where it may be ionized or excited into states that would cause it to emit detectable radiation.

This general possibility—excitation of atoms or molecules in space by nearby stars—gives rise to a number of interesting speculations. For example, it has been calculated that a hydrogen atom raised to a highly excited state (that is, with its electron in a very high orbit) will emit energy in the microwave range, because the energy difference between adjacent orbits at that height is small. Within the past year Soviet radio astronomers have reported the detection of radio emissions by highly excited hydrogen atoms in the Omega nebula at 5,763.6 and 8,872.5 megacycles. We are waiting with interest for confirmation of these observations.

The Secondary Recovery of Petroleum

The life of an oil well can be extended by techniques that push oil out of the rock after the decline of natural pressure. Such methods may account for half of the U.S. output of oil by 1980

by Noel de Nevers

A common sight in the early days of the oil industry was the "gusher," a new well that spouted oil over the surrounding countryside. Gushers are rare today because modern well-drilling equipment is designed to avoid these wasteful and hazardous episodes. Of course the natural cause of gushers still exists; petroleum in an untapped natural reservoir is generally under high pressure from the overlying rock. When a drill penetrates the reservoir, the pressure causes the oil to flow to the surface unassisted.

In most wells oil will continue to flow to the surface for several years; all the operator has to do is collect the oil, separate it from the natural gas that comes with it, ship it and sell it. Eventually, however, he finds that the pressure in the reservoir is no longer sufficient to lift the oil to the surface. He then installs a pump in the well. Thereafter he will continue to recover oil for many years; the natural pressure still makes the oil flow out of the rock into the well, so that all he has to do is pump it to the surface.

The two methods of production described thus far are called "primary recovery." They fall far short of total recovery of the oil that is in a field. In the early days of the oil industry the primary-recovery techniques were all a driller had available; when the natural pressure in the oil-bearing rock declined until it no longer forced fluid out of the rock into the well and pumping ceased to be productive, the operator would plug the well and move on to new fields. As a rule in such a situation it could be calculated from the area, thickness and porosity of the abandoned field that he was leaving behind as much as 80 percent of the oil originally present. If he had extracted more than a third of the oil, he had done well.

There were several incentives to look for ways to extract more of the oil. The operator knew where it was, owned the mineral rights to it and had the wells and surface facilities for processing it already in place. These incentives led oilmen to develop several techniques that are called "secondary recovery." The four basic possibilities in such recovery are mining, squeezing, pushing and sucking.

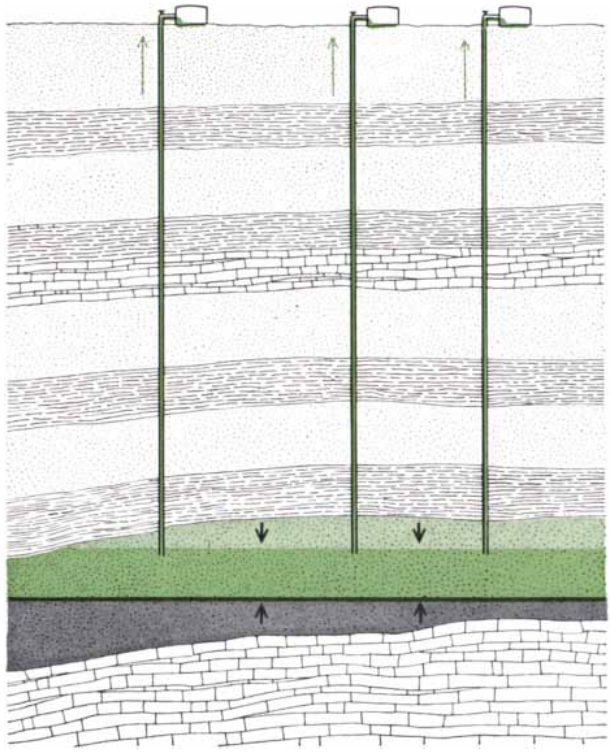
Mining—physically removing the oil-bearing rock from its position hundreds or thousands of feet below the surface of the earth and bringing it to the surface for processing as an ore—is in most cases impractical. Crude oil is a cheap substance, selling in the U.S. for less than a cent a pound. Moreover, its concentration in the "ore" is rather low, and the depth of most deposits makes mining difficult. For these reasons secondary recovery by mining is usually uneconomical.

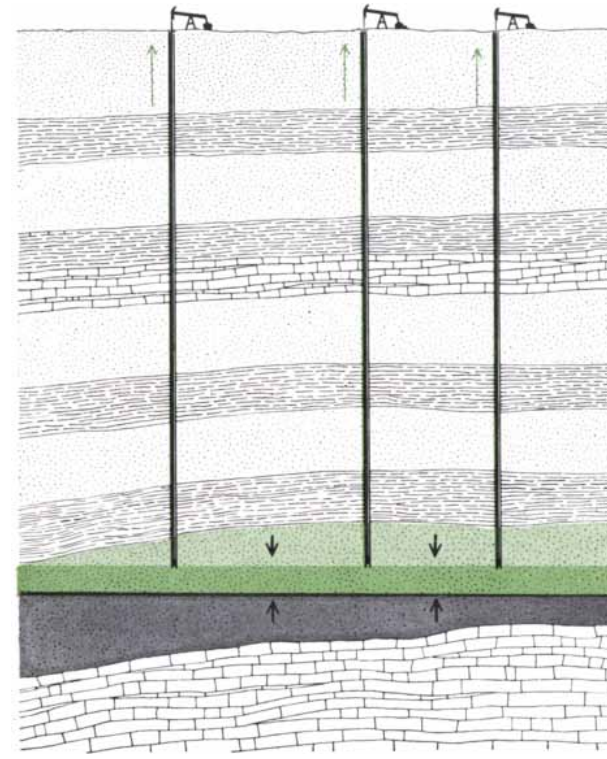
Squeezing means forcing the particles of oil-bearing rock closer together and thus pressing out the oil. Most oil-bearing rock is much too strong to be compressed by any available force. A rare exception obtains in Long Beach, Calif., where the oil-bearing sand under the city is highly compressible. In fact, the city slowly began to sink when the oil was first extracted, and the sinking effect could have been used as a natural squeeze to achieve a high recovery of oil. The location of Long Beach at the edge of the ocean, however, meant that large expenditures would have been needed for dikes and seawalls to keep the city from being flooded. As a result the squeeze procedure is not used at Long Beach; the oil that is removed is replaced with salt water so that the city does not sink.

Pushing involves directly displacing the oil from the rock with some other substance. It has so far been the most successful method of secondary recovery. The pushing must be done with a fluid, because there is no known way to force solids into the tiny spaces in the porous rock in which the oil is held. Since the pushing fluid remains in the rock after the operation is over, it must be far less valuable than the oil recovered. This requirement restricts the choice of pushing fluids to water, air and natural gas. Pushing oil out with water is called water flooding; pushing it with gas or air is called gas drive or air drive.

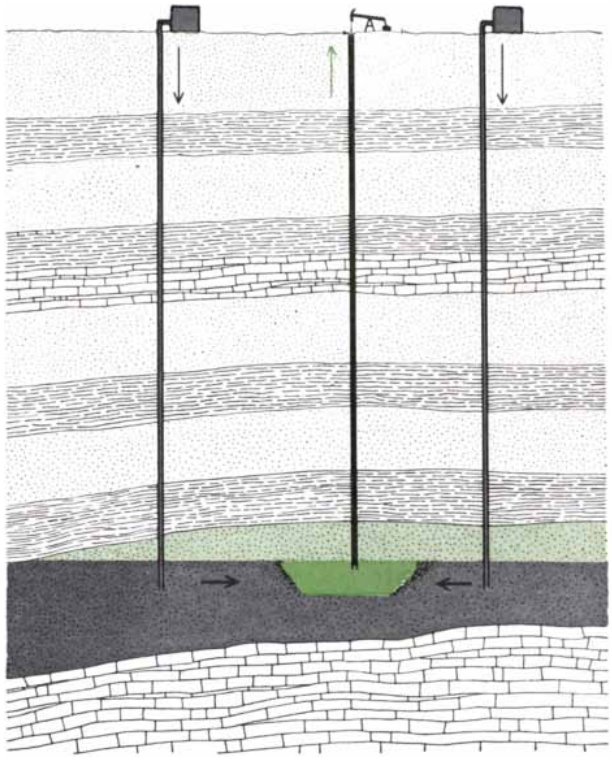
Sucking oil out is a variation of pushing: one uses the air in the atmosphere as a pusher. Sucking is an inefficient type of pushing. The effectiveness of a push is proportional to the difference in pressure between the pushing fluid entering the oil field and the oil being removed from the field. In a sucking operation the maximum pressure difference available is about 15 pounds per square inch, which is the pressure of atmospheric air. Far higher pressure

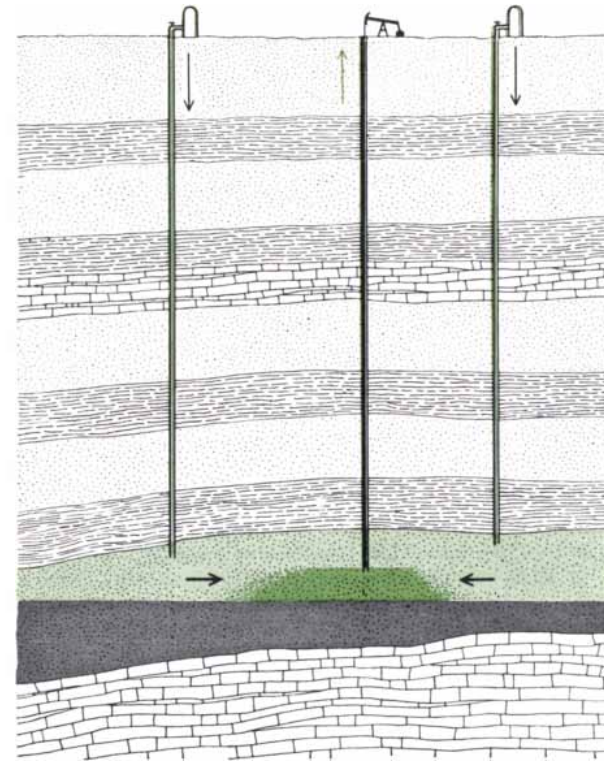
PUSHING FLUID is tested in a glass-bead model (*opposite page*) at the laboratory of the California Research Corporation, a subsidiary of the Standard Oil Company of California. Model represents an oil-bearing stratum of coarse texture; spaces between beads are filled with a viscous oil. A less viscous fluid, in this case a lighter oil dyed so that its behavior can be seen clearly, is introduced as a test of how effectively it sweeps out the reservoir oil. A pushing fluid less viscous than reservoir oil develops a fingering effect and hence bypasses some of the reservoir oil. In this test the pushing is from the lower left corner to the other three.





PRIMARY-RECOVERY SITUATIONS are depicted schematically. Panel at left represents a natural reservoir soon after it has been tapped; pressure from overlying gas and underlying water is sufficient to force the oil up the well without any outside assistance. At right the situation is shown after the pressure has declined and pumps are necessary to lift the oil to the surface. The pressure is still high enough to force the oil out of the reservoir rock and into the well; this too is regarded as primary recovery.





SECONDARY-RECOVERY SITUATIONS arise when the natural pressure falls so low that it can no longer force oil out of the reservoir rock. Then the oil must be pushed out of the rock toward the production well by a substance injected from the surface. The panel at left shows secondary recovery by the injection of water. In panel at right the substance being injected as a pushing fluid is gas.

differences can be obtained by pushing with a fluid that is itself under high pressure.

Secondary recovery by pushing operations involving water or gas is now practiced extensively on a commercial scale in the U.S. A survey made by the Interstate Oil Compact Commission, an agency formed by several states that produce oil and gas, found that in 1960 there were 5,734 fluid-injection projects and that they accounted for about 30 percent of the 2.4 billion barrels of oil produced in the U.S. that year. (One barrel equals 42 U.S. gallons.) The economic possibilities in such an operation are suggested by a recent report of a water-flooding project in Texas. The project covers 8,110 acres with 199 producing wells. An investment of approximately $4 million in the venture, in addition to the money already spent for primary recovery, is expected to achieve an ultimate additional recovery of some 67.5 million barrels of oil, worth $169 million at current prices.

Almost all the commercially successful secondary-recovery operations involve pushing with water or gas. These operations decrease the amount of oil left behind when the field is finally abandoned from about 80 percent of what was there initially to about 50 percent. Secondary-recovery operations of this kind are now routine; the more interesting aspect of secondary recovery is the work the major oil companies and a few university laboratories are doing to improve its efficiency. This work has two goals: (1) increasing the area swept out by the pushing fluid and (2) increasing the oil-recovering efficiency of the pushing fluid.

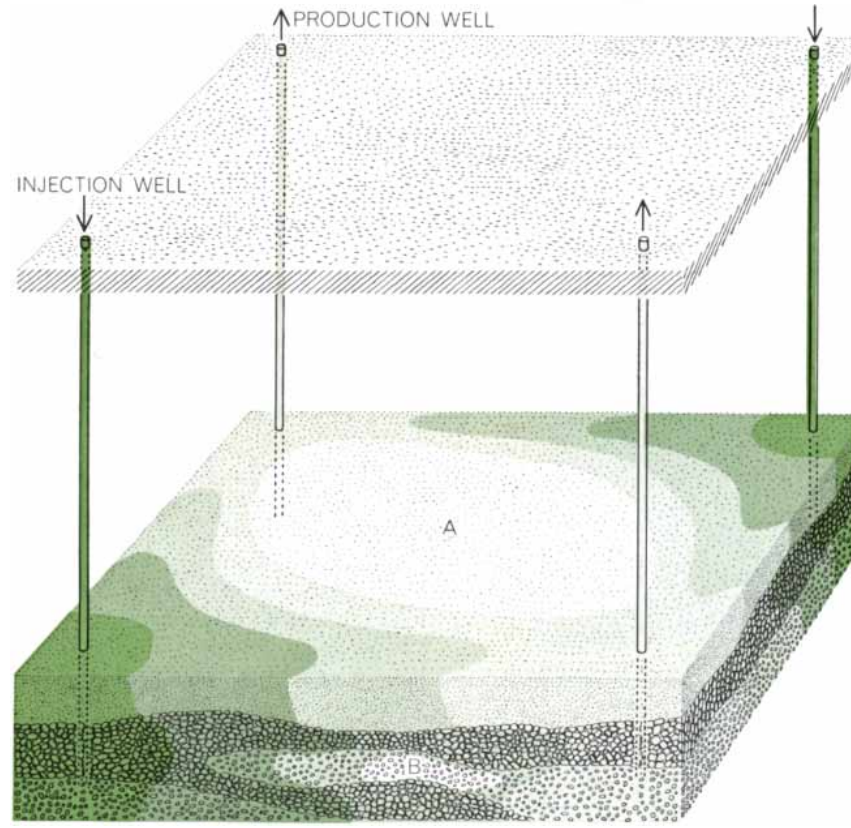
The reason for trying to increase the swept area can be seen if one examines a standard pushing operation, which is carried out by establishing a well pattern resembling a checkerboard. In a typical pattern four injection wells, into which the pushing fluid is inserted, surround one production well, from which the pushed oil emerges. The objective is to have pushing fluid from the injection wells penetrate the oil-bearing rock strata, moving the oil ahead of it into the production well [*see illustration at right*]. If the pushing fluid flowed in a uniform front through all the oil-bearing rock, most of the oil would be removed. In an oil-field situation, however, the pushing fluid bypasses about half of the oil-bearing areas.

This bypassing occurs partly because the various layers of oil-bearing rock were laid down on ocean bottoms during successive geologic periods and differ widely in the size of their constituent sand or limestone grains. As a result they also differ widely in their resistance to the flow of a fluid. One can imagine an oil reservoir containing at its bottom a layer of boulders, on top of the boulders a layer of gravel and on top of the gravel a layer of sand. Most of the fluid from an injection well will go through the layer of boulders because that layer has less resistance to flow than the gravel and sand layers. The fluid thus bypasses much of the oil in these layers. Once the boulder layer has been swept free of oil, most of the water will flow through that layer. Further pumping of water through the field will remove so little of the oil from the sand and gravel layers that the operation becomes uneconomical.

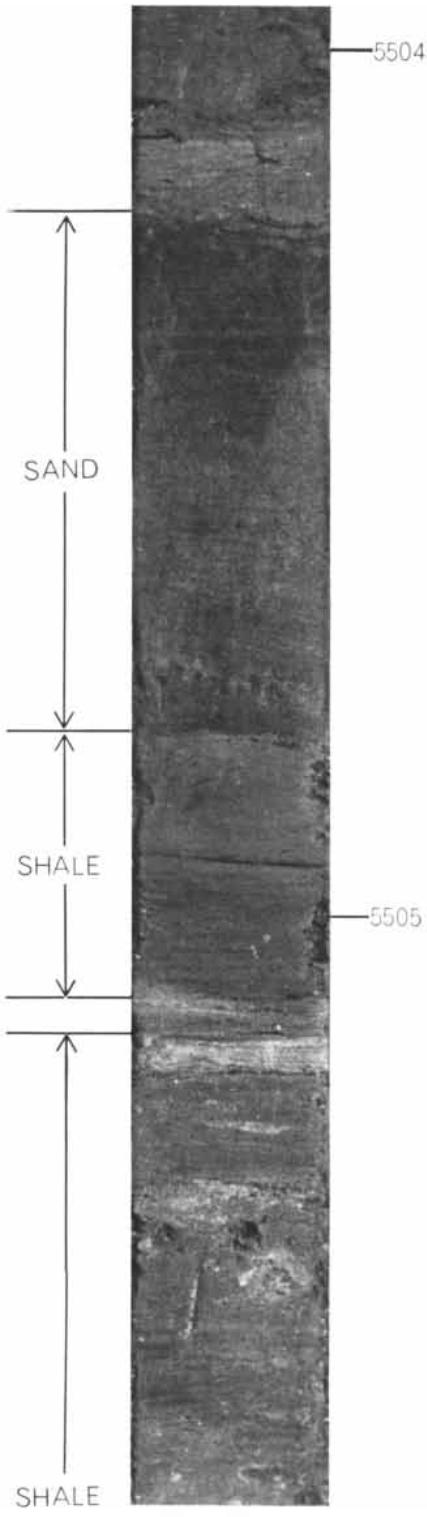
In such a situation it would be desirable to plug the coarser layer after it has been cleared of oil. It is fairly easy to plug with cement any chosen layer adjacent to a production well or an injection one. Unfortunately such sealing off of a layer is effective only in a limited area immediately adjacent to the well. If the boulder layer is plugged, the pushing fluid at first flows into the gravel layer. After traveling beyond the plug, however, the fluid goes back into the boulder layer because of the lower resistance to flow in it. Plugging would thus be successful only if there were a plug that would move slowly through the plugged layer so that the entire layer above it could be swept by the pushing fluid. No such plug has been devised.

A more promising way to enlarge the area of a reservoir swept by a pushing fluid is to increase the viscosity of the pushing fluid. Viscosity is a measure of a fluid's resistance to flow. The viscosity of water is 50 times that of air; the viscosity of molasses, 1,000 times that of water. Crude oils have viscosities ranging from that of water to 1,000 times that of molasses; most have viscosities between those extremes.

The advantage of making the pushing fluid more viscous is shown by considering what would happen if one tried to use water to push axle grease, a highly viscous fluid, out of a tube. If the



WATER FLOODING of oil-bearing rock strata pushes oil (*light color*) through the rock to production wells. Contour lines indicate the yearly advances of the water, which develops a fingering pattern because of its tendency to move through the coarser layers and to take a direct line from injection to production wells. Hence it bypasses some areas, as at *A* and *B*.



OIL-BEARING STRATUM is shown in a sample obtained by coring at a depth of 5,505 feet in an oil field. The core has been cut in half longitudinally. Both the sand and the shale in such a sample contain oil, but only the oil in the sand is recoverable because the shale is so impermeable that the oil cannot flow out of it. About 25 percent of the sand layer consists of void space, and of that space approximately 70 percent is normally occupied by oil and 30 percent by water. Some of the oil cannot be recovered.

tube were originally full of axle grease and water were forced in at one end, pure axle grease would initially emerge at the other end. Soon, however, the water would channel through the grease and break through at the other end. Thereafter the output at the downstream end would be mostly water. If the procedure were reversed, so that axle grease was forced into the end of a tube full of water, nearly all the water would be driven out before the grease appeared at the downstream end. As the example shows, a pushing fluid that is more viscous than the pushed fluid is much less likely to bypass the pushed fluid than if the viscosities are reversed.

There are no known additives that bring about large changes in the viscosity of air or gas, so the idea of increasing the viscosity of the pushing fluid does not apply to air drives or gas drives. It is possible, however, to increase the viscosity of water dramatically with small quantities of various natural and synthetic gums. For example, the addition of 1 percent of methyl cellulose will make water 500 times more viscous. Nonetheless, for such an additive to be useful in secondary recovery it must meet a formidable list of requirements. Among them are low cost and the ability to bring about a large change in viscosity at low concentrations, preferably less than .5 percent. The additive must be much more soluble in water than in oil so that the oil will not steal it from the water. This is a severe problem in the case of natural and synthetic gums that have hydrocarbon subunits. Another requirement is that the additive must not be degraded by being subjected for as long as two years to the moderately high temperatures (up to 200 degrees Fahrenheit) encountered in oil reservoirs. Finally, the additive must not be adsorbed onto the surface of reservoir rocks. Inasmuch as such rocks have an internal surface area of up to one acre per gallon of contained fluid, this is probably the most demanding of all the requirements.

As yet no additive for increasing the viscosity of a pushing fluid is in use on a commercial scale because none has adequately met these requirements. Some additives, however, have shown promise in the laboratory. They are principally the synthetic polymers containing polyacrylamide. The most promising additives are now being subjected to field trials.

The second approach to improving the effectiveness of secondary-recovery processes is to make the pushing fluid do a more thorough job of sweeping oil out of the areas the fluid actually enters. Even under the best of conditions in a laboratory unlimited flushing of an oil-saturated specimen of sandstone with an ordinary pushing fluid will remove only about 70 percent of the oil. The remainder is trapped in nooks and crannies where the pushing fluid cannot get behind it. There are two broad possibilities for improving the performance of a pushing fluid: solvent or chemical techniques, which dissolve oil into the pushing fluid or change its properties to make it easier for the pushing fluid to pick it up, and thermal techniques, which involve heating the oil to reduce its viscosity and thereby make it easier to push.

A solvent technique depends on making the pushing fluid miscible with the reservoir oil. If natural gas is the pushing fluid, it can be made miscible with oil by the addition of liquid petroleum gas—a mixture of ethane, propane and butane that is usually called LPG. In the laboratory this miscible displacement process will sweep a rock clean. In an oil field the process has some disadvantages. One is that some of the LPG, which is expensive, is invariably left behind in the reservoir. In addition the mixture of LPG and crude oil formed in the reservoir is less dense than crude oil and tends to float on it, thereby bypassing some of the reservoir. In spite of its disadvantages miscible displacement is slowly working its way into the reservoir engineer's list of procedures for secondary recovery; a survey this spring found 12 miscible-displacement projects of commercial significance in the U.S. It appears, however, that the process will prove more economical than water flooding and gas driving only in special situations, such as fields in which the oil has a very low viscosity.

Miscible displacement is commercially successful, but it requires large amounts of expensive LPG. Therefore cost-conscious engineers tried using less LPG than was needed to make the pushing gas miscible with oil. The result is called an "enriched-gas drive." An enriched gas does not contain enough LPG to form a one-phase, or miscible, system but does have enough so that when it comes in contact with oil, the oil extracts LPG from it. As a result of absorbing the LPG the oil swells and decreases in viscosity. Both

effects contribute to the increased recovery of oil. Swelling causes part of the oil to be forced out of the nooks and crannies and into the main stream, where the LPG can push it along to the production well. Lower viscosity causes the oil to flow more easily, decreasing the tendency for the LPG to bypass it. Thus an enriched-gas drive can recover more oil than a plain-gas drive but not so much as a miscible-displacement process.

The economics of an enriched-gas drive depends on how much LPG must be left behind in the oil that is never recovered. The loss of LPG can be minimized by inserting the LPG as a "slug" at the beginning of the process and then driving it forward with unenriched natural gas. In this case the LPG left behind in unrecovered oil will be extracted back from the oil into the driving gas because the gas is poor in LPG. The LPG thereby extracted is then carried forward by the gas until it meets oil that is poor in LPG, whereupon the oil extracts LPG from the gas. In this way LPG is reused many times in a cycle.

The slug process is closely analogous to the gas-liquid chromatograph that is widely used for chemical analysis [*see illustrations at right*]. In such an analytic chromatograph a gas, normally helium or hydrogen, drives a sample of some unknown material through a chromatographic column: a thin tube filled with a solid matrix, such as crushed firebrick, that is saturated with a nonvolatile liquid solvent such as silicone oil. The various chemical compounds of the sample migrate at different speeds through the column, the speed of each compound depending on its affinity for the solvent. At the outlet of the column a detecting device records the emergence of each different compound; the chemical composition of the sample is calculated from the signals of the detector.

The analogy should not be pushed too far, because there is a significant difference between gas-liquid chromatography and the use of a slug of LPG to promote the secondary recovery of petroleum. Gas-liquid chromatographs are designed to prevent the movement of the liquid phase. In oil recovery the objective of the operation is to increase the movement of the liquid phase. In addition, this moving-liquid phase makes the mathematics and the experimental technique of an enriched-gas drive substantially more difficult than those of a gas-liquid chromatograph.

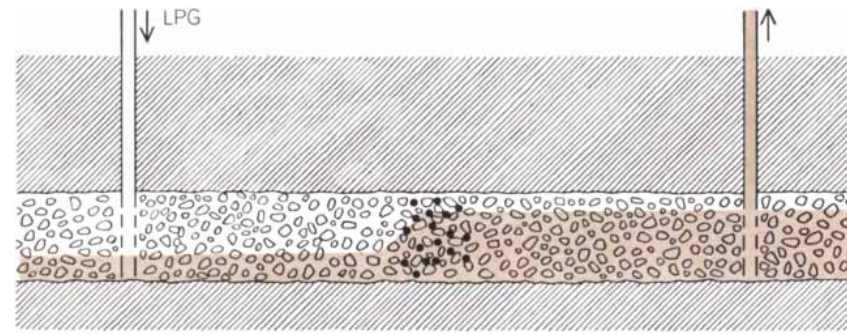
Experiments with enriched-gas drive are under way in the laboratories of most of the large petroleum companies. A few field tests have been started. If an enriched-gas drive has been undertaken on a commercial scale, it is not a matter of public record. Oil companies, being highly competitive, seldom publicize their research work until it has proved successful.

The solvent techniques discussed thus far improve the efficiency of secondary recovery if natural gas is the pushing fluid. What if water is the pushing fluid? There are additives such as acetone that will enable oil and water to form a miscible system, but they would have to be used in prohibitively large concentrations. Oil and water can be emulsified; an example is mayonnaise, which is made by vigorously beating together an oily substance (salad oil) and an aqueous substance (egg yolks and vinegar). No one, however, has found a way to accomplish the neces-

CHROMATOGRAPHY

GAS-LIQUID CHROMATOGRAPH, to which several secondary-recovery processes are analogous, is represented schematically. A gas drives a sample of a substance through a column containing a liquid solvent distributed on a granular solid. Sample's compounds migrate at different speeds according to their affinity for the solvent; the chromatograph thus performs a chemical analysis. Some analogies to secondary recovery are shown below.

ENRICHED-GAS DRIVE



ENRICHED-GAS DRIVE is a secondary-recovery process in which a gas drives a "slug" of liquid petroleum gas (LPG) through a stratum of oil-bearing rock. The LPG is a solvent that collects oil from the rock, and the gas then drives the oil toward a production well.

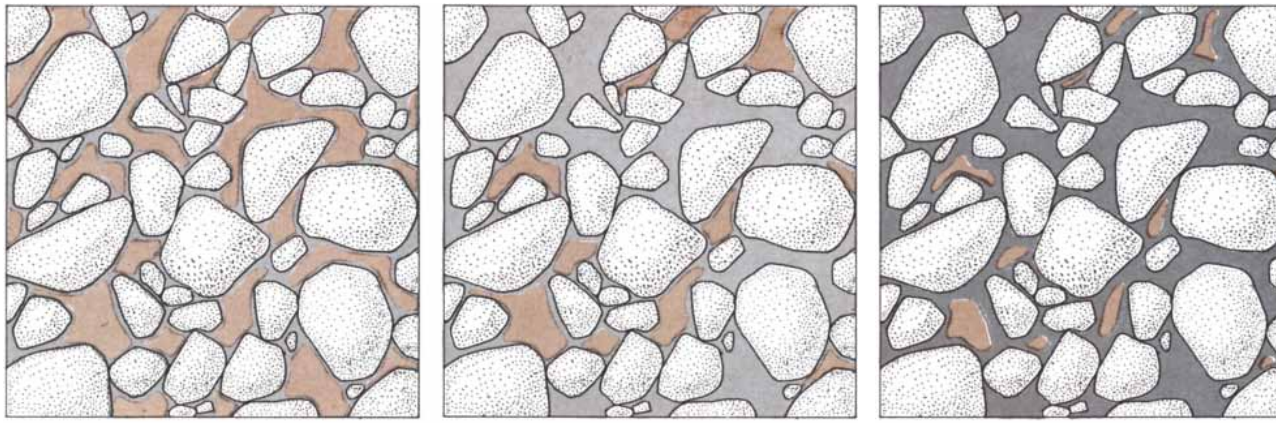
PROCESS	SOLID SUPPORT	LIQUID ADSORBED ON SUPPORT	CARRIER	MATERIAL MOVED BY CHROMATOGRAPHY
GAS-LIQUID CHROMATOGRAPH	CRUSHED FIREBRICK	SILICONE OIL, ETC.	HELIUM OR NITROGEN	SAMPLE
ENRICHED-GAS DRIVE	RESERVOIR ROCK	OIL	NATURAL GAS	LPG
CARBONATED WATER FLOOD	RESERVOIR ROCK	OIL	WATER	CARBON DIOXIDE
UNDERGROUND COMBUSTION	RESERVOIR ROCK		AIR	HEAT PRODUCED BY BURNING OIL

SECONDARY-RECOVERY processes are similar to the gas-liquid chromatograph. A basic difference is that in chromatography the liquid phase is stationary, whereas in a secondary-recovery process the aim is to move the liquid phase, which is the petroleum in the rock.

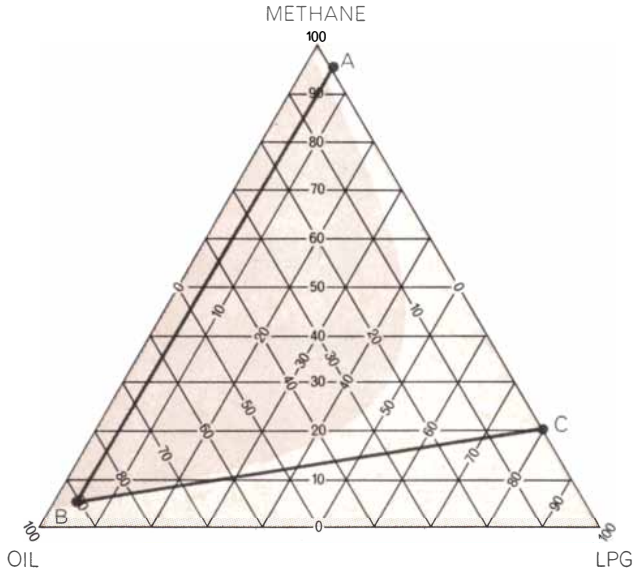
sary beating deep inside a reservoir rock.

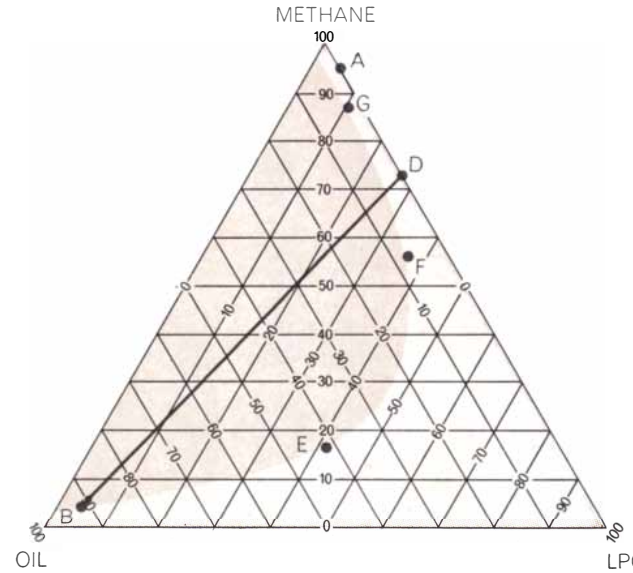
One interesting way to improve the efficiency of water as a pushing fluid is to add to it a chemical that will transfer back and forth from water to oil the way LPG does in an enriched-gas drive. The best additive for this purpose appears to be carbon dioxide. It dissolves to a moderate extent in water, as is apparent in carbonated beverages, and it is highly soluble in most oils. When carbon dioxide is dissolved in an oil, it causes the oil to swell and lowers its viscosity—the same effects produced by LPG in an enriched-gas drive. The secondary-recovery process using this technique is called a carbonated-water flood. Up to now carbonated-water flooding has been confined to laboratory experiments and small-scale field tests. A larger-scale field test is under way in Wyoming, but the results will not be published for several years.

Another possible way to improve the oil-sweeping properties of water is to add small quantities of detergent to the water. In the laboratory a detergent will increase the amount of oil that water can flush out of sandstone, just as a household detergent will remove oily substances from dishes and laundry. To be usable in secondary recovery, however, detergents face the same problems met by the substances employed to increase the viscosity of water: they are absorbed by oils and adsorbed on rock surfaces. After long effort most in-



RECOVERY PROBLEMS are indicated as they would appear if a fragment of oil-bearing sandstone were seen under considerable magnification. At left is the situation before drilling: oil (*color*) and water fill the space between sandstone particles. At center water (*light gray*) has been flooded into the sandstone to push some of the oil out. At right carbon dioxide (*dark gray*), which is soluble in most oils, has been added to the water; it makes oil swell and also reduces its viscosity, thereby aiding oil recovery.





PHASE DIAGRAMS show proportions in which oil, methane and liquid petroleum gas will mix into a single fluid. Mixtures in darker colored region form two phases, gas and oil; those in lighter area form a single phase. At left, for example, in a diagram representing the miscibility of natural gas and oil, point *A* represents a typical natural gas: mostly methane with a little LPG. *B* represents a typical reservoir oil, which contains some methane and LPG. *C* represents a natural gas highly enriched with LPG. Line *AB* shows that most mixtures of gas and oil have compositions that lie in the two-phase region; the gas and oil are not miscible. Line *BC* shows that if a gas is sufficiently enriched with LPG, all compositions obtained by mixing it with oil are in the one-phase region; the oil and gas mix. Phase diagram at right shows an enriched-gas drive system. *A* and *B* are the same as in diagram at left; *D* is a natural gas slightly enriched with LPG. As the gas passes through a sand containing oil the oil extracts LPG from it, producing a composition like that at *E* (for the oil) and *F* (for the gas). Gas that gives up LPG and moves on has composition shown at *G*.

vestigators have concluded that adding detergents to flood water will be commercially impractical until new types of detergents that are not easily adsorbed on rock surfaces become available.

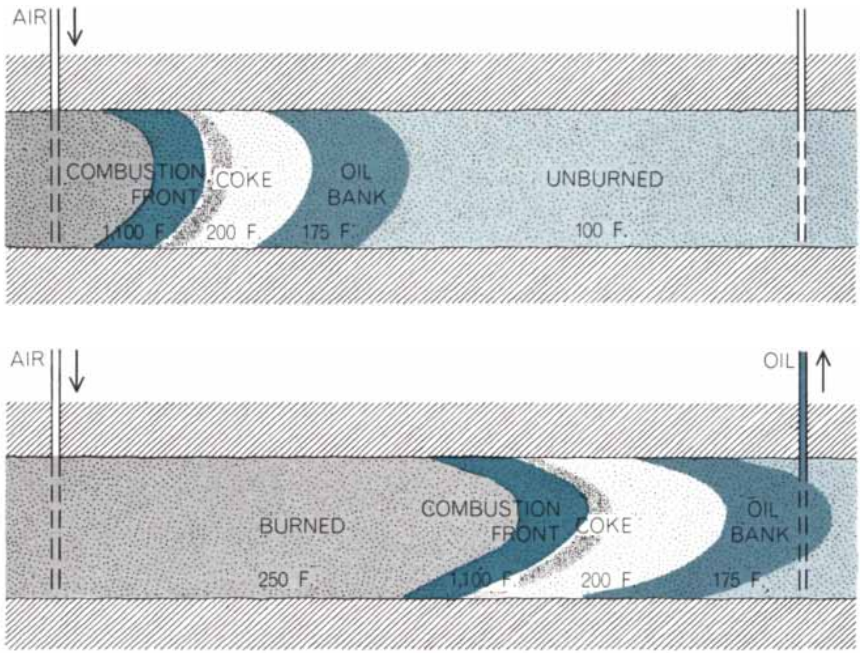
Thermal processes for improving secondary recovery include the interesting scheme of burning some of the oil in a reservoir, a process called *in situ* combustion. This scheme appears wasteful at first glance, but it is not. Every pushing process leaves some oil in the reservoir, never to be recovered. It is reasonable to burn a portion of the unrecoverable oil in order to increase the amount of oil that is recovered.

Burning requires air, both to support combustion and to act as a pushing fluid. Air used to be the least practical of the three possible pushing fluids, because of the explosion hazard resulting from the mixture of air and hydrocarbons in oil-gathering equipment. A burning process converts the pushing air into a hot stream of combustion gases, which will heat the oil. Heating swells the oil and reduces its viscosity, providing a secondary-recovery process akin to that in an enriched-gas drive or a carbonated-water flood. The underground burning also uses up the oxygen in the air, so that there is no explosion hazard in the oil-gathering equipment.

A fire started deep underground in an injection well advances from the well in a roughly circular pattern, moving only a few inches a day. Its fuel is a coke formed from the heavier oil by the advancing flame. The flame is more akin to a glowing charcoal briquet than to the blaze of a campfire [*see top illustration on this page*].

Because the reservoir rock has a much higher capacity for heat storage than the injected air has, the heat from the flame is stored mostly in the rock upstream of the flame. As a result a high-temperature region, much like the LPG slug in an enriched-gas drive, passes through the reservoir. Here, however, the analogy to a gas-liquid chromatograph is stretched because there is no liquid adsorbent and because the material being transported chromatographically is heat.

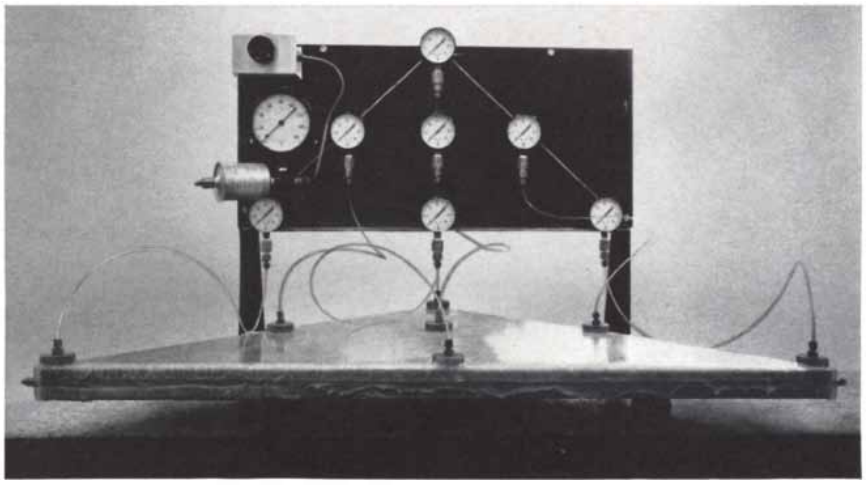
Engineers are even further from a full understanding of what happens in the underground combustion process than they are from a full understanding of the enriched-gas drive or carbonated-water flood processes. In a combustion process the burning and the leakage of heat out of the rock greatly increase the physical and mathematical complexity



***IN SITU* COMBUSTION is a secondary-recovery process that involves burning part of the oil in a reservoir. Air injected through an injection well (*left*) supports combustion and also drives oil through the formation to the production well (*right*). Heat of combustion swells the oil and reduces its viscosity. Combustion converts some of the oil into coke, which is burned. Numbers give approximate temperatures in degrees Fahrenheit. Oil bank is a region of high oil content that undergoes heating. The flame advances a few inches a day; the upper illustration shows the situation after one year, the lower one after five years.**

of the process. As a result of such problems *in situ* combustion remained a laboratory process for many years. Then, beginning in the early 1950's, several oil companies tested it in the field. Now two commercial-scale projects are under way in California. The size of these operations can be appreciated from the fact that the compressors supplying air for one of the projects pump into the reservoirs 43 million cubic feet of air per day. That volume is about equivalent to the amount of natural gas consumed daily by a city of 100,000 people in cold weather.

Considerable research has gone into two other thermal techniques for secondary recovery: (1) using heated water as a pushing fluid; (2) using steam to heat the reservoir and following the

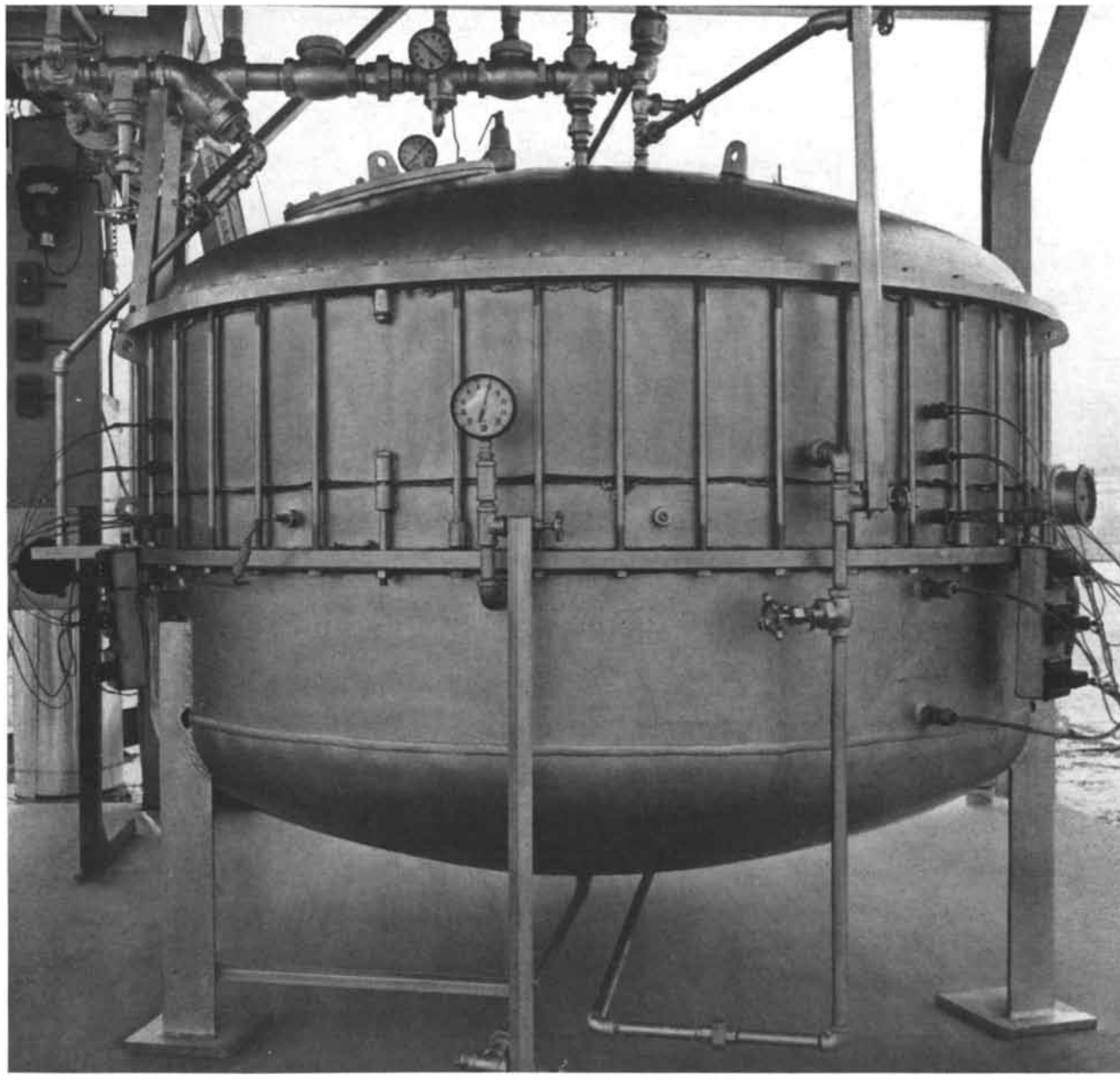


RESERVOIR MODEL at California Research Corporation contains a porous sandstone like that found in a reservoir. The void spaces in the sandstone are filled with oil; various pushing fluids such as propane are injected to see how much of the oil they will force out. Taps at various places on the model measure pressure, which is indicated on the gauges.

steam with water as a pusher. Both processes recover more oil in the laboratory than does pushing with plain water or an inert gas. These methods are now being tested commercially.

The processes I have described as chromatographic-transport modifications of standard secondary-recovery operations are now for the most part at the stage of field trial. At about the time they enter commercial practice on a wide scale techniques using nuclear explosives probably will be under investigation in research laboratories. The tests of nuclear explosives for peaceful purposes conducted under Project Plowshare suggest that the heat from such an explosion might be used in secondary recovery. In addition the mechanical effect of an underground explosion in breaking up rock formations may allow a pushing fluid to flow more uniformly through an oil-bearing rock.

Whatever the technique, secondary recovery is a big business. The Department of the Interior recently reported that in 1950 less than a fifth of the domestic production of crude oil came from secondary recovery and that in 1963 the figure had risen to a third. By 1980, the department estimated, secondary recovery will account for half of the production. Putting the matter another way, the report said: "The United States may be considered to have had on January 1, 1962, recoverable reserves of crude oil in known deposits of at least 31 billion barrels in proved reserves. This figure could probably be increased to 47 billion by the installation of additional secondary-recovery units and possibly to as much as 87 billion with anticipated improvement of newer methods of secondary recovery."



THERMAL PROCESSES of secondary recovery are studied in this tank at the California Research Corporation. The tank is filled with sand. The central portion of the tank simulates an oil-bearing stratum in an oil field. When hot water or steam is injected into this part of the tank, some of the heat reduces the viscosity of the oil, making it more easily pushed out of the sand, and some heat is lost. The apparatus, which stands about eight feet high, is used to measure both the recovery of oil and the extent of heat loss.

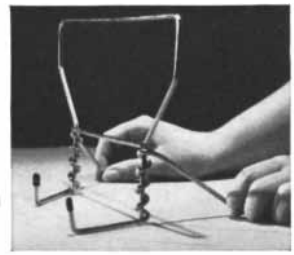
Kodak advertises:

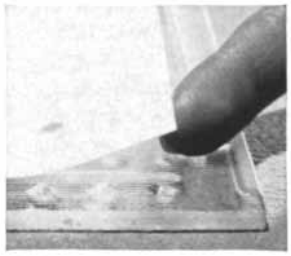
how to make a thin-layer chromatogram . . . getting a project off the ground by lightening the budget . . . color vision instead of brain strain

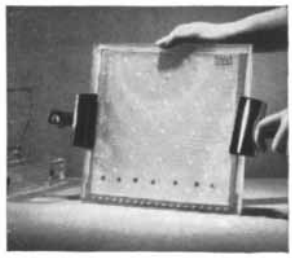
The chemist's lot eases further

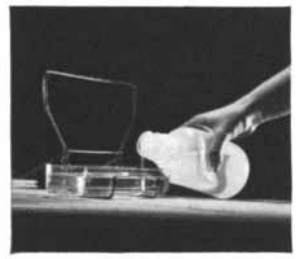
Just before 5 on the afternoon of March 1, 1965, a strong movement toward the ballroom of the Penn-Sheraton was under way. These people must have done their homework by giving careful advance scrutiny to the program of the great Pittsburgh Conference on Analytical Chemistry and Applied Spectroscopy. Within the hour, what we had done to thin-layer chromatography was animating conversation in lines waiting for tables in restaurants all over the Golden Triangle.

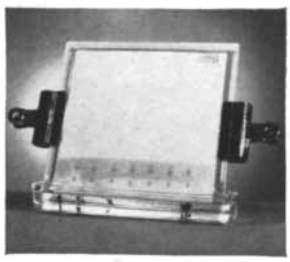
What we had done was to break the news that EASTMAN CHROMAGRAM Sheet had arrived to take the mess out of TLC by providing silica gel properly coated on a snippable base of poly(ethylene terephthalate). What we had failed to do was to provide a minimal-volume, fast equilibrating chamber in which to develop the stuff. This we have put on the market as EASTMAN CHROMAGRAM Developing Apparatus. Place order *now*, along with your order for the sheet. Upon delivery, proceed as follows:











Orders are accepted, prices are quoted, literature and samples of the sheet are supplied by Distillation Products Industries, Rochester, N. Y. 14603 (Division of Eastman Kodak Company).

The sheet can be ordered with or without fluorescent indicator incorporated in the 100μ of polyvinyl-alcohol-bonded silica gel. With it in, a dark spot seen under ultraviolet on the developed chromatogram locates a compound that either has absorption bands in the exciting radiation or quenches the fluorescence photochemically, or a spot brighter than background marks a fluorescence-enhancing compound.

Tourists adore it

This is a 5× enlargement of 8mm movie film in the format in which it has served admirably to bring joy to the grandparents of a significant fraction of the population now 30 years of age or under. Do not scoff at the importance of bringing joy to grandparents.

Right now the most up-to-date families are making their 8mm movies on this format. This spring we introduced a line of KODAK INSTAMATIC Movie Cameras and Projectors which, thanks to the precision with which inexpensive items can now be manufactured in great quantities, can devote far more of the film to picture. The 50% increase in picture area, coupled with today's sharper KODACHROME Film, has markedly upgraded the quality of 8mm movies. This fact deserves the attention of scientific workers who would sooner swallow three pounds of marshmallows than serve as official photographer for a post-kindergarten birthday party but who want to use movies for data-gathering in work where the new "Super 8" quality level and the lightening effect of 8mm on the film budget (as compared with professional-grade costs) might get the project off the ground.

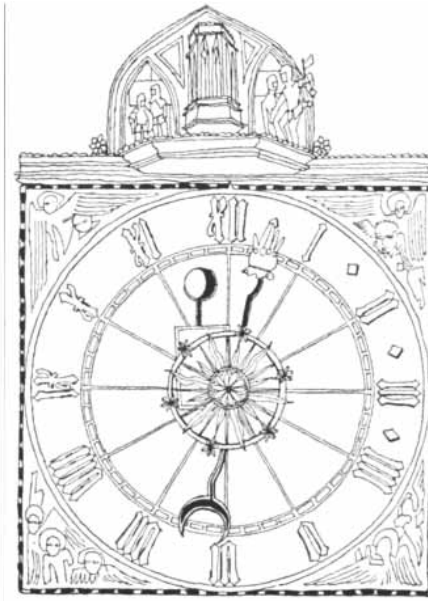
The KODAK INSTAMATIC *M6 Movie Camera does both its viewfinding and its automatic exposure control through the lens, which can also zoom and accept close-up attachments. Drop-in loading. No threading. The whole 50 feet of film in the* KODAPAK *Movie Cartridge can be battery-driven through the film gate (which is in the camera, not the cartridge) at one run, or frames can be exposed as few as one at a time. The M70 projector affords a choice of seven speeds—forward, backward, and zero. Other* INSTAMATIC *camera and projector models are stocked by camera shops at less than half the prices of these. To this writing nobody has yet discovered how to load any of them wrong, but lots of people are working on it.*

When many channels vex the mind

Avoid brain strain. Why waste high-powered deductive reasoning on tasks just as well accomplished through mere color vision? Primates like you generally have use of color vision free of charge. We (Instrumentation Products Division, Eastman Kodak Company, Rochester, N. Y. 14650) can supply name of nearby dealer who now accepts orders for KODAK LINAGRAPH 705 Paper. Comes in widths 3⅜-in. to 12-in., lengths to 400 feet.

Blue light, as through a KODAK WRATTEN Filter No. 34, falling on this paper leaves a red trace; green light, as through WRATTEN No. 12, a blue-green trace; white light (or an overlap of the blue and green light), a nearly neutral black one—after running the paper through an oscillogram processor wherein the four regular baths have been replaced from a KODAK LINAGRAPH 705 Processing Kit that you had better order from the dealer while you are at it.

We mention oscillogram processors. If the place where you work takes much interest in physical phenomena at frequencies above one cycle/sec, there may well be an oscillogram processor around. Take it over. Tell them to have it replaced with a new KODAK EKTALINE Processor, which will get them out of their hole by at least quadrupling the speed at which they handle the paper pouring out of the oscillographs that abound on the premises. Let's hope their channels don't overlap as often as yours.



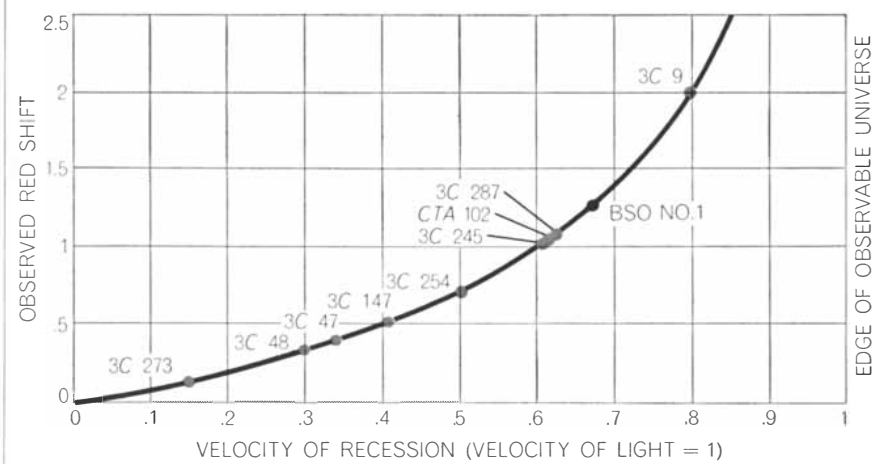
Closer to the Big Bang

The 200-inch telescope on Palomar Mountain continues to provide exciting information about the most distant objects in the universe. And at Holmdel, N.J., an engineering investigation much like the one that initiated radio astronomy 35 years ago has uncovered what may be the first radiation emitted by the newly born universe.

On Palomar Mountain, Maarten Schmidt has determined the velocity of recession of five more quasi-stellar radio sources (quasars), making a total of nine for which velocities are known. All together about 40 quasars have been identified since the first ones were discovered early in 1963. First recognized as intense sources of radio energy, quasars turned out to be among the most luminous and distant objects in the universe. A typical quasar emits as much energy as 100 galaxies.

As part of the general expansion of the universe all distant objects appear to be receding; the more distant the object, the higher its recession velocity. The effect of such recession is to stretch out the wavelength of the object's emitted radiation, producing a "red shift" in its spectrum as recorded on the earth. To determine the red shift one must be able to identify in the object's spectrum the emission lines associated with atoms in some specific state of excitation. Two years ago Schmidt succeeded in making such an identification for the first time in the spectrum of a quasar. He found that the spectral lines of the quasar 3*C* 273 had been red-shifted by 16 percent. This implies that 3*C* 273 is receding at about 15 percent of the velocity of light, which corresponds to a distance of about two billion light-years. Last year Schmidt found that the quasar 3*C* 147 is receding at a velocity of 76,000 miles per second—41 percent of the velocity of light—which made it the most distant object then known. The exact distance associated with such high velocities is uncertain, but it is undoubtedly several billion light-years.

Now Schmidt reports that the quasar 3*C* 9 is receding at a velocity of 149,000 miles per second, or 80 percent of the velocity of light (*see illustration below*). He also finds that four other quasars—3*C* 254, 3*C* 245, *CTA* 102 and 3*C* 287—are receding at velocities between



Recession velocities of nine quasars and blue stellar object (BSO) No. 1

THE CITIZEN

50 and 61 percent that of light. Schmidt estimates that the light now recorded from 3C 9 left it only a few billion years after the universe was born; the universe is estimated to be 10 billion to 15 billion years old. When the light left 3C 9, the universe was only a third as large as it is today.

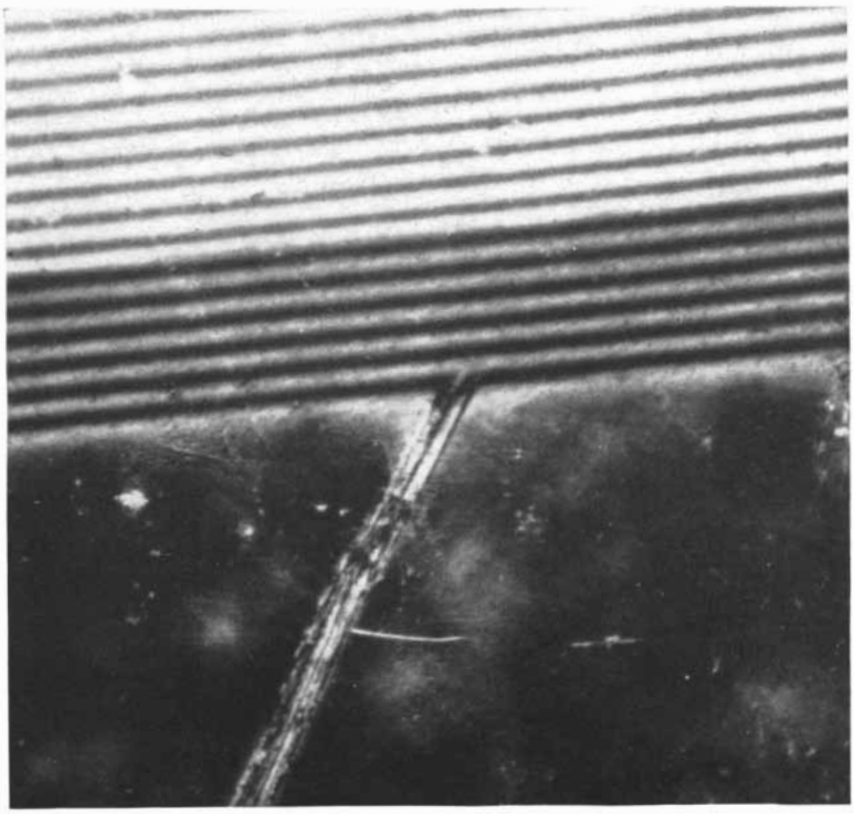
In an independent study Allan R. Sandage has used the 200-inch telescope to identify a new class of objects that resemble quasars but do not emit radio energy. Originally Sandage had called them interlopers; he renamed them quasi-stellar blue galaxies, or blue stellar objects (BSO). Until now astronomers had thought that they were simply blue stars in the outer regions of our galaxy. It turns out from a study of their spectra, in which Schmidt collaborated, that they have red shifts comparable to those seen in the spectra of quasars.

Sandage and Schmidt have now measured the velocities of three of the blue stellar objects. BSO No. 1 is receding at 125,000 miles per second, a velocity that is second only to that of 3C 9. The other two BSO's are receding at 16,000 and 24,000 miles per second.

What makes the BSO's so important is that they appear to be about 500 times more plentiful than quasars. Sandage estimates that there are more than 100,000 BSO's down to the 19th apparent magnitude, a magnitude bright enough to produce good spectra with the 200-inch telescope. "The quasi-stellar blue galaxies are so numerous and reach so far into space," says Sandage, "that we should be able to determine the effects of space curvature and the slowing down of the expansion of the universe within the next few years. ... The clues indicate that our universe is a finite, closed system originating in a 'big bang,' that the expanding universe is slowing down, and that it probably pulsates perhaps once every 80 billion years."

Meanwhile, at the Bell Telephone Laboratories in Holmdel, the first radiation to become disengaged from matter after the big bang may have been detected by Arno A. Penzias and Robert W. Wilson. The evidence for such radiation was found during a series of precision measurements made at a wavelength of 7.3 centimeters with the large horn-reflector antenna used for commu-

Solid-State "Genetics"



nication-satellite experiments. The objective was to account for all known sources of radio noise from the earth and its atmosphere, from space and from the antenna and its receiving equipment. After exhaustive tests Penzias and Wilson found that there still was an unexplained residuum of noise at the wavelength they were studying.

The Bell Laboratories workers subsequently learned that an all-pervasive radiation extending to that wavelength had been predicted in a new hypothesis proposed by R. H. Dicke, P. J. Peebles, P. G. Roll and D. T. Wilkenson of Princeton University. The Princeton group had calculated that if the universe had originated with a big bang, it should have reached a temperature of 10 billion degrees centigrade or more. At such high temperatures radiation would be absorbed almost as fast as it was emitted. Finally, after the temperature had dropped to 100 million degrees, radiation would be "decoupled" from matter and set free to begin a journey that still continues. The Princeton workers had predicted that this primordial radiation would have consisted chiefly of gamma rays and that its wavelength would have been shifted into the centimeter range by its long journey through an expanding universe. They were building receiving equipment to look for it at a wavelength of three centimeters when they heard of the Holmdel discovery. The Princeton group is completing its equipment and will seek to confirm the existence of the predicted radiation.

U.S. Population Problem

The U.S., like many underdeveloped nations, has a population problem. In essence it is that many of the country's poor and uneducated people are deprived of what has come to be considered a basic human right: the right to limit one's family to the number of children wanted, when they are wanted. This is the conclusion of a committee on population of the National Academy of Sciences–National Research Council. The committee reports that the vast majority of Americans approve of and practice family planning. Some 85 to 90 percent of all couples, of all major religions, use some form of contraception. About 20 percent of these couples do so ineffectively, however. And 10 to 12 percent of American couples do not try to limit their families. As in other countries, the burden of excess fertility falls in great disproportion on the poor and the uneducated. It is not that they want more children than other people do; they simply do not have the information or the resources to practice contraception effectively. The high birthrate among the impoverished is not a major threat to the prosperity of a rich country such as the U.S., according to the committee's chairman, William D. McElroy of Johns Hopkins University, "but it is one of the factors that put that prosperity out of the reach of millions of our citizens."

The committee made a number of recommendations to make the "freedom to limit family size" effective throughout U.S. society. In the first place, it urged increased research in all disciplines bearing on population, including the physiology of reproduction and the social, psychological and economic aspects of family planning. Second, the committee called for improved education in fertility control for physicians and for the training of more family-planning administrators. Finally, the committee urged that family planning be considered an integral part of maternal health services in public health programs and individual medical care, as well as a component of the welfare programs of nonmedical public and private agencies.

Cheapest Atomic Power

Great Britain, which has built more nuclear power plants than any other country, has selected a design for a second generation of plants that will produce power that is cheaper than that obtainable from fossil fuels. By the end of this year Britain will have 19 nuclear power reactors in operation, accounting for more than 60 percent of the world's nuclear power capacity.

The design for the second generation of plants, involving a projected total cost of more than one billion dollars, was the subject of intense competition. Seven rival reactor designs were carefully evaluated; four were British versions of American water-cooled reactors and three were advanced versions of the graphite-moderated, gas-cooled type of reactor that Britain had selected for its initial power program. The winning design was a gas-cooled reactor submitted by a consortium known as Atomic Power Constructions Limited, which had not won a contract before. It is estimated that the winning design will produce electricity for about .45 pence (about half a cent) per kilowatt-hour, or slightly below the .5 pence that would make nuclear power competitive with conventional power.

The new reactor will differ in several important ways from the original gas-cooled reactors. Instead of using natural uranium as fuel it will use uranium slightly enriched in the fissionable isotope U-235. The fuel will consist of uranium oxide pellets instead of uranium metal and it will be canned in stainless steel instead of in a magnesium alloy. As a result of these changes the reactor will be able to produce steam matching in temperature and pressure that from the most modern conventional boilers. The net efficiency of the proposed nuclear generating station will be well over 40 percent, surpassing that of any other power plant in the world. The new station, to be built at Dungeness in southeastern England, is scheduled to be operating in 1970; it will produce about 1,200 megawatts of electric power from two reactors.

Enzyme in Three Dimensions

X-ray crystallographers have succeeded for the first time in determining the three-dimensional structure of an enzyme: the protein known as lysozyme. Originally discovered in tears, where it acts as a mild antiseptic, lysozyme has the ability to dissolve the mucopolysaccharides found in the walls of certain bacteria. The determination of its structure has already led to experiments identifying the regions in the lysozyme molecule that appear to be involved in its wall-destroying activity. An account of lysozyme's structure was recently published in *Nature* by C. C. F. Blake, D. F. Koenig, G. A. Mair, A. C. T. North, D. C. Phillips and V. R. Sarma, who completed a project that had been started by R. J. Poljack in 1960. The work was done at the Royal Institution in London.

The method used for determining the structure of lysozyme is the same as that previously used to establish the structure of hemoglobin, the oxygen-carrying molecule in blood, and myoglobin, the oxygen-storing molecule in muscle (see "The Hemoglobin Molecule," by M. F. Perutz; SCIENTIFIC AMERICAN, November, 1964). In brief, X rays are beamed at a single crystal that contains many molecules of the protein in an orderly array. The planes of atoms in the crystal reflect the X rays in various preferred directions depending on the spacing between planes. In some samples of the crystal the protein molecules have been modified to incorporate atoms of a heavy metal—for example mercury or uranium—at known positions. By altering the X-ray pattern in characteristic

ways the metal atoms provide essential information about the spacing of atomic planes.

In the study of the lysozyme molecule more than 9,000 X-ray reflections were measured and fed into a computer, which calculated the density of electrons at a series of reference planes through the molecule. The electron density in turn indicates the location of small groups of atoms and, in some cases, individual atoms. The X-ray study provides a picture of the lysozyme molecule with a resolution of two angstrom units, which is about twice the spacing between carbon and hydrogen atoms in typical organic compounds. The lysozyme molecule contains about 2,200 atoms fitted into a space measuring roughly 45 by 30 by 30 angstroms.

In order to construct a visual model the electron-density values were plotted in the form of contour lines on 60 plastic sheets, representing parallel slices two-thirds of an angstrom apart through the molecule. When the sheets were assembled in a stack, with a half-inch spacing between sheets, the Royal Institution workers had no difficulty tracing out a continuous, intricately folded ribbon of high density—the path of the central chain of the lysozyme molecule. Previous models based on resolutions of six and five angstroms had not revealed the path of this chain unambiguously.

It had been known from independent chemical studies that the molecule is composed of 129 amino acid units strung together. It was also known that the chain must be intricately folded because it is cross-linked by sulfur-sulfur bonds at four places. The sulfur atoms for the bonds are provided by eight widely separated units of the amino acid cysteine. For example, a cysteine at site No. 6 forms a sulfur-sulfur bond with a cysteine at site No. 127. Similar bonds link together sites 30 and 115, 64 and 80, and 76 and 94. Neither myoglobin nor hemoglobin, the only other protein molecules known in comparable detail, contain such linkages. With the two-angstrom model all 129 amino acid units in the lysozyme molecule can be accurately located.

The most fascinating question about an enzyme is how it promotes a specific type of chemical reaction; in the case of lysozyme the reaction involves the splitting (by hydrolysis) of cell-wall mucopolysaccharides. It was known that lysozyme's activity can be inhibited by coupling it with certain sugar-like compounds. The assumption is that these compounds occupy the molecule's active sites, but it is difficult to learn much

about these sites from chemical methods alone.

One of the Royal Institution workers, Phillips, and Louise N. Johnson used X-ray diffraction to study the structure of lysozyme to which various inhibitors had been attached. They found that the inhibiting molecules fit neatly in a prominent cleft in the surface of the lysozyme structure, thus supporting the lock-and-key hypothesis of enzyme action that has long seemed plausible.

Housefly's Knell?

Two substances that cause female houseflies to feed vigorously in clusters have been isolated from yeast and casein by a group of Department of Agriculture Research Service experimenters; the finding may provide another opportunity to eliminate an insect pest by means of a biological control. W. E. Robbins, R. T. Yamamoto, T. J. Shortino and M. J. Thompson, who were studying housefly nutrition and reproduction, noted that the addition of yeast or casein hydrolyzates to a basic laboratory diet brought a sharp increase both in the feeding activity and in the food uptake of the flies. Analyzing the additives, they found the two most active components to be guanosine monophosphate in the yeast and the amino acid leucine in the casein. As little as 10 micrograms of either substance on a bit of filter paper was enough to stimulate vigorous feeding and clustering among caged flies. Curiously 98 percent of the flies in the clusters were females. Why the substances are sex-specific is not yet understood; the investigators suggest a possible correlation with the female cycle of ovarian development.

Neither substance acts as a long-range attractant; the observed clustering is evidently the result of random contact on the part of an individual fly and the subsequent attraction of others as a result of the discoverer's vigorous feeding. Because yeast and protein hydrolyzates in bulk have been successfully used as bait in the eradication of the Mediterranean fruit fly, use of the more potent isolated substances to trap houseflies appears to be a promising possibility.

Geologic Disaster

In October, 1963, a hydroelectric project in the Dolomite Alps near Italy's northeastern frontier was the scene of a freakish disaster that cost more lives than any other dam failure in history. George A. Kiersch of Cornell University has recently summarized in *Civil Engineering* the geological events that were climaxed when a massive rockslide suddenly expelled more than 100 million cubic yards of water from the reservoir behind the Vaiont dam with a resulting death toll of 2,600.

The rocks of the Dolomites are Jurassic and Cretaceous sediments—mainly thick strata of limestone interbedded with clays and marls—that have undergone heavy folding, faulting, uplift and erosion until today they form some of the most spectacular alpine scenery. In Pleistocene times glaciers scoured the Vaiont valley into a smooth trough; thereafter a swift mountain stream cut a 600-to-900-foot gorge in the valley floor. Since its formation the limestone bedrock has been steadily leached away by groundwater, developing sinkholes and other weak areas; the clays and marls are slippery when wet.

The rocks of the valley are plainly unstable; periodic seismic shocks show that adjustment to internal stress and continued mountain uplift are both in progress. One major rockslide had blocked the valley in prehistoric times; in 1960, after the construction of the dam, a lesser slide alarmed engineers to such an extent that they established networks of geodetic stations along the reservoir's steep sides. Observations soon revealed that sections of the southern slope were "creeping" downhill at the rate of one centimeter per week. The rise in the water table as the reservoir filled further lessened stability. Before April, 1963, the water level was kept at 680 meters; in mid-September, after the level was raised by another 20 meters, the rate of creep accelerated to one centimeter per day.

The last week of September and the first week of October were rainy. On October 8 the engineers realized that the rate of creep (which had increased to 10 centimeters per day) did not involve isolated rock zones as they had previously supposed but that instead nearly three square kilometers of hillside was moving. They immediately began to lower the level of the reservoir, but heavy runoff from the rains made this slow work. On October 9 more than 240 million cubic yards of hillside and underlying rock abruptly slumped into the reservoir, filling it in about 60 seconds. Preceded by an atmospheric shock wave, a wall of water jetted over the dam; more than 200 feet high as it advanced, it obliterated downstream farms and villages for miles. Oddly enough neither the main shell nor the abutments of the dam suffered injury, although the combination of shock wave and flooding destroyed all its interior installations, including the powerhouse.

Intelligent Mongoloid

Down's syndrome, or mongolism, is a congenital malformation characterized by mental retardation and certain physical signs; it has been associated since 1959 with specific anomalies of the chromosomes, the structures in the cell nucleus that contain the genetic material. Now investigators at the University of Alabama Medical Center have discovered a patient with the physical but not the mental stigmata of mongolism and with a partial form of one of the typical chromosomal anomalies. Further study of this case and similar ones might locate the genes responsible for the various aspects of mongolism.

The error most commonly seen in the cells of mongoloids is "trisomy": a small chromosome, No. 21 of the 46 in the normal human cell, is present in three copies instead of the usual two (one from each parent). The extra No. 21 material thus implicated in mongolism can also be present in the form of a "translocation": one chromosome No. 21 is attached to a different chromosome, ordinarily one of the group numbered 13 to 15. A "carrier" parent of a translocation mongoloid has only 45 chromosomes, the translocated material being compensated for by the absence of one No. 21 chromosome. When the random sorting of genetic material during germ-cell formation brings two normal No. 21 chromosomes into the fertilized egg along with a translocation chromosome, however, there is an "essential trisomy" of No. 21, and mongolism results.

The patient described (in *The New England Journal of Medicine*) by Sara C. Finley, Wayne H. Finley, Clarence J. Rosecrans and Carey Phillips is a six-year-old boy with the special hand and finger anatomy and facial characteristics of Down's syndrome. Nevertheless, he is alert and articulate, and his IQ—variously scored as 86, 82 and 85—is substantially above the highest previously reported for mongoloid children. Chromosome analysis showed a translocation of what appeared to be the 21st chromosome to a member of the 13-to-15 group. Analysis of cells taken from the patient's normal mother, grandmother and sister, however, showed 46 chromosomes, not 45. There was a tiny fragment, a "partial 21," in addition to the normal No. 21. In other words, only a part of chromosome No. 21 is translocated in this patient and in his carrier relatives; he is therefore trisomic

for only a part of the chromosome. If the Alabama workers could be sure just what part of the chromosome is present in excess, they would know what part of the chromosome is responsible for the physical—as distinct from the mental—signs of Down's syndrome.

Dreamless Sleepwalking

Just before Lady Macbeth comes onstage for her memorable sleepwalking scene one of her ladies-in-waiting says: "Since his Majesty went into the field, I have seen her rise from her bed, throw her nightgown upon her, unlock her closet, take forth paper, fold it, write upon 't, read it, afterwards seal it, and again return to bed; yet all this while in a most fast sleep." In that speech and in the action that follows, when Lady Macbeth talks more or less coherently and tries to wash imaginary blood from her hands, Shakespeare expressed a view still widely held: that sleepwalking is essentially the acting out of a dream. Now four investigators at the University of California at Los Angeles have found that "the long-assumed relation between dreaming and sleepwalking is highly questionable." The investigators—Allan Jacobson, Anthony Kales, Dietrich Lehmann and J. R. Zweizig—describe in *Science* an experiment in which for several nights they obtained continuous electroencephalographic recordings of nine known sleepwalkers. The apparatus was arranged so that the subjects could move without hindrance. The experimenters "arbitrarily defined a sleepwalking incident as anything from sitting upright in bed to more complex behavior, including sleepwalking."

Within this definition the investigators recorded 74 incidents. They found that "all incidents began during periods of slow-wave sleep rather than during periods of rapid eye movement," which is associated with dreaming. Somnambulism appeared to occur separately from dreaming. On the other hand, it did not seem to affect the total amount of time the subjects spent dreaming. In contrast to the kind of behavior exhibited by Lady Macbeth the investigators found such characteristics as these: "The sleepwalkers appeared to be aware of their environment but indifferent to it. Their eyes were open, expressions blank, and movements somewhat rigid.... If spoken to, the subjects answered monosyllabically as if annoyed.... There was complete amnesia for the incidents when they awakened...."

THE PINEAL GLAND

The function of this small organ near the center of the mammalian brain has long been a mystery. Recent studies indicate that it is a "biological clock" that regulates the activity of the sex glands

by Richard J. Wurtman and Julius Axelrod

Buried nearly in the center of the brain in any mammal is a small white structure, shaped somewhat like a pinecone, called the pineal body. In man this organ is roughly a quarter of an inch long and weighs about a tenth of a gram. The function of the pineal body has never been clearly understood. Now that the role of the thymus gland in establishing the body's immunological defenses has been demonstrated, the pineal has become perhaps the last great mystery in the physiology of mammalian organs. This mystery may be nearing a solution: studies conducted within the past few years indicate that the pineal is an intricate and sensitive "biological clock," converting cyclic nervous activity generated by light in the environment into endocrine—that is, hormonal—information. It is not yet certain what physiological processes depend on the pineal clock for cues, but the evidence at hand suggests that the pineal participates in some way in the regulation of the gonads, or sex glands.

A Fourth Neuroendocrine Transducer

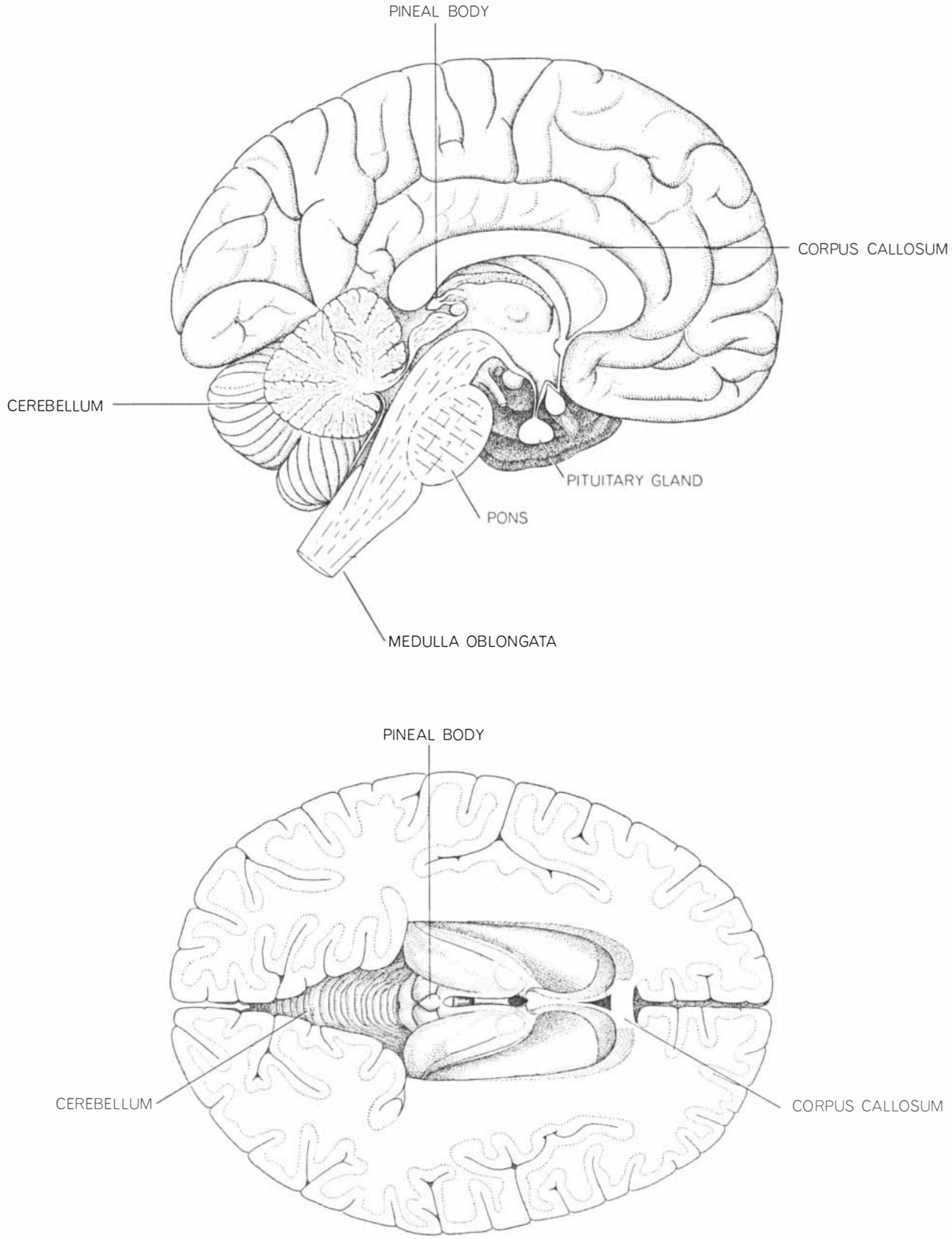
Until quite recently most investigators thought that the mammalian pineal was simply a vestige of a primitive light-sensing organ: the "third eye" found in certain cold-blooded vertebrates such as the frog. Other workers, noting the precocious sexual development of some young boys with pineal tumors, had proposed that in mammals the pineal was a gland. When the standard endocrine tests were applied to determine the possible glandular function of the pineal, however, the results varied so much from experiment to experiment that few positive conclusions seemed justified. Removal of the pineal in young female rats was frequently followed by an enlargement of the ovaries, but the microscopic appearance of the ovaries did not change consistently, and replacement of the extirpated pineal by transplantation seemed to have little or no physiological effect. Most experimental animals could survive the loss of the pineal body with no major change in appearance or function.

In retrospect much of the difficulty early workers had in exploring and defining the glandular function of the pineal arose from limitations in the traditional concept of an endocrine organ. Glands were once thought to be entirely dependent on substances in the bloodstream both for their own control and for their effects on the rest of the body: glands secreted hormones into the blood and were themselves regulated by other hormones, which were delivered to them by the circulation. The secretory activity of a gland was thought to be maintained at a fairly constant level by homeostatic mechanisms: as the level of a particular hormone in the bloodstream rose, the gland invariably responded by decreasing its secretion of that hormone; when the level of the hormone fell, the gland increased its secretion.

In the past two decades this concept of how the endocrine system works has proved inadequate to explain several kinds of glandular response, including changes in hormone secretion brought about by changes in the external environment and also regular cyclic changes in the secretion of certain hormones (for example, the hormones responsible for the menstrual cycle and the steroid hormones that are produced on a daily cycle by the adrenal gland). Out of the realization that these and other endocrine responses must depend in some way on interactions between the glands and the nervous system the new discipline of neuroendocrinology has developed.

In recent years much attention has centered on the problem of locating the nervous structures that participate in the control of glandular function. It has been known for some time that special types of organs would be needed to "transduce" neural information into endocrine information. Nervous tissue is specialized to receive and transmit information directly from cell to cell; according to the traditional view, glands are controlled by substances in the bloodstream and dispatch their messages to target organs by the secretion of hormones into the bloodstream. In order to transmit information from the nervous system to an endocrine organ a hypothetical "neuroendocrine transducer" would require some of the special characteristics of both neural and endocrine tissue. It should respond to substances (called neurohumors) released locally from nerve endings, and it should contain the biochemical machinery necessary for synthesizing a hormone and releasing it into the bloodstream. Three such neurosecretory systems have so far been identified. They are (1) the hypothalamus–posterior-pituitary system, which secretes the antidiuretic hormone and oxytocin, a hormone that causes the uterus to contract during labor; (2) the pituitary-releasing-factor system, also located in the hypothalamus, which secretes polypeptides that control the function of the pituitary gland, and (3) the adrenal medulla, whose cells respond to a nervous input by releasing adrenaline into the bloodstream.

The advent of neuroendocrinology has provided a conceptual framework



TWO VIEWS of the human brain reveal the central position of the pineal body. Section at top is cut in the median sagittal plane and is viewed from the side. Section at bottom is cut in a horizontal plane and is viewed from above; an additional excision has been made in this view to reveal the region immediately surrounding the pineal. In mammals the pineal is the only unpaired midline organ in the brain. The name "pineal" comes from the organ's resemblance to a pinecone, the Latin equivalent of which is *pinea*.

that has been most helpful in characterizing the role of the pineal gland. On the basis of recent studies conducted by the authors and their colleagues at the National Institute of Mental Health, as well as by investigators at other institutions, it now appears that the pineal is not a gland in the traditional sense but is a fourth neuroendocrine transducer; it is a gland that converts a nervous input into a hormonal output.

A Prophetic Formulation

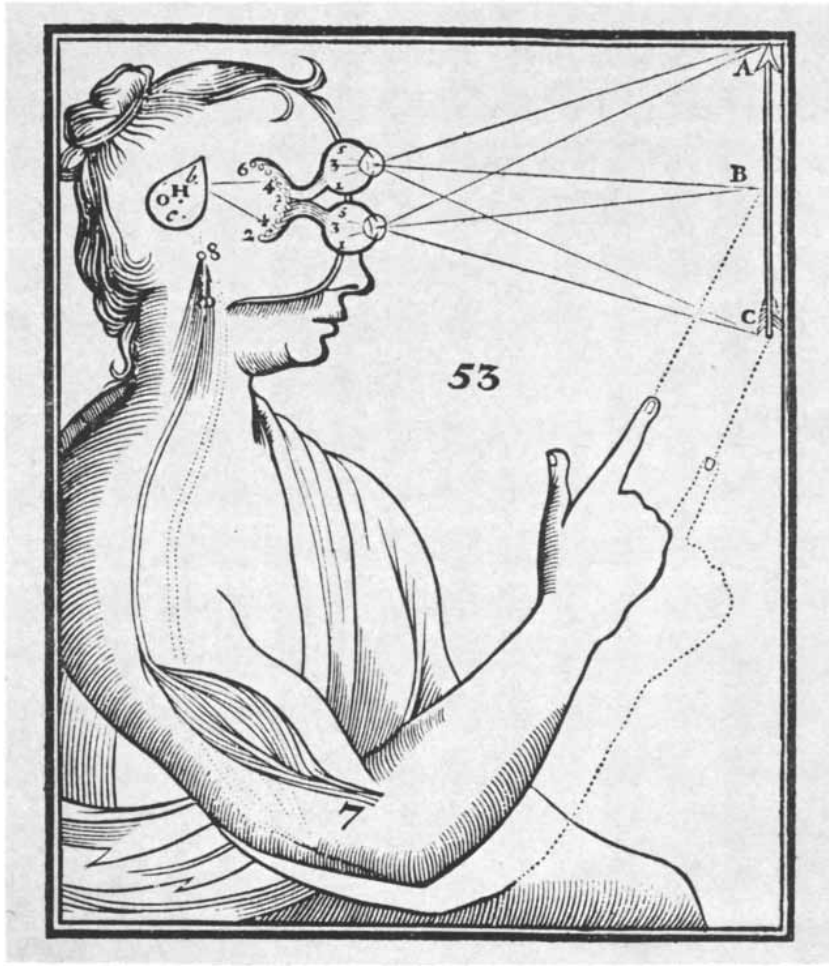
The existence of the pineal body has been known for at least 2,000 years. Galen, writing in the second century A.D., quoted studies of earlier Greek anatomists who were impressed with the fact that the pineal was perched atop the aqueduct of the cerebrum and was a single structure rather than a paired one; he concluded that it served as a valve to regulate the flow of thought out of its "storage bin" in the lateral ventricles of the brain. In the 17th century René Descartes embellished this notion; he believed that the pineal housed the seat of the rational soul. In his formulation the eyes perceived the events of the real world and transmitted what they saw to the pineal by way of "strings" in the brain [*see illustration below*]. The pineal responded by allowing humors to pass down hollow tubes to the muscles, where they produced the appropriate responses. With the hindsight of 300 years of scientific development, we can admire this prophetic formulation of the pineal as a neuroendocrine transducer!

In the late 19th and early 20th centuries the pineal fell from its exalted metaphysical state. In 1898 Otto Heubner, a German physician, published a case report of a young boy who had shown precocious puberty and was also found to have a pineal tumor. In the course of the next 50 years many other children with pineal tumors and precocious sexual development were described, as well as a smaller number of patients whose pineal tumors were associated with delayed sexual development. Inexplicably almost all the cases of precocious puberty were observed in boys.

In a review of the literature on pineal tumors published in 1954 Julian I. Kitay, then a fellow in endocrinology at the Harvard Medical School, found that most of the tumors associated with precocious puberty were not really pineal in origin but either were tumors of supporting tissues or were teratomas (primitive tumors containing many types of cells). The tumors associated with delayed puberty, however, were in most cases true pineal tumors. He concluded that the cases of precocious puberty resulted from reduced pineal function due to disease of the surrounding tissue, whereas delayed sexual development in children with true pineal tumors was a consequence of increased pineal activity.

The association of pineal tumors and sexual malfunction gave rise to hundreds of research projects designed to test the hypothesis that the pineal was a gland whose function was to inhibit the gonads. Little appears to have resulted from these early efforts. Later in 1954 Kitay and Mark D. Altschule, director of internal medicine at McLean Hospital in Waverly, Mass., reviewed the entire world literature on the pineal: some 1,800 references, about half of which dealt with the pineal-gonad question. They concluded that of all the studies published only two or three had used enough experimental animals and adequate controls for their data to be analyzed statistically. These few papers suggested a relation between the pineal and the gonads but did little to characterize it. After puberty the human pineal is hardened by calcification; this change in the appearance of the pineal led many investigators to assume that the organ was without function and further served to discourage research in the field. (Actually calcification appears to be unrelated to the pineal functions we have measured.)

As long ago as 1918 Nils Holmgren, a Swedish anatomist, had examined the pineal region of the frog and the dogfish with a light microscope. He was surprised to find that the pineal contained distinct sensory cells; they bore a marked resemblance to the cone cells of the retina and were in contact with nerve cells. On the basis of these obser-



SEAT OF THE RATIONAL SOUL was the function assigned to the human pineal (*H*) by René Descartes in his mechanistic theory of perception. According to Descartes, the eyes perceived the events of the real world and transmitted what they saw to the pineal by way of "strings" in the brain. The pineal responded by allowing animal humors to pass down hollow tubes to the muscles, where they produced the appropriate responses. The size of the pineal has been exaggerated in this wood engraving, which first appeared in 1677.

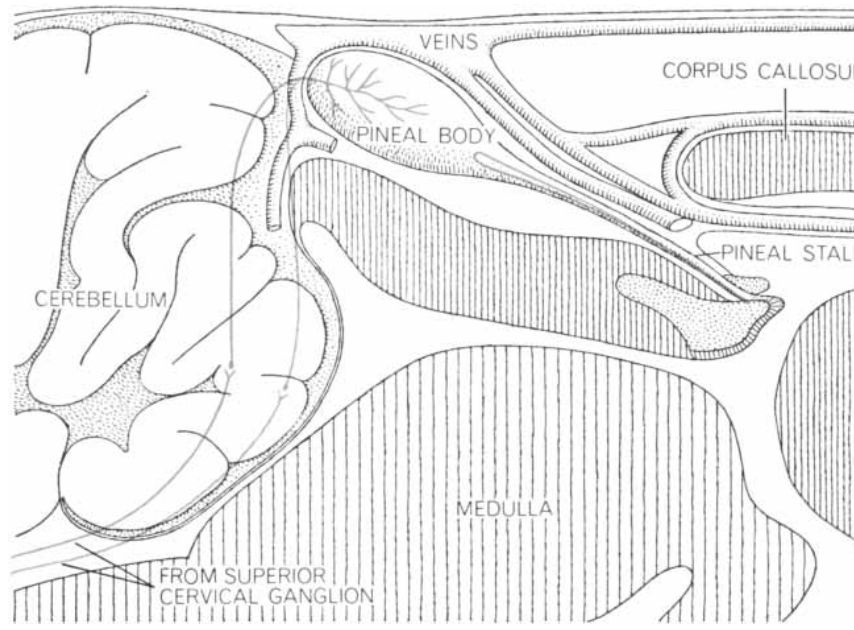
vations he suggested that the pineal might function as a photoreceptor, or "third eye," in cold-blooded vertebrates. In the past five years this hypothesis has finally been confirmed by electrophysiological studies: Eberhardt Dodt and his colleagues in Germany have shown that the frog pineal is a wavelength discriminator: it converts light energy of certain wavelengths into nervous impulses. In 1927 Carey P. McCord and Floyd P. Allen, working at Johns Hopkins University, observed that if they made extracts of cattle pineals and added them to the media in which tadpoles were swimming, the tadpoles' skin blanched, that is, became lighter in color.

Such was the state of knowledge about the pineal as late as five or six years ago. It appeared to be a photoreceptor in the frog, had something to do with sexual function in rats and in humans (at least those with pineal tumors) and contained a factor (at least in cattle) that blanched pigment cells in tadpoles.

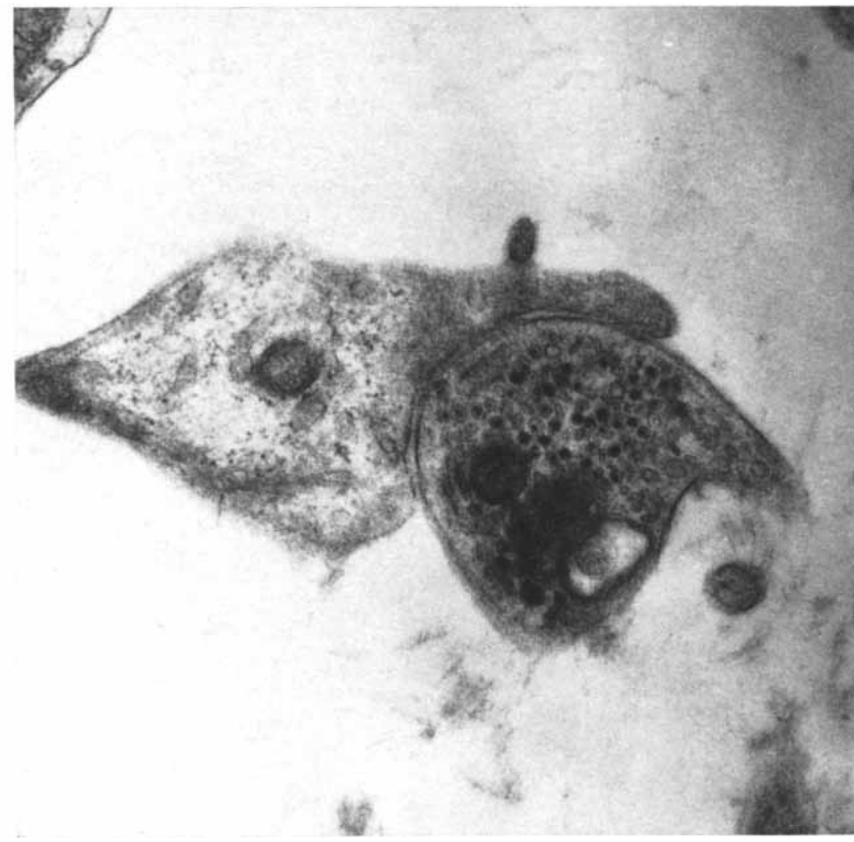
The Discovery of Melatonin

Then in 1958 Aaron B. Lerner and his co-workers at the Yale University School of Medicine identified a unique compound, melatonin, in the pineal gland of cattle [see "Hormones and Skin Color," by Aaron B. Lerner; SCIENTIFIC AMERICAN, July, 1961]. During the next four years at least half a dozen other major discoveries were made about the pineal by investigators representing many different disciplines and institutions. Lerner, a dermatologist and biochemist, was interested in identifying the substance in cattle pineal extracts that blanched frog skin. He and his colleagues prepared and purified extracts from more than 200,000 cattle pineals and tested the ability of the extracts to alter the reflectivity of light by pieces of excised frog skin. After four years of effort they succeeded in isolating and identifying the blanching agent and found that it was a new kind of biological compound: a methoxylated indole, whose biological activity requires a methyl group (CH_3) attached to an oxygen atom [*see illustration on next two pages*].

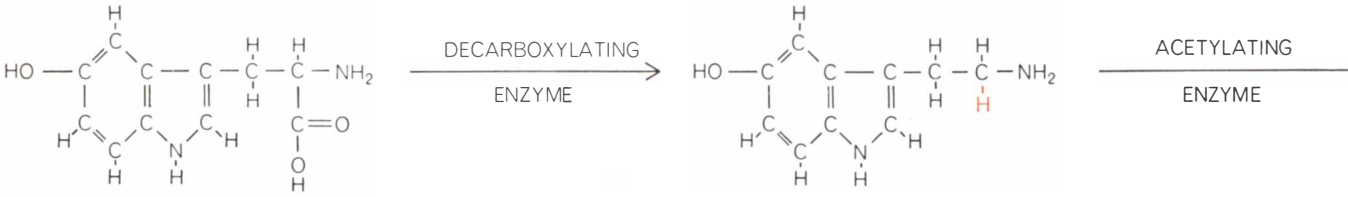
Methoxylation had been noted previously in mammalian tissue, but the products of this reaction had always appeared to lose their biological activity as a result. The new compound, named melatonin for its effect on cells containing the pigment melanin, appeared to lighten the amphibian skin by causing



INNERVATION OF RAT PINEAL was the subject of a meticulous study by the Dutch neuroanatomist Johannes Ariëns Kappers in 1961. He demonstrated that the pineal of the adult rat is extensively innervated by nerves from the sympathetic nervous system. The sympathetic nerves to the pineal originate in the neck in the superior cervical ganglion, enter the skull along the blood vessels and eventually penetrate the pineal at its blunt end (*top*). Aberrant neurons from the central nervous system sometimes run up the pineal stalk from its base, but these generally turn and run back down the stalk again without synapsing. The pineal is surrounded by a network of great veins, into which its secretions probably pass. According to Ariëns Kappers, the innervation of the human pineal is quite similar.



SYMPATHETIC NERVE terminates directly on a pineal cell, instead of on a blood vessel or smooth muscle cell, in this electron micrograph of a portion of a rat pineal made by David Wolfe of the Harvard Medical School. The nerve ending is characterized by dark vesicles, or sacs, that contain neurohumors. Magnification is about 12,500 diameters.



SYNTHESIS OF MELATONIN in the rat pineal begins with the removal of a carboxyl (COOH) group from the amino acid 5-hydroxytryptophan by the enzyme 5-hydroxytryptophan decarboxylase. Serotonin, the product of this reaction, is then enzymatically

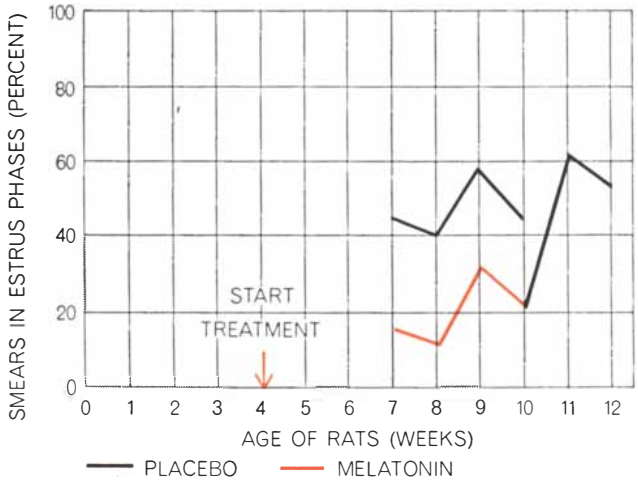
the aggregation of melanin granules within the cells. It was effective in a concentration of only a trillionth of a gram per cubic centimeter of medium. No influence of melatonin could be demonstrated on mammalian pigmentation, nor could the substance actually be identified in amphibians, in which it exerted such a striking effect. It remained a biological enigma that the mammalian pineal should produce a substance that appeared to have no biological activity in mammals but was a potent skin-lightening agent in amphibians, which were unable to produce it!

Both aspects of the foregoing enigma have now been resolved. Subsequent research has shown that melatonin does in fact have a biological effect in mammals and can be produced by amphibians. Spurred by Lerner's discovery of this new indole in the cattle pineal, Nicholas J. Giarman, a pharmacologist at the Yale School of Medicine, analyzed pineal extracts for their content of other biologically active compounds. He found that both cattle and human pineals contained comparatively high levels of serotonin, an amine whose molecular structure is similar to melatonin and whose function in nervous tissue is largely unknown. Studies by other investigators subsequently showed that the rat pineal contains the highest concentration of serotonin yet recorded in any tissue of any species.

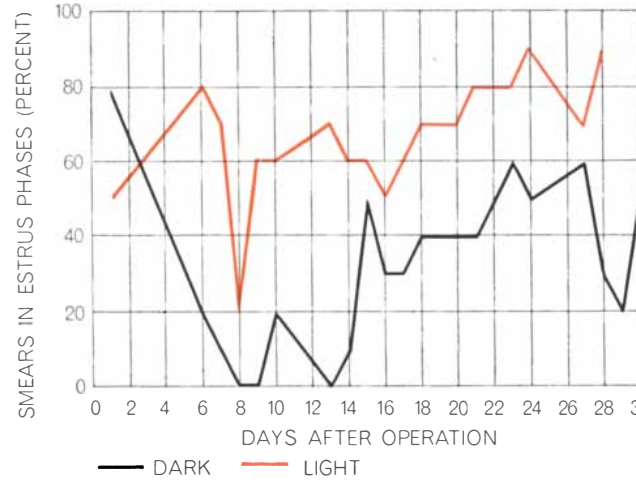
A year before the discovery of melatonin one of the authors (Axelrod) and his co-workers had identified a methoxylating enzyme (catechol-O-methyl transferase) in a number of tissues. This enzyme acted on a variety of catechols (compounds with two adjacent hydroxyl, or OH, groups on a benzene ring) but showed essentially no activity with respect to single-hydroxyl compounds such as serotonin, the most likely precursor of melatonin. In 1959 Axelrod and Herbert Weissbach studied cattle pineal tissue to see if it might have the special enzymatic capacity to methoxylate hydroxyindoles. They incubated N-acetylserotonin (melatonin without the methoxyl group) with pineal tissue and a suitable methyl donor and observed that melatonin was indeed formed. Subsequently they found that all mammalian pineals shared this biochemical property but that no tissue other than pineal could make melatonin. Extensive studies of a variety of mammalian species have confirmed this original observation that only the pineal appears to have the ability to synthesize melatonin. (In amphibians and some birds small amounts of melatonin are also manufactured by the brain and the eye.) Other investigators have found that the pineal contains all the biochemical machinery needed to make melatonin from an amino acid precursor, 5-hydroxytryptophan, which it obtains from the bloodstream. It was also found that circulating melatonin is rapidly metabolized in the liver to form 6-hydroxymelatonin.

Anatomy of the Pineal

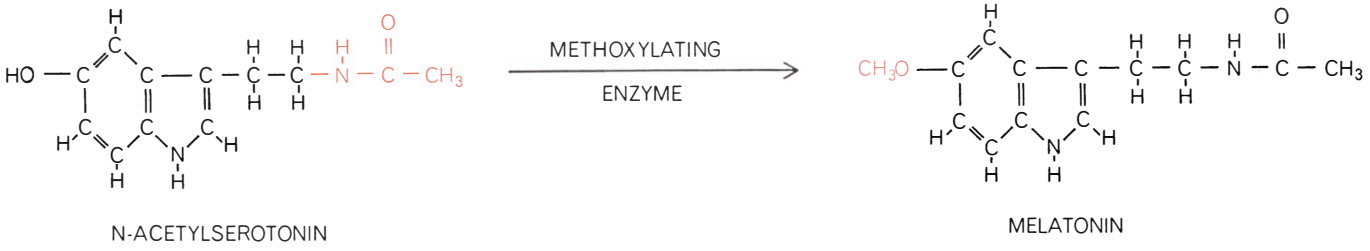
While these investigations of the biochemical properties of the pineal were in progress, important advances were being made in the anatomy of the pineal by the Dutch neuroanatomist Johannes Ariëns Kappers and by several



EFFECT OF MELATONIN on the estrus cycles of female rats is depicted here. Rats that had been given daily injections of melatonin starting in their fourth week of life developed a longer estrus cycle than rats that had been similarly treated with a placebo. When the melatonin-treated animals were 10 weeks old, a placebo was substituted for the melatonin and the estrus cycle returned to normal.



EFFECTS OF LIGHTING on the estrus cycles of three groups of female rats are shown in the graphs on these two pages. The groups, each consisting of about 20 rats, were subjected respectively to a sham operation (*left*), removal of their superior cervical ganglion (*middle*) and removal of their eyes (*right*). Each group was then further subdivided, with about half being placed in constant light



acetylated to form N-acetylserotonin. This compound in turn is methoxylated by the enzyme hydroxyindole-O-methyl transferase (HIOMT) to yield melatonin. In mammals HIOMT is found only in the pineal. Changes in basic molecule are indicated by color.

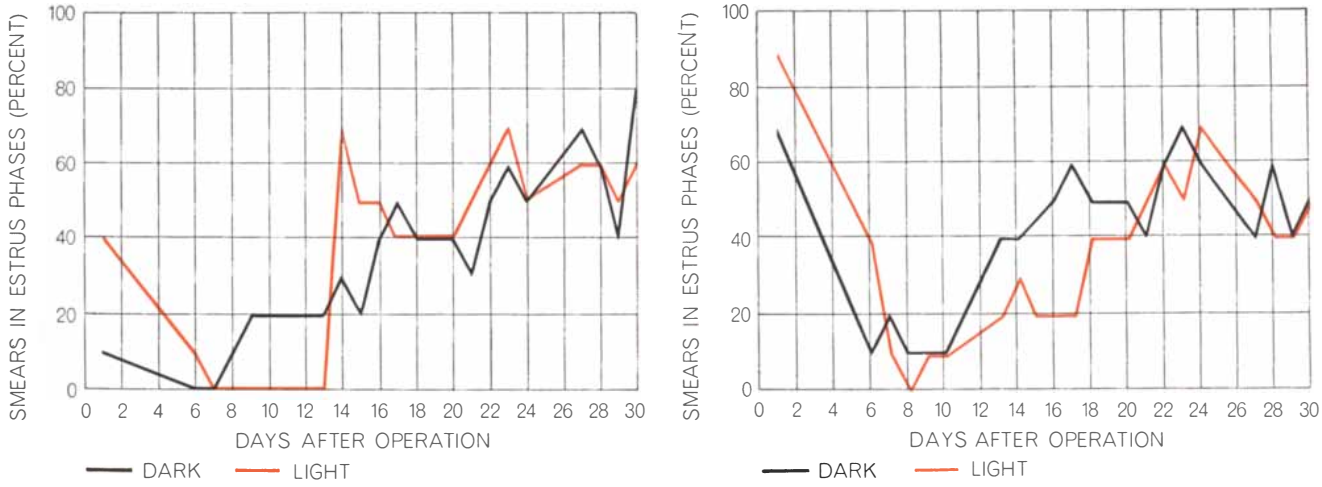
American electron microscopists, including Douglas E. Kelly of the University of Washington, Aaron Milofsky of the Yale School of Medicine and David Wolfe, then at the National Institute of Neurologic Diseases and Blindness. In 1961 Ariëns Kappers published a meticulous study of the nerve connections in the rat pineal. He demonstrated clearly that although this organ originates in the brain in the development of the embryo, it loses all nerve connections with the brain soon after birth. There is thus no anatomical basis for invoking "tracts from the brain" as the pathway by which neural information is delivered to the pineal.

Ariëns Kappers showed that instead the pineal of the adult rat is extensively penetrated by nerves from the sympathetic portion of the autonomic nervous system. The sympathetic nervous system is involuntary and is concerned with adapting to rapid changes in the internal and external environments; the sympathetic nerves to the pineal originate in the superior cervical ganglion in the neck, enter the skull along the blood vessels and eventually penetrate the pineal [*see top illustration on page 53*]. Electron microscope studies later showed that within the pineal many sympathetic nerve endings actually terminate directly on the pineal cells, instead of on blood vessels or smooth-muscle cells, as in most other organs [*see bottom illustration on page 53*]. Among endocrine structures the organization of nerves in the mammalian pineal appeared to be most analogous to that of the adrenal medulla, one of the three demonstrated neuroendocrine transducers.

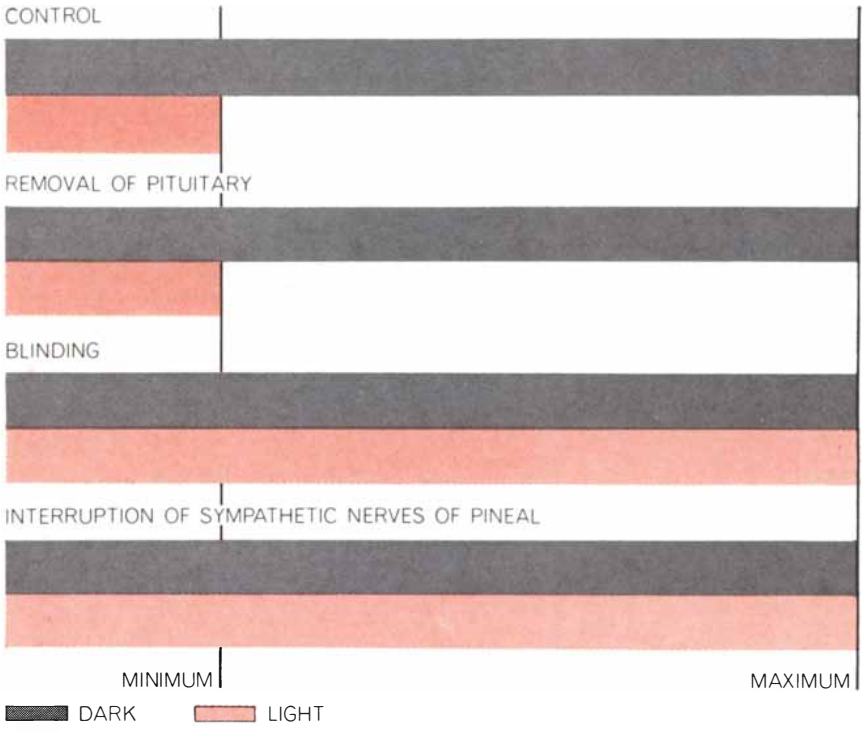
Meanwhile electron microscope studies by other workers on the pineal regions of frogs had confirmed many of Holmgren's speculations. It was found that the pineal cells of amphibians contained light-sensitive elements that were practically indistinguishable from those found in the cone cells of the retina, but that the pineal cells of mammals did not contain such elements. By 1962 it could be stated with some assurance that the mammalian pineal was not simply a vestige of the frog "third eye," since the "vestige" had undergone profound anatomical changes with evolution.

The Melatonin Hypothesis

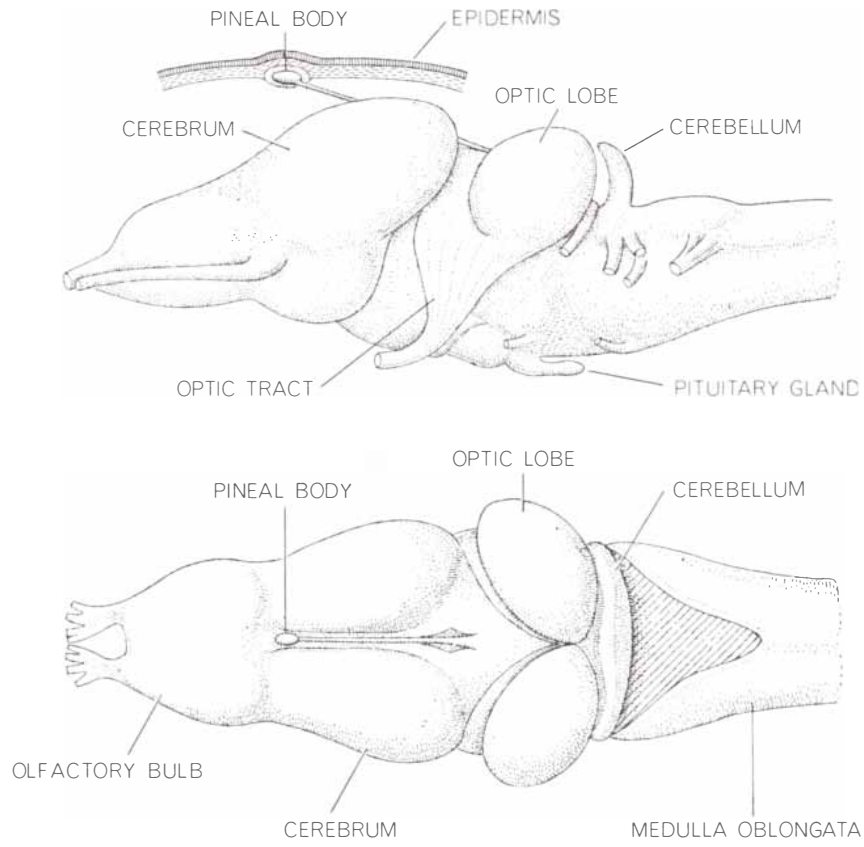
Even though the mammalian pineal no longer seemed to respond directly to light, there now appeared good evidence that its function continued to be related somehow to environmental light. In 1961 Virginia Fiske, working at Wellesley College, reported that the exposure of rats to continuous environmental illumination for several weeks brought about a decrease in the weight of their pineals. She had been interested in studying the mechanisms by which the exposure of rats to light for long periods induces changes in the function of their gonads. (For example, continous light increased the weight of the ovaries and accelerated the estrus cycle). At the same time one of the authors (Wurtman, then at the Harvard Medical School), in collaboration with Altschule and Willard Roth, was studying the conditions under which the administration



and the other half in constant darkness beginning one day after their respective operations. Daily vaginal smears were taken on the first day and on the sixth through the 30th days after the operations. Results were plotted as the percentage of all the smears in a treatment group showing estrus phases each day. In general it was found that interference with the transmission of light information to the pineal gland (either by blinding or by cutting the sympathetic nerves) also abolished most of the gonadal response to light. These findings supported the authors' melatonin hypothesis, which holds that one mechanism whereby light is able to accelerate the estrus cycle in normal animals is by inhibiting the synthesis in the pineal of melatonin, a compound that in turn inhibits estrus.



RESPONSE OF MELATONIN-FORMING ENZYME hydroxyindole-O-methyl transferase (HIOMT) to continuous light or darkness is shown under four different circumstances. In the control, or normal, animal continuous darkness induces an increase in HIOMT activity, whereas exposing the animal to continuous light has the opposite effect. The ability of the pineal gland to respond to environmental lighting is unaffected by the removal of the pituitary gland but is abolished following blinding or sympathetic denervation of the pineal.



PINEAL EYE is a primitive photoreceptive organ found in certain cold-blooded vertebrates such as the frog. Frog's brain is shown from the side (*top*) and from above (*bottom*).

of cattle pineal extracts decreased ovary weight and slowed the estrus cycle.

We soon confirmed Mrs. Fiske's findings, and we were also able to show that the exposure of female rats to continuous light or the removal of their pineals had similar, but not additive, effects on the weight of their ovaries. These experiments suggested that perhaps one way in which light stimulates ovary function in rats is by inhibiting the action of an inhibitor found in pineal extracts. It now became crucial to identify the gonad-inhibiting substance in pineal extracts and to see if its synthesis or its actions were modified by environmental lighting.

In 1962 we began to work together on isolating the anti-gonadal substance present in pineal extracts. Our plan was to subject extracts of cattle pineal glands to successive purification steps and test the purified material for its ability to block the induction by light of an accelerated estrus cycle in the rat. Before undertaking the complicated and time-consuming procedure of isolating the active substance in the pineal glands of cattle, we first tested a mixture of all the constituents that had already been identified in this tissue. The mixture was found to block the effects of light on the estrus cycle.

Next we tested melatonin alone, since it was apparently the only compound produced uniquely by the pineal. To our good fortune we found that when rats were given tiny doses (one to 10 micrograms per day) of melatonin by injection, starting before puberty and continuing for a month thereafter, the estrus cycle was slowed and the ovaries lost weight—just as though the animals had been treated with pineal extracts. In later studies we found that this effect of melatonin was chemically specific: it was simulated by neither N-acetylserotonin, the immediate precursor of melatonin, nor 6-hydroxymelatonin, the major product of its metabolism. Moreover, it was possible to accelerate the estrus cycle by removal of the pineal and to block this response by the injection of melatonin.

On the basis of these studies, performed in collaboration with Elizabeth Chu of the National Cancer Institute, we postulated that melatonin was a mammalian hormone, since it is produced uniquely by a single gland (the pineal), is secreted into the bloodstream and has an effect on a distant target organ (the vagina and possibly also the ovaries). We were not able to identify the precise site of action of melatonin in affecting the gonads. The

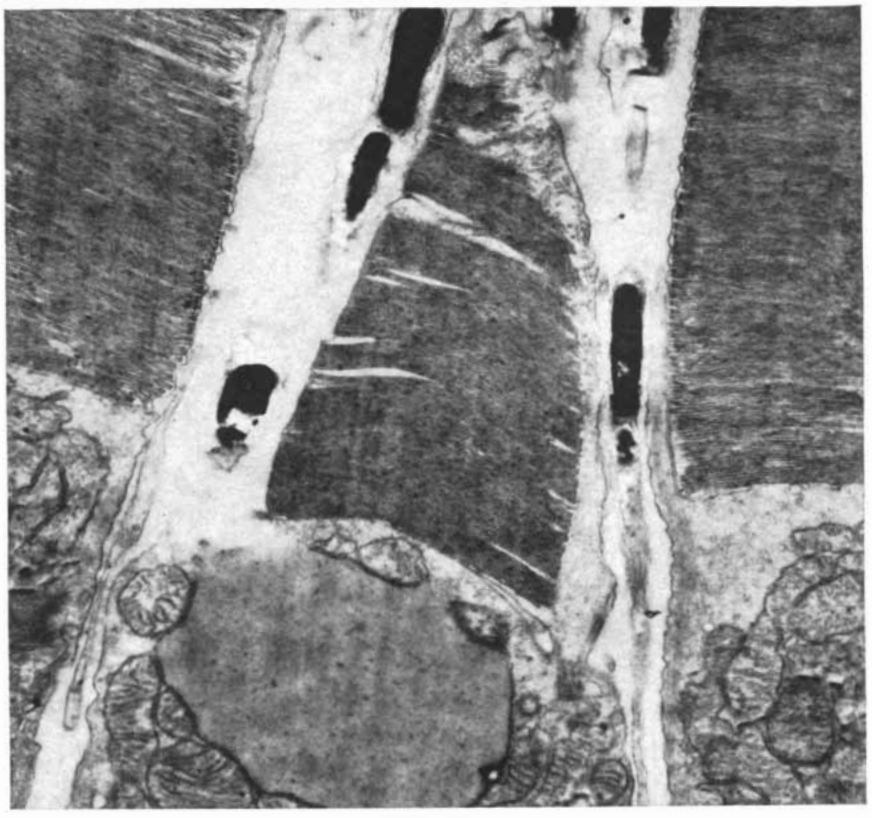
slowing of the estrus cycle could be produced by actions at any of several sites in the neuroendocrine apparatus, including the brain, the pituitary, the ovaries or the vagina itself. When melatonin was labeled with radioactive atoms and injected into cats, it was taken up by all these organs and was selectively concentrated by the ovaries.

William M. McIsaac and his colleagues at the Cleveland Clinic have confirmed the effects of melatonin on the estrus cycle and have identified another pineal methoxyindole—methoxytryptophol—that has similar effects. It appears likely that pineal extracts contain a family of hormones: the methoxyindoles, all of which have in common the fact that they can be synthesized by the methoxylating enzyme found only in the mammalian pineal.

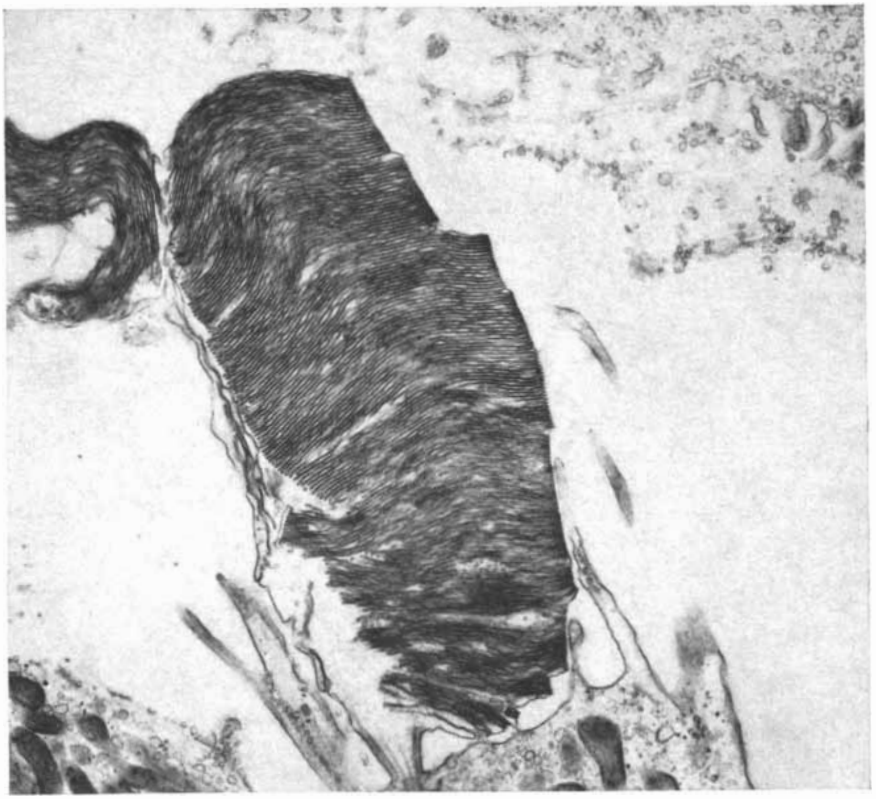
We next set out to determine whether or not these effects of injected melatonin were physiological. Could the rat pineal synthesize melatonin and, if so, in what quantities? When rat pineal glands were examined for their ability to make melatonin, we were disappointed to find that the activity of the melatonin-forming enzyme (hydroxyindole-O-methyl transferase, or HIOMT) in the rat was much lower than in most other species; the maximum amount of melatonin that the rat could make was probably on the order of one microgram per day. Our disappointment was soon relieved, however, when we realized that the low activity of this enzyme made it likely that it was controlling the rate-limiting step in melatonin synthesis in the intact animal. Knowing that continuous exposure to light decreased pineal weight, as well as the amount of ribonucleic acid (RNA) and protein in the pineal, we next explored what effect illumination might have on HIOMT activity and thus on melatonin synthesis.

Since the rat pineal gland was so small (about a milligram in weight) and had so little enzymatic activity, it was necessary to devise extremely sensitive techniques to measure this activity. When rats were subjected to constant light for as short a period as a day or two, the rate of melatonin synthesis in their pineals fell to as little as a fifth that of animals kept in continuous darkness. Since this effect of illumination or its absence could be blocked by agents that interfered with protein synthesis, it appeared that light was actually influencing the rate of formation of the enzyme protein itself.

How was information about the state of lighting being transmitted to the rat pineal? Three possible routes suggested



RETINAL CONE CELL from the eye of an adult frog is shown in this electron micrograph made by Douglas E. Kelly of the University of Washington. The photoreceptive outer segment of the cell (*top center*) consists of a densely lamellated membrane. Parts of two larger rod photoreceptors can be seen on each side of cone. Magnification is about 13,000 diameters.



PINEAL CONE CELL from the pineal eye of an adult frog is shown in this electron micrograph made by Kelly at approximately the same magnification as the micrograph at top. The lamellated outer segment of the pineal cell is practically indistinguishable from that of the retinal photoreceptor. Part of the membrane has torn away from the cell (*top left*).

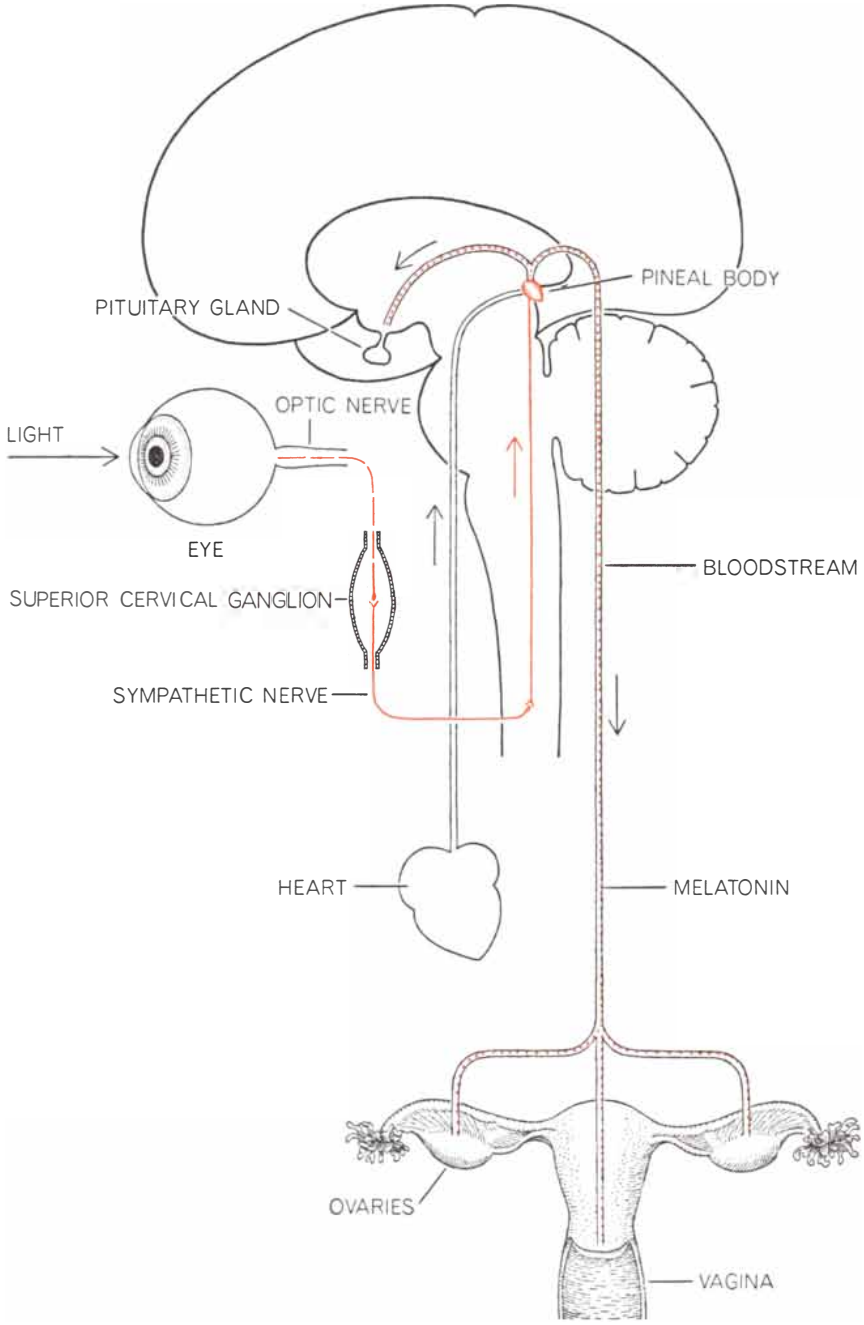
themselves. The first was that light penetrated the skull and acted directly on the pineal; W. F. Ganong and his colleagues at the University of California at Berkeley had already shown that significant quantities of light do penetrate the skulls of mammals. This hypothesis was ruled out, however, by demonstrating that blinded rats completely lost the capacity to respond to light with changed HIOMT activity; hence light had to be perceived first by the retina and was not acting directly on the pineal.

The second possibility was that light altered the level of a circulating hormone, perhaps by affecting the pituitary gland, and that this hormone secondarily influenced enzyme activity in the pineal gland. This hypothesis was also ruled out by demonstrating that the removal of various endocrine organs, including the pituitary and the ovaries, did not interfere with the response of pineal HIOMT to light.

The third possibility was that information about lighting was transmitted to the pineal by nerves. Fortunately Ariëns Kappers had just identified the nerve connections of the rat pineal as coming from the sympathetic nervous system. We found that if the sympathetic pathway to the pineal was interrupted by the removal of the superior cervical ganglion, the ability of melatonin-forming activity to be altered by light was completely lost. Thus it appeared that light was stimulating the retina and then information about this light was being transmitted to the pineal via sympathetic nerves. Within the pineal the sympathetic nerves probably released neurohumors (noradrenaline or serotonin), which acted on pineal cells to induce (or block the induction of) HIOMT; this enzyme in turn regulated the synthesis of melatonin.

Since one way light influences the gonads is by changing the amount of melatonin secreted from the pineal, we reasoned that the effects of light on the gonads might be blocked if the transmission of information about light to the pineal were interrupted. This could be accomplished by cutting the sympathetic nerves to the pineal—a procedure much less traumatic than the removal of the pineal itself. To test this hypothesis we placed groups of rats whose pineals had been denervated along with blinded and untreated animals in continuous light or darkness for a month. Vaginal smears were checked daily for evidence of changes in the estrus cycle, and pineals were tested for melatonin-synthesizing ability at the end of the experiment. It was found that interrupting the transmission of light information to the pineal (by cutting its sympathetic nerves—a procedure that does not interfere with the visual response to light) also abolished most of the gonadal response to light.

Incidentally, the observation that sympathetic nerves control enzyme synthesis in the pineal has provided, and should continue to provide, a useful tool for studies in a number of other biological disciplines. For example, studying the changes in brain enzymes produced by environmental factors offers a useful method for tracing the anatomy



SUGGESTED PATHWAY by which light influences the estrus cycle in the rat is depicted in this schematic diagram. Light stimuli impinge on the retinas and cause a change in the neural output of the superior cervical ganglion by way of an unknown route. This information is then carried by sympathetic nerves to the pineal gland, where it causes a decrease in the activity of HIOMT and in the synthesis and release of melatonin. This decrease in turn lessens the inhibiting effect of the circulating melatonin on the rate of the estrus cycle. The precise site of action of melatonin in influencing the gonads is unknown; the slowing of the estrus cycle could be produced by actions at any one of several sites in the neuroendocrine apparatus, including the brain, the pituitary, the ovaries and the vagina.

of the nerve tracts involved. The observation that the activity of at least one part of the sympathetic nervous system (the superior cervical ganglion) is affected by environmental lighting raises the possibility that other regions of this neural apparatus are affected similarly. If so, physiological studies of the effects of light on other sympathetically innervated structures (for example the kidneys and fat tissue) may be profitable.

We have also found that light influences the serotonin-forming enzyme in the pineal gland but not in other organs. In contrast to HIOMT, the activity of this enzyme increases when rats are kept in constant light and decreases in darkness. When rats are blinded or when the sympathetic nerves to the pineal are cut, the effect of light and darkness on the serotonin-forming enzyme is also extinguished. Furthermore, certain drugs that block the transmission of sympathetic nervous impulses also abolish the effect of illumination on this enzyme. The fact that lighting influences pineal weight and at least two enzyme systems in this organ suggests that it may regulate many additional, undiscovered biochemical events in the pineal, via the sympathetic nervous system.

Diurnal and Circadian Rhythms

The pineal had been shown to respond and function under quite unusual conditions; for example, when an experimental animal was exposed to continuous light or darkness for several days. In nature, of course, animals that live in the temperate and tropical zones are rarely subjected to such conditions. It became important to determine if the pineal could also respond to naturally occurring changes in the environment.

In nature the level of light exposure changes with both diurnal and annual cycles. Except in polar regions every 24-hour day includes a period of sunlight and a period of darkness; the ratio of day to night varies with an annual rhythm that reaches its nadir at the winter solstice and its zenith on the first day of summer. Lighting cycles have been shown to be important in regulating several types of endocrine function: the increase in sunlight during the winter and spring triggers the annual gonadal growth and breeding cycles in many birds and some mammals that breed yearly, and the daily rhythm of day and night synchronizes a variety of roughly daily rhythms in mammals, such as the cycle of adrenal-steroid secretion. Such rhythms are called cir-

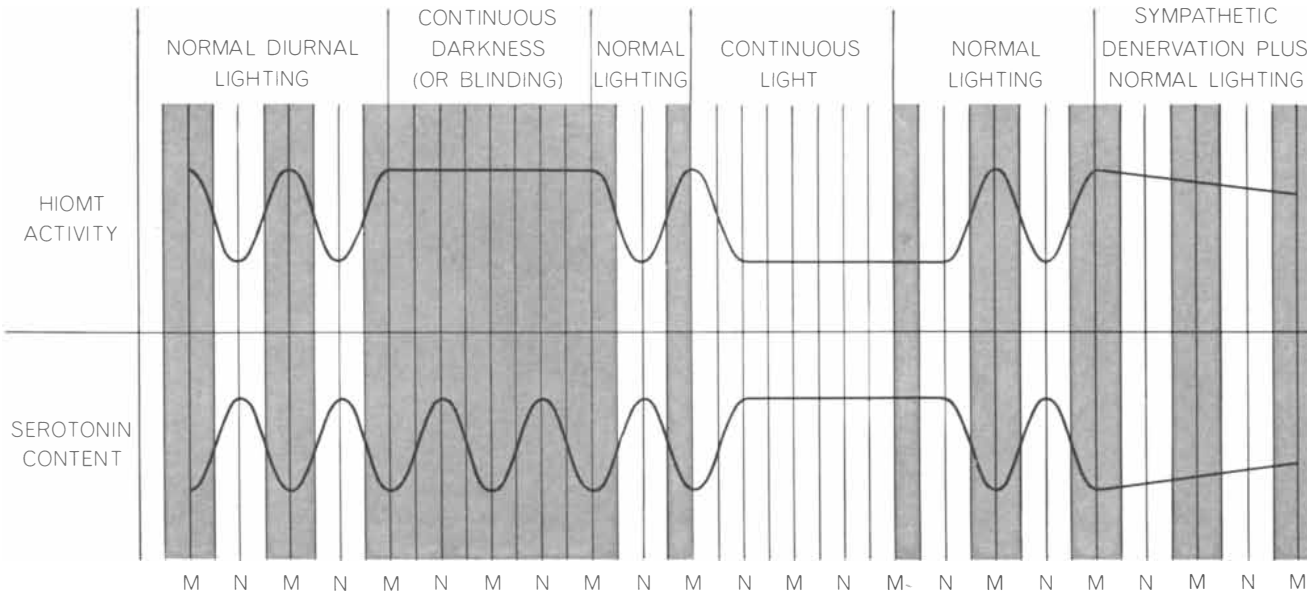
cadian, from the Latin phrase meaning "about one day." Could the pineal respond to natural diurnal lighting shifts? If so, it might function to synchronize the endocrine apparatus with these shifts.

In order to determine if normal lighting rhythms influenced the pineal, we kept a large population of rats under controlled lighting conditions (lights on from 7:00 A.M. to 7:00 P.M.) for several weeks and then tested their pineals for melatonin-forming ability at 6:00 A.M., noon, 6:00 P.M. and midnight. In the five hours after the onset of darkness (that is, by midnight) this enzymatic capacity increased between two and three times. Moreover, pineal weight also changed significantly during this period, again indicating that light was affecting many more compounds in the pineal than the single enzyme we were measuring.

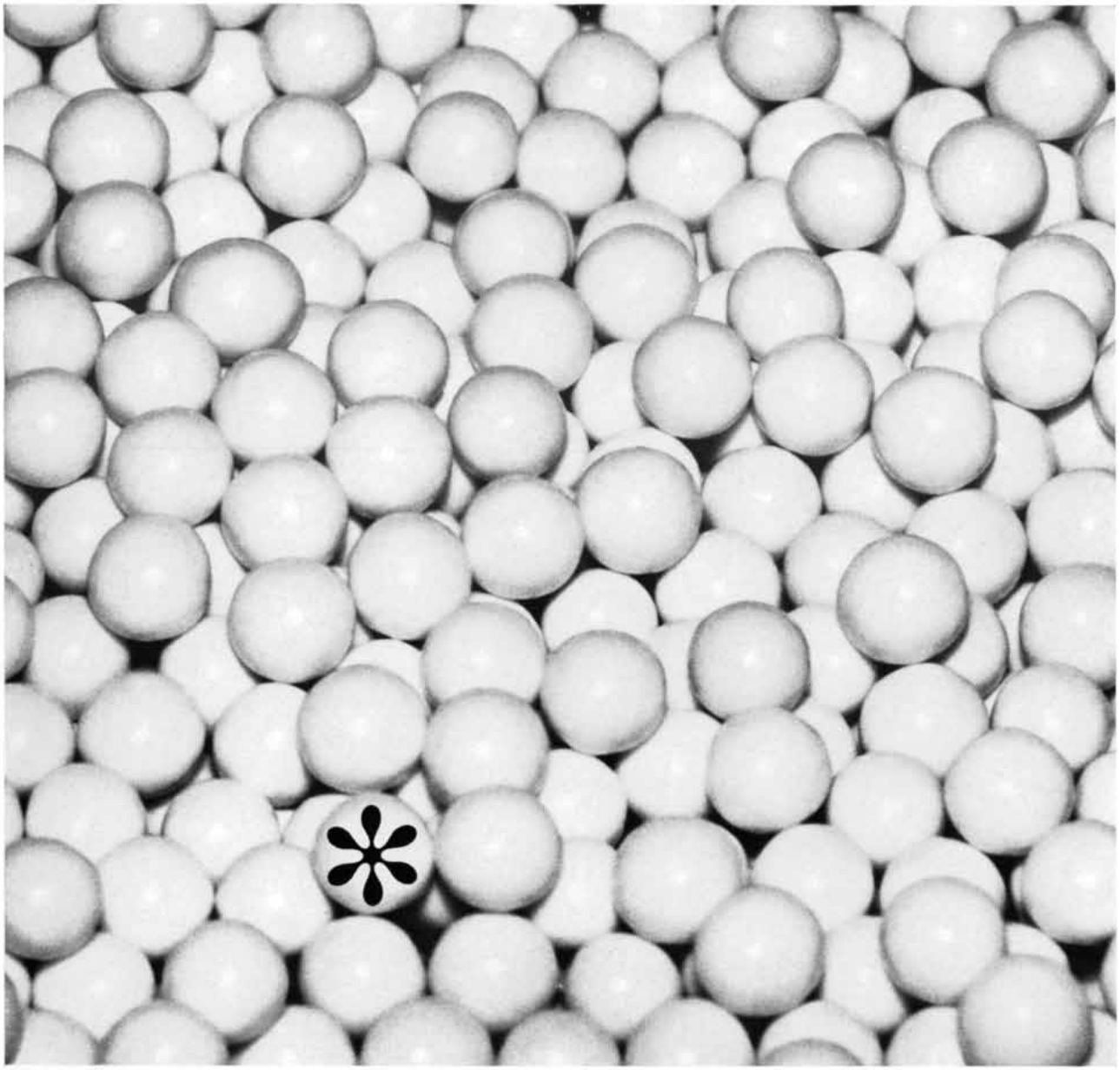
All circadian rhythms studied up to this stage had in common the ability to persist for some weeks after animals were deprived of environmental lighting cues (by blinding or being placed in darkness). These rhythms no longer showed a period of precisely 24 hours, but they did fall in a range between 22 and 26 hours and hence were thought to be regulated by some internal mechanism not dependent on, but usually synchronized with, environmental lighting. Such endogenous, or internally regulated, circadian rhythms in rodents include motor activity and rectal temperature, as well as the rhythm in adrenal-steroid secretion. When we blinded rats or placed them in continuous light or darkness, the pineal rhythm in melatonin-forming activity was rapidly extinguished. If instead of turning off the lights at 7:00 P.M. illumination was continued for an additional five hours and pineals were examined as usual at midnight, the expected rise in melatonin-forming activity was completely blocked. This pineal rhythm in HIOMT activity thus appears to be truly exogenous, or externally regulated, and is entirely dependent on shifts in environmental lighting. Hence this enzyme rhythm may be more important in carrying information about light to the glands than other circadian rhythms that do not depend on light for their existence.

Recently Wilbur Quay of the University of California at Berkeley has found that the content of serotonin in the rat pineal also undergoes marked circadian rhythms. The highest levels of this amine are found in pineals at noon and the lowest levels at midnight. Serotonin content falls rapidly just at the time that melatonin-forming activity is rising. In collaboration with Solomon Snyder we studied the mechanism of the serotonin cycle. When rats are kept in continuous light, the serotonin cycle is extinguished. To our surprise, however, when rats are kept continuously in darkness or blinded, this rhythm persists, unlike the rhythm in the melatonin-forming ezyme. When the sympathetic nerves to the pineal are cut, the serotonin and HIOMT cycles are both suppressed. When the nerves from the central nervous system to the superior cervical ganglion are interrupted, the serotonin rhythm is also abolished [*see illustration below*]. Hence the serotonin rhythm in the pineal gland is similar to most other circadian rhythms (and differs from the HIOMT cycle) in that it is endogenous and depends on environmental light only as an external synchronizer. The mechanism that controls the serotonin rhythm appears to reside within the central nervous system. The pineal gland thus contains at least two distinct biological clocks, one totally dependent on environmental lighting and the other originating within the brain but cued by changes in lighting.

At present little is known about what organs are dependent on the pineal clock for cues. The ability of melatonin to modify gonadal function suggests, but does not prove, that its secretion may have something to do with the timing of the estrus and menstrual cycles—two phenomena about whose mechanisms of control very little is known. One is tempted to argue teleologically that any control mechanism as complicated and sensitive as that found in the mammalian pineal gland must have some place in the economy of the body.



BIOCHEMICAL RHYTHMS in the pineal gland of the rat were recorded under various lighting and other conditions. Normally both the content of serotonin and the activity of the melatonin-forming enzyme (HIOMT) vary with a 24-hour cycle. The serotonin content is greatest at noon (*N*), whereas the HIOMT activity is greatest at midnight (*M*). The HIOMT cycle is completely dependent on environmental lighting conditions: it disappears when animals are kept in continuous light or darkness, or when they are blinded. The serotonin cycle persists in continuous darkness or after blinding but can be abolished by keeping the rats in continuous light. Both cycles are depressed when the sympathetic nerves to the pineal gland are cut (*extreme right*). Gray areas signify darkness.

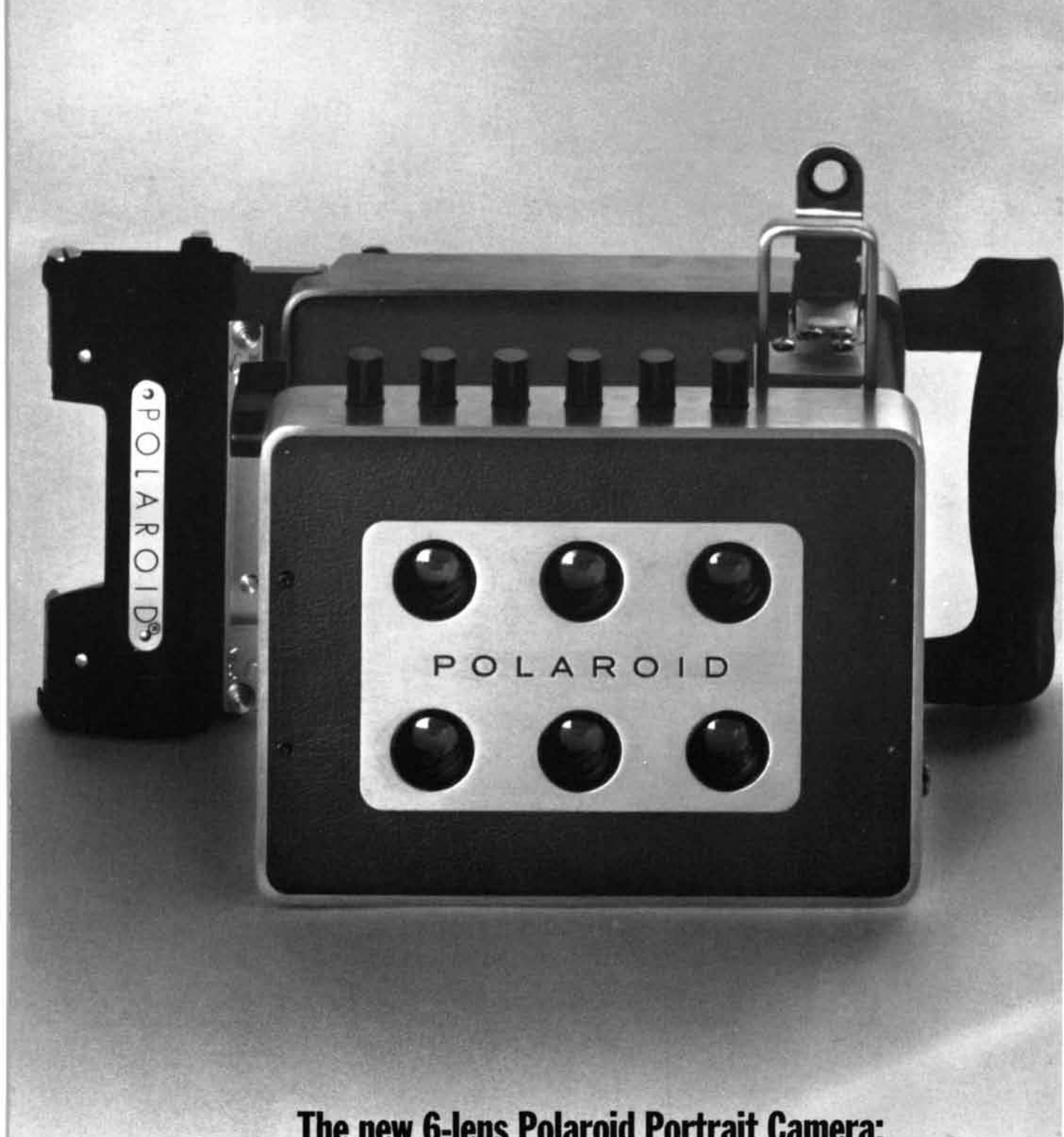


*One in a million and never easier to find

Automated information handling of microfilm records is here—for businesses large and small—with modern Recordak Microfilm Systems. Let you retrieve any microfilmed record—even get a paper-print—in seconds with push-button ease. So if you are concerned about the information explosion (and what it now costs your business) write: Recordak Corp., Dept. PP-7, 770 Broadway, New York, N.Y. 10003.

RECORDAK® SUBSIDIARY OF EASTMAN KODAK COMPANY

Kodak



The new 6-lens Polaroid Portrait Camera:

Some funny things happen to people when they have to sit through the stop-and-go mechanics of conventional portrait-taking. (Get the pose, get the lighting, get the focus, get the picture. Reload, re-pose, reshoot. And so on.) They're apt to become self-conscious, or fidgety, or just plain bored. Not the best faces to put forward.

We had good reason to believe we could help out. We already had a system of photography that offered immediacy and excitement, and an inherent rapport between photographer and subject. Taking Polaroid Land photography a few steps further, we developed the 6-lens Polaroid Portrait Camera.

With this camera and Polaroid Land positive/negative film, the photographer can catch his subject in six different poses without interruption. And 20 seconds later, the subject gets six proofs while the photographer gets six negatives. This system makes the meeting between photographer and subject far more spontaneous. And makes it easier to get good pictur of people.

On-the-spot proofs keep the subject i terested. They can also save him a retu trip to the studio for reshooting; if he a the photographer are unhappy with t first six pictures they can try some mo then and there. And once the best pro has been picked, the subject can order t final enlargement right away. It's a wrapped up in one session.

The photographer is freed from nea

"POLAROID"®

6 positives and 6 negatives from one piece of film in 20 seconds.

all the mechanics of the conventional portrait sitting and can give full attention to directing pose and expression. He doesn't have to reload as often because he gets six pictures on one piece of film. He doesn't have to focus because each lens is fix-focused for between five and six feet.

And he can keep shooting in rapid succession—pushing shutter release buttons as fast as the electronic flash unit can recycle (about 10 seconds). Because he doesn't have to pause more than a few seconds between shots, just about everything he sees and likes in the viewfinder he can get on film.

But all this wouldn't count for much if it didn't produce a technically perfect finished portrait. Polaroid Land Type 55 P/N film makes the system complete. Although each negative measures only 1⅛ x 1⅜ in., fine-grain and sharpness characteristics are such—resolution is up around 150 lines per mm—that 8 x 10 or 11 x 14 in. enlargements can be made with no loss of detail or appearance of grain. Even with some cropping.

Right now, the Polaroid Portrait System is being used by England's largest organization of portrait studios. It will be introduced here in a few months. Perhaps you'll see it in action next time you, or some member of your family, have a session with the portrait photographer. If so, we think you'll be intrigued by the system. We're sure you'll like the pictures.
Polaroid Corporation, Cambridge, Mass.

IMPLOSION EXPERIMENT, photographed at the Los Alamos Scientific Laboratory in New Mexico, can create magnetic fields of more than 10 million gauss for a period of about a microsecond. The highest fields attained by continuously operating magnets are about 250,000 gauss; a pulsed-field magnet has reached 750,000 gauss for a few milliseconds. The experiment shown in the photograph was conducted by C. M. Fowler, W. B. Garn and R. S. Caird. A diagram of the implosion device appears on the opposite page.

ULTRASTRONG MAGNETIC FIELDS

Defined as fields above 10 million gauss, they can be produced by clever use of explosives. They may make it possible to study matter under pressure conditions found inside stars and planets

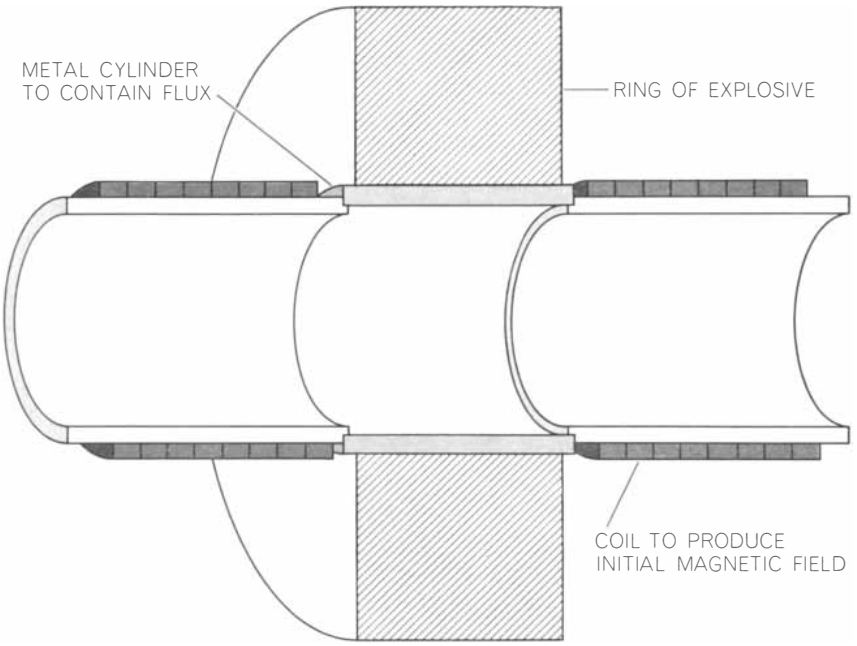
by Francis Bitter

In the April issue of *Scientific American* my colleagues Henry H. Kolm and Arthur J. Freeman presented an article entitled "Intense Magnetic Fields," in which they described the methods employed at the National Magnet Laboratory at the Massachusetts Institute of Technology to produce steady magnetic fields with strengths of between 100,000 and 250,000 gauss. In this article I should like to carry the discussion forward and describe methods that have already created for brief periods fields in the neighborhood of 10 million gauss and that appear potentially capable of achieving fields exceeding 100 million gauss. These methods have been pioneered at the Los Alamos Scientific Laboratory by C. M. Fowler, R. S. Caird, W. B. Garn and D. B. Thomson. I have had the pleasure and benefit of discussing with them many of the ideas presented in this article.

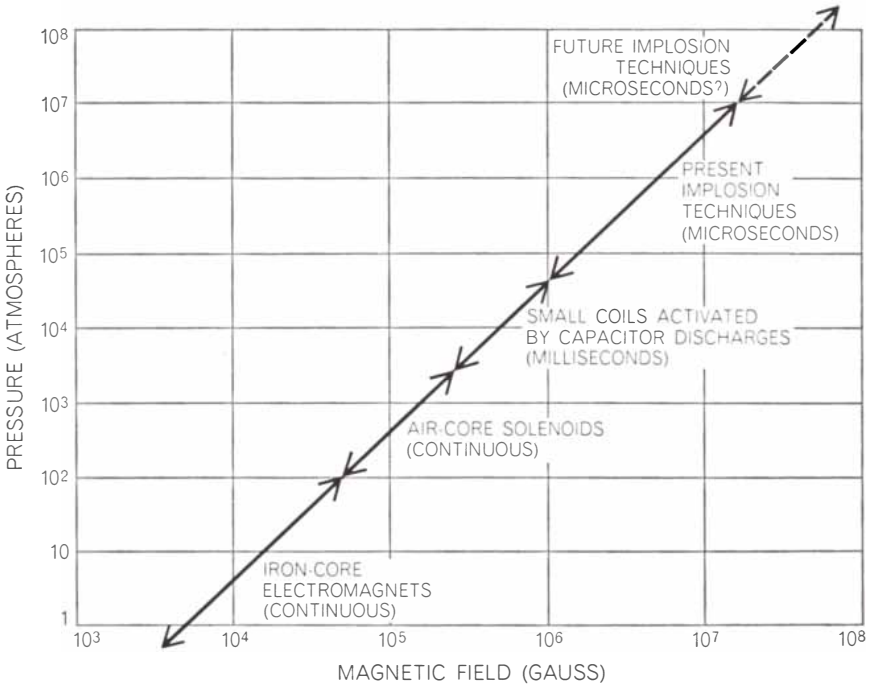
What purpose might be served by producing ultrastrong magnetic fields? My own belief, which is shared by a number of my colleagues, is that such fields can provide new insights into the structure of matter. Ultrastrong magnetic fields should distort matter in various characteristic ways, and one would like to see if the distortions predicted by theory are actually observed under experimental conditions. Beyond that one may well encounter new and previously unsuspected phenomena.

The techniques I shall discuss, however, present an opportunity to do more than simply explore the response of matter to magnetic fields of a new order of intensity. Magnetic fields have a little-appreciated property that can be exploited along with the production of ultrastrong fields: they transmit forces with the velocity of light. As we shall see, the only method currently available for creating ultrastrong magnetic fields employs explosive charges arranged in a ring so as to produce an implosion and thereby compress a preexisting magnetic field. Ultrastrong fields produced in this way can probably exert or withstand pressures of many millions of atmospheres. Transmitted with the speed of light, they can momentarily provide a shielded experimental chamber in the midst of an explosion, and they can also provide a window through which the events taking place in the chamber can be observed with precision and in detail. In other words, the fields should make it possible to study matter under far more extreme conditions than can be provided by magnetic fields alone—conditions that normally exist only in the interior of planets and stars. This seems to me the most exciting prospect presented by the effort to create ultrastrong magnetic fields.

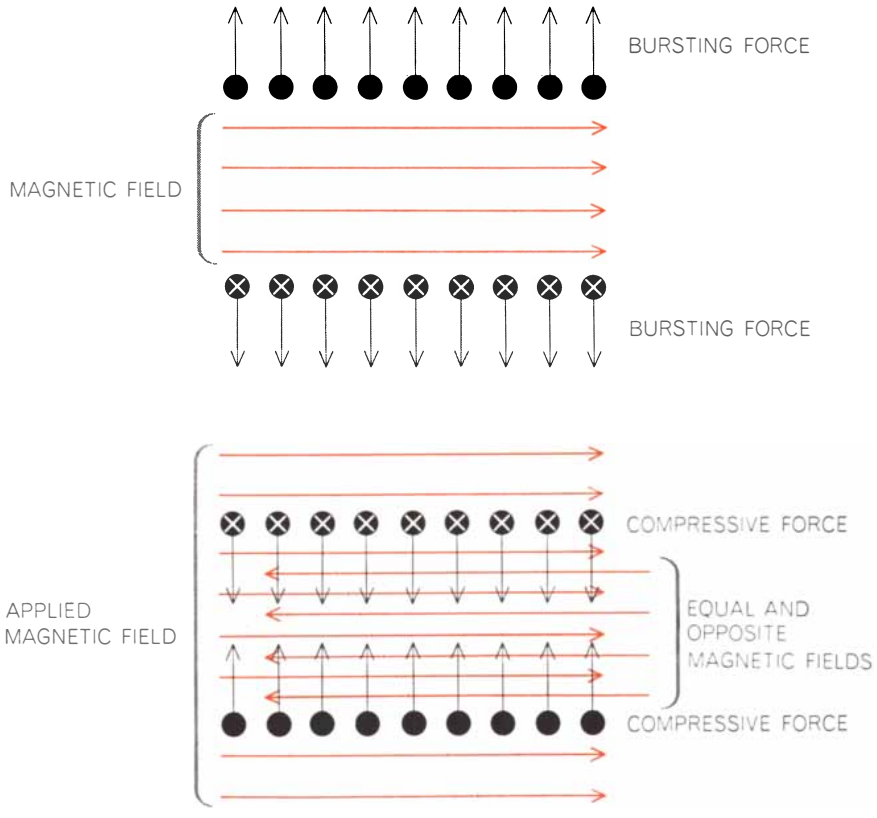
Let us analyze how it is that magnetic fields can exert a pressure. If you were to put your hand into a steady magnetic field, even a very strong one, you would not experience any sensation at all. This is because the magnetic field penetrates tissue and bone. A magnetic field exerts a pressure only if it



IMPLOSION DEVICE employed by Los Alamos workers consists of a metal cylinder about four inches in diameter with coils at each end that can create a pulsed magnetic field of about 30,000 gauss. A ring containing some 10 to 20 pounds of high explosive surrounds the center of the cylinder and is timed to go off when the pulsed field reaches its peak. The subsequent implosion of the cylinder compresses the magnetic flux, thereby raising the field to several million gauss. The principal of flux compression is illustrated on page 67.



METHODS FOR CREATING MAGNETIC FIELDS of increasing strength are indicated, together with the pressures exerted by fields of various intensities. One million atmospheres, attainable by the implosion technique, is about a third of the pressure at the center of the earth. Sustained fields are limited to about 250,000 gauss because the corresponding pressure of 2,500 atmospheres (35,000 pounds per square inch) approaches the yield point of the conductors in air-core solenoids. Such fields require several megawatts of power.



MAGNETIC FIELDS EXERT PRESSURE when confined to one side of a barrier they cannot penetrate. In the diagram at top an energized coil produces a magnetic field that is trapped within the coil and therefore exerts an outward pressure. In the diagram at bottom the coil is surrounded inside and out by an applied magnetic field. If the coil is then energized so that the interior field is canceled, the applied magnetic field will tend to compress the coil. Crosses in the coils indicate that the direction of current flow is into the page.

encounters a barrier beyond which it is excluded; the barrier then "feels" the force of the field pressing against it.

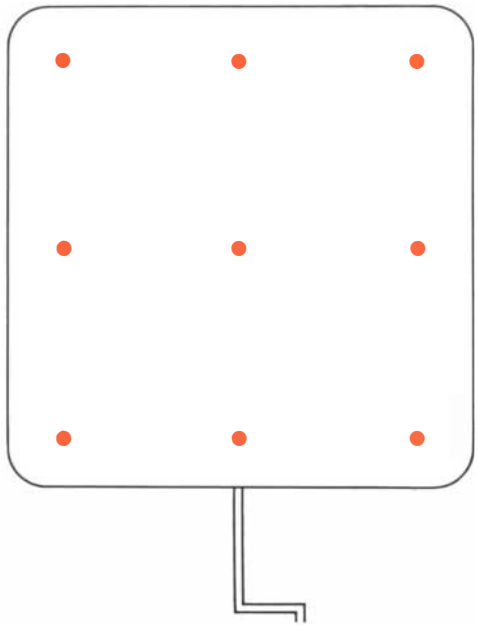
A pressure of this kind is developed inside a solenoid, the familiar device that consists of wire wrapped in the form of a cylindrical coil. When an electric current is passed through the coil, a magnetic field is generated inside the solenoid. Because the field cannot penetrate the coil it exerts an outward pressure, and if the field is sufficiently strong, the coil will burst. By the same token one can compress the solenoid by placing it in a uniform magnetic field and passing a current through it just large enough, and in the proper direction, to cancel the part of the external field that is inside the solenoid [*see bottom illustration at left*].

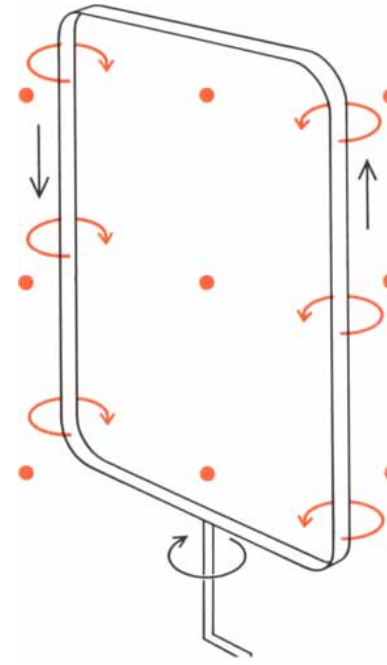
By expressing magnetic fields in terms of their pressure equivalents one can convey something of their magnitude. At the earth's surface the earth's magnetic field ranges from about half a gauss at the Equator to about one gauss at the magnetic poles. A solenoid that is adjusted to exactly cancel the earth's field inside it would be subjected to an external pressure of a small fraction of an ounce per square inch—the pressure of a gentle breeze. A large iron-core electromagnet of the type found in many laboratories can develop a field of about 30,000 gauss, which exerts a pressure of about 40 atmospheres in the gap between its pole pieces. (One atmosphere is 14.7 pounds per square inch. To translate a magnetic field B, measured in gauss, into pressure measured in atmospheres multiply B^2 by 4×10^{-8}.) A summary of the means for producing magnetic fields of various strengths, together with their pressure equivalents, is given in the top illustration at the left.

The strongest magnetic fields yet measured are those created by the implosion technique developed at Los Alamos by Fowler and his associates. Their method depends on "flux compression," which is implicit in the phenomenon of electromagnetic induction discovered independently in 1831 by Michael Faraday of England and by Joseph Henry of the U.S. Faraday and Henry showed that a current is generated when a coil of wire is rotated between the poles of an iron magnet. When the phenomenon is examined closely, it can be seen that the number of lines of magnetic flux passing through the open loop of the coil tends to stay constant when the loop is rotated inside the field [*see upper illustration on opposite page*]. As the loop is ro-

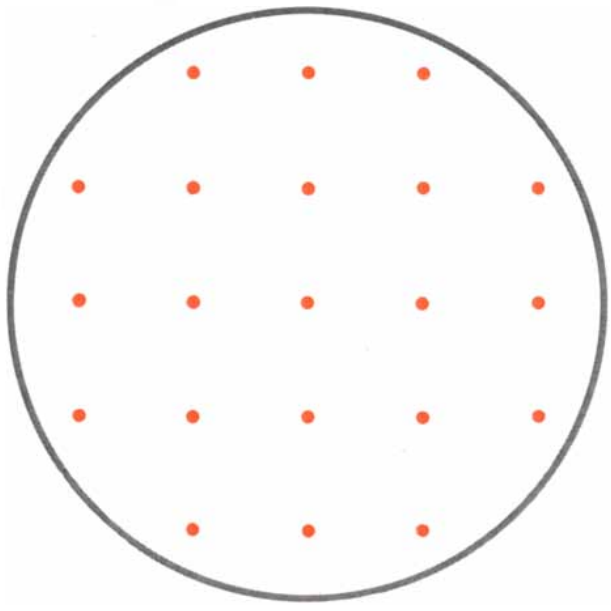
tated from a position perpendicular to the lines of flux, it cuts through the lines and leaves some of them outside the loop, but for every line of flux left outside a new one is created inside the loop by the current that is induced to flow in the loop. This is of course the ideal situation; it can be attained only if none of the energy of the induced current is converted into heat by resistance in the wire.

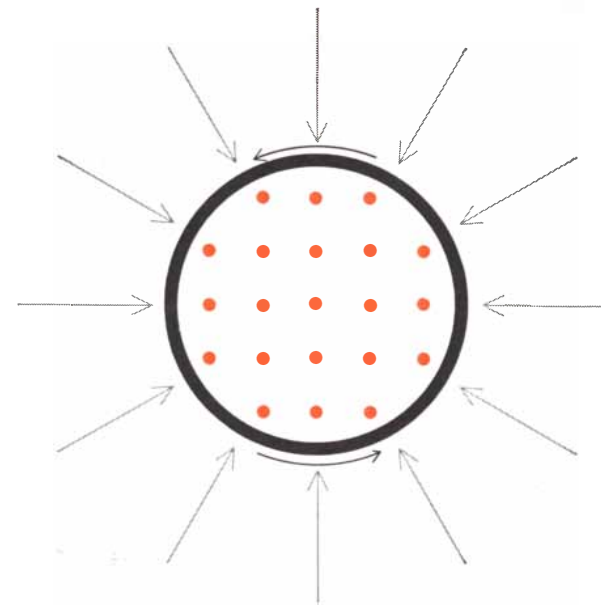
The Los Alamos workers have found a way to exploit flux compression on a dramatic scale. Instead of using a coil of wire to contain the flux they employ a metal ring a few inches in diameter. A strong magnetic field is created inside the ring by coils that are energized by a relatively slow discharge from a bank of capacitors. The magnetic flux generated by the coils penetrates the ring; when it reaches its peak, the ring is suddenly compressed by an encircling charge of high explosive. As the ring



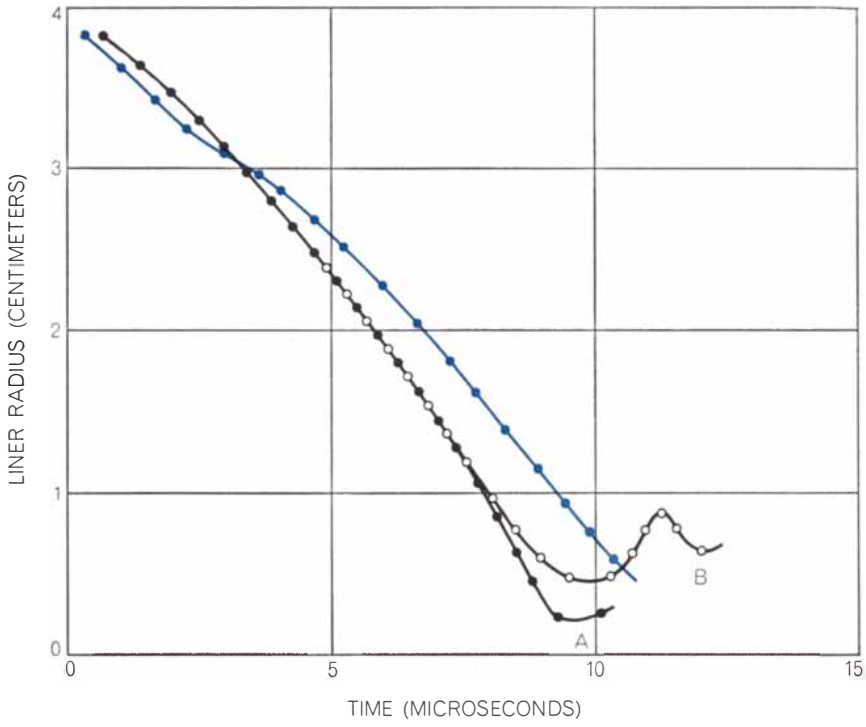


CONSTANCY OF MAGNETIC FLUX is demonstrated by a coil of wire placed in a magnetic field. When the coil is perpendicular to the field (*left*), a certain number of lines of flux pass through it. (Here nine lines of flux are seen end on; the magnetic field is directed outward from the page.) As long as the coil is motionless no current flow is induced. When the coil is rotated (*right*), however, it cuts through some of the flux lines and current begins to flow. The direction of current flow in the coil (counterclockwise) creates new lines of magnetic flux (*curved arrows*) that tend to keep the total number of lines of magnetic flux inside the coil constant.

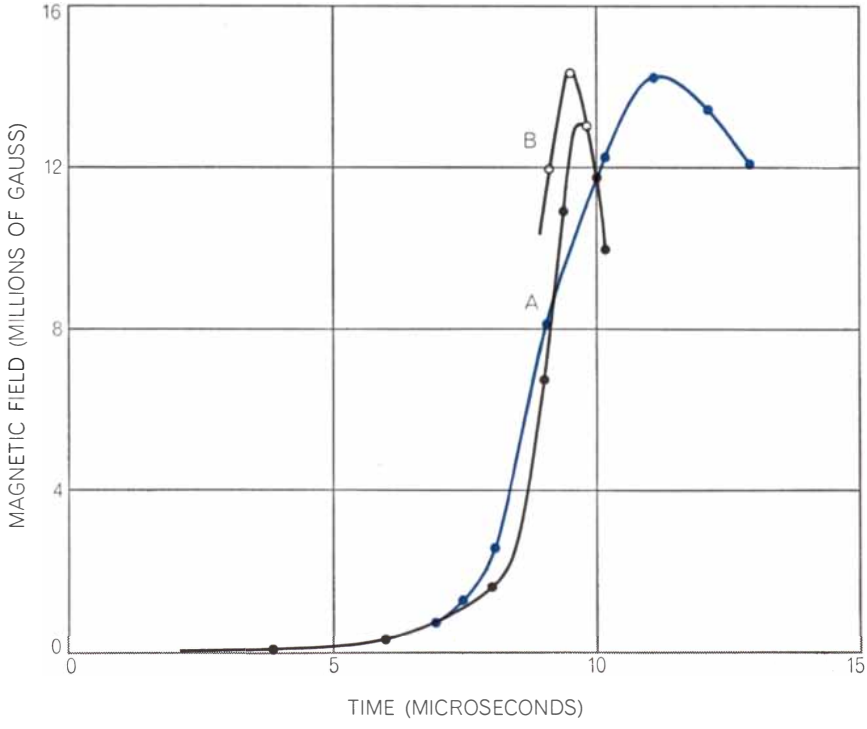




FLUX COMPRESSION, another example of the principle illustrated at top of page, is used in creating ultrastrong magnetic fields. At left a magnetic field of a certain flux density is produced inside a metal cylinder. When the cylinder is imploded (*right*), a current is induced in the cylinder that tends to keep the magnetic flux constant. Magnetic-field strength is proportional to the flux per unit area. Thus if the cylinder is compressed to 1 percent of its original cross section, the magnetic field will be increased by a factor of 100.



IMPLOSION OF CYLINDER as recorded in a Los Alamos experiment (*colored curve*) is compared with theoretical values computed by H. C. Hoyt and P. Kazek. Curve *A*, which assumes an initial magnetic field of 50,000 gauss, shows a final radius of .209 centimeter for the cylinder, or liner. The final field for curve *A* was computed to be 16.6 million gauss. Curve *B* assumes an initial field of 150,000 gauss; the final radius was computed to be .436 centimeter and the final field 11.5 million gauss. In the actual experiment the initial field was 40,000 gauss and the last measured radius was about .5 centimeter. Evidently the radius continued to shrink and smashed a magnetic probe before the peak field was reached.



GROWTH OF MAGNETIC-FIELD STRENGTH in another implosion experiment (*colored curve*) is compared with values calculated by Hoyt and Kazek. The initial magnetic fields assumed for curves *A* and *B* respectively were 100,000 and 75,000 gauss. In the experiment the initial field was 90,000 gauss and the peak field measured 14.3 million gauss. The two sets of curves on this page are from a paper published in 1960 by Fowler, Garn and Caird.

shrinks rapidly in size, a current is induced in it that is sufficient to maintain the magnetic flux it originally contained. Magnetic-field strength is proportional to the flux per unit area; therefore if the ring originally contained a field of 50,000 gauss in an area of 100 square centimeters, the field would be increased by a factor of 100, or to five million gauss, if the ring area could be compressed to one square centimeter.

Flux compression of this kind is possible only if the heat produced by the current in the ring remains low as compression takes place. If the resistance of the ring is large, the work done by compression degenerates quickly into heat, the induced currents die away and the flux escapes through the ring. Whether or not flux is compressed in an implosion experiment depends on just how high the resistance of the ring is after the tremendous currents required to produce millions of gauss in a single-turn coil have been generated.

In order to develop a field of 10 million gauss in a cylindrical ring one must induce in it a current of about 10 million amperes per centimeter of its length. What temperature would the ring reach in only a microsecond? And what would happen to the metal in the ring when it was being squeezed by pressures of millions of atmospheres? Before the experiment was tried no one knew. Fowler and his co-workers showed conclusively that the electrical conductivity of various metal rings was sufficiently high so that flux could be trapped and flux densities of millions of gauss could be achieved by implosion techniques. Dare one extend these results to higher fields and higher pressures? The possibility of using implosions to create fields higher than 10 million gauss has been questioned, but I feel that such pessimism is unwarranted. I can see no evidence from the Los Alamos experiments that the implosion technique has begun to approach limits imposed by nature.

To give the reader an idea of what has actually been tried and achieved I shall describe in somewhat more detail the type of experiment that has been conducted many times during the past five years by Fowler and his group and has now also been undertaken elsewhere. A diagram of the apparatus is shown on page 65. A metal ring three or four inches (7.5 to 10 centimeters) in diameter is surrounded by 10 to 20 pounds of explosive. Coaxial with the ring are coils activated by a capacitor discharge that can create an initial mag-

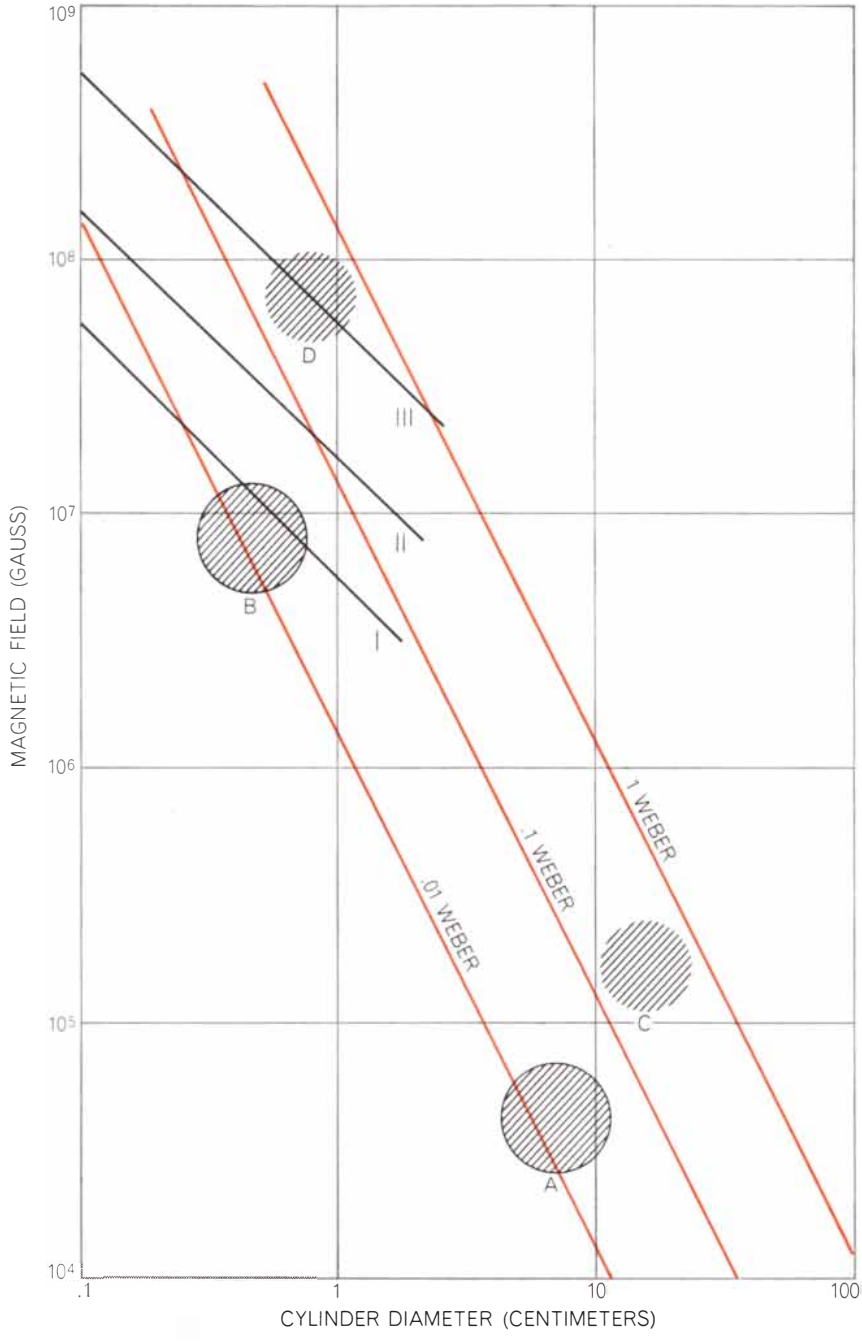
netic field of about 30,000 gauss. (In some experiments fields as high as 100,000 gauss have been used.) The initial field is established in about a millisecond. When the field reaches its peak, a rim of detonators triggers an explosion that collapses the metal ring and raises the intensity of the trapped magnetic field to between five million and 10 million gauss. The imploding ring travels inward at the rate of about four centimeters in 10 microseconds, which is about four kilometers per second, or 8,000 miles per hour.

The internal magnetic pressure builds up as the diameter of the ring shrinks. When the implosive pressure is balanced by the magnetic pressure, the motion stops and is then reversed. The magnetic pressure drives the ring outward again, but this is usually not observed because the measuring equipment at the center of the apparatus has by now been destroyed—perhaps as the result of slight asymmetries in the compression of the ring. In a typical experiment the last measurement indicates that the ring is compressed to a diameter of a few millimeters [*see illustrations on opposite page*]. The energy per unit length of the compressed field, which is supplied by the explosive, is proportional to the square of the field and inversely proportional to the square of the final radius.

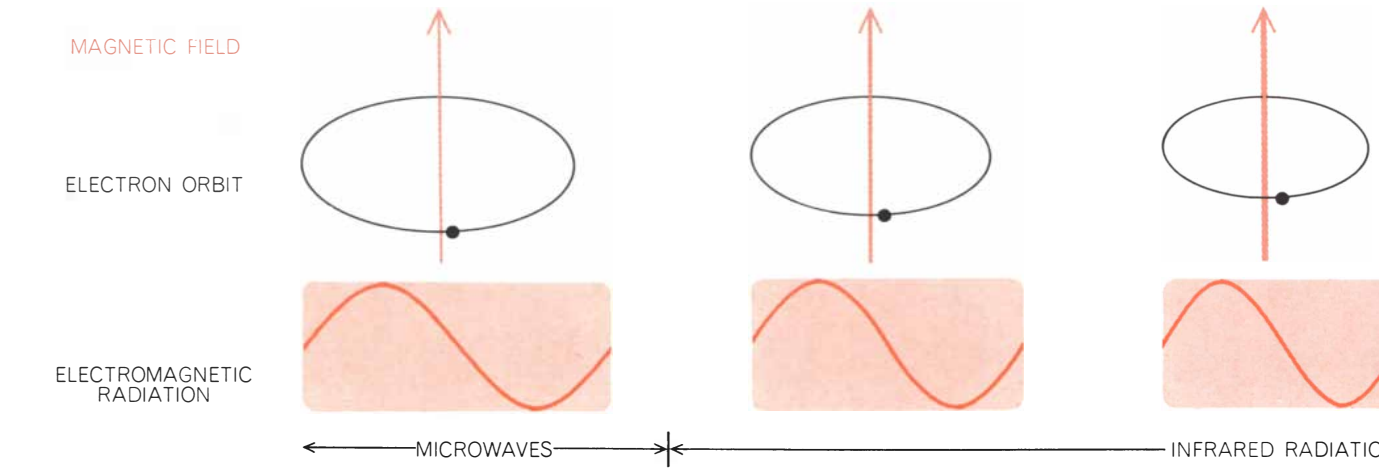
Let us now consider what would be required to create fields significantly stronger than 10 million gauss in regions appreciably larger than a few millimeters in diameter. Although it might seem easier and less costly to produce ultrastrong fields in a region measured in tenths of a millimeter, the task of producing such conditions repeatably would almost certainly be difficult and perhaps even impossible.

The illustration at the right is a diagram relating magnetic-field intensity to the diameter of the final working space. The diagram introduces the weber, the unit of flux used in the meter-kilogram-second (mks) system. One weber is the flux in a field of 10,000 gauss occupying an area of a square meter. This is a large amount of flux. It is easy to achieve in a big electromagnet that has an iron core but very difficult in an air-core magnet, which is obviously the type needed in the implosion technique. In the Los Alamos experiments the flux amounted to only a few hundredths of a weber, equivalent to between 30,000 and 50,000 gauss in a cylinder four inches in diameter. This amount of flux, indicated by region *A* in the diagram, is compressed by implosion to a few million gauss in a cylinder a few millimeters in diameter, shown by region *B*. In order to reach region *D*, which represents fields of between 50 million and 100 million gauss in regions of about a centimeter in diameter, one must compress about half a weber of flux. If the apparatus is to be of reasonable size and cost, the initial diameter of the flux-containing ring should be less than a foot, which means that it must contain an initial field of about 200,000 gauss, shown by region *C*.

Obtaining half a weber of flux in a



FLUX-COMPRESSION DIAGRAM shows the magnetic fields and cylinder diameters corresponding to given amounts of magnetic flux. One unit of flux, expressed in webers, is the flux in a field of 10,000 gauss in an area of a square meter. The Los Alamos experiments had initial values in region *A* and reached *B* after flux compressions. Line *I* corresponds to the magnetic energy per unit length of field that has been generated experimentally by a few tens of pounds of explosive. Lines *II* and *III* represent successive increments of a factor of 10 in magnetic energy. The goal is region *D:* a magnetic field of between 50 million and 100 million (10^8) gauss in a region whose diameter is about a centimeter. To reach it one must combine the initial conditions shown in *C* with an energy input represented by *III*.



MAGNETIC FIELD (GAUSS)	10^5	10^6	10^7
RADIUS OF ELECTRON ORBIT (CENTIMETERS)	$\sim 10^{-6}$	$\sim 4 \times 10^{-7}$	$\sim 10^{-7}$
WAVELENGTH OF EMITTED RADIATION (CENTIMETERS)	$\sim 10^{-1}$	$\sim 10^{-2}$	$\sim 10^{-3}$

"MAGNETIC ATOMS" could be produced by exposing electrons to an ultrastrong magnetic field. The term signifies that an electron, which is obliged to travel in a circular orbit in a magnetic field, would have an orbit of atomic dimensions if placed in a very high field. When traveling in orbits, electrons characteristically emit radiation, known as "cyclotron" radiation. The smaller the orbit, the

cylinder eight inches in diameter (region *C*) will be difficult but probably quite possible. The magnetic energy stored in an eight-inch coil of useful length would be of the order of tens of megajoules. That much energy might be obtained by discharging a large bank of capacitors or by extracting the energy stored in a large flywheel attached to an electric generator. Capacitor banks that can deliver a megajoule or so of energy now exist in several laboratories. The cost and engineering difficulties involved in building a bank to deliver tens of megajoules would have to be compared with the cost and difficulties of supplying the same energy by the flywheel method.

Fairly good cost estimates for the flywheel method have already been made by Crawford Adams, the engineer who was responsible for the detailed design and construction of the National Magnet Laboratory. A 100-megawatt generator capable of delivering a million amperes at 100 volts for a few seconds could be built and installed for about $1,250,000. Such a generator could produce a weber of flux in a coil 10 inches in diameter, provided that the coil was precooled in liquid nitrogen or cooled by a large surge of water. We can therefore assume that the required flux can be produced, even though a detailed study is needed to demonstrate just which of several methods is the most practicable.

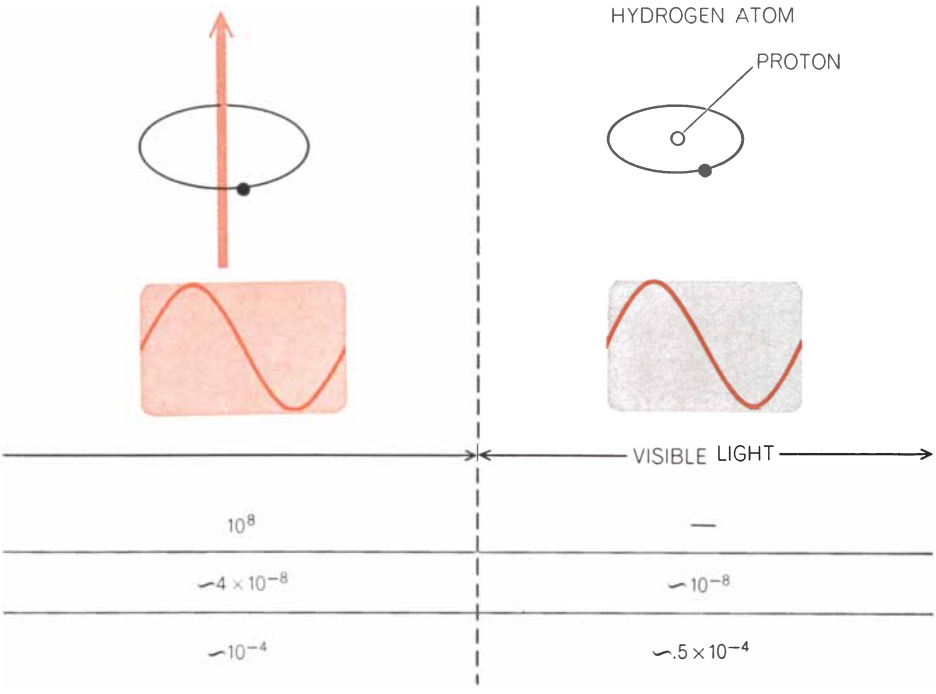
The production of enough flux to reach region *C* in the diagram on the preceding page, however, solves only the first half of the problem. The flux must also be compressed by an implosion. This aspect of the problem is illustrated by the lines labeled *I*, *II* and *III*, which represent constant energy per unit length in the field. If the magnetic energy per unit length has some particular value along line *I*, its value along line *II* is 10 times greater and along line *III* 100 times greater. The determination of the amount of explosive needed to reach region *C* is again a challenging engineering problem, but one can at least make a preliminary guess as to the requirements. Since tens of pounds of explosive are needed to reach region *B*, roughly 100 times as many, or thousands of pounds, will probably be needed to reach region *D*.

Clearly these proposals involve large experimental facilities and large explosions. The actual limits are difficult to set, but fields of the order of tens of millions of gauss in volumes of the order of cubic centimeters, and of hundreds of millions of gauss in volumes of cubic millimeters, are conceivable using ordinary explosives. Eventually one may wish to consider the use of nuclear explosives.

If ultrastrong magnetic fields can be generated in this manner, what might be learned from them? I shall discuss first the direct magnetic effects of ultrastrong fields and then describe how they might be used to create extremely high pressures.

It is generally stated that a magnetic field produces only one very simple change in an atom: it makes the atom rotate around the lines of magnetic force at a rate proportional to the strength of the field. This is called Larmor precession, after the British physicist Joseph Larmor. The rate of precession can be measured by observing how electromagnetic radiation interacts with the rotating atoms. When the rate of precession and the frequency are matched, the atoms absorb electromagnetic energy. For magnetic fields of the strength commonly available in physics laboratories the resonant frequencies for atomic nuclei are in the radio region of the spectrum, in which wavelengths range from somewhat less than a meter to several hundred meters. For electrons the resonant frequencies are in the microwave region, with wavelengths from about one millimeter to a meter.

Larmor precession is only part of the story. As the magnetic field is increased the atom is distorted. The rotation induced by the field can be regarded as a current that prevents the applied field from penetrating the interior of the atom, with the result that the atom is compressed. To be more precise, the electrons of the atom are driven closer



higher the rotation frequency and the shorter the wavelength of the emitted radiation. At 100 million gauss the orbit is only slightly larger than that of an electron in a hydrogen atom (in a low-energy state) and the emitted radiation approaches the visible region.

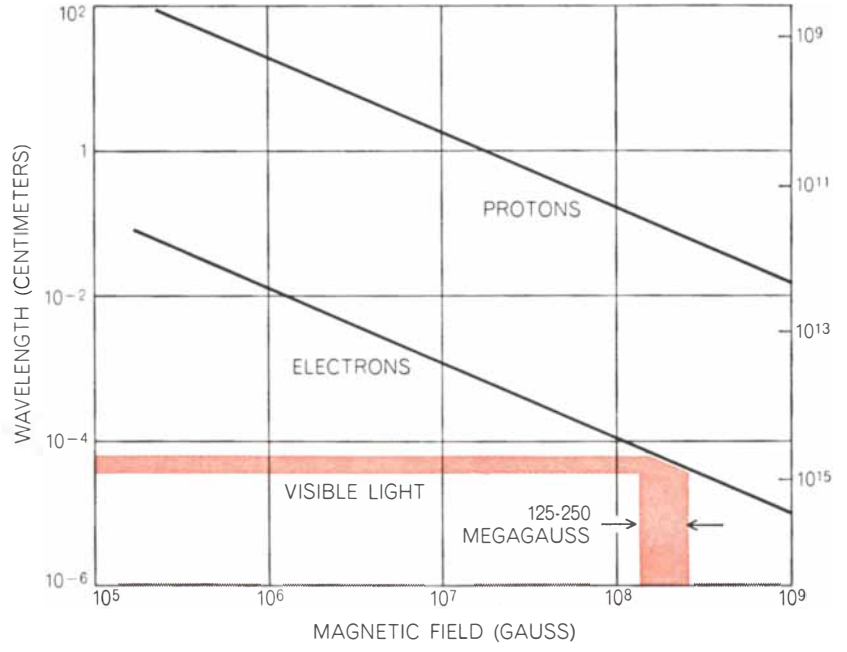
to its nucleus. In weak fields these compressive magnetic forces are small compared with the electrostatic forces that hold the atom together. In fields above 50 million gauss, however, atoms tend to be radically deformed.

The extent of these deformations is difficult to calculate accurately, but semiquantitative estimates are fairly simple to make. Instead of thinking in terms of atoms, with their various numbers of electrons, let us consider instead how free electrons and protons behave in a magnetic field. When a charged particle is placed in a magnetic field, it is deflected at right angles to the field, so that it begins traveling in an orbit around the lines of magnetic force at a rate known as the cyclotron frequency. For electrons this frequency turns out to be just twice the Larmor frequency in a field of the same strength. The cyclotron frequency for electrons in cycles per second is obtained by multiplying the field B, measured in gauss, by 2.8×10^6. For protons, which are 1,835 times more massive than electrons, the frequency is proportionately lower: 1.5×10^3 multiplied by B.

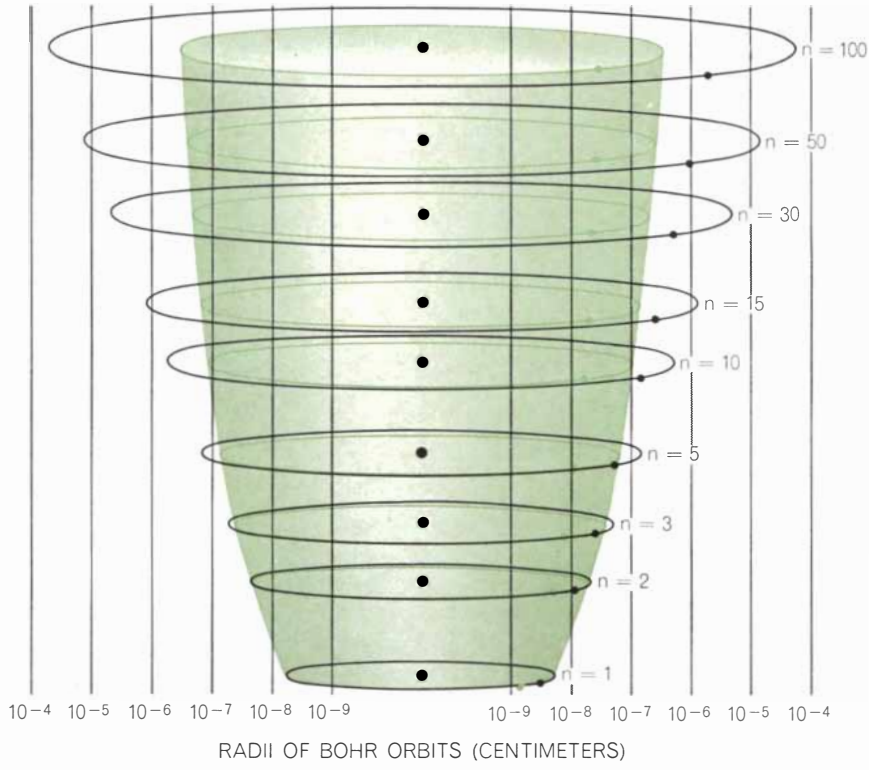
In the case of electrons magnetic fields in the region between a million and 100 million gauss lead to very high cyclotron frequencies indeed. As the intensity of the field increases, the radius of the electron's orbit becomes steadily smaller until it is comparable to the radius of the orbit of an electron in an atom. The actual radius of orbit for an electron in a magnetic field, as for an electron in an atom, depends on the electron's angular momentum, which is expressed in units of a quantum number, n. (In the presence of a magnetic field the ordinary mechanical angular momentum must be replaced by the "canonical" angular momentum, which takes account of the energy circulating in the field.) The illustration at the left shows—for magnetic fields of 100,000, one million, 10 million and 100 million gauss—the orbital radius for electrons that have the minimum amount of angular momentum (that is, n equal to one).

When an electron travels in a magnetic field, it characteristically emits quanta of electromagnetic radiation. The wavelength of the emitted radiation is inversely proportional to the cyclotron frequency: the higher the frequency, the shorter the wavelength. For fields between 125 million and 250 million gauss the emitted radiation falls in the visible region of the spectrum [*see illustration below*]. Thus one can see that the properties of free electrons can be modified by ultrastrong magnetic fields to resemble those of electrons that are bound in atoms.

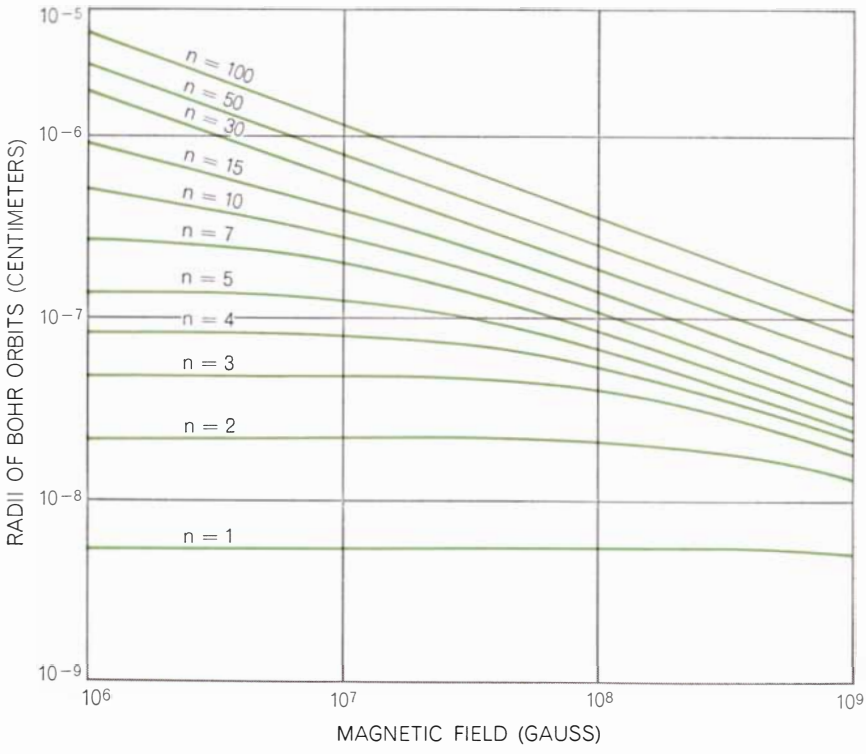
Ultrastrong magnetic fields should also drive bound electrons closer to atomic nuclei and thereby compress the atoms into the shape of sausages with their long axes aligned parallel to the direction of the magnetic field. For a rough calculation of the extent of magnetic compression one can use the simple Bohr model of the hydrogen atom and assume that the electron executes a circular orbit at right angles to



WAVELENGTH OF CYCLOTRON RADIATION is about 2,000 times longer for protons than for electrons in the same magnetic field. In the region between 125 million and 250 million gauss the cyclotron radiation of electrons lies in the visible part of the spectrum.



COMPRESSION OF HYDROGEN ATOM by an ultrastrong magnetic field depends on the atom's state of excitation, expressed by a quantum number *n*, which represents angular momentum. In the absence of a magnetic field the radius of the atom increases as the square of the value of *n*, as shown by the electron orbits depicted in black. If a magnetic field of 100 million gauss were applied, the orbits would have approximately the dimensions shown in color. The scale of the drawing is arbitrary below a radius of 10^{-9} centimeter.



MAGNETIC COMPRESSION of the hydrogen atom is shown for various energy states and for magnetic fields that range in strength from a million to a billion gauss. For hydrogen in its lowest energy state (when *n* equals one), compression is slight even at a billion gauss.

the applied magnetic field. It turns out that the amount of reduction in the size of the orbit depends on the orbit's initial radius, which is proportional to the electron's angular momentum, *n*. The radii of electron orbits for several values of *n* in the absence of a magnetic field and in the presence of a field of 100 million gauss are compared in the illustration at the left. The companion illustration at the bottom of the page presents a family of curves relating orbit radius to magnetic-field strength. When a hydrogen atom is in its lowest energy state (*n* equal to one), it is hardly distorted at all, even in a field of a billion gauss. When it is in an excited state, the Bohr orbits become larger and the magnetic forces exerted by fields as low as 10 million gauss begin to be important. When the atom is in a highly excited state, with *n* equal to between 10 and 100, the magnetic forces are dominant and the cyclotron orbits they produce are only slightly modified by the electrostatic forces exerted by the centrally located proton.

What does all this mean in terms of experiments that can be performed and knowledge that might be gained? It will take years to answer this question, in view of our sketchy knowledge of what to expect. The experiments themselves would presumably involve the observation of either solids or plasmas (assemblages of electrons and other charged particles). One would like to record, for example, the infrared and visible radiation emitted when a plasma interacts with an ultrastrong magnetic field. One might expect atomic spectra as we know them to be supplemented by entirely new patterns representing the Larmor frequency and its harmonics.

As for solids, one may expect intense magnetic fields to change their optical reflectivity, including the way they reflect polarized light. The magnetic properties of atoms, including their magnetic moment, will be sharply modified by ultrastrong fields. Such modifications should produce changes in chemical bonding and crystal structure. The Los Alamos group has already measured changes in the magnetic properties of solids exposed to ultrastrong magnetic fields. It remains to follow up and refine the methods pioneered by Fowler and his associates.

Let me turn finally to the use of ultrastrong magnetic fields for exploring the properties of matter under the conditions of extreme pressure found in the earth's interior. Although mechanical presses for creating static pressures are

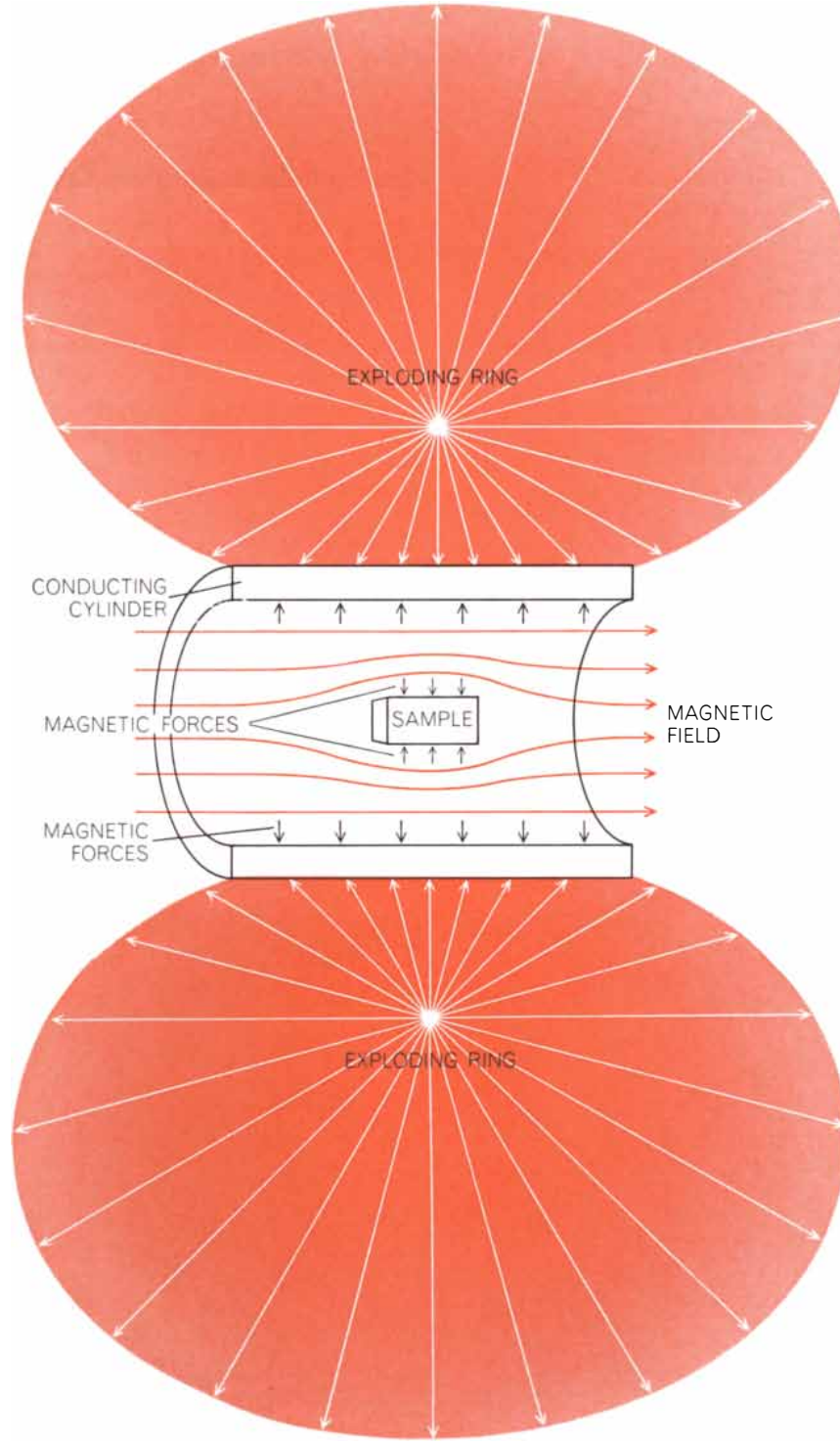
being steadily improved, the maximum pressure they can exert seems limited to the region below 500,000 atmospheres. To achieve pressures of a million to three million atmospheres, which exist deep within the earth, there appears to be no alternative but to use explosives. It is here that new techniques for transmitting the energy of an explosion to a sample are much to be desired.

The standard method for transmitting the energy of explosives to a sample is simply to make use of the shock wave produced by an explosion. In this method the condition of interest—maximum pressure—is not created until the shock wave arrives, and then it is quickly over. If a magnetic field were used for pressure transmission, however, the pressure of the explosion would be transmitted virtually as soon as it began to build up and it would rise steadily until the peak value was reached. As Fowler and his colleagues pointed out in 1960, the ability of a magnetic field to transmit forces at the velocity of light should substantially increase the kinds of observations that can be made.

The illustration at the right shows schematically how a magnetic field might be used to transmit the intense pressures created by an explosion. As in the earlier Los Alamos work, a magnetic field is generated inside a cylinder and is compressed by an implosion. Inside the cylinder there is another chamber that houses the sample to which pressure is to be transmitted. The chamber consists of a metal box that excludes the ultrastrong magnetic field created by the implosion; consequently the pressure of the magnetic field is exerted continuously against the metal box as the field climbs to a peak.

It is not possible to predict the ultimate value of this technique. It may, in fact, turn out to be valueless because of unforeseen difficulties. If it works as hoped, however, it should offer many advantages for experimentation. Perhaps the biggest advantage is that the magnetic field is completely transparent and should thus provide a clear window through which the sample can be observed by high-speed photography during its final collapse. In addition, the optical properties of the sample can be followed and measured. Another potential advantage is that one might be able to vary the rate at which pressure is applied by varying the rate of increase of the magnetic field. For instance, one might construct a nest of nonconducting cylinders inside the outer conducting cylinder that compresses the magnetic flux when it collapses. These nonconducting cylinders would not interfere with the compression of flux, but they would add mass to the assembly and so decrease the velocity of implosion.

Such speculations serve merely to illustrate the need for a carefully conceived program of experimentation. The goal is the development of techniques that seem to offer a unique opportunity for studying the distortion of matter at extreme pressures.



MAGNETIC TRANSMISSION OF PRESSURE might be achieved by an implosion device of this general design. The magnetic flux trapped in the conducting cylinder is compressed. The sample is contained in a metal box that excludes the magnetic field and so the box too is subjected instantaneously to the full pressure generated within the collapsing cylinder. The magnetic field provides a clear window through which the sample can be observed.

The Role of Chlorophyll in Photosynthesis

The pigments of plants trap light energy and store it as chemical energy. They do this by catalyzing an oxidation-reduction process in which hydrogen atoms are boosted from water to organic matter

by Eugene I. Rabinowitch and Govindjee

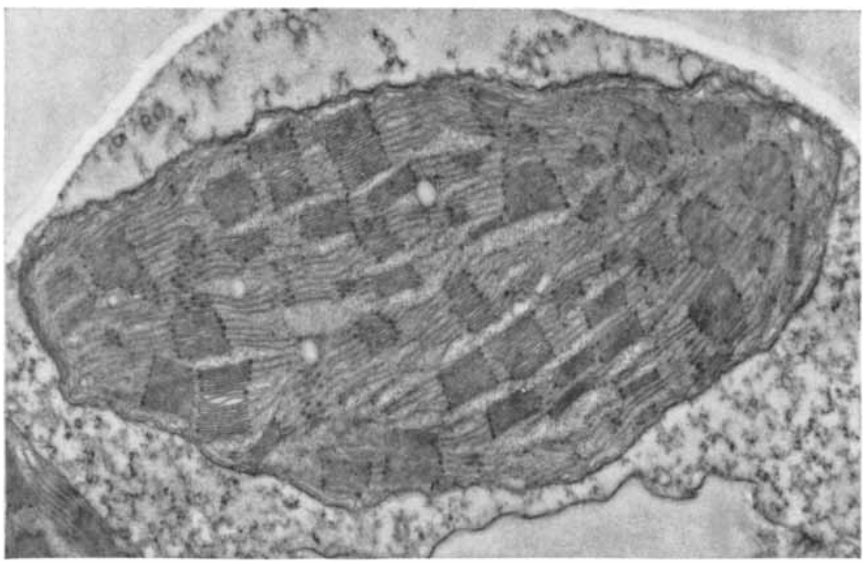
Any effort to understand the basis of life on this planet must always come back to photosynthesis: the process that enables plants to grow by utilizing carbon dioxide (CO_2), water (H_2O) and a tiny amount of minerals. Photosynthesis is the one large-scale process that converts simple, stable, inorganic compounds into the energy-rich combination of organic matter and oxygen and thereby makes abundant life on earth possible. Photosynthesis is the source of all living matter on earth, and of all biological energy.

The overall reaction of photosynthesis can be summarized in the following equation: $CO_2 + H_2O + \text{light} \rightarrow (CH_2O) + O_2 + 112{,}000$ calories of energy per mole. (CH_2O) stands for a carbohydrate; for example, glucose: $(CH_2O)_6$. "Mole" is short for "gram molecule": one gram multiplied by the molecular weight of the substances in question—in this case carbohydrate and oxygen.

When one of us first summarized the state of knowledge of photosynthesis 17 years ago, the whole process was still heavily shrouded in fog [see "Photosynthesis," by Eugene I. Rabinowitch; Scientific American, August, 1958]. Five years later investigation had penetrated the mists sufficiently to disclose some of the main features of the process [see "Progress in Photosynthesis," by Eugene I. Rabinowitch; Scientific American, November, 1953]. Since then much new knowledge has been accumulated; in particular the sequence of chemical steps that convert carbon dioxide into carbohydrate is now understood in considerable detail [see "The Path of Carbon in Photosynthesis," by J. A. Bassham; Scientific American, June, 1962]. The fog has also thinned out in other areas, and the day when the entire sequence of physical and chemical events in photosynthesis will be well understood seems much closer.

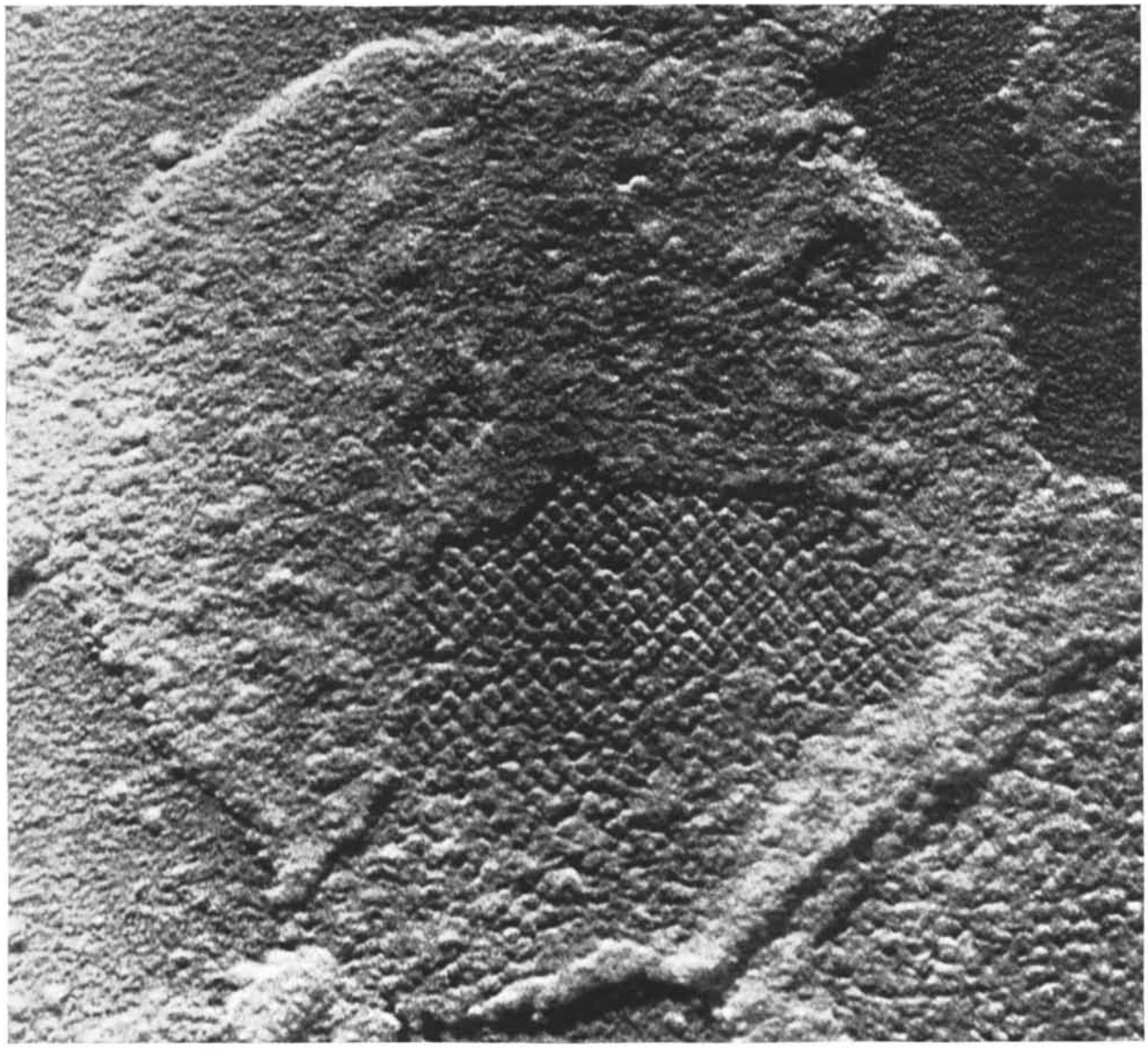
The photosynthetic process apparently consists of three main stages: (1) the removal of hydrogen atoms from water and the production of oxygen molecules; (2) the transfer of the hydrogen atoms from an intermediate compound in the first stage to one in the third stage, and (3) the use of the hydrogen atoms to convert carbon dioxide into a carbohydrate [*see illustration on page 76*].

The least understood of these three stages is the first: the removal of hydrogen atoms from water with the release of oxygen. All that is known is that it entails a series of steps probably requiring several enzymes, one of which contains manganese. The third stage—the production of carbohydrates from carbon dioxide—is the best understood, thanks largely to the work of Melvin Calvin and his co-workers at the University of California at Berkeley. The subject of our article is the second stage: the transfer of hydrogen atoms from the first stage to the third. This is the energy-storing part of photosynthesis; in it, to use the words of Robert Mayer, a discoverer of the law of the conservation of energy, "the fleeting sun rays are fixed and skillfully stored for future use."



CHLOROPLAST is the organelle in a plant cell within which photosynthesis takes place. The chlorophyll is contained in the "grana," stacks of membranous sacs called lamellae, seen here in cross section. A maize-cell chloroplast is enlarged 19,000 diameters in this electron micrograph made by A. E. Vatter of the University of Colorado Medical Center.

The light energy to be converted into chemical energy by photosynthesis is first taken up by plant pigments, primarily the green pigment chlorophyll. In photosynthesis chlorophyll functions as a photocatalyst: when it is in its energized state, which results from the absorption of light, it catalyzes an energy-storing chemical reaction. This



PHOTOSYNTHETIC UNITS may be the small elements, looking somewhat like cobblestones, visible in this electron micrograph made by Roderic B. Park and John Biggins of the University of California at Berkeley. In the micrograph a single lamella and a part of another one are shadowed with chromium and enlarged 175,000 diameters. Where the membrane is torn away one can see an ordered array of the units, which Park and Biggins call quantasomes and calculate could contain 230 chlorophyll molecules each.

reaction is the primary photochemical process; it is followed by a sequence of secondary "dark"—that is, nonphotochemical—reactions in which no further energy is stored.

Once it was thought that in photosynthesis the primary photochemical process is the decomposition of carbon dioxide into carbon and oxygen, followed by the combination of carbon and water. More recently it has been suggested that the energy of light serves primarily to dissociate water, presumably into hydroxyl radicals (OH) and hydrogen atoms; the hydroxyl radicals would then react to form oxygen molecules. It is better than either of these two formulations to say that the primary photochemical process in photosynthesis is the boosting of hydrogen atoms from a stable association with oxygen in water molecules to a much less stable one with carbon in organic matter. The oxygen atoms "left behind" combine into oxygen molecules, an association also much less stable than the one between oxygen and hydrogen in water. The replacement of stable bonds (between oxygen and hydrogen) by looser bonds (between oxygen and oxygen and between hydrogen and carbon) obviously requires a supply of energy, and it explains why energy is stored in photosynthesis.

The transfer of hydrogen atoms from one molecule to another is called oxidation-reduction. The hydrogen atom is transferred from a donor molecule (a "reductant") to an acceptor molecule (an "oxidant"); after the reaction the donor is said to be oxidized and the acceptor to be reduced. The transfer of an electron can often substitute for the transfer of a hydrogen atom: in an aqueous system (such as the interior of the living cell) there are always hydrogen ions (H^+), and if such an ion combines with the electron acceptor, the acquisition of an electron becomes equivalent to the acquisition of a hydrogen atom (electron + H^+ ion $\rightarrow$ H atom).

The chain of oxidation-reduction reactions in photosynthesis has some links that involve electron transfers and

others that involve hydrogen-atom transfers. For the sake of simplicity we shall speak of electron transfers, with the understanding that in some cases what is actually transferred is a hydrogen atom. Indeed, the end result of the reactions undoubtedly *is* the transfer of hydrogen atoms.

In the oxidation-reduction reactions of photosynthesis the electrons must be pumped "uphill"; that is why energy must be supplied to make the reaction go. The tiny chlorophyll-containing chloroplasts of the photosynthesizing plant cell act as chemical pumps; they obtain the necessary power from the absorption of light by chlorophyll (and to some extent from absorption by other pigments in the chloroplast). It is important to realize that the energy is stored in the two products organic matter and free oxygen and not in either of them separately. To release the energy by the combustion of the organic matter (or by respiration, which is slow, enzyme-catalyzed combustion) the two products must be brought together again.

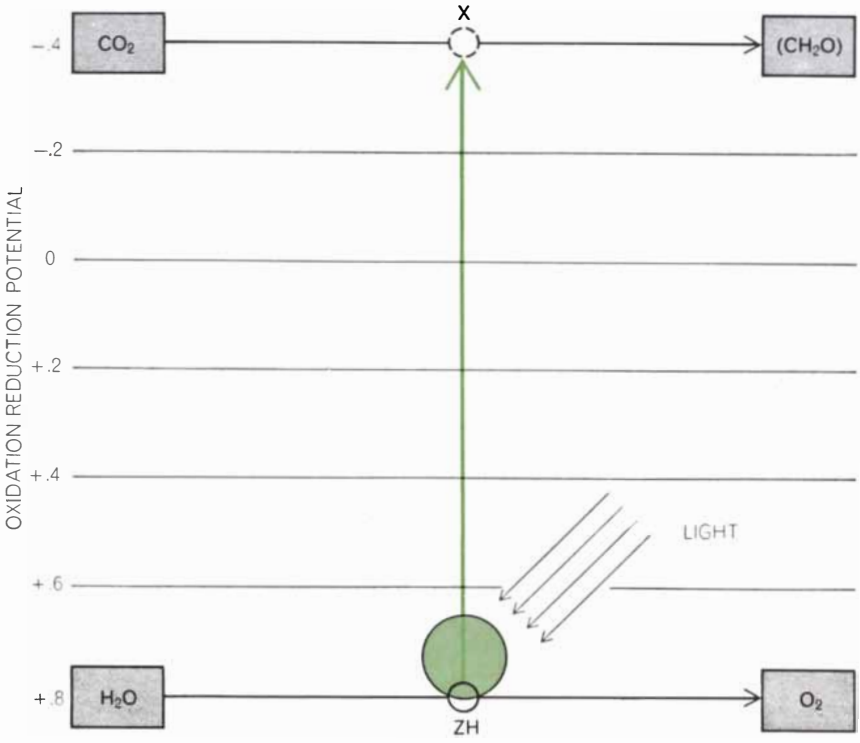
How much energy is stored in the transfer of electrons from water to carbon dioxide, converting the carbon dioxide to carbohydrate and forming a proportionate amount of oxygen?

Oxidation-reduction energy can conveniently be measured in terms of electrochemical potential. Between a given donor of electrons and a given acceptor there is a certain difference of oxidation-reduction potentials. This difference depends not only on the nature of the two reacting substances but also on the nature of the products of the reaction; it is characteristic of the two oxidation-reduction "couples." For example, when oxygen is reduced to water (H_2O) its potential is +.81 volt, but when it is reduced to hydrogen peroxide (H_2O_2) the potential is +.27 volt. The more positive the potential, the stronger is the oxidative power of the couple; the more negative the potential, the stronger is its reducing power.

When two oxidation-reduction couples are brought together, the one containing the stronger oxidant tends to oxidize the one containing the stronger reductant. In photosynthesis, however, a weak oxidant (CO_2) must oxidize a weak reductant (H_2O), producing a strong oxidant (O_2) and a strong reductant (a carbohydrate). This calls for a massive investment of energy. The specific amount needed is given by the difference between the oxidation-reduction potentials of the two couples involved in the reaction: oxygen-water and carbon-dioxide–carbohydrate. The oxygen-water potential is about +.8 volt; the carbon-dioxide–carbohydrate potential, about −.4 volt. The transfer of a single electron from water to carbon dioxide thus requires +.8 minus −.4, or 1.2, electron volts of energy. For a molecule of carbon dioxide to be reduced to CH_2O—the elementary molecular group of a carbohydrate—*four* electrons (or hydrogen atoms) must be transferred; hence the total energy needed is 4.8 electron volts. This works out to 112,000 calories per mole of carbon dioxide reduced and of oxygen liberated. In short, the pumping of electrons in the second stage of photosynthesis entails the storage of 112,000 calories of energy per mole for each set of four electrons transferred.

We know the identity of the primary electron donor in photosynthesis (water) and of the ultimate electron acceptor (carbon dioxide), but what are the intermediates involved in the transfer of electrons from the first stage to the third? This has become the focal problem in recent studies of the photosynthetic process. As a matter of fact, it is not yet definitely known what compound releases electrons from the first stage, and what compound receives them in the third; that is why these compounds are respectively labeled *ZH* and *X* in the illustration at the left. About the donor, *ZH*, we have almost no information; the following considerations suggest the possible nature of the primary acceptor, *X*.

From the study of the mechanism of respiration we are familiar with an important oxidation-reduction catalyst: nicotinamide adenine dinucleotide phosphate, or NADP (formerly known as triphosphopyridine nucleotide, or TPN). NADP has an oxidation-reduction potential of about −.32 volt, thus in itself it is not a strong enough reductant to provide the −.4-electron-volt potential needed to reduce carbon dioxide to carbohydrate. NADP can achieve this feat, however, if it is supplied with additional energy in the form of the high-energy compound adenosine triphosphate, or ATP. A molecule of ATP supplies about 10,000 calories per mole when its terminal phosphate group is split off, and this is enough to provide the needed boost to the reducing power of NADP. Furthermore, we know that NADP is reduced when cell-free preparations of chloroplasts are illuminated. Put together, these two facts led to the now widely accepted hypothesis that the sec-



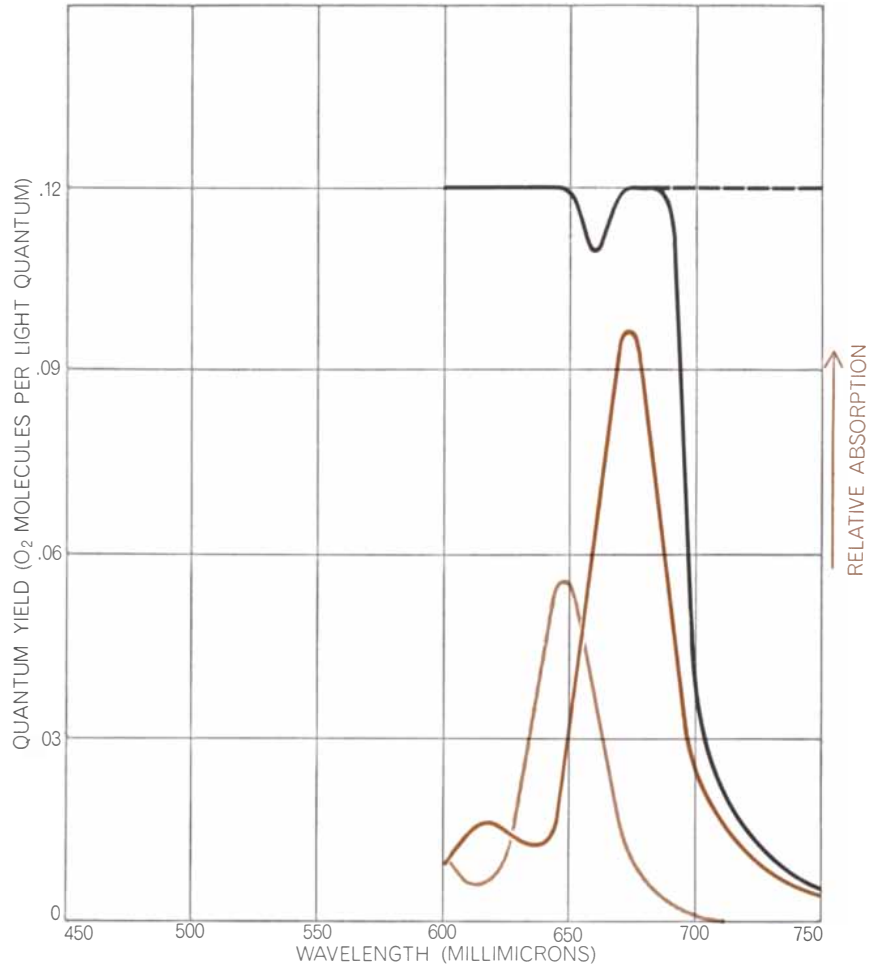
THREE STAGES of photosynthesis are the removal of hydrogen from water with the release of oxygen (*bottom arrow*), the transfer (*vertical arrow*) of the hydrogen by energy from light trapped by chlorophyll (*color*) and the use of the hydrogen to reduce carbon dioxide to carbohydrate (*top arrow*). In this scheme the oxidation-reduction potentials involved are indicated by the scale at the left, and the hypothetical "primary reductant" and "primary acceptor" intermediates are designated as *ZH* and *X* respectively.

ond stage of photosynthesis manufactures both ATP and reduced NADP and feeds them into the third stage.

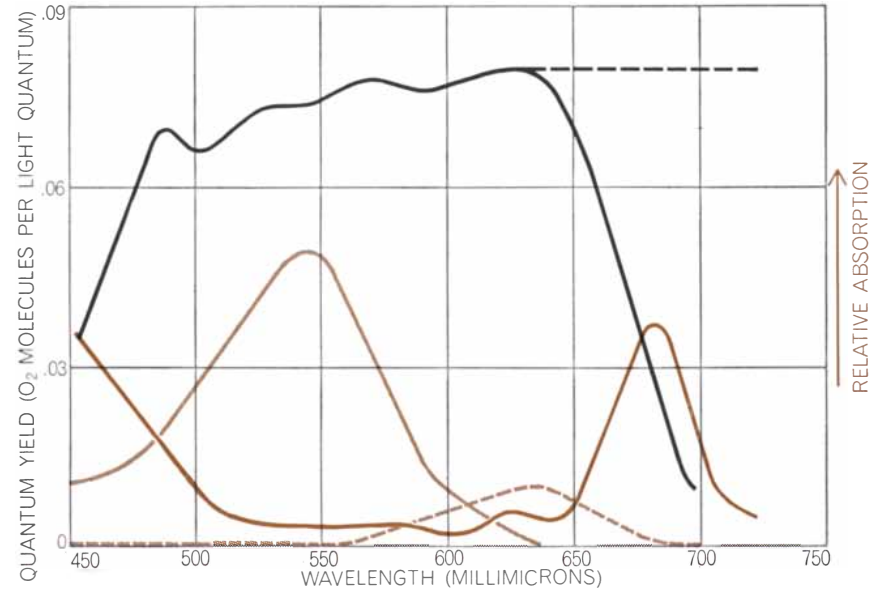
At first it was assumed that NADP is identical with *X*, the primary acceptor in our scheme. Subsequent experiments by various workers—notably Anthony San Pietro at Johns Hopkins University and Daniel I. Arnon and his colleagues at the University of California at Berkeley—suggested, however, that NADP is preceded in the "bucket brigade" of electron transfer by ferredoxin, a protein that contains iron. This compound has an oxidation-reduction potential of about —.42 volt; therefore if it is reduced in light it can bring about the reduction of NADP by a "dark" reaction requiring no additional energy supply.

More recently Bessel Kok of the Research Institute for Advanced Studies in Baltimore has found evidence suggesting that compound *X* may be a still stronger reductant, with a potential of about —.6 volt. If this is so, plants have the alternatives of either applying this stronger reductant directly to the reduction of carbon dioxide or letting it reduce first ferredoxin and then NADP and using reduced NADP to reduce carbon dioxide. It seems a roundabout procedure to create a reductant sufficiently strong for the task at hand, then to sacrifice a part of its reducing power and finally to use ATP to compensate for the loss. It is not unknown, however, for nature to resort to devious ways in order to achieve its aims.

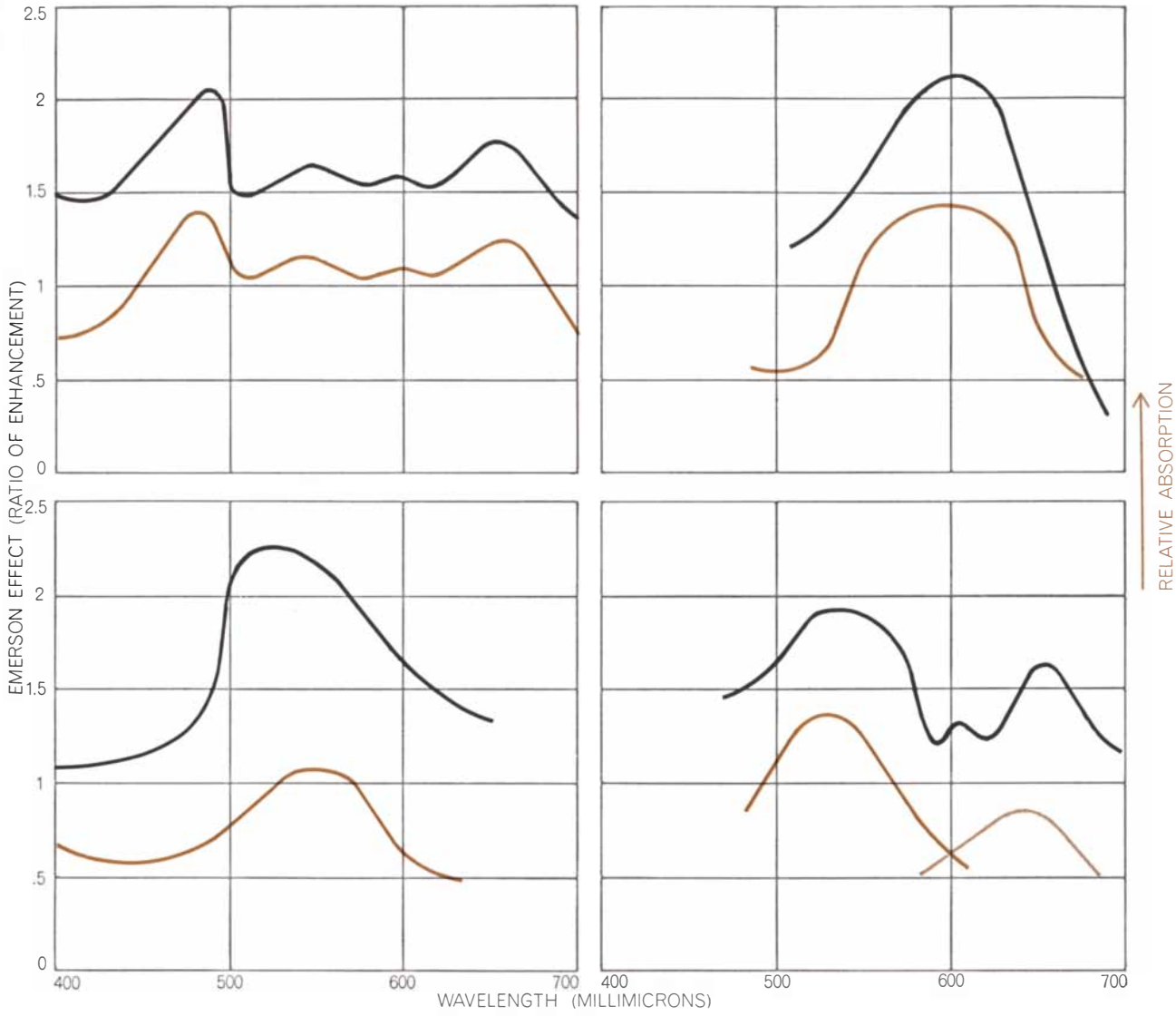
For photosynthesis to be a self-contained process the required high-energy phosphate ATP must be itself manufactured by photosynthesis. The formation of ATP has in fact been detected in illuminated fragments of bacteria by Albert W. Frenkel of the University of Minnesota and in chloroplast fragments by Arnon and his co-workers [see "The Role of Light in Photosynthesis," by Daniel I. Arnon: Scientific American, November, 1960]. As a matter of fact, ATP is needed not only to act as a booster in the reduction of an intermediate in the carbon cycle by reduced NADP but also for another step in the third stage of photosynthesis. According to a sequence of reactions worked out in 1951 by Andrew A. Benson and his colleagues at the University of California at Berkeley, carbon dioxide enters photosynthesis by first reacting with a "carbon dioxide acceptor," a special sugar phosphate called ribulose diphosphate. It turns out that the production of this compound from its precursor—



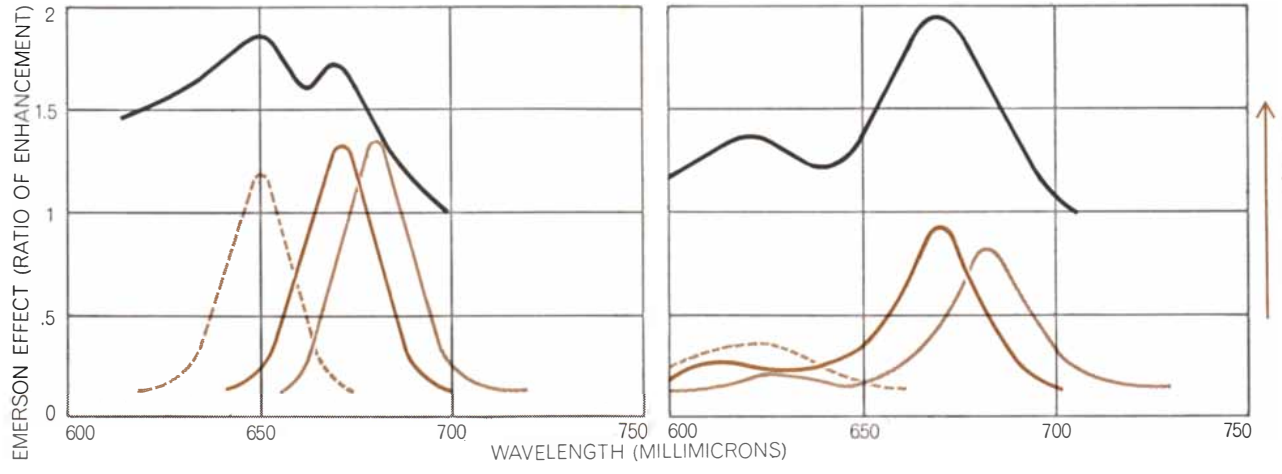
"RED DROP," the drop in quantum yield (*black curve*) of oxygen in photosynthesis under long-wave illumination, is demonstrated in the green alga *Chlorella pyrenoidosa*. Peak efficiency is restored (*broken line*) by supplementary shorter-wave illumination. Absorption curves of chlorophylls *a* (*solid color*) and *b* (*light color*) are also shown. This illustration and the next two are based on data of the late Robert Emerson of the University of Illinois.



QUANTUM YIELD is similarly affected by long-wave illumination in the red alga *Porphyridium cruentum* (*black curve*). Yield drops to less than half of its maximum when absorption by chlorophyll *a* (*solid-color curve*) is at its peak. Absorption of the pigments phycoerythrin (*light-color curve*) and phycocyanin (*broken-color curve*) are also shown.



EMERSON EFFECT is shown for *Chlorella* (*top left*), the blue-green alga *Anacystis nidulans* (*top right*), *Porphyridium* (*bottom left*) and the diatom *Navicula minima* (*bottom right*). In each case the black curve shows the action spectrum of the Emerson effect, or the degree of enhancement in quantum yield as the wavelength of the supplementary illumination is varied. The curve of the action spectrum turns out to be parallel to the absorption curves (*color*) of the various accessory pigments: chlorophyll *b* in *Chlorella*, phycocyanin in *Anacystis*, phycoerythrin in *Porphyridium* and fucoxanthol (*solid color*) and chlorophyll *c* (*light color*) in *Navicula*.



DETAILED ACTION SPECTRA of the Emerson effect reveal the presence of chlorophyll *a* 670 in *Chlorella* (*left*) and *Navicula* (*right*). The Emerson-effect peaks coincide with the absorption peaks of chlorophyll *a* 670 (*solid color*) as well as of chlorophyll *b* in *Chlorella* and chlorophyll *c* in *Navicula* (*broken curves*). The chlorophyll *a* absorption curve is also shown (*light-color curve*).

ribulose monophosphate—calls for a molecule of ATP.

ATP is produced both in chloroplasts and in mitochondria, the tiny intracellular bodies that are the site of the energy-liberating stage of respiration in animals as well as plants. The mitochondria produce ATP as their main function, exporting it as packaged energy for many life processes. The chloroplasts, on the other hand, make ATP only as an auxiliary source of energy for certain internal purposes. The energy of the light falling on the chloroplasts is stored mostly as oxidation-reduction energy by the uphill transfer of electrons. Only a relatively small fraction is diverted to the formation of ATP, and this fraction too ultimately becomes part of the oxidation-reduction energy of the final products of photosynthesis: oxygen and carbohydrate.

Let us now consider the uphill transport of electrons in greater detail. Recent investigations have yielded considerable information about this stage. Apparently the pumping of the electrons is a two-step affair, and among the most important intermediates in it are the catalysts called cytochromes.

The idea of a two-step electron-transfer process grew from a consideration of the energy economy of photosynthesis. Precise measurements, particularly those made by the late Robert Emerson and his co-workers at the University of Illinois, showed that the reduction of one molecule of carbon dioxide to carbohydrate, and the liberation of one molecule of oxygen, requires a minimum of eight quanta of light energy. The maximum quantum yield of photosynthesis, defined as the number of oxygen molecules that can be released for each quantum of light absorbed by the plant cell, is thus 1/8, or 12 percent. Since the transfer of four electrons is involved in the reduction of one carbon dioxide molecule, it was suggested that it takes two light quanta to move each electron.

Emerson and his colleagues went on to determine the quantum yield of photosynthesis in monochromatic light of different wavelengths throughout the visible spectrum. They found that the yield, although it remained constant at about 12 percent in most of the spectrum, dropped sharply near the spectrum's far-red end [*see illustrations on page 77*]. This decline in the quantum yield, called the "red drop," begins at a wavelength of 680 millimicrons in green plants and at 650 millimicrons in red algae.

There are two chlorophylls present in green plants: chlorophyll *a* and chlorophyll *b*. Only chlorophyll *a* absorbs light at wavelengths longer than 680 millimicrons; the absorption of chlorophyll *b* rises to a peak at 650 millimicrons and becomes negligible at about 680 millimicrons. Emerson found that the quantum yield of photosynthesis at the far-red end of the spectrum beyond 680 millimicrons can be brought to the full efficiency of 12 percent by simultaneously exposing the plant to a second beam of light with a wavelength of 650 millimicrons. In other words, when light primarily absorbed by chlorophyll *a* was supplemented by light primarily absorbed by chlorophyll *b*, both beams gave rise to oxygen at the full rate. This relative excess in photosynthesis when a plant is exposed to two beams of light simultaneously, as compared with the yield produced by the same two beams separately, is known as the Emerson effect, or enhancement.

On the basis of his discovery Emerson concluded that photosynthesis involves two photochemical processes: one using energy supplied by chlorophyll *a*, the other using energy supplied by chlorophyll *b* or some other "accessory" pigment. Experimenting with various combinations of a constant far-red beam with beams of shorter wavelength, and using four different types of algae (green, red, blue-green and brown), Emerson's group found that the strongest enhancement always occurred when the second beam was absorbed mainly by the most important accessory pigment (the green pigment chlorophyll *b* in green cells, the red pigment phycoerythrin in red algae, the blue pigment phycocyanin in blue-green algae and the reddish pigment fucoxanthol in brown algae). Such results suggested that these other pigments are not mere accessories of chlorophyll *a* but have an important function of their own in photosynthesis [*see top illustration on opposite page*].

Certain findings concerning the behavior of pigments in living plant cells, however, seemed to make this conclusion untenable. Illuminated plant cells fluoresce; that is, pigment molecules energized by the absorption of light quanta reemit some of the absorbed energy as fluorescent light. The source of fluorescence can be identified, because each substance has its own characteristic fluorescence spectrum. The main fluorescing pigment in plants always proves to be chlorophyll *a*, even when the light is absorbed by another pigment. This had first been shown for brown algae in a study conducted in 1943 by H. J. Dutton, W. H. Manning and B. B. Duggar at the University of Wisconsin; later the finding was extended to other organisms by L. N. M. Duysens of the University of Leiden. Known as sensitized fluorescence, the phenomenon indicates that the initial absorber has transferred its energy of excitation to chlorophyll *a;* the transfer is effected by a kind of resonance process. Careful measurements have shown that certain accessory pigments—chlorophyll *b*, phycoerythrin, phycocyanin and fucoxanthol—pass on to chlorophyll *a* between 80 and 100 percent of the light quanta they absorb. For some other accessory pigments—for example carotene—the transfer is less efficient.

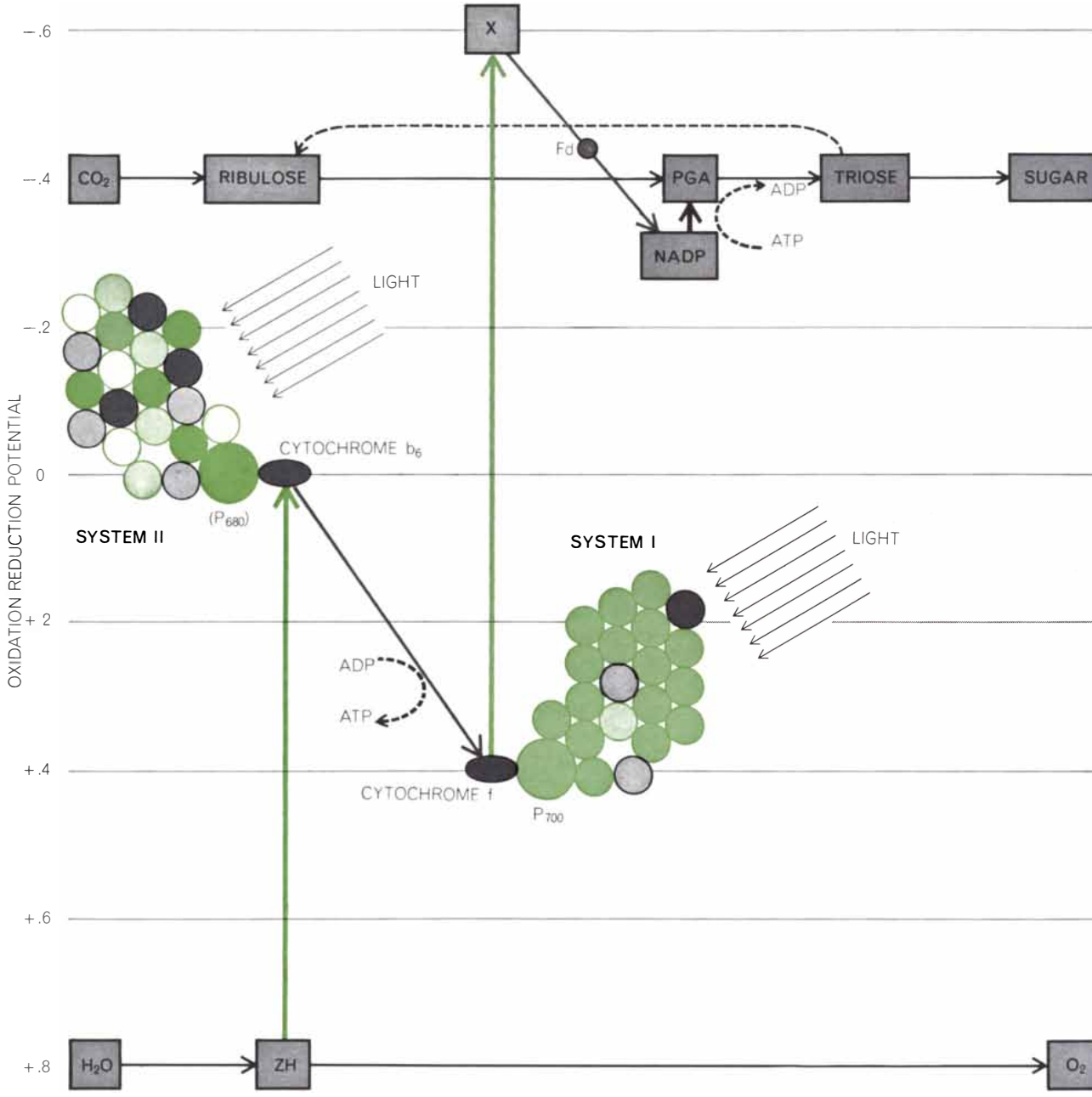
This puts accessory pigments back in the role of being mere adjuncts to chlorophyll *a*. True, they can contribute, by means of resonance transfer, light energy to photosynthesis, thereby improving the supply of energy in regions of the spectrum where chlorophyll *a* is a poor absorber. Chlorophyll *a*, however, collects all this energy before it is used in the primary photochemical process. Why, then, the enhancement effect? Why should chlorophyll *a* need, in order to give rise to full-rate photosynthesis, one "secondhand" quantum obtained by resonance transfer from another absorber in addition to the one quantum it had absorbed itself?

A better understanding of this paradox resulted from the discovery that there apparently exist in the cell not only chlorophyll *a* and chlorophyll *b* but also two forms of chlorophyll *a*. These two forms have different light-absorption characteristics, and they probably also have different photochemical functions.

In the living cell chlorophyll *a* absorbs light most strongly in a broad band with its peak between 670 and 680 millimicrons. In our laboratory at the University of Illinois we undertook to plot the Emerson effect more carefully than before as a function of the wavelength of the enhancing light. We found that for green and brown cells the resulting curve showed, in addition to peaks corresponding to strong absorption by the accessory pigments, a peak at 670 millimicrons that must be due to chlorophyll *a* itself [*see bottom illustration on opposite page*]. It was this finding that suggested the existence of two forms of chlorophyll *a*. The form that absorbs light at the longer wavelengths—mainly above 680 millimicrons—seemed to belong to one pigment system, now often called System I. The form that absorbs at 670 millimicrons seemed to

belong to another pigment system: System II. In the second system the form of chlorophyll *a* that absorbs at 670 millimicrons is strongly assisted by accessory pigments, probably by resonance transfer of their excitation energy. Careful analysis of the absorption band of chlorophyll *a* by C. Stacy French of the Carnegie Institution of Washington's Department of Plant Biology, and also in our laboratory, confirmed that the band is double, with one peak near 670 millimicrons (at 668 millimicrons) and another band at 683 millimicrons [*see bottom illustration on page 78*].

If chlorophyll *a* is extracted from living plants, there is only one product; we must therefore assume that in the living cell the two forms differ in the way molecules of chlorophyll *a* are clumped together, or in the way they are associated with different chemical partners (proteins, lipids or other substances). Be this as it may, the important implication of the new finding is that photosynthesizing cells possess two light-absorbing systems, one containing a form of chlorophyll *a* absorbing around 683 millimicrons and the other a form absorbing around 670. The latter system includes chlorophyll *b* (in green-plant cells) or other accessory pigments (in brown, red and blue-green algae). Further investigation—particularly of red algae—has suggested, however, that the distribu-



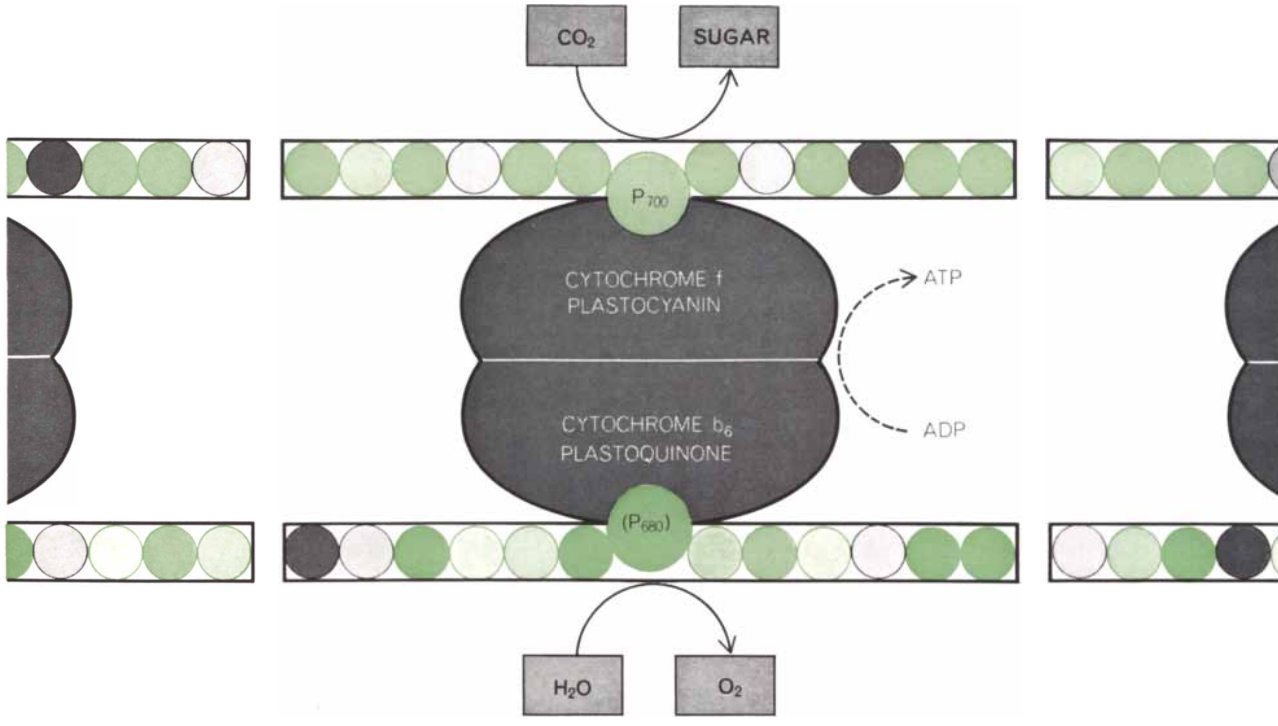
HYDROGEN TRANSFER in photosynthesis is now conceived of as a two-step process involving two pigment systems. Hydrogen atoms (or electrons) from the donor (*ZH*) are boosted to cytochrome b_6 by energy collected in System II and trapped by a hypothetical "pigment 680" (*P 680*). The pigments of System II include chlorophyll *a* 670 and such accessory pigments as chlorophyll *b* or *c*, phycoerythrin or phycocyanin, depending on the plant. The electrons are passed "downhill" to cytochrome *f*, synthesizing adenosine triphosphate (*ATP*) in the process. Energy from System I (primarily chlorophyll *a*, with some accessory pigments), trapped by pigment 700 (*P 700*), boosts the electrons to a receptor (*X*), whence they move via ferredoxin (*Fd*) to nicotinamide adenine dinucleotide phosphate (*NADP*). Energy from ATP helps to move the electrons to phosphoglyceric acid (*PGA*) and into the carbon cycle.

tion of these two components in the two systems may be less clear-cut. In red algae a large fraction of the chlorophyll *a* absorbing at 670 millimicrons seems to belong to System I rather than System II.

In all likelihood the two systems provide energy for two different photochemical reactions, and efficiency in photosynthesis requires that the rates of the two reactions be equal. What are these reactions? This question brings us to another significant finding, which suggested the participation of cytochromes in photosynthesis.

Cytochromes are proteins that carry an iron atom in an attached chemical group. They are found in all mitochondria, where they serve to catalyze the reactions of respiration. Robert Hill and his co-workers at the University of Cambridge first found that chloroplasts also contain cytochromes—two kinds of them. One, which they named cytochrome *f*, has a positive oxidation-reduction potential of about .4 volt. The other, which they named cytochrome b_6, has a potential of about 0 volt. In 1960 Hill, together with Fay Bendall, proposed an ingenious hypothesis as to how the two cytochromes might act as intermediate carriers of electrons and connect the two photochemical systems [*see illustration on opposite page*]. They suggested that cytochrome b_6 receives an electron by a photochemical reaction from the electron donor *ZH;* the electron is then passed on to cytochrome *f* by a "downhill" reaction requiring no light energy. (The oxidation-reduction potential of cytochrome *f* is much more positive than that of cytochrome b_6.) A second photochemical reaction moves the electron uphill again, from cytochrome *f* to the electron-acceptor *X* in the third stage of photosynthesis. In this sequence the photochemical reactions store energy and the reaction between the two cytochromes releases energy. Some of the released energy, however, can be salvaged by the formation of an ATP molecule; this occurs in the transfer of electrons among cytochromes in respiration. In this way ATP is obtained without spending extra light quanta on its formation, which the tight energy economy of photosynthesis does not allow.

Experiments by Duysens and his associates confirmed this hypothesis, by showing that the absorption of light by System I causes the oxidation of a cytochrome, whereas the absorption of light by System II causes its reduction. This is exactly what we would expect. The illustration on the opposite page shows that the light reaction of System II



HYPOTHETICAL ARRANGEMENT of pigments in a chloroplast lamella would have System I in a monomolecular layer at the top and System II at the bottom. The space between the two pigment layers might contain the compounds responsible for the transport of hydrogen atoms (or electrons). The water-to-oxygen cycle would then be linked to System II and the carbon cycle to System I.

should flood the intermediates between the two photochemical reactions with electrons taken from *ZH;* the light reaction of System I should drain these electrons away, sending them up to the acceptor *X* and into the third stage of photosynthesis.

This, then, describes in a general way the oxidation-reduction process by which the chloroplasts store the energy of light in photosynthesis. Several other investigators have contributed evidence for the two-step mechanism; notable among them are French, Kok, Arnon, Horst Witt of the Max-Vollmer Institute in Berlin and their colleagues. In detail the process probably is much more complex than our scheme suggests. Its "downhill" central part seems to include, in addition to the two cytochromes, certain compounds of the group known as quinones and also plastocyanin, a protein that contains copper.

What is known of the submicroscopic structure in which the reactions of the second stage of photosynthesis take place? There is much evidence that the photosynthetic apparatus consists of "units" within the chloroplasts, each unit containing about 300 chlorophyll molecules. This picture first emerged from experiments conducted in 1932 by Emerson and William Arnold on photosynthesis during flashes of light; it was later supported by various other observations. The pigment molecules are packed so closely in the unit that when one of them is excited by light it readily transfers its excitation to a neighbor by resonance. The energy goes on traveling through the unit, rather as the steel ball in a pinball machine bounces around among the pins and turns on one light after another. Eventually the migrating energy quantum arrives at the entrance to an enzymatic "conveyor belt," where it is trapped and utilized either to load an electron onto the belt or to unload one from it. (The steel-ball analogy should not be taken literally; the migration of energy is a quantum-mechanical phenomenon, and the quantum's location can only be defined in terms of probability; its entrapment depends on the probability of finding it at the entrance to the conveyor belt.)

How is the quantum trapped? The trap must be a pigment molecule with what is called a lower excited state; the migrating quantum can stumble into such a molecule but cannot come out of it. Kok has found evidence that System I contains a small amount of a special form of chlorophyll called pigment 700 because it absorbs light at a wavelength of 700 millimicrons; this pigment could serve as a trap for the quantum bouncing around in System I. There seems to be a proper amount of pigment 700: about one molecule per unit. Furthermore, experiments suggest that pigment 700 is oxidized by light absorbed in System I and reduced by light absorbed in System II. It has an oxidation-reduction potential of about +.4 volt. All these properties fit the role we have assigned pigment 700 in our scheme: collecting energy from a 300-molecule unit in System I, using it to transfer an electron to the acceptor *X* and recovering the electron from cytochrome *f* [*see illustration on page 80*].

One suspects that there should be a counterpart of pigment 700 in System II, but so far none has been convincingly demonstrated. We believe, however, that a pigment we have tentatively named pigment 680—from the anticipated position of its absorption band—does serve as an energy trap in System II. Its existence is supported by the discovery of a new fluorescent emission band of chlorophyll at 693 millimicrons, which is compatible with absorption at 680 millimicrons. This band is emitted by certain algae when they are exposed to strong light of the wavelengths absorbed by System II.

What is the spatial organization of the pigment systems in the electron-boosting mechanism of the second stage of photosynthesis? It seems that the two systems may be arranged in two mono-

molecular layers, with a protein layer between them containing the enzymatic conveyor belt [*see illustration on opposite page*]. The chloroplasts are known from electron microscope studies to consist of a set of lamellae: thin alternating layers of protein and fatty material piled one atop the other. Each layer appears to consist of particles arrayed rather like cobblestones in a pavement. The particles were first observed in electron micrographs made by E. Steinmann of the Technische Hochschule in Zurich; subsequently Roderic B. Park and John Biggins of the University of California at Berkeley made clearer micrographs of the particles and named them quantasomes [*see illustration on page 75*]. The units comprising Systems I and II may operate independently or they may be sufficiently close together to exchange energy by resonance, when such exchange is needed to maintain a balanced rate of operation by the two systems.

The picture of the energy-storing second stage of photosynthesis presented in this article is, of course, still only a working hypothesis. Alternative hypotheses are possible, one of which we shall briefly describe. For many years the late James Franck, who shared the Nobel prize in physics for 1925, tried to develop a plausible physicochemical mechanism of photosynthesis. In 1963 he proposed, together with Jerome L. Rosenberg of the University of Pittsburgh, a concept according to which the two consecutive photochemical steps occur in one and the same energy trap. In other words, according to Franck, the same chlorophyll molecule that takes the electron away from the initial donor *ZH* and transfers it to a cytochrome then supplies energy for the transfer of the electron from the cytochrome to the acceptor *X*. In the first transfer, Franck suggested, the chlorophyll *a* molecule functions in the short-lived "singlet" excited state (in which the valence electrons have opposite spins); in the second transfer it functions in the long-lived "triplet" state (in which the valence electrons have parallel spins). Franck's hypothesis avoids certain difficulties of the "two trap" theory, but new difficulties arise in their place. On balance the two-trap picture seems to us the more plausible one at present.

No doubt this picture will change as more information emerges. It is merely a first effort to penetrate the inner sanctum of photosynthesis, the photocatalytic laboratory in which the energy of sunlight is converted into the chemical energy of life.

ANCIENT JERUSALEM

A sprawling city today, Jerusalem was nothing but a fortified ridge in the desert when David seized it 3,000 years ago. New excavations are revealing its successive stages of development

by Kathleen M. Kenyon

Over the centuries there has been a succession of cities on the site of Jerusalem. The city was a Bronze Age Jebusite stronghold until its capture by David about 1000 B.C., and the seat of the Jewish monarchy until its destruction by the Babylonians in 586 B.C. It was a prosperous capital again under the Herods just before and at the start of the Christian era, only to be destroyed by the Roman conqueror Titus in A.D. 70 and obliterated by the Roman city Hadrian built in the second century A.D. Most of the walls that today enclose the Old City of Jerusalem were constructed in the time of Suleiman the Magnificent in the 16th century A.D. With each successive era the city has shifted its geographical boundaries; modern Arab Jerusalem is built mainly to the north of the Old City, and Jewish Jerusalem runs around the Old City's western walls.

Organized investigations of Jerusalem's history and prehistory have been in progress for at least 100 years; as an example, the Palestine Exploration Fund celebrates its centenary this year, and the first object of its explorations was Jerusalem. It may be asked why after nearly a century of studies the British School of Archaeology in Jerusalem (in association with the École Biblique et Archéologique de Saint-Étienne and the Royal Ontario Museum) should have undertaken several seasons' work in this well-explored area. The answer is simply that the excavation techniques of the pioneers of the 19th century could not produce scientific evidence at a supremely difficult site such as Jerusalem. Even in the 1920's the techniques used in Near Eastern excavations, which lagged considerably behind those employed in European countries, still could not produce sound archaeological evidence.

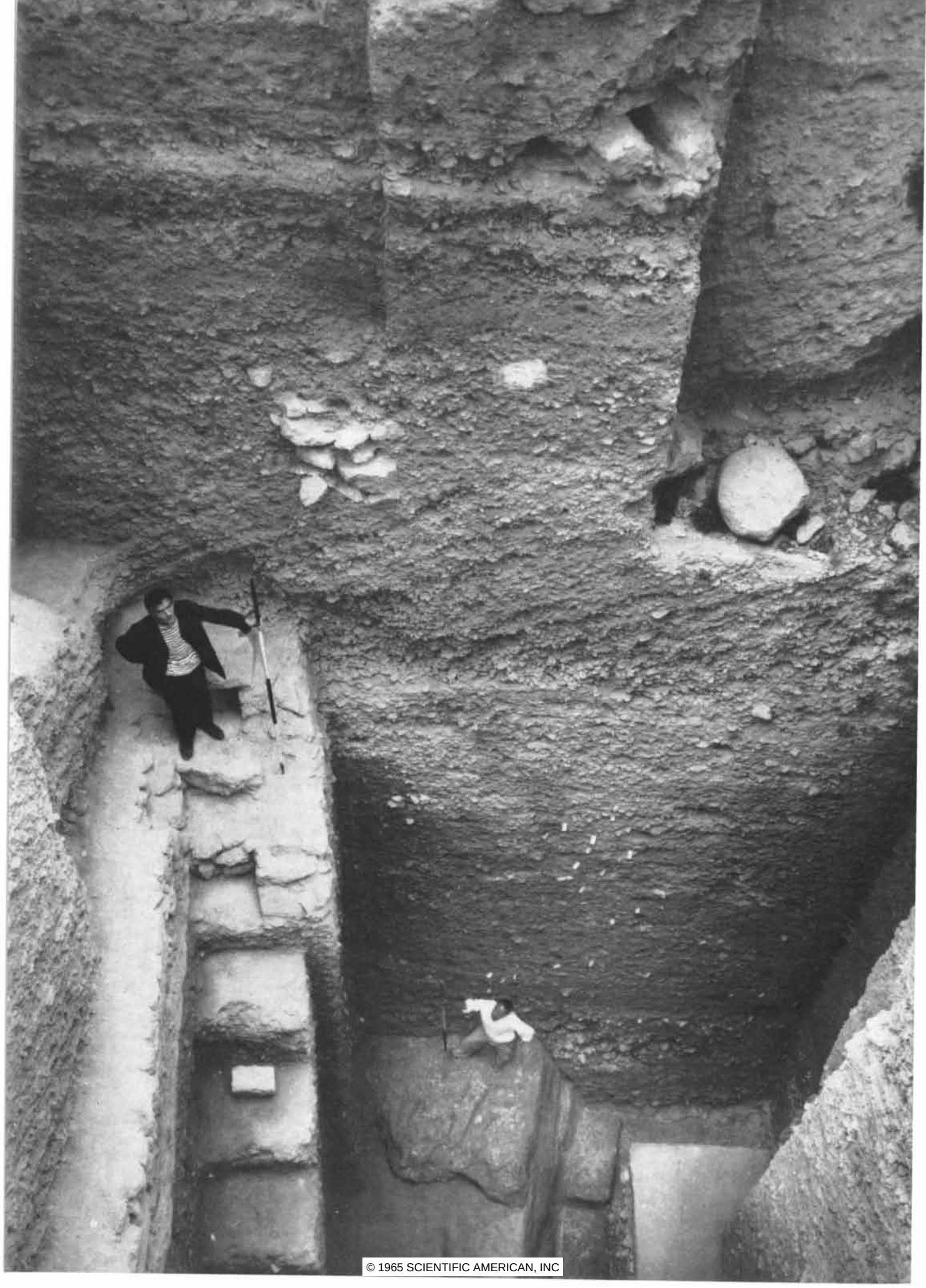
The earlier excavators could follow the lines of ancient walls and produce city plans, but their work had two flaws. First, their methods did not provide stratigraphic evidence of how the walls they uncovered were connected with the levels of adjacent soil that contained objects of known age. Second, there had yet to be accumulated a knowledge of the varieties of pottery—which constitute the major evidence concerning the age of various excavated strata—to be found at Jerusalem. In the absence of such data the tendency was to estimate the age of walls and other structures on the basis of historical probabilities, and concerning these probabilities there could be and were divergent opinions. It was with the object of settling which, if any, of these opinions were correct and of providing an answer to problems and difficulties associated with all of them that our excavations were begun in 1961.

To understand the history of ancient Jerusalem it is first necessary to be familiar with the geography of the site. Seen from the south, the Old City appears to be a plateau, bounded on the east and west by deep valleys [*see illustration on page 86*]. As one walks through the Old City one realizes that the apparent plateau has a considerable dip in the middle. This dip is in fact a third valley; the Old City actually straddles the northern portions of two spurs that extend southward from the higher ground to the north. The Old City is bounded to the east by the valley of the Kidron and to the west by that of the Hinnom (or Gehenna), which curves around to join the Kidron. Between them and running out to their point of junction is the third valley, called the Tyropoeon by Josephus, the Jewish historian of the first century A.D. This central valley is now considerably silted up; originally it was almost as deep as the other two [*see illustration at left on page 88*].

The consensus when we began work was that the original town had stood on the easternmost of the two spurs, a ridge called Ophel. This town was founded by the Jebusites, one of the Amorite tribes that had been established in Palestine since late in the third millennium B.C. It was strong enough to resist the first stages of Israelite infiltration; it is specifically stated in the Book of Joshua that the Jebusites dwelt in the midst of the children of Judah, who could not drive them out.

The Jebusite town was strong even at the stage—around 1000 B.C.—in which David was engaged in uniting all the Israelite tribes. The inhabitants jeered at the attacking Israelites and maintained that even the blind and the maimed could defend the town against them. It was essential, however, that it

DEEP SHAFT shown on opposite page, dug in the heart of Jerusalem's Old City, provides an example of the fresh evidence regarding historical questions that can be obtained by modern archaeology. In this case the question was whether or not the site of the Church of the Holy Sepulcher had been outside Jerusalem's north wall in the time of Herod the Great. Excavation at a point south of the church revealed a thick stratum of fill before bedrock was reached 50 feet below surface. This rock showed signs of quarrying as early as the seventh century B.C., but the fill dated only to the second century A.D., when a Roman city was built over the ruins of Jerusalem. Thus the quarry, and therefore the more northerly site of the church as well, evidently were outside the city's north wall in Herod's day.



should be captured; its control of the route along the central backbone of the hill country meant that only when it was in David's hands could north and south be united. David was successful, and Jebusite Jerusalem became the City of David.

The reason for the consensus that the original settlement was on the eastern ridge was partly the physical character of Ophel, which makes it eminently defensible, and partly that it is adjacent to the only perennial water supply, the spring Gihon (or the Virgin's Fountain) in the Kidron valley at its foot. This consensus held in spite of the fact that Josephus called the western ridge Mount Zion and therefore presumably believed it was the City of David.

Running along the crest of the Ophel ridge is a massive wall that had been revealed by earlier excavations. The best-preserved part of the wall is a well-built tower; the excavators of the 1920's believed David built it as an addition to an original wall attributed to the Jebusites. At the southern end of the ridge this eastern wall is met by another, which excavators in the 1890's had traced across the central valley and up onto the western ridge. Again on the basis of historical probability this east-west extension was attributed to the period of the Jewish monarchy, although whether it had been erected during the rule of Solomon or that of some later monarch was a subject for argument. It can now be said that only one of these three identifications and assumptions has survived the test of modern archaeology: the belief that the original city stood on the Ophel ridge.

The first problem we set out to tackle in 1961 was the vital one of the relation of the Ophel ridge defenses (attributed to the period of the Jebusites and of David) to the water supply in the valley of the Kidron. A water supply is useful to a beleaguered city only if it is accessible from within the walls. Associated with Gihon is a series of water channels, the earliest of which consists of a gallery that runs back into the side of the ridge to the foot of a vertical shaft. This shaft comes to the surface not inside the line of wall attributed to Jebusite times by the early excavators but nearly 30 yards to the east of it; it would have been of little or no use to the defenders of that wall. Moreover, a rather cryptic scriptural passage suggests that the way David captured the city was by sending a picked detachment to climb up the water channel and take the defenders of the Jebusite walls from the rear; it would certainly have been pointless for the raiding party to emerge 30 yards outside the Jebusite defenses.

The accurate dating of the defenses on the east side of the Ophel ridge was therefore the first objective of our group. With this in mind we laid out a trench 11 meters wide extending to a point on the slope east of the line of walls ascribed to David and the Jebusites [*see top illustration on page 90*]. The angle of the slope was nearly 45 degrees and the difficulties of excavation were horrible, but the very first discovery

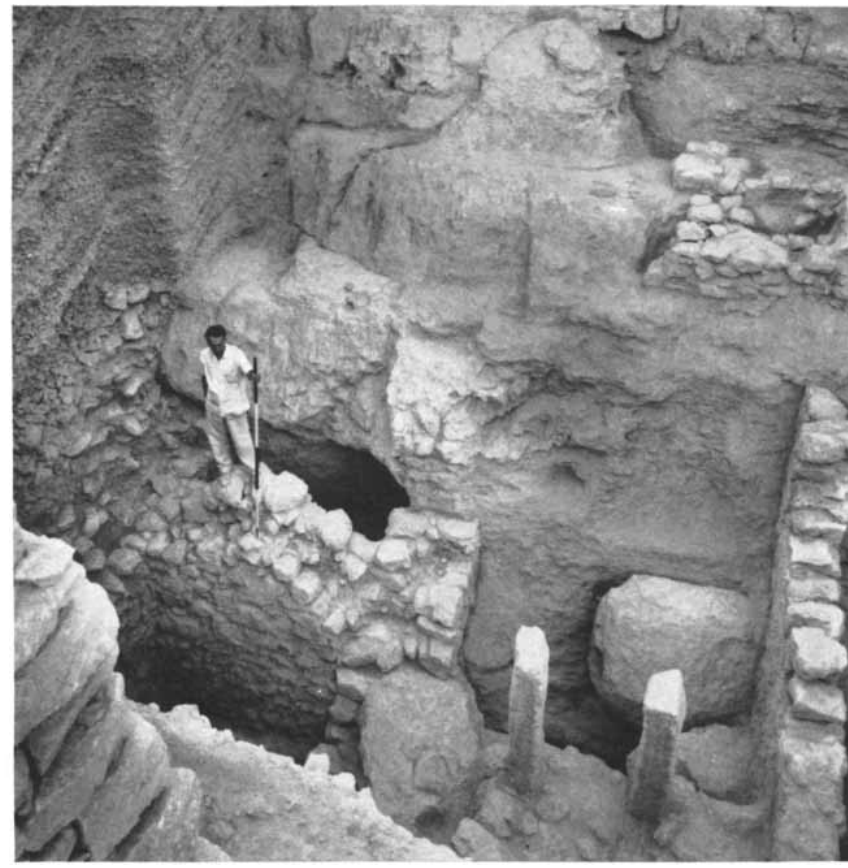
MODERN JERUSALEM spreads far beyond its medieval bounds, in general those of the walled Old City (*center of photograph*). This view looks northward; at the Old City's southeast corner is the Dome of the Rock. Afternoon shadows half fill the valley of the Kidron, beyond the Old City's eastern wall. To the west the valley of the Hinnom curves around the thinly built-up southern part of the modern city to join the Kidron (*bottom center*). Still a third valley, the Tyropoeon, meets this juncture and can then be traced northward into the Old City. Ophel, the steep ridge between the Tyropoeon and the Kidron valleys, south of the Dome of the Rock, was the site of the original Jebusite stronghold that was seized by David 3,000 years ago (*see illustrations on page 88*).

was ample reward. Under no circumstances could the tower attributed to David have been built in his time of about 1000 B.C., because it stood atop the ruins of houses that had been destroyed more than 300 years later in the seventh century B.C. In fact, the layer of earth running up to the foot of David's supposed tower proved to be 500 years younger than that; it belonged to the second century B.C. The tower is therefore a construction of the Maccabean period.

Excavation of the trench down the slope continued to be a difficult business. To a depth of four meters, and deeper in places, the underlying material consisted largely of stones deposited by the collapse of buildings that had once stood higher up the slope. Eventually we reached solid structures, but it was not until we came to the extreme end of the trench we had planned to dig in 1961 that anything like a town wall was exposed. When the area of excavation was extended in 1962, this structure proved not only to be the town wall but also to date to about 1800 B.C.!

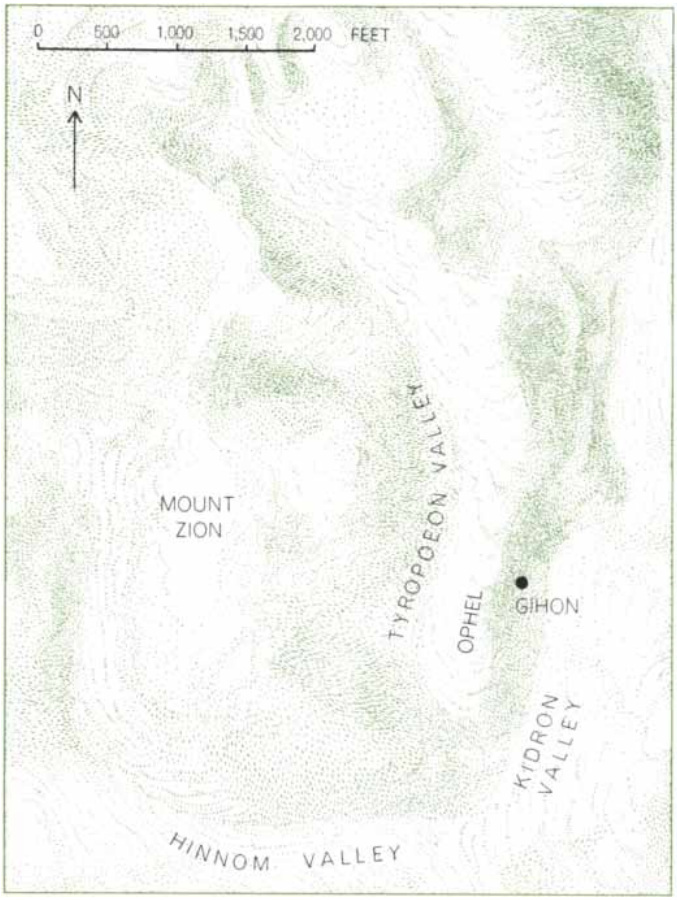
This wall must have defended Jebusite Jerusalem; it is in a position such that the water shaft from Gihon comes to the surface within its bounds. This is also the wall of the next Jerusalem, the City of David; there is clear evidence that it continued in use until the seventh century B.C., when it was succeeded by another wall built slightly higher up the slope. Only short lengths of these two successive walls were exposed in 1961 and 1962, but they were sufficient to show that, instead of being confined to a strip on the summit of the Ophel ridge barely 100 yards in width at its narrowest point, the Jerusalem of the Jebusites and David extended far down the slope on the eastern side. Whether this was also the case on the western side of the ridge remains to be discovered, but already 50 yards have been added to the eastern part of the site.

These 50 yards, however, are on very steeply sloping ground; the present surface is at an angle of 45 degrees and the rock under it is only slightly less steep. The houses of the Middle Bronze Age town that stood here climbed up the hill, following the steep slope of the rock; they are on a small scale and ill preserved. Then a great town-planning operation was carried out in the Late Bronze Age, probably about the 13th century B.C. This work consisted of a complex of masonry platforms, with re-

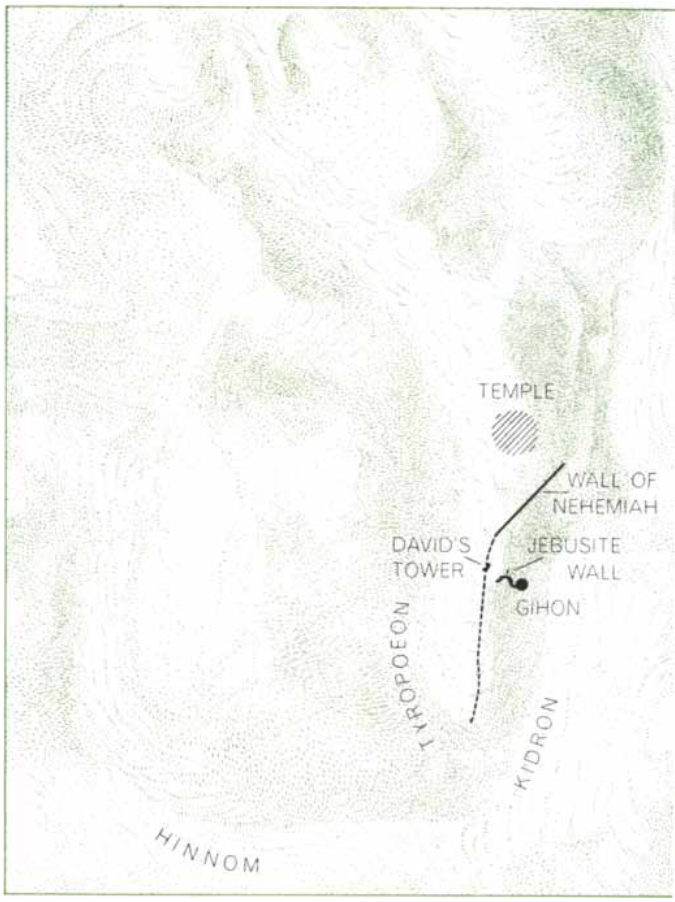


PRECAPTIVITY SHRINE was unearthed by the author outside the earliest eastern wall found on the slope of the Ophel ridge. It was associated with a natural cave (*opening in face of rock to right of man*). The pair of standing monoliths (*bottom right*) may have been cult objects; the roughly square platform of stones (*top right*) may have formed an altar.

CACHE OF POTTERY, of the kind produced in Jerusalem about 800 B.C., was found inside the cave. At first it was thought that the cache might indicate a burial site, but no bodies were unearthed. Later discovery in the outer room of a door giving narrow access to the rock face suggests the possibility that libations of some kind were offered at the shrine.



SITE OF JERUSALEM, marked by three principal valleys, was fortified by Jebusite tribesmen about 1800 B.C. They built walls on the Ophel ridge, the spur of high ground running southward between the Tyropoeon and Kidron valleys, and cut a shaft and channel so that the defenders could draw water from the Gihon spring. The terrain shown is that of today's city, with a few modifications.



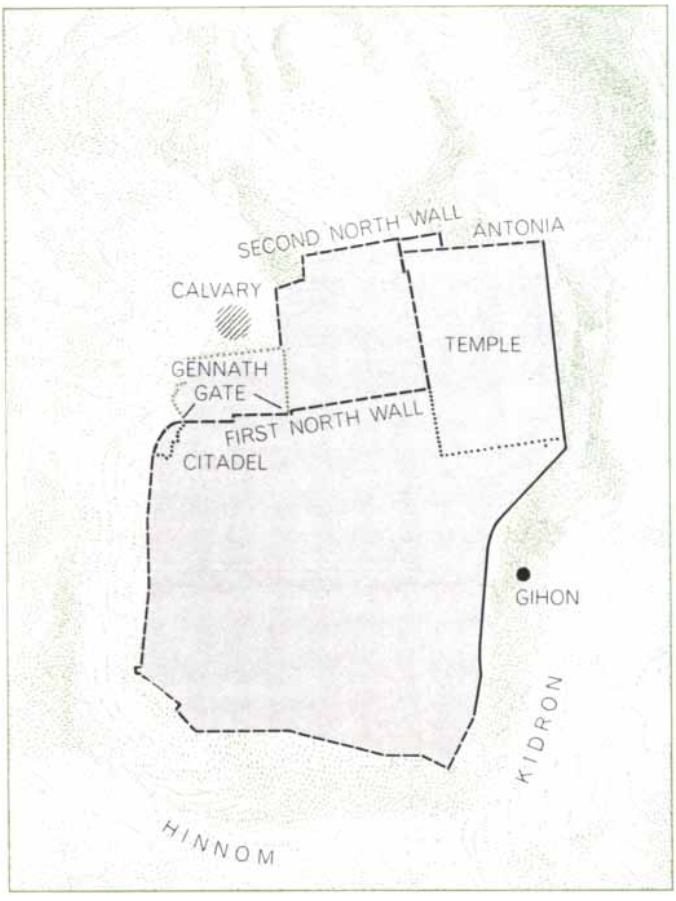
JERUSALEM AFTER DAVID still stood for the most part on the Ophel ridge, although its walls probably were soon extended north to reach the temple Solomon built. The author's excavations have uncovered a part of the original Jebusite east wall, downhill from the wall (*broken line*) and tower attributed to David. The latter, like Nehemiah's wall, were actually erected hundreds of years later.

taining walls that ran at right angles to the pitch of the slope; the fill behind each wall was divided into compartments. The platforms supported a series of terraces. The earliest platforms are Jebusite but no Jebusite houses survive on them; any disturbance, whether natural (such as an earthquake or a torrential rain) or the work of enemies, would have brought about the collapse of such precarious structures. Indeed, many of the platforms show signs of repair and reconstruction prior to the time—in the seventh century B.C.—when the only surviving remains were built. The Israelites not only made use of the Jebusite platforms but also added to them by building out more massive platforms farther down the slope. Scriptural references to the work done by David, Solomon and their successors on Millo—the "Filling"—can probably be interpreted as evidence of their care in maintaining the substructure on which much of ancient Jerusalem was built.

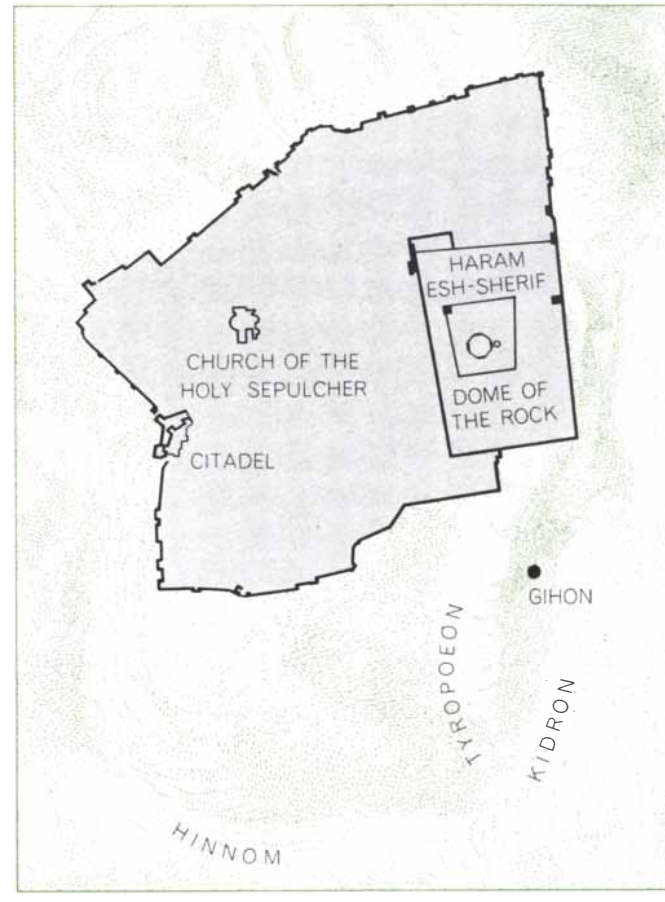
The surviving houses are not particularly impressive. Built of rough stone and no doubt originally covered with mud plaster, they follow the familiar Iron Age tripartite plan: two small rooms and one larger room, divided by two rows of standing monoliths. The houses contain pottery typical of the last stages of the Jewish monarchy. The great quantity of debris overlying the houses dates from the Babylonian destruction of Jerusalem in 586 B.C., when the inhabitants of Judah were led into captivity. It has always been clear from the Biblical account how disastrous this was for Judah and the Jews; the extent of the physical disaster for Jerusalem has only become apparent as a result of these excavations.

Although the location of the town that David made his capital has now been firmly established on the Ophel ridge, the position of its northern limit at that time is not yet certain. Today the southeastern corner of the Old City is dominated by the great Moslem sanctuary Haram esh-Sherif, familiarly known as the Dome of the Rock. This structure occupies the platform where the Israelite temple stood in the days of Herod the Great; Herod's temple, in turn, was traditionally said to stand atop Solomon's. The temple of Solomon is known from Scripture to have been built on the threshing floor of Araunah the Jebusite, and this threshing floor was presumably located outside the limits of the Jebusite settlement. Thus the original northern limit of Jerusalem may lie not far to the north of our trench and almost certainly lies between the trench and the Dome of the Rock.

The first major expansion of early Jerusalem may have followed on Solomon's building of his temple; it seems logical to suppose that soon thereafter the town's northern boundary would have been shifted to link up with the temple. Another expansion of the early town probably occurred when Mount Zion—the western ridge—was enclosed by the southerly wall traced by excavators during the 1890's [*see illustration*



JERUSALEM AFTER HEROD was a large city, dominated by the great temple he raised over the site of Solomon's structure. The city's northernmost wall, outside which Calvary must have stood, ran from the fortress Antonia to the Gennath gate, the location of which is now unknown. Dotted lines (*color*) show two ways in which Calvary's traditional site could have been outside this wall.



OLD CITY TODAY is surrounded by 16th-century A.D. walls, which closely follow the bounds of the Roman city that was built over the northern ruins of Herodian Jerusalem after its sack by Titus in A.D. 70. The Church of the Holy Sepulcher lies within these walls, but the author has proved that its site, traditionally that of Calvary, was to the north of the city wall during the reign of Herod the Great.

at left above]. This expansion, formerly ascribed to one or another period of the Jewish monarchy, has turned out to be late rather than early: the town on the western ridge dates only to the first century A.D. and was the work of Herod Agrippa.

The most interesting find belonging to the period of the Jewish monarchy that our work has unearthed so far is a structure lying outside the wall on the eastern slope. There the 1962 excavations revealed a shallow cave in the face of the rock, its entrance surrounded by massive stone walls. During the early stages of excavation we thought we had uncovered a tomb; this seemed even more probable when further work revealed a large deposit of pottery vessels of the kind one would expect to find in a tomb [*see lower illustration on page 87*]. Complete investigation, however, revealed no burials. When the structure was traced farther in 1963, features suggesting a ceremonial area emerged. In an adjacent room were two standing monoliths, perhaps memorial pillars or cult objects; on top of a setback in the rock face was a square structure that may have been an altar. Most mysterious of all, in the back wall of the room containing the monoliths was a blocked doorway; when this was cleared, it proved to lead to a space only a foot wide between the wall and the rock face. The wall rests on an enormous boulder; when the boulder is removed, the purpose of this narrow access—perhaps the pouring of libations or something of that kind—may become evident.

Although the captivity of the Jews ended about 530 B.C., when Babylon's new Persian rulers allowed the first Jewish exiles to return to Jerusalem, the city's walls were not rebuilt until Nehemiah arrived as governor about 440 B.C. It has always been assumed that the rebuilt walls followed the line of the pre-exile ones, but apparently this does not apply to the eastern wall. The destruction on this side of the city, involving the collapse of the intricate system of platforms and terraces in a cascade of stones into the valley of the Kidron, was such that Nehemiah did not even attempt a restoration. Moreover, the number of Jews who returned from exile was a mere fraction of the population of the earlier Jerusalem. A town concentrated on the summit of the ridge provided adequate residential space; here, built along the top of a low cliff, we have now identified the postexile wall of Nehemiah [*see illustration at right on opposite page*].

Jerusalem after the exile thus consisted of a narrow strip in this restricted area of the Ophel ridge. In the subsequent Maccabean period the tower hitherto mistakenly known as the tower of David was added to the wall of Nehemiah. The Maccabean rulers of Jerusalem, who were attempting once more to create a Jewish state, presumably carried out a number of other works of reconstruction; it may be that

BROAD TRENCH, descending the abrupt eastern slope of the Ophel ridge (*diagonal from upper left to center of photograph*), was dug through a jumble of masonry fallen from the wreckage of platforms that had once been the foundations of buildings on the slope. Tradition mistakenly located the Jebusite east wall near the ridge crest, well uphill from the perennial spring, Gihon (*which is sheltered by the lower building, center*).

LABORIOUS EXCAVATION into the 45-degree slope revealed the wall plans of several small houses built by the Israelites in the seventh century B.C. Many of the platforms supporting these houses had first been constructed centuries earlier by the Jebusite tribesmen who founded Jerusalem. All this part of the ancient city showed evidence of the wide destruction wrought by the Babylonians who conquered Jerusalem in 586 B.C.

the northern end of the western ridge was included in the city in their time, although this has yet to be proved by excavation.

This northern expansion had certainly taken place by the reign of Herod the Great in the last third of the first century B.C. To Herod belong the only imposing structures in the Old City that have survived from the time of the Gospels. The greatest of them is the temple platform. Today at the southwest corner of this platform Herod's masonry still rises to a height of about 70 feet above the present surface, and soundings show that it extends another 80 feet down before reaching bedrock. Yet this was only one part of the grandiose rebuilding of Jerusalem that Herod undertook. A viaduct spanned the whole city, from the temple on the east to Herod's palace on the western ridge; a portion of one of the arches that supported this elevated roadway can still be seen. A considerable part of one of the towers of the present citadel on the western ridge is a structure that was originally part of Herod's fortress.

A puzzle that confronts today's visitor to the Old City belongs to the days of Herod. In the fourth century A.D. the Byzantine empress Helena built the Church of the Holy Sepulcher so that it reputedly covered both Calvary and the site of the tomb from which Jesus rose. Yet this church is located in the heart of the Old City, whereas Calvary certainly lay outside the walls of Herod's Jerusalem. The boundaries of the Old City are not those of Herod's Jerusalem; instead they are in general those of Aelia Capitolina, the city Hadrian built in the second century A.D. to obliterate the Jewish town. The question therefore arises: Where was Jerusalem's north wall in the time of Herod? If the Church of the Holy Sepulcher indeed lies inside the line of Herod's wall, then the tradition that guided Helena in her pious work was mistaken.

Our current excavation program has thrown some light on this question. The present north wall of the Old City may be near the line of the north wall built in A.D. 40–44 by Herod Agrippa; this is the most northerly of three city walls described by Josephus at the time of the attack by Titus in A.D. 70. The earlier and innermost of the other two walls is believed to have crossed the valley from the temple platform to the western citadel; the intermediate wall Josephus describes as running from the fortress Antonia at the northwest corner

of the temple enclosure to the gate Gennath in the innermost wall. Unfortunately there is no clue to the position of this gate. Most scholars have been inclined to place it near the western citadel; a wall following a normal course from Antonia to that point would enclose Helena's church and so mean that its site is not authentic. It has also been suggested that the gate in question was near the center of the innermost wall, or alternatively that the intermediate wall followed a zigzag course. Both suggestions would leave the church outside the boundaries of Herod's Jerusalem.

Although the whole of the Old City is now very closely built up, there was fortunately one site due south of the Church of the Holy Sepulcher where it was possible to excavate. In spite of the small area available, this site—after three years' work—has produced conclusive results. Under relatively shallow deposits belonging to the Arab and Byzantine periods our excavations revealed a deep zone of fill that appears to date from the time when Hadrian's city was laid out in the second century A.D. At a depth of 50 feet from the present surface, bedrock was at last reached. This rock exposure proves to have been quarried in the seventh century B.C. It is quite clear that from this period down to the second century A.D. the quarry was located outside the north wall of the city. Since the quarry site is south of the Church of the Holy Sepulcher, this area must also have been outside the north wall. Thus although the excavation does not prove that the church is built on Calvary, it does at least make it quite possible that Helena was guided by an authentic tradition when she decided to have the church built there.

Just as David's Jerusalem was destroyed by the victorious Babylonians in 586 B.C., so the Jerusalem of the Gospels was destroyed by Titus in A.D. 70. Everywhere in the area we have excavated to the south of the Old City there are signs of this destruction: occupation comes to an abrupt end. More than a century later the obliteration of ancient Jerusalem was completed by the construction of Hadrian's city. The fortunate chance that this covers only the northern part of the ancient Jewish capital has meant that in the area to the south we still have the opportunity to search out the beginnings of the city's history. We must work fast, however; modern Jerusalem is now expanding in this direction.

FLIES AND DISEASE

Do flies spread disease? Surprisingly the evidence is still inconclusive. Efforts to provide an answer have nonetheless yielded significant information on the nature of infection

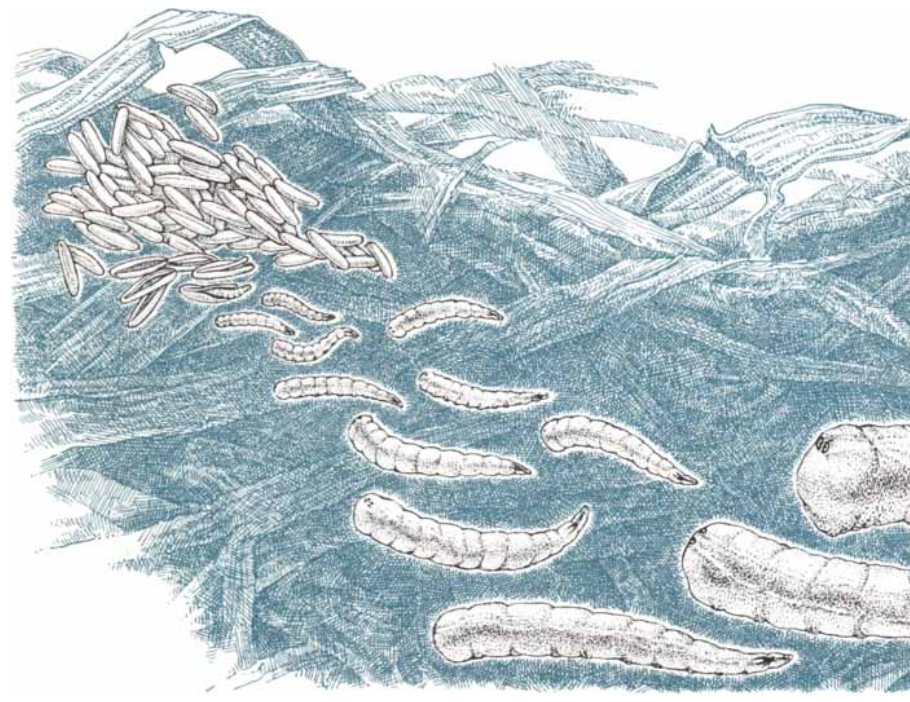
by Bernard Greenberg

From the Antarctic Circle to the Arctic the "domestic" flies have found a favorable cohabitation with the human species. Since the beginnings of civilization men have considered flies to be certainly noxious and possibly dangerous. These commonplace insects, being highly adaptive, have survived every threat to their existence, including the recent assaults man has made on them with a variety of sophisticated chemicals. Indeed, preoccupation with the control of flies has often overshadowed the question of their role in the spread of disease. For the majority of the world's people, however, deciding the question is a vitally important issue.

The Bible notes that the house of Pharaoh was plagued by "a grievous swarm of flies." The Israelites' contemptuous nickname Beelzebub ("Lord of the Flies") for the Chaldean god Baal is probably derived from the clouds of flies that buzzed around the Chaldeans' animal sacrifices. Among the earliest examples of sanitary engineering are the tiled lavatories and sanitary drains of Ur and Kish in Mesopotamia, dating from 3000 B.C., and Moses' edict to the Israelites to bury their excreta; such measures suggest firsthand knowledge of the habits of flies. Greek and Roman writings contain speculations that flies spread dysentery by falling into food. In later centuries various observers expressed similar suspicions. In 1498 Bishop Knud of Denmark noted the increase of flies as being one of the first signs of the approach of plague. Thomas Sydenham, the great 17th-century English physician, observed that "if swarms of insects, especially houseflies, were abundant in the summer, the succeeding autumn was unhealthy."

With the birth of the discipline of bacteriology investigators began to compile incriminating evidence against flies. At a time when the public was just becoming aware of germs and infectious disease the unclean habits of flies and their predilection for man's food aroused public concern. In the summer of 1871 Sir John Lubbock—reformer, student of insects, popularizer of science and a member of Parliament—exhorted the House of Commons on the subject of flies: "Far from looking upon them as dipterous angels dancing attendance on Hygeia, regard them rather in the light of winged sponges speeding hither and thither to carry out the foul behests of Contagion."

By now flies have been found to harbor well over 100 different species of pathogenic organisms, and the list of human and animal diseases they are charged with transmitting now stands at more than 65, including dysentery (amebic and bacillary), typhoid, cholera, salmonellosis, tuberculosis, leprosy, anthrax, poliomyelitis, Coxsackie virus infection, infectious hepatitis, yaws, trachoma, conjunctivitis and various infections by parasitic worms. Yet the evidence is still only circumstantial. The reputation of the domestic flies is in the position of a man charged with homicide because he is found standing beside the victim with a loaded gun in his hand. In most cases it can-



LIFE CYCLE of the common housefly (*Musca domestica*) is depicted. At left are eggs, laid in manure. Within a day the eggs hatch into maggots, which constitute the larval stage.

not be proved conclusively that the flies in question fired the gun.

There are insect-borne diseases in which the culprit has been identified beyond doubt; it is known, for example, that human malaria is transmitted only by injections from *Anopheles* mosquitoes and that typhus is spread by the louse. The domestic flies, however, are ambiguous suspects. The infections they are accused of spreading can actually be spread by any of four different agents: food, fingers, feces and flies. Moreover, we know that flies are not essential for the spread of a disease; outbreaks of dysentery and poliomyelitis sometimes take place in winter, when the fly population is dormant.

Be that as it may, the flies must surely be considered prime suspects, because the diseases to which they have been linked are particularly common in warm regions and underdeveloped countries, where flies are most abundant and have easy access to food exposed in outdoor markets. The crucial question is: Precisely how important a factor are they, and under what conditions are they most likely to be significant transmitters of disease? Plainly information on these points should be of great help in the control of fly-borne diseases.

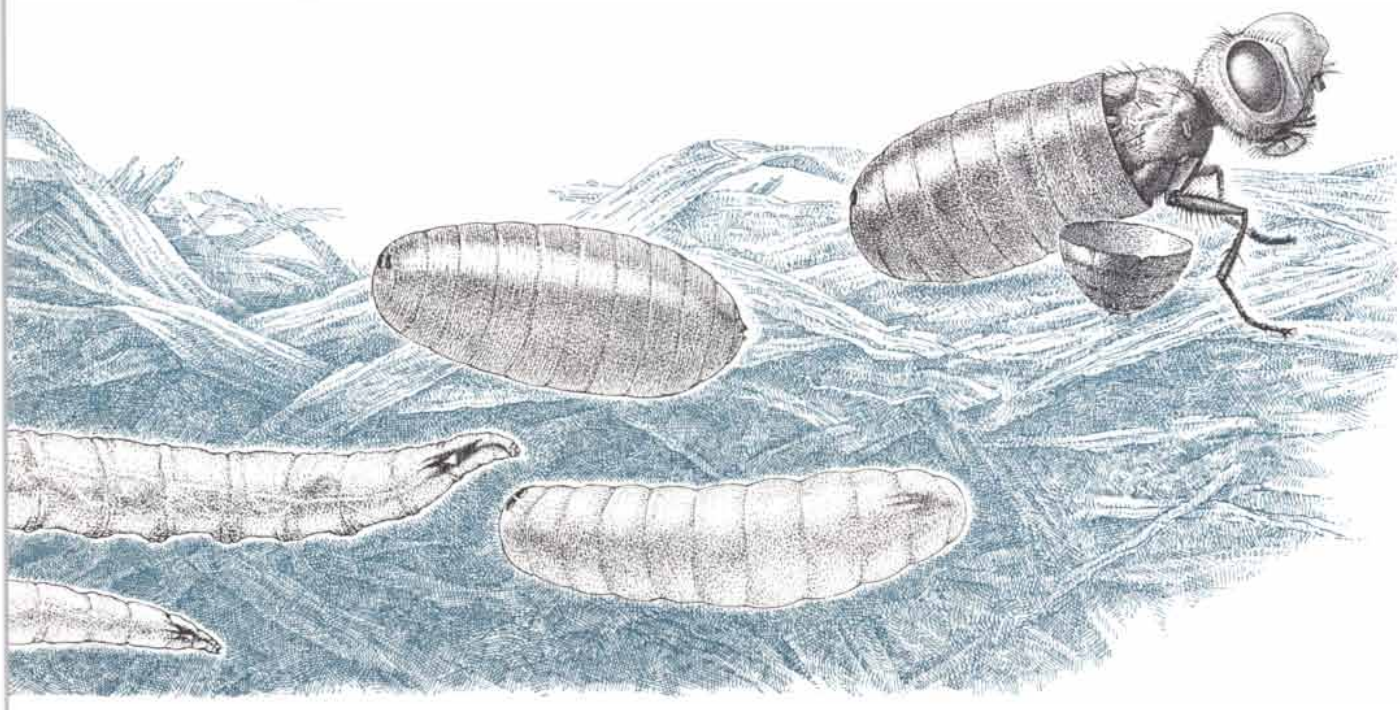
Let us first define the species we are concerned with. The order of Diptera (two-winged insects) includes mosquitoes, midges, gnats, blackflies and a great variety of other flying insects. We are discussing here, however, only the "synanthropic" flies—those that have set up housekeeping in intimate relationship with man or with his domestic animals. There are about 200 species of these insects. The most important and most common are the housefly (*Musca domestica*), the stable fly (*Stomoxys calcitrans*) and certain species of blowflies, including the bluebottle fly, whose iridescent blue body is often seen around windowpanes.

The housefly is noted for, among other things, its habit of feeding on excrement. It is not a biter; its spongy proboscis cannot pierce the skin. Lubbock described it accurately as a "winged sponge." The blowflies, frequenters of garbage, decaying fruit and meat, also are a nonbiting breed. Although the nonbiting flies do not themselves have direct access to the blood of animals, some of them associate with biting flies and crowd around a wound when the skin is pierced; in this way they may pick up pathogenic organisms from an infected animal. Notable among the biters, as bathers at the beach are commonly reminded, is the stable fly.

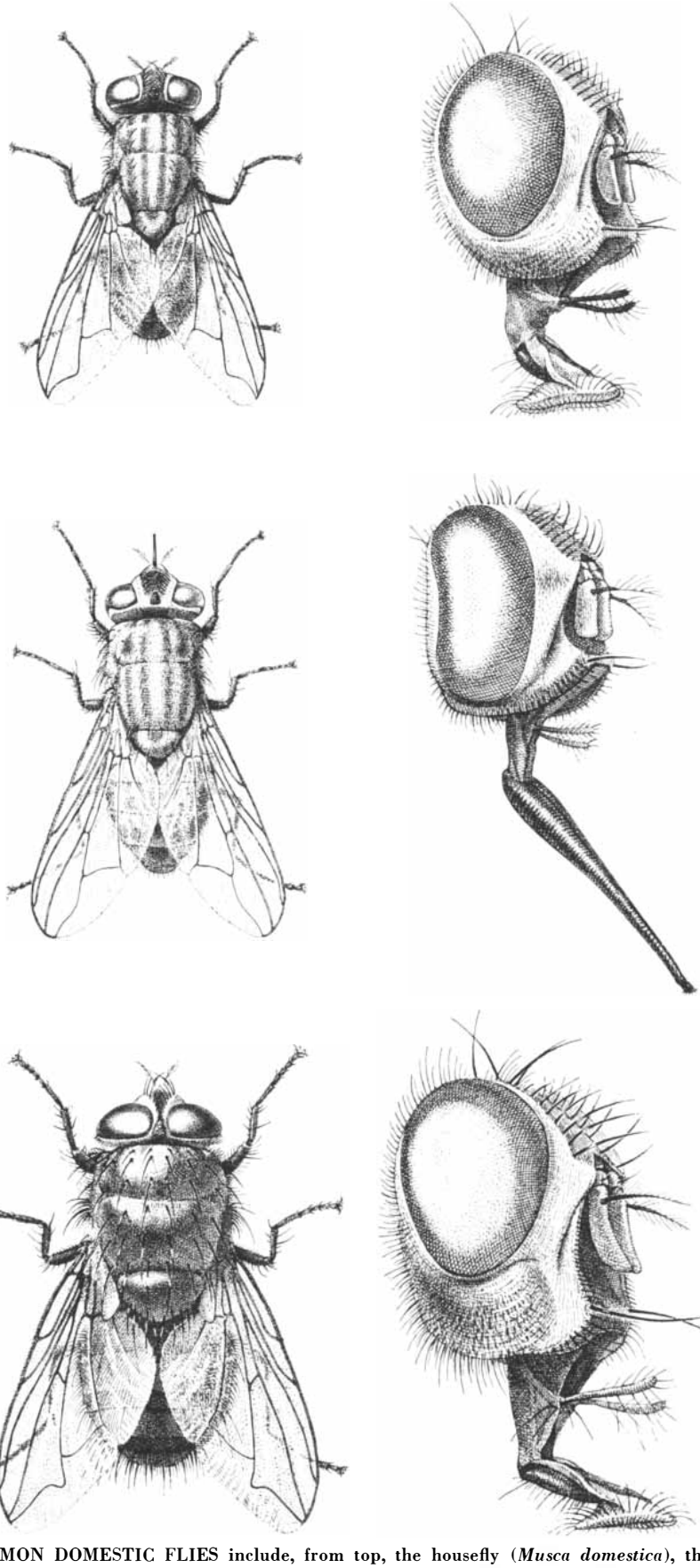
One approach in the investigation of the flies' role in disease is environmental. This entails studies of their behavior and activities in the ecological milieu. Many investigations have shown, for instance, that outbreaks of diarrheal diseases correspond closely with seasonal rises in the size of the fly population. This, however, is only a general and indirect clue; what is wanted is more detailed information relating the flies' behavior to the actual transmission of a disease in specific situations.

A good example of a search for such information is concerned with the dissemination of poliomyelitis. In 1941 the Yale University epidemiologist John R. Paul and his co-workers reported that they had found poliomyelitis virus to be present in flies in the ordinary environment. Later Joseph L. Melnick of the Yale group showed that among the various species carrying the virus the principal ones were the blowflies. Dorothy M. Horstmann and others in the same group proceeded to a still more specific discovery. Conducting an intensive examination of flies in Arizona and later in Costa Rica, where the human population had been treated with the Sabin poliomyelitis vaccine consisting of attenuated live viruses, they found the flies carrying these virus strains, which they had evidently picked up from the vaccinated population. All of this demonstrated that people can infect flies, but it still fell short of showing how and to what extent flies can spread poliomyelitis to people.

Our group at the University of Illi-



Maggots grow rapidly, undergoing two molts. At about the fifth day the maggot stops feeding, thereafter contracting somewhat (*bottom right*) before entering the pupal stage, represented just above the prepupal maggot. The entire cycle takes about two weeks.



COMMON DOMESTIC FLIES include, from top, the housefly (*Musca domestica*), the stable fly (*Stomoxys calcitrans*) and the bluebottle blowfly (*Calliphora vicina*). Shown adjacent to each fly is an enlargement of its head and mouth parts. The proboscises of the housefly and the blowfly, which are nonbiting insects, are adapted for sucking. The proboscis of the stable fly, which is a stinging insect, is equipped to pierce an animal's skin.

nois, in collaboration with investigators in Mexico, has looked into the role of flies in transmitting salmonellosis, in which the *Salmonella* bacterium produces symptoms of food poisoning, gastroenteritis or systemic infection. The disease seems to be on the increase in North and South America and in Europe and is becoming a worldwide problem in this day of shrinking distances and the interchange of prepared foods between regions. Contaminated fish meal from Peru may infect poultry in the Netherlands, and powdered eggs from the Orient may bear *Salmonella* to breakfasts served by U.S. airlines.

Suspecting that flies might have a good deal to do with the contamination of food by *Salmonella*, we undertook a joint study in the summer of 1962 with Gerardo Varela, director of the Institute of Health and Tropical Diseases in Mexico City, and two of his associates, Alex Arroyo and Homero Hernandez. The locale was a village in Mexico that has a slaughterhouse at the edge of town. We found that the flies around the slaughterhouse harbored *Salmonella* bacteria. In the 10 most abundant fly species we detected 12 types of the bacteria—three times as many types as were carried by the cattle and hogs themselves or by rats in the area. The leading carriers were the housefly and several common blowflies.

Having established that the flies bore the infection, our next question was: How far might they spread it? Experiments performed many years ago had shown that houseflies raised in laboratory cages would travel as much as 13 miles when they were released. This, however, was not necessarily an indication of how flies might behave in nature, because the laboratory flies were all about the same in age and flying experience and might conceivably have developed a strong migratory urge as a result of their confinement. In the summer of 1963 we measured the migratory propensities of the free slaughterhouse flies in a less artificial way. We simply marked the flies flying around the slaughterhouse and then tried to trace their travels. The marker was a fluorescein dye, with which we sprayed more than 200,000 flies on the premises. When the flies were recaptured and put under ultraviolet light, the tag would show up as a yellow-green fluorescence.

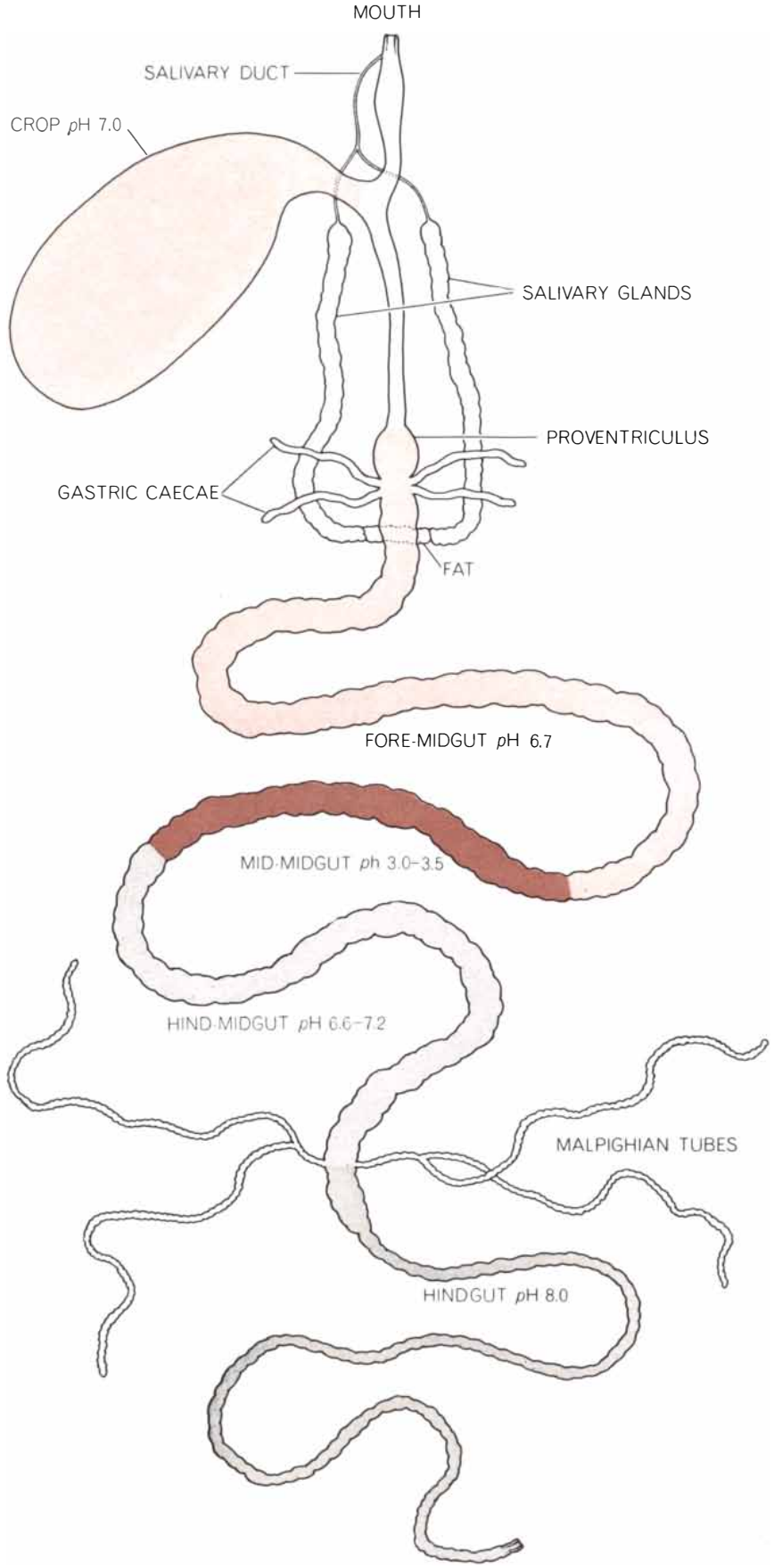
The slaughterhouse is at the edge of a semiarid plain. There are no nearby attractions to compare with the slaughterhouse grounds, which provide the flies with carcasses, offal, a rich manure

and resting places in the weeds and in piles of hogs' hair. Yet the day after the flies were marked some of them were found as far away as the other side of town. Within a few days a total of 543 marked flies were recovered at various points: in the village marketplace, in a dairy, in residential areas and in a neighboring village three miles away. Clearly the slaughterhouse flies did considerable wandering and invaded all kinds of neighborhoods, from the poorest to the most exclusive.

The flies, then, had demonstrated two attributes necessary for vectors of disease: mobility and a burden of bacteria. It remained to be shown that they delivered enough bacteria to produce infection in man. Tests have shown that the development of infection in the human body requires an introduction of at least 100,000 viruses in the case of poliomyelitis and a million or more bacteria in the case of salmonellosis. Flies certainly could not bring about any such delivery into the body merely by direct contact. Could they deliver a sufficient number of organisms to the intestine by way of contamination of food? We exposed houseflies to the feces of a dog infected with *Salmonella* and then gave the flies access to beakers of atole, a Mexican drink containing corn meal. Ten human volunteers then drank the atole. Six of the 10 later showed traces of *Salmonella* in their stools. None, however, developed any symptoms of disease or had enough bacteria to produce an antibody response.

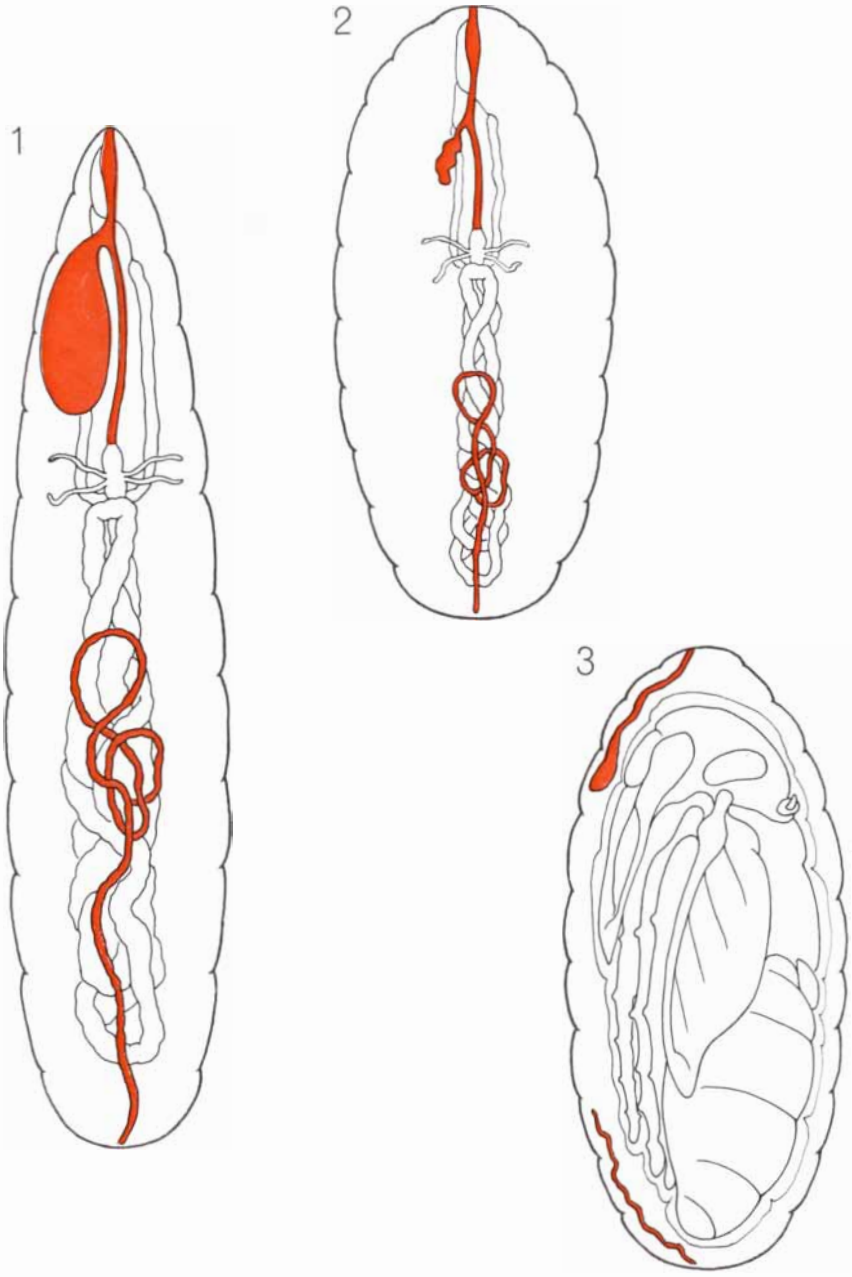
Our counts of the bacterial population in each link of the chain indicated that the flies did not carry a heavy contamination of pathogens to begin with. It seems likely, however, that even a small deposit of bacteria on food such as milk, soup or meat would multiply rapidly and soon reach the threshold level required for infection if the food were left unrefrigerated. The ordinary heat of light cooking may not destroy *Salmonella:* it has been found to survive in scrambled eggs cooked to the soft state. The atole would have incubated a disease-producing dose of bacteria, without any change in its taste, if we had left it standing longer. It was not our purpose, however, to cause illness in the volunteers. Therefore it remains a distinct possibility that flies are capable of delivering disease by way of food and drink.

In 1946 and 1947 the disease-carrying potential of flies was submitted to a large-scale test of a different kind by James Watt and Dale R. Lindsay, then in the U.S. Public Health Service. In



MAGGOT GUT is shown about 20 times natural size. The acidity and alkalinity at various regions are shown by *p*H numbers; a number below 7 represents an acid condition; above 7, alkaline. Normal bacteria occur throughout gut; *Salmonella*, which adult flies may transmit to humans, ordinarily are abundant ahead of acid mid-midgut and infrequent elsewhere.

Hidalgo County in Texas, after determining the incidence of acute diarrheal infection among children, they divided the county's towns into two groups; one group of towns was sprayed with DDT to destroy the flies, whereas the other was left untreated. In the untreated towns the disease rate remained unchanged, but in those where the fly density was diminished by DDT the rate of dysentery caused by *Shigella* bacteria declined (although the salmonellosis rate was not much affected). After a year and a half the treatment was reversed, with DDT spraying now applied in the former control towns and discontinued in the towns that had received it. The dysentery rates then began to reverse, declining in the DDT group and rising in the towns no longer sprayed. Unfortunately before the experiment could run its full course the flies developed resistance to DDT, so that the final results were not clear-cut.



INTERNAL MOLT occurring at the pupal stage in the life cycle of the fly is a mechanical means by which the fly reduces its population of bacteria, so that the newly emerged adult is often virtually sterile. In a maggot (*1*) the gut parts (*color*), which contain the bacteria, are of substantial size. A newly formed pupa (*2*) has a contracted gut because the larva stopped feeding in the prepupal stage. In a molt early in pupal stage the maggot's foregut and hindgut linings are cast aside, as shown (*3*) in a pupa halfway through pupal stage.

There is another approach to investigating the flies-and-disease question; this is to turn the flies inside out, so to speak, in order to examine the life of the microbes within them. Our group has carried out a number of studies along this line with a view to unraveling the relation between flies and bacteria.

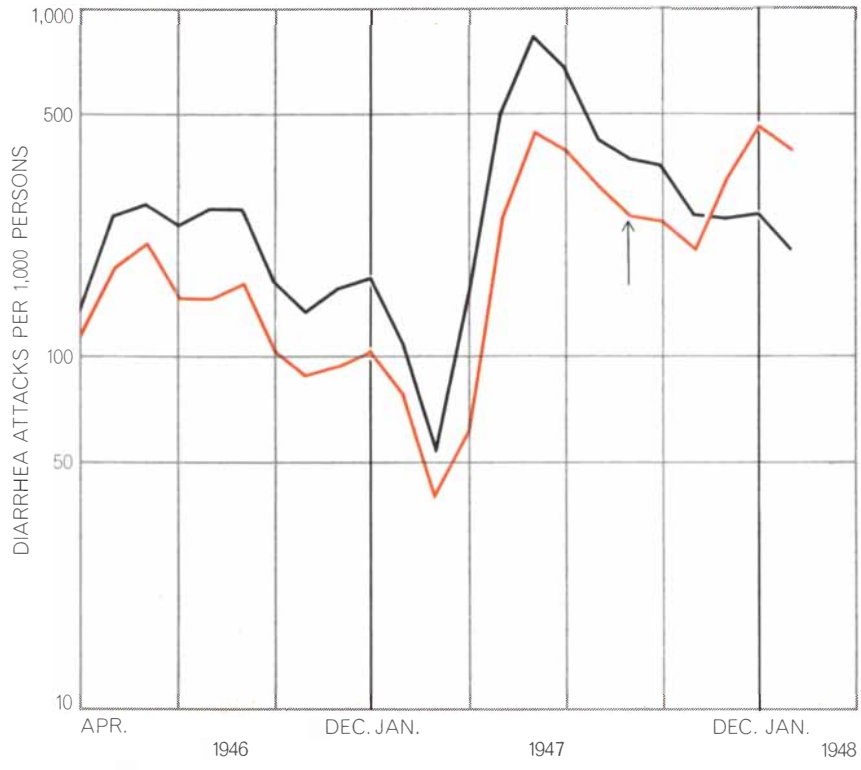
The life of a fly, like that of other highly evolved insects, has four distinct stages: egg, larva, pupa and adult. A female domestic fly usually deposits a batch of 100 to 150 eggs. Houseflies and stable flies commonly lay the eggs in manure, blowflies in carrion. Within a day the egg hatches into a tiny larva, or maggot. The larva grows with phenomenal speed. A newly hatched bluebottle larva weighs .1 milligram; five days later it weighs 84 milligrams—a total increase of more than 800 times. Unable to ingest particles of any appreciable size, the larva must obtain its food in liquid form. The French naturalist Jean Henri Fabre discovered that blowfly larvae excrete enzymes capable of reducing solid proteins to a broth. The standard medium now used by laboratories the world over for breeding houseflies consists of a mixture of brewers' grain, wheat bran, yeast extract and alfalfa meal moistened with water. Fly eggs deposited in this mixture will advance through their successive larval and pupal stages and emerge as adult flies in about 14 days, the time depending on the incubation temperature.

In nature the domestic-fly larva begins life in a milieu teeming with microorganisms. When we undertook our experiments, we set out to learn what role bacteria play in the housefly larva's diet. Using the standard laboratory food medium, we set up three different conditions. For one culture we sterilized the medium as soon as it was prepared and then pipetted disinfected fly eggs into the flasks. For the second and third cultures we incubated bacteria in the flasks at 37 degrees centigrade (the human body temperature) before sterilizing the medium and seeding in the germ-free eggs. In one set of flasks the incubation period before the microorganisms were killed was one day; in the other set, two days.

The three different conditions produced three distinctly different results in larval development. In the first set of flasks, where the medium had the lowest number of microbes, the larvae failed to grow beyond the first molt. In the second set, where bacteria, yeasts and molds had grown for one day before the eggs were introduced, the lar-

vae developed up to the second molt. In the third medium, affected by two days of microbial growth, the larvae developed normally into adult flies. Thus the experiment demonstrated that fly larvae require microorganisms or some product of microorganisms for growth. Zvi Levinson, an investigator in Israel, has since shown that fly maggots will grow normally when they are supplied with a dried concentrate of the bacterium *Escherichia coli*.

It is clear, then, that bacteria play an important part in the nutrition of the larva of the housefly, and probably of other domestic flies. Apparently it is only in the larval stage that flies depend on bacteria for feeding. In our laboratory A. M. Burkman and I have found that, as adults, germ-free houseflies live just as long as their counterparts that are fed a mixed bacterial flora.



DIARRHEA RATES of Hidalgo County in Texas showed a correlation with fly population when the U.S. Public Health Service experimented with fly control. At first the fly population was kept down with insecticides in one group of towns (*color*); in another (*black*) flies were uncontrolled. The treatments were reversed at the time shown by the arrow.

What is the fate of the pathogenic organisms a fly maggot ingests? Do they make a permanent home in the insect and survive in it when it emerges as an adult fly? Some 50 years ago G. S. Graham-Smith and other British entomologists investigated this question. They made the interesting discovery that, when they virtually drenched housefly maggots with *Salmonella* and *Shigella*, the adults emerging from the pupal stage were free of these bacteria! In recent years, with new techniques, our group and others have found that some *Salmonella* do survive the metamorphosis to the adult fly; among them are a species responsible for food poisoning in man and another that is the agent of the deadly pullorum disease of chickens. It remains true, however, that in most cases flies get rid of infectious organisms by the time they become adults.

In part this is simply a mechanical process. As the larva approaches the pupal stage it stops feeding and its gut gradually contracts. Most of the microbes, natural and pathogenic, are eliminated. Clusters of them remain sequestered in folds and crypts of the digestive tract, but these clusters are eliminated when the larva molts within its pupal case; in the fly larva the integument that is shed extends into the foregut and hindgut. The emerging fly, as it squirms free, leaves all or nearly all the bacteria behind in the pupal case. In some species this process of autosterilization is surprisingly complete. We have found 54 percent of stable flies, 17 percent of houseflies and 37 percent of greenbottle flies to be completely germfree at the moment of emergence from the pupal case. The phenomenon of elimination of bacteria at adulthood may be widespread among Diptera: it has been observed not only in the domestic flies but also in horn flies, midges and mosquitoes.

The mechanical events are not the full story of how the fly gets rid of its juvenile load of microbes. Invited by Giuseppe Penso and Giuseppe Saccá to work at the Higher Institute of Health in Rome in 1960 and 1961, Vincenzo Miggiano and I made a quantitative study of the elimination of bacteria by the bluebottle fly (*Calliphora vicina*). We doused maggots of the fly with a streptomycin-resistant strain of a *Salmonella* species (*S. typhimurium*) and compared the survival rate of this microbe with that of the maggot's normal flora. In the two days just before it changed into a pupa the maggot reduced the population of some 100 million bacteria in its digestive tract to fewer than one million. In the same period it completely wiped out the *Salmonella* bacteria, amounting to more than a million, that were in the tract.

What killed the *Salmonella*? Were they attacked by the maggot's normal bacteria? We tested this hypothesis by eliminating all possible competitors and exposing the maggots to *Salmonella* alone. The medium consisted only of chick embryos that had been removed aseptically from hens' eggs; these were injected with *Salmonella* and dropped into a flask with some sterile sawdust. Germ-free fly eggs were then deposited in the flask. With all competition from other organisms removed, the *Salmonella* survived in the maggot's gut, although not as well as the normal flora would have. Next we introduced as a competitor *Escherichia coli*, a normal resident of the human intestinal tract. The *Salmonella* still thrived, proliferating more than the *E. coli*. When we tried the test with an *E. coli* strain that produces colicine, a substance that usually destroys bacteria that do not make it, the *Salmonella* bacteria again survived. They finally yielded when we exposed them to a strain of the natural fly bacterium *Proteus mirabilis*, which had been freshly isolated from a fly. The contest started with the *Proteus* slightly outnumbering the *Salmonella*. As the competition proceeded, the *Salmonella* bacteria did so poorly that in the maggot they were soon reduced to a ratio of one *Salmonella* to 11,500 *Proteus*. However, when *Proteus* bacteria were pitted against *Salmonella* in the maggot's absence—that is, in a

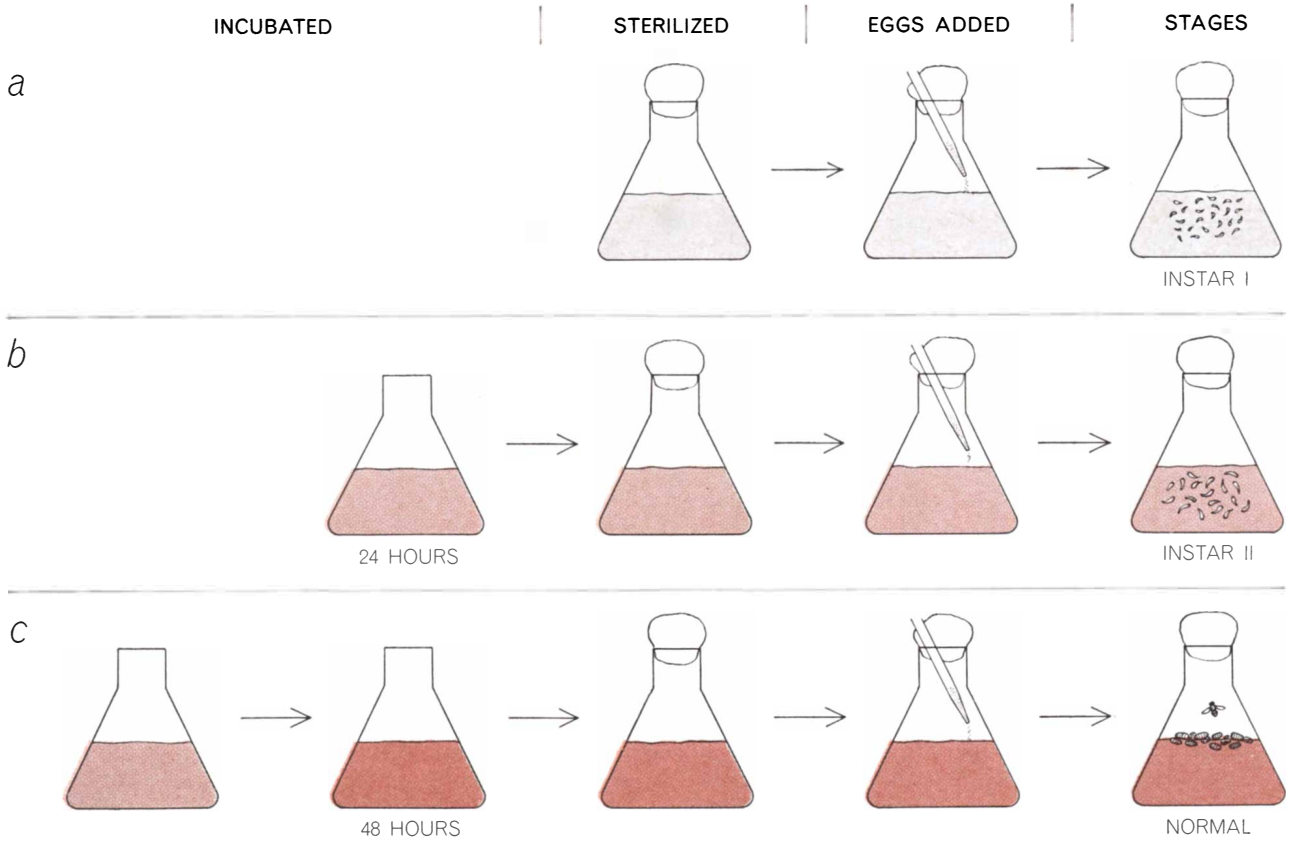
flask containing broth—the *Proteus* was not so potent; it outgrew *Salmonella* only about 20 to 1. This suggested that some factor in the maggot's gut was partly responsible for the inhibition of *Salmonella*.

It has been suspected for some time that the digestive tract of flies contains bactericides. Exploring this hypothesis, we tried several experiments but were unable to find any evidence of the presence of a bactericide in the tissues or fluids of a maggot's tract. By various stratagems we kept maggots in the larval stage and exposed bacteria to their digestive tracts for months, without discovering any noticeable inhibiting effect on the bacteria.

Aware that clues to the complete elimination of *Salmonella* must be sought within the digestive tract of the maggot itself, we undertook a closer study. As in higher animals, the tract is much longer than the maggot and occupies most of the interior as a mass of coils and loops held together by tracheal tubes. The foregut sucks liquids from the larval medium by means of a muscular pump; peristalsis keeps food moving into either the crop or the midgut. The crop is absent in moderate feeders such as housefly maggots. Blowflies, however, engorge far beyond their gut capacity, and their crops swell with the excess. At the peak of feeding the crop is a miniature sausage dominating the forward half of the maggot. Two days after the maggot ceases to feed the crop is a collapsed sac [*see illustration on page 96*]. The midgut is the center for digestion and absorption, although some absorption occurs in the hindgut. By introducing suitable dyes into the food and dissecting a maggot three or four hours later, we found that the entire tract is primarily alkaline except for the mid-midgut, which has a *p*H of 3 to 3.5. These *p*H characteristics are also found in the tracts of germ-free maggots, indicating that the *p*H is mostly independent of bacteria. We do have evidence, however, that bacteria contribute to the rather high *p*H of 8.5 in the rectum.

During the past two years we have been studying the fate of *Salmonella* and other bacteria in various regions of the digestive tract of maggots, prepupae and adults. We have introduced a single species of bacteria and quantitatively recorded its survival, then two species and finally combinations of three species. We have compared each organism's survival in the host with its response in a broth and in a chick embryo.

We found that the number of organisms present in the maggot's crop and entering the midgut depended on their density in the larval medium. *Proteus* always outnumbered *Salmonella* by 10 to 100 times in both places; *Pseudomonas* had a tenfold superiority over *Salmonella*. We were not surprised to find that the acid midgut destroys bacteria, but the rate and extent to which this occurs are phenomenal. During active feeding millions of organisms constantly pour into this region, but only a few hundred survive passage through the "valley of death." Taken collectively, the normal bacteria appeared to be more rugged than our test species. Nevertheless, a drop from 25 million to as few as 50,000 represents considerable destruction. The sojourn of organisms in this acid section probably lasts less than five minutes. Materials can pass through the entire tract and arrive in the rectum



FLASK EXPERIMENT demonstrated a correlation between the growth of maggots and the degree of microbial proliferation in the growth medium. Sterilized eggs put into a sterilized medium (*a*) hatched but the maggots failed to grow beyond the first instar, or the stage between the egg and the first molt. When the medium was incubated at warm temperature for a day (*b*) so that microbes could develop in it, the maggots grew to the second instar. Incubation of medium for two days (*c*) produced normal development.

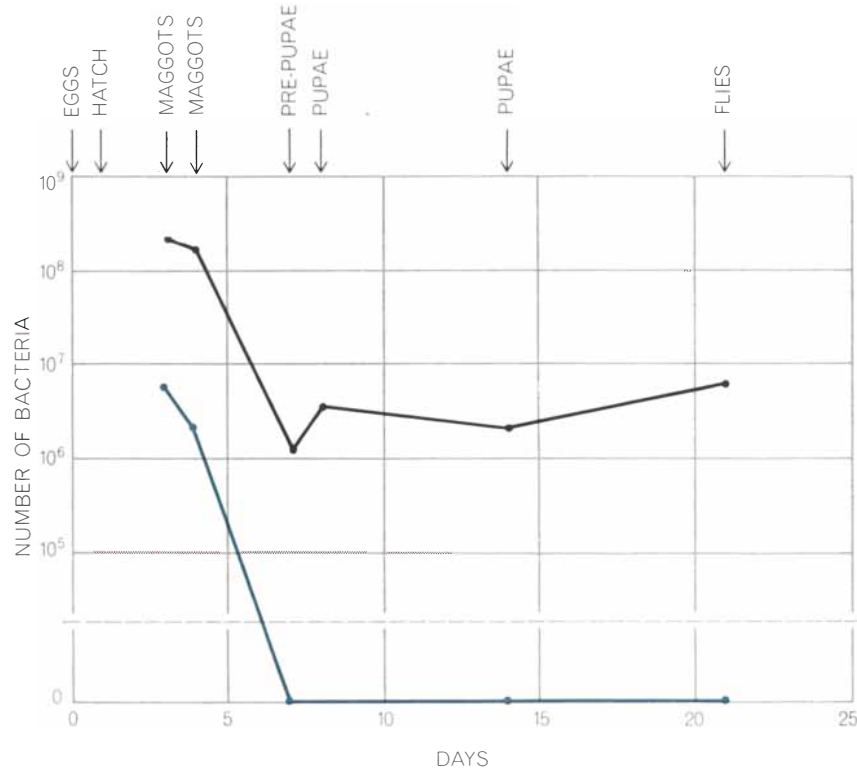
in only 30 minutes. The acid midgut is about a tenth of the tract's total length, but this brief exposure is enough to eliminate entire populations of *Salmonella* and other bacteria. Beyond this region the survivors may multiply. Some species, *Salmonella* included, rarely recover, whereas *Proteus* usually regains its dominance.

The nature of the acid in the maggot's tract is unknown. We have compared the survival of several kinds of bacteria in varying concentrations of phosphoric and hydrochloric acid. At *p*H readings comparable to that in the acid midgut we have found that a similar kill is obtained only after an exposure of several hours to half a day.

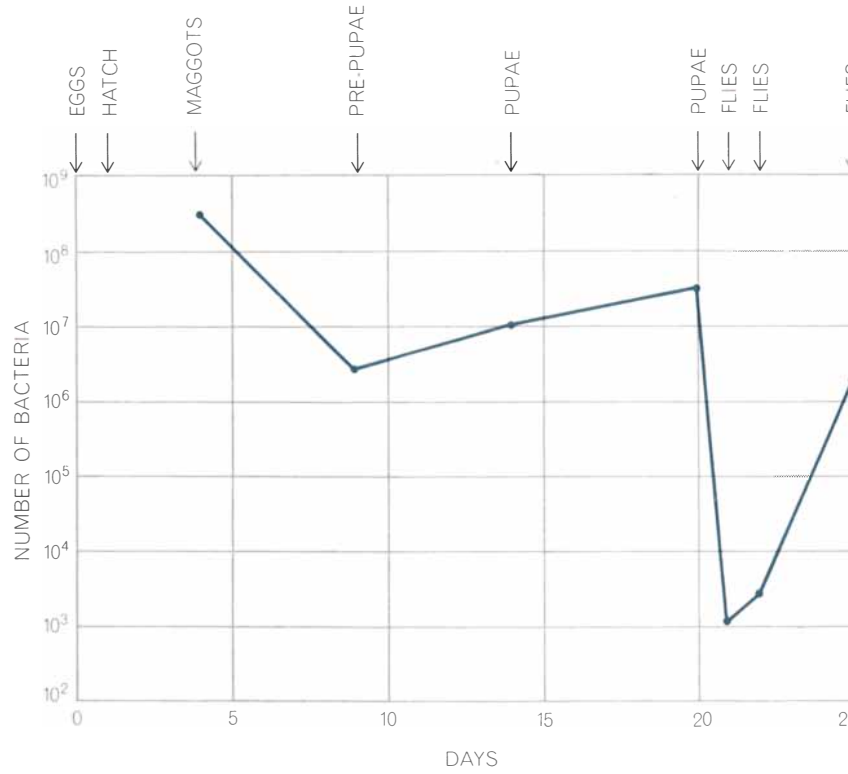
Bacteria do better in the monocontaminated state; lacking competition, they are present in maximum numbers for the journey through the tract. In spite of this advantage, *Salmonella* fares poorly in the maggot compared with *Proteus*. We know that *Proteus* is more acid-resistant than *Salmonella* and that its population potential is slightly greater. These and other differences between organisms add to the complexity of microbial interactions in so dynamic an environment as the fly's digestive tract.

Another provocative question is posed by the inhibition of *Salmonella* by the *Proteus* flora. There is good reason to believe that intestinal flora play a part in resistance to infection in many animals. Rolf Freter of the Jefferson Medical College and Samuel B. Formal and his co-workers at the Walter Reed Army Institute of Research have demonstrated the phenomenon in mice and guinea pigs. By changing the intestinal flora they have made normally resistant animals susceptible to cholera and dysentery. On the other hand, it is known that a dog or cat, perhaps by virtue of its intestinal flora, will often show no effects from a dose of 15 billion *Salmonella*, which would make a human severely ill.

The interaction of microorganisms deserves more research attention than it has received. It may shed a good deal of light on the fundamental processes of infection and resistance to infection. For investigating this matter the flies are unusually convenient laboratory subjects. They can easily be bred in the germfree state and observed under exposure to measured concentrations of various types of microorganisms. Thus the symbiosis of flies and microorganisms, which is so dangerous to man, may lead in the laboratory to a better understanding of infectious disease.



AUTOSTERILIZATION IN MAGGOT was measured in an experiment in which the author participated. Maggots of the bluebottle fly were drenched with *Salmonella* bacteria. In two days before becoming a pupa the maggot reduced its *Salmonella* population from more than a million to zero (*colored curve*). Count of insect's normal bacteria is also shown (*black*).



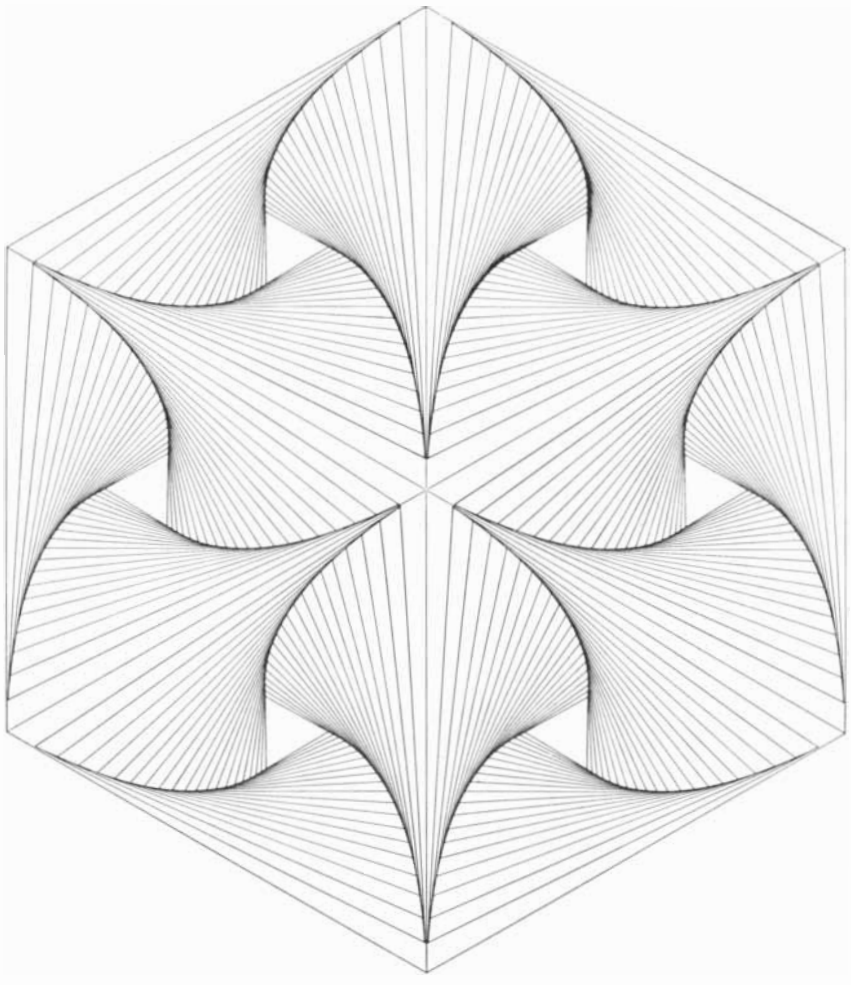
SURVIVAL OF *SALMONELLA* occurred when maggot's normal bacteria were eliminated so that *Salmonella* existed without competition. Later experiments indicated that *Salmonella* in a normal maggot are suppressed by *Proteus mirabilis*, part of the fly's normal flora.

MATHEMATICAL GAMES

On the relation between mathematics and the ordered patterns of Op art

by Martin Gardner

Op (for "optical") has topped Pop (for "popular") as the fashionable gallery art of 1965; its patterns quiver in advertisements and on dresses, bathing suits, ties, stockings, window shades, draperies, wallpaper, floor coverings, package designs and what have you. Op art, as everyone surely knows by now, is the new name for a form of hard-edge abstractionism that has been around for half a century. Its distinguishing feature is a strong emphasis on mathematical order. Sometimes it is accompanied by effects intended to dazzle and wrench the eye: vivid colors that generate strong afterimages when the eye shifts, optical illusions, striped and dotted patterns that torture the brain like the retinal scintillations of migraine. One branch of Op art deals with moiré patterns of the type described in this magazine by Gerald Oster and Yasunori Nishijima [see "Moiré Patterns," May, 1963] and by C. L. Stong ["The Amateur Scientist," November, 1964]. Indeed, Oster's shimmering patterns have been exhibited in several New York art galleries.



Design based on the "three-bug problem"

The Op trend, many critics have been saying, is more than just a rebellion (like Pop) against the randomness of abstract expressionism; it reflects the growing extent to which mathematics, science and technology press on our lives. This magazine, it has been observed, has been presenting Op art for years. Consider the following covers: "Perfect" Rectangle, November, 1958; Reactor Fuel Elements, February, 1959; "Graeco-Latin" Square, November, 1959; "Visual Cliff" (with its distorted checkerboards, a popular Op motif), April, 1960; Spark Chamber, August, 1962; Moiré Pattern, May, 1963, and Afterimage Test Pattern, October, 1963. These covers are almost pure Op. They leave little doubt about Op's close kinship with modern science.

Although Op art is sometimes rich and warm with colors, its appeal seems to lie more in its cold, rigid, precise, unemotional and impersonal qualities. Its astonishing popularity revives ancient questions about art and mathematics. To what extent is art ruled by mathematical laws? To what extent can pure mathematical structure arouse aesthetic emotions? "The chief forms of beauty are order and symmetry and precision," wrote Aristotle in his *Metaphysics* (Book 13), "which the mathematical sciences demonstrate in a special degree." "A mathematician . . . ," declared G. H. Hardy in *A Mathematician's Apology,* "is a maker of patterns. . . . [His] patterns, like the painter's or the poet's, must be *beautiful;* the ideas, like the colors or the words, must fit together in a harmonious way. Beauty is the first test: there is no permanent place in the world for ugly mathematics."

We are surrounded on all sides, say the defenders of Op, by hard-edge squares and circles, ellipses and rectangles. The windows of a skyscraper, the streets of a city, the fronts of file cabinets all form orthogonal patterns like a checkerboard. Why should these basic geometric designs not be reflected in our art? Opponents counter: But we want to escape from, not be reminded of, the low-order curves and 90-degree angles of a technological culture. Our eyeballs ache for random curves, impure colors and soft edges; for the patterns of leaves and clouds and water in motion. Who can write an equation for the shape of an oak tree? The mathematical structure is still there, but in

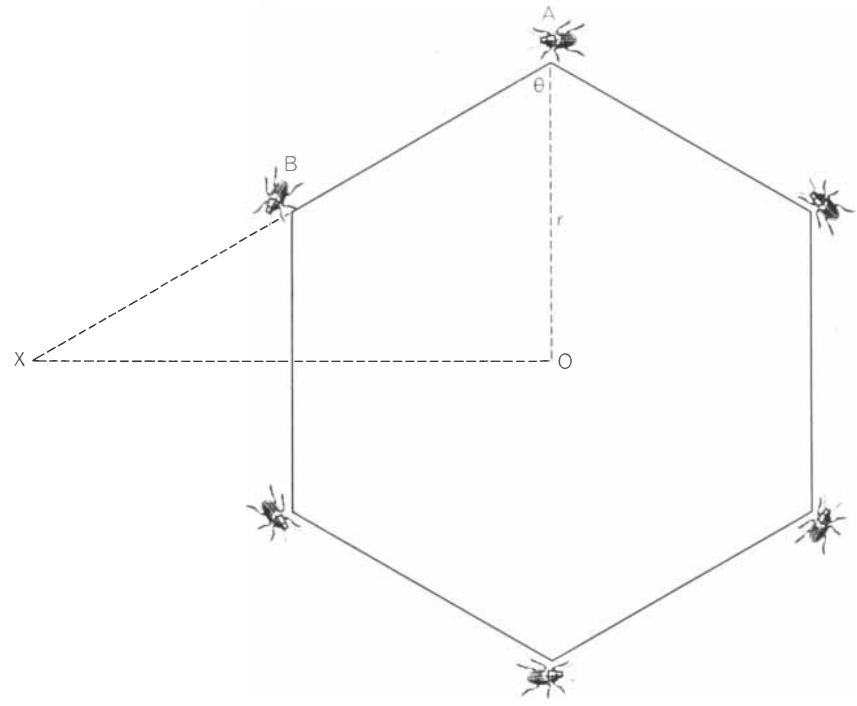
nature, as in less rigid abstract art, it is more complex, more careless, and—say Op's detractors—aesthetically less boring.

Whatever one's attitude toward Op, there is no denying its fascination. Nor is it surprising that many Op patterns are closely related to problems of recreational mathematics. Consider, for example, the nested and rotating squares (or rectangles) that appear in so many Op paintings and fabric designs and that whirl inward on the cover of this month's SCIENTIFIC AMERICAN. The pattern can be interpreted as an illustration for the well-known "four-bug problem," which appeared in this department in November, 1957. Four bugs at the corners of a square start to crawl clockwise (or counterclockwise) at a constant rate, each moving directly toward its neighbor. At any instant, as the bugs march toward a meeting point at the center, they mark the corners of a square, and as they crawl the square they delineate both diminishes and rotates. Each bug travels on a logarithmic spiral with a length exactly equal to the side of the original square.

If n bugs start at the corners of any regular n-sided polygon, their positions at any instant during their march will mark the corners of a similar polygon. Like the square, this polygon will shrink and turn as the bugs spiral inward. A design based on the triangular case is shown in the illustration on the opposite page, originally drawn for an old issue of *Scripta Mathematica* by Rutherford Boyd. The picture contains nothing but triangles, but they are hard to see because the eye is so strongly dominated by the spiral curves. In this case each logarithmic spiral is 2/3 of the original triangle's side.

For regular polygons of more than four sides the length of each bug's path is greater than a side. As J. Charles Clapham proved in the now defunct *Recreational Mathematics Magazine* (August, 1962), the length of the path of a bug starting at corner A can be found trigonometrically by extending a side AB [*see illustration on this page*] and locating on it a point X such that the angle AOX is 90 degrees. The distance AX—which is equal to r times the secant of angle θ—is the distance the bug travels. As the illustration shows, on a hexagon each bug's path is twice the length of a side.

Clapham's simple formula also applies to the square and triangular cases, and even to the degenerate "two-sided polygon"—a straight line with a zero angle θ and bugs at each end that tramp toward each other until they bump head on. At the other extreme, the circle can be considered a degenerate "infinite-sided polygon" with bugs at an infinite number of "corners." These bugs march forever around the circle like the Pine Processionary caterpillars in a famous experiment of Jean Henri Fabre's, which trailed each other for eight days around the rim of a large vase. When we apply Clapham's right triangle to the circle, sure enough, angle θ is 90 degrees and the hypotenuse is infinite.

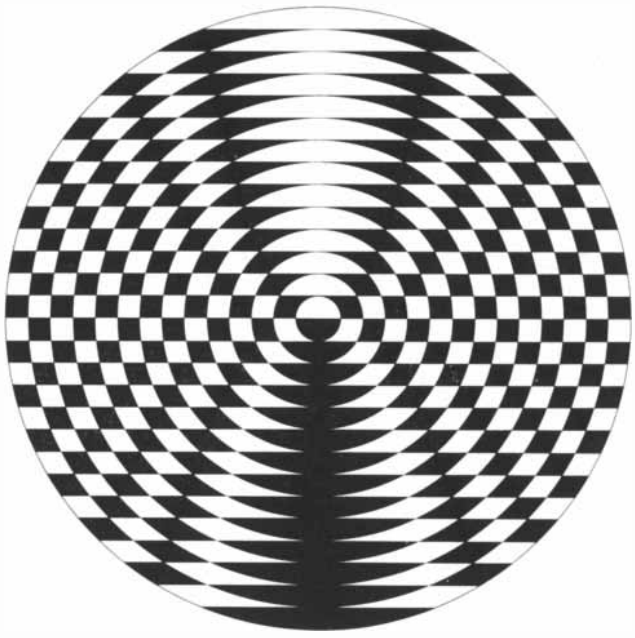


Calculating the length of the bug's path

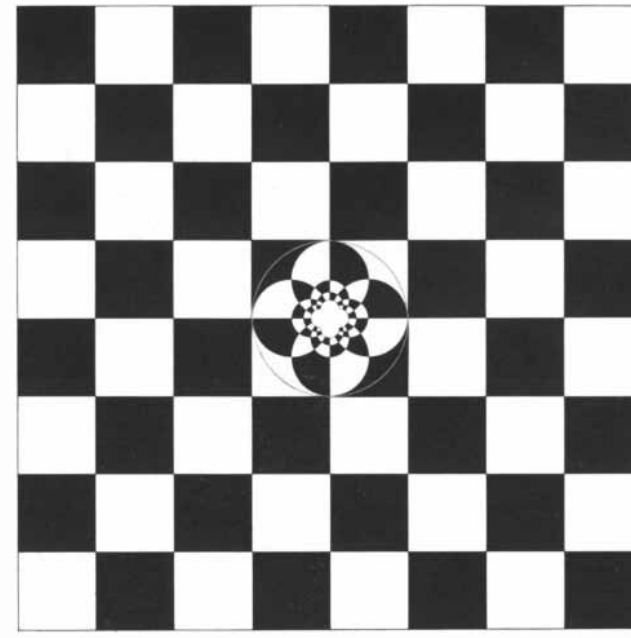
One suspects that Op painters both here and abroad have yet to discover the thousands of eye-twisting patterns that lie buried in scientific and mathematical textbooks and back copies of academic journals. Early issues of *Scripta Mathematica,* for example, vibrate with exciting pre-Op. The illustration at top left on the next page shows a striking pattern the mathematician Hermann Baravalle obtained by ruling parallel lines across concentric circles and then coloring the regions in checkerboard fashion. One might think that this pattern is topologically the same as a square checkerboard—in other words, that a square checkerboard on a rubber sheet could be continuously deformed to produce the pattern. This is not the case, but it suggests a pretty puzzle to be answered next month. Can you cut the pattern into two parts with one straight cut so that each part is topologically equivalent to a square checkerboard?

In the illustration at top right on the next page Baravalle has inverted every point P that lies outside the circle on the checkerboard into a corresponding point P' inside the circle, such that $OP \times OP' = r^2$, where O is the circle's center and r its radius. Every point on the plane outside the circle is thus put into one-to-one correspondence with every point inside. A line extending outward from the board to infinity corresponds to a line inside the central white space, extending inward toward the center but never reaching it.

Inversion geometry can, of course, be applied to three-space as easily as to the plane. An old mathematical joke says that to catch a lion you just build a cage and perform an inversion operation on the beast. The cosmos itself can be inverted and compressed inside a tennis ball. In this country during the 1870's a religious cult was actually founded on the belief that such an inverted three-space reflects the true state of affairs. Cyrus Reed Teed's "Koreshanity" put the entire universe *inside* the earth. We imagine ourselves on the outside of the earth looking out at gigantic stars scattered through an infinite space; the truth, said Teed, is that we are on the inside of a hollow earth looking *in* at small stellar bodies moving in a space that is the geometrical inverse of the space of orthodox astronomy.



Baravalle's circular "checkerboard"



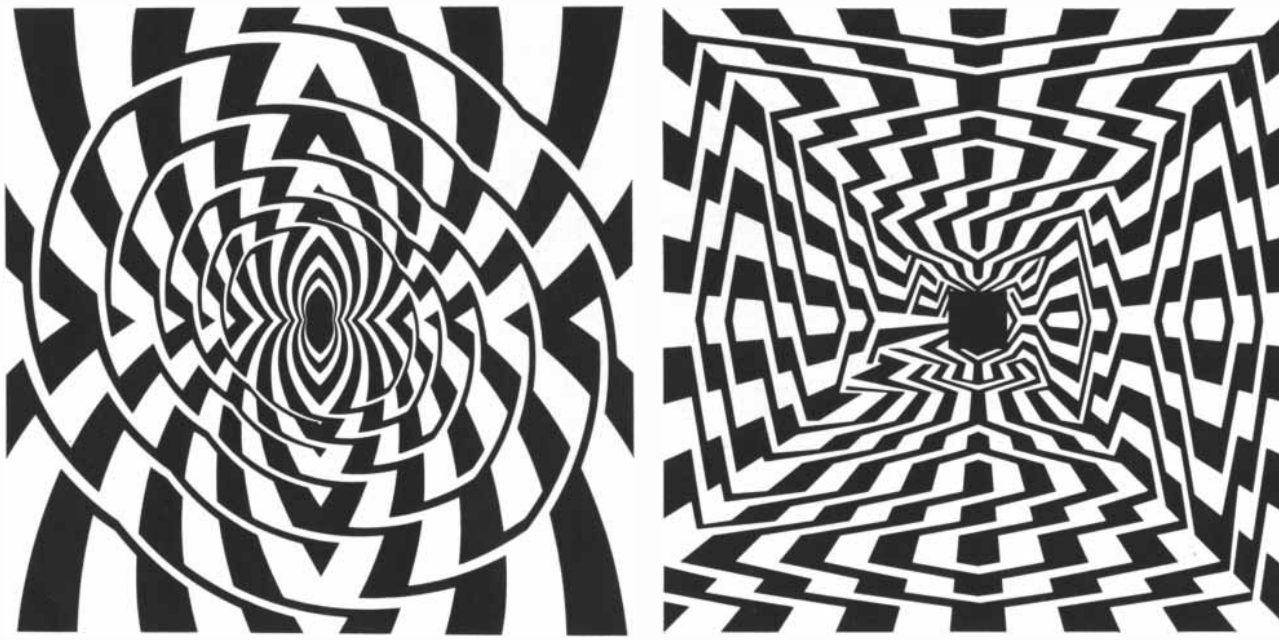
Checkerboard inversion pattern

Teed defended his views in many books and articles; years later his ideas attracted a following in Nazi Germany.

The illustrations below are examples of many vertigo-inducing patterns that were studied by psychologists more than 50 years ago. They are known as "twisted-cord illusions" because they were first discovered by twisting black and white string into a single cord that was then arranged in various ways on differently patterned backgrounds. The figure on the left consists of concentric circles (as you can prove with a compass); in the one on the right a spiral is made up of straight horizontal and vertical "cords" (as you can prove with a ruler).

Tesselations of the plane created by fitting together replicas of the same basic shape have long been used in design and are now turning up in many of the latest Op fabrics. The cross-pentomino appears on an Op dress advertised by Bonwit Teller. *All* polyominoes and polyiamonds (polyiamonds are formed by joining equilateral triangles instead of squares) of order six or less will fit together to cover the plane, but so far I have seen only the cross-pentomino and the *L*-tromino (the latter on a scarf sold by Gimbels in New York City) on Op fabrics. The reader can easily create his own new Op patterns by finding ways to tile the plane with each of the 12 pentominoes and the 12 hexiamonds (for the hexiamond shapes see this department for December, 1964).

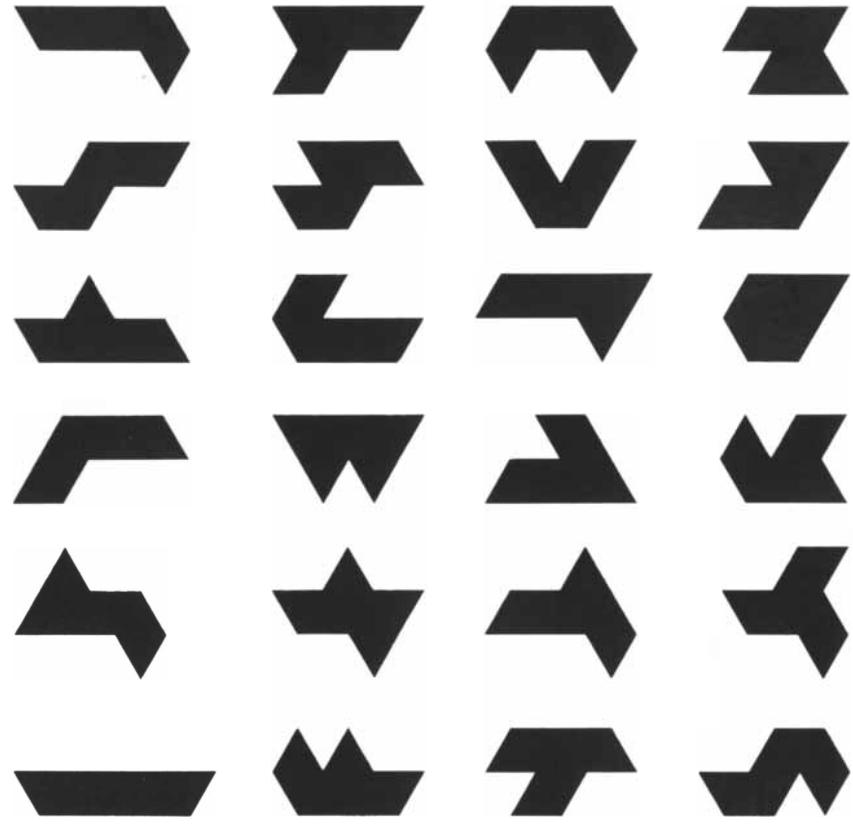
Most, but not all, of the 108 hep-



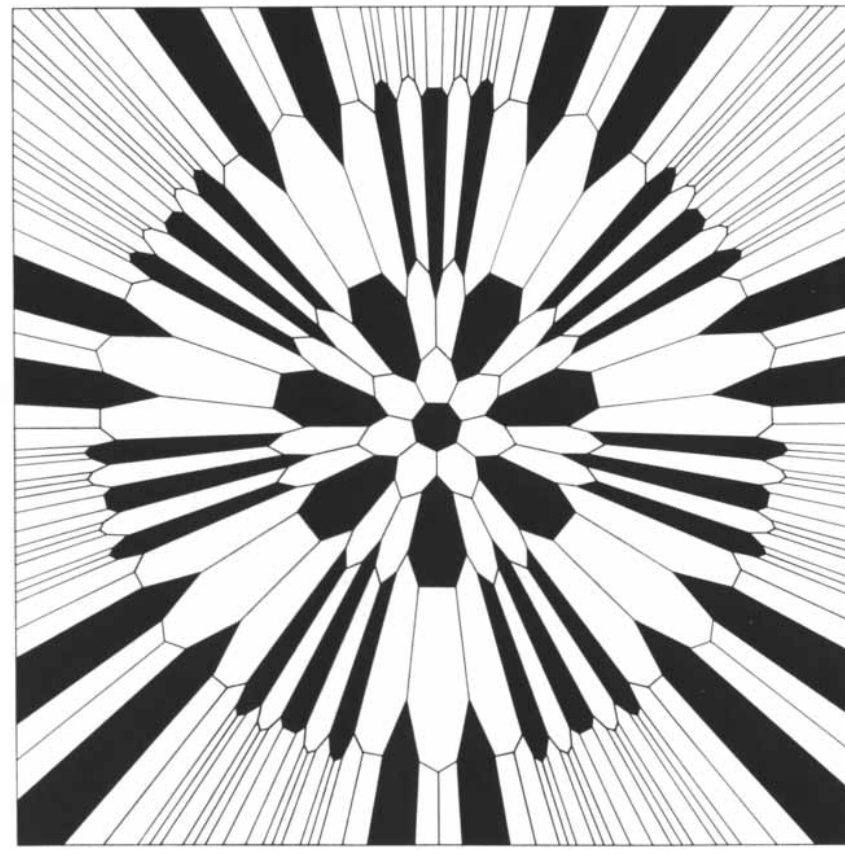
"Twisted cord" concentric circles (left) *and spiral* (right)

tominoes (for their diagrams see Solomon W. Golomb's recent book *Polyominoes*, pages 108–109) will tile the plane. Several British mathematicians are currently working on the difficult question of which of them are *not* plane-fillers. The corresponding problem for the 24 heptiamonds [*see illustration at top right*] was proposed by T. H. O'Beirne of Glasgow and was solved this year by Gregory J. Bishop of Boston. Only one of the 24 shapes is not a plane-filler. Can the reader identify it and prove that it cannot tesselate the plane? Bishop's simple proof will be given next month.

The Op pattern that covers the plane with convex noncongruent heptagons [*bottom right*] embodies a curious paradox that twiddles the brain even more than the eye. If this pattern is repeated infinitely, what is the average angle in it? Since the plane contains nothing but heptagons, and since the interior angles of any heptagon sum to 900 degrees, it follows that the average angle is 900/7 or 128⅐ degrees. Note, however, that every point on the pattern is a meeting of three angles. This surely requires that the average angle be 360/3, or 120 degrees. Explain! The source of this problem and its answer will be given next month.



The 24 heptiamonds. Which cannot tile a plane?



Tesselation of convex heptagons

Last month's problem of finding a minimum-length path covering all eight streets in a square area three blocks on a side, making the minimum number of turns, can be solved with as few as 10 turns, as shown in the top illustration on the next page. The solution is unique except for reflections and rotations. To prove that 10 is minimal, note first that the network has eight vertices where an odd number of paths meet. According to well-known rules, the graph cannot be traversed by one continuous path (without going over any portion of the graph twice) unless the odd vertices are reduced to two or none. This can be accomplished by doubling segments of the graph, but we must do it in a way that adds as little as possible to the total length of the lines. It is easy to see that the shortest path is obtained by doubling three segments as shown in the second illustration from the top on the next page. The doubled segments indicate portions of the original graph that must be traversed twice. This minimal path is 27 blocks long, with *A* and *B* (the two remaining odd vertices) as its end points.

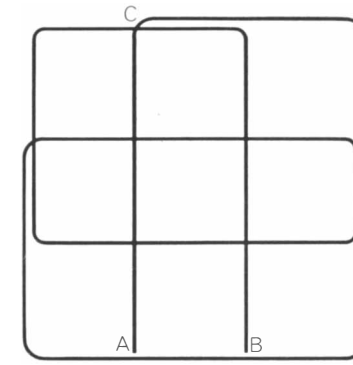
Five streets can be traversed their full length without a turn. If we call

any segment traveled without a turn a "move," it is clear that these five streets demand at least five moves. Each of the remaining three streets requires at least two moves because each has a middle block that must be traversed twice. Therefore any continuous path from *A* to *B* must have at least 11 moves, which is the same as saying it must have at least 10 turns. Suppose we start at *A* and proceed to *C*. We cannot turn left at *C* because then two moves would be necessary to complete the right two-thirds of the top street, making three moves in all for this street, whereas the minimum-turn path limits this street to two. So we must turn right. Continuing in this way, analyzing all alternatives at each juncture, we find that only two travel patterns complete the trip in 10 turns. One pattern is a mirror image of the other. The third illustration from the top at the right shows a 27-block path with the maximum number of turns: 26. This too is unique except for reflections and rotations.

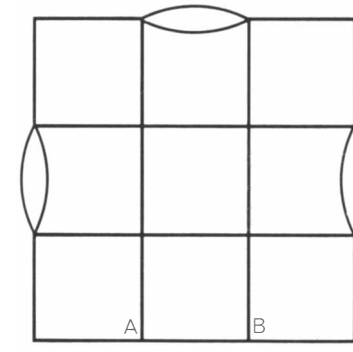
The longest path for visiting the row of 10 houses, in the second of last month's problems, is shown below. It has a length of 49 units. When the number of houses is even, the length of the "worst" path is $\frac{1}{2}(n^2 - 2)$; when it is odd, the length is $\frac{1}{2}(n^2 - 3)$. For the derivation of both formulas see problem No. 64 in Hugo Steinhaus' *One Hundred Problems in Elementary Mathematics* (Basic Books, 1964).

Smith's house number is 239, in a row of 169 houses. Jones's is 408, in a row of 288 houses. The solution for Smith involves finding integral solutions of $2x^2 - 1 = y^2$; for Jones, integral solutions of $2x^2 + 2x = y^2$, where x is the number of houses and y the house number. Both Diophantine equations have an infinity of solutions, but we were told that the number of houses in each case is between 50 and 500. This restricts each equation to one pair of values for x and y.

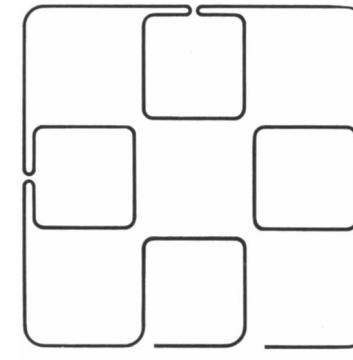
The probability that exactly one letter will go into the wrong envelope, if four are inserted at random into four envelopes, is zero, because it is impossible for three letters to match their envelopes and the remaining one not to match! The quotation, "Neither snow nor rain nor heat nor gloom of night . . . ," which is carved on the façade of New York City's General Post Office Building at Eighth Avenue and 33rd Street, is from the Greek historian Herodotus.



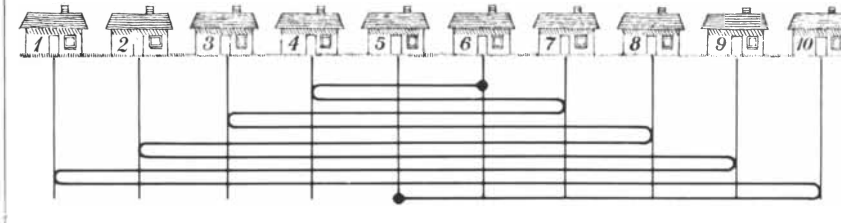
Minimum-turn solution



Minimum-turn proof



Maximum-turn solution



Answer to the "worst path" problem

THE AMATEUR SCIENTIST

The semiconducting "Hall device" opens up a broad new field of experimentation

Conducted by C. L. Stong

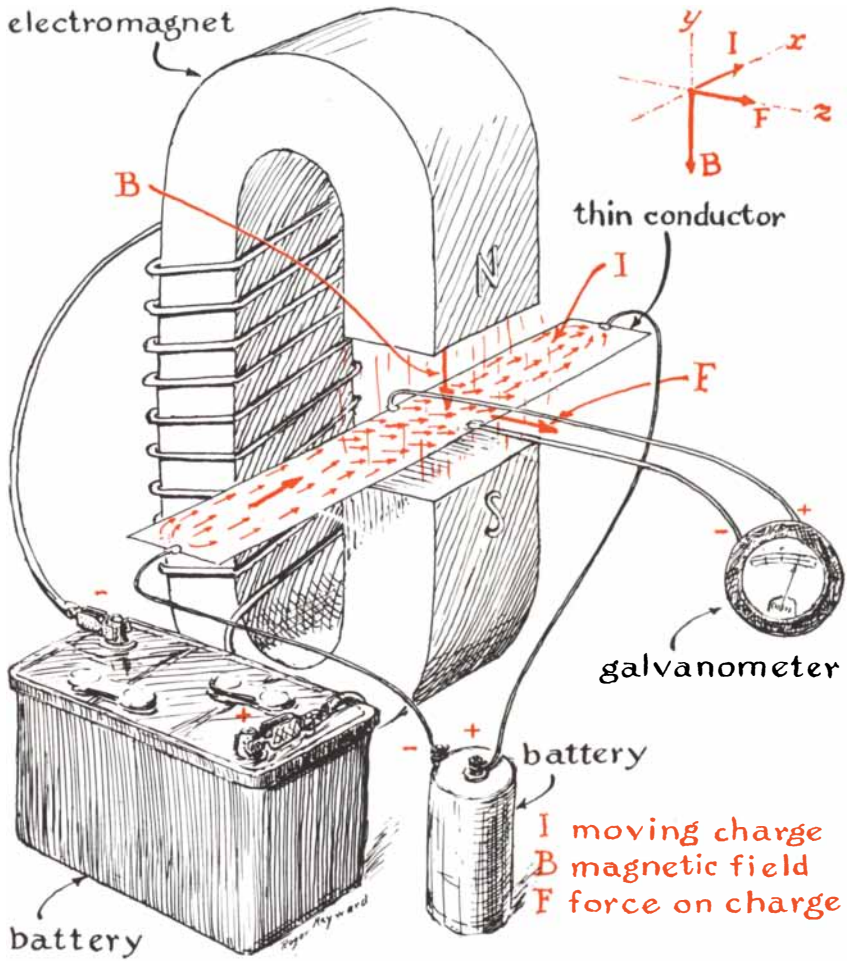
A recently developed semiconducting device that controls electric current by means of an electromagnet has opened to amateurs a new field of experimentation that is potentially as interesting as the one introduced by the transistor. The principle on which the new device is based dates from work done in 1879 by Edwin H. Hall at Johns Hopkins University. Hall, who had placed between the poles of a magnet a thin strip of copper that carried an electric current, observed that a voltage appeared across the width of the strip as long as the magnet was in place. He found that he could reverse the polarity of the voltage by reversing the magnet.

Hall explained the effect by assuming that the magnetic field caused the current to crowd toward one edge of the strip, much as a jetty diverts the flow of water in a river, and he concluded that the voltage appeared as a result of the crowding. Little came of the experiment, apart from its academic interest, because the potential difference Hall observed was less than a millionth of a volt. Recently, however, it was discovered that the Hall effect can be increased more than 100,000 times by substituting for Hall's copper strip a thin wafer of semiconducting material such as an alloy of indium and arsenic. An apparatus that makes use of this principle is called a Hall device.

According to Edwin D. Sisson of Columbus, Ohio, who has contributed to the development of numerous Hall devices, the discovery of the augmented Hall effect has solved some of the experimenter's most stubborn circuit problems. For example, a new instrument that incorporates a Hall device and has no moving parts can measure the strength of magnetic fields directly and as easily as voltages and currents have been measured in the past. When the Hall device is equipped with an electromagnet, it can do several things: multiply one current by another; generate trigonometric and other mathematical functions; convert direct current to alternating current; modulate one alternating current by another; measure current, power, linear displacements, physical and electrical angles, and perform numerous sensing functions. Sisson writes:

"The output voltage of the Hall device varies in proportion to the product of the current in the semiconducting material and the current that energizes the electromagnet; the voltage varies inversely with the thickness of the semiconductor. For this reason thickness is minimized; in a typical device it amounts to only a few thousandths of an inch. The thinness of construction makes possible the use of the device for measuring magnetic fields in magnets where the gap between poles is narrow. The area of the device can also be made small, which improves its resolution as a probe for exploring the uniformity of magnetic fields.

"On the other hand, semiconducting materials are relatively brittle and be-

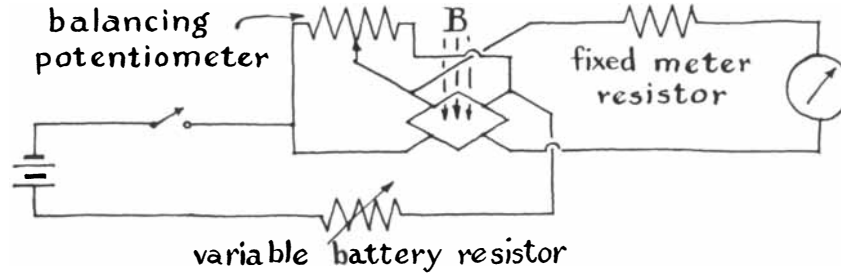


A schematic arrangement of a Hall device

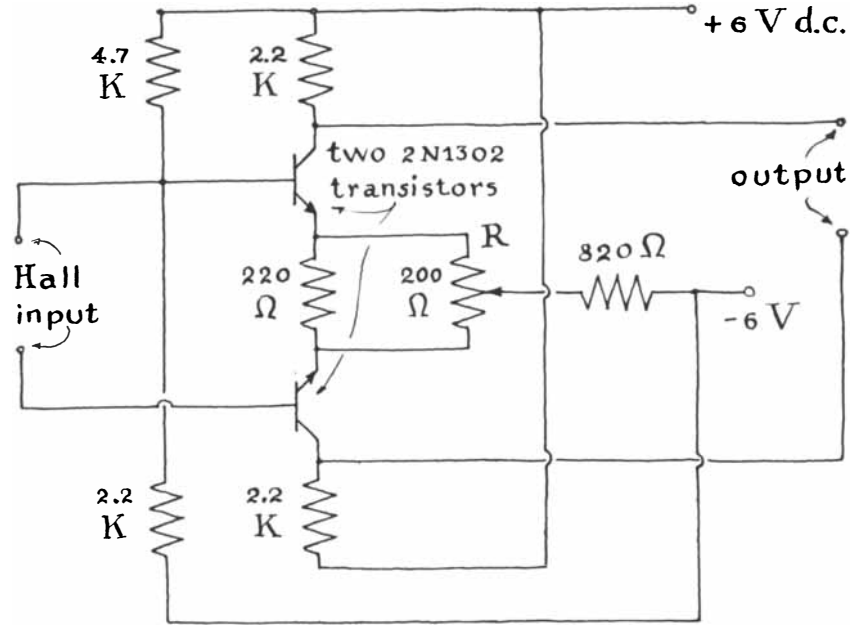
come fragile if they are made too thin. Moreover, reduction in thickness is always accompanied by a corresponding reduction in the maximum current the wafer can carry without overheating. Both the physical strength of the device and its capacity for transmitting current can be increased, however, by mounting the wafer on a substrate that provides physical support and conducts heat away from the unit.

"In typical circuits the Hall device must be supplied by current from a source such as a battery. In most applications the device is mounted between the poles of an electromagnet energized by a separate current. In some applications both currents may be under investigation or measurement; in others the electromagnet can be equipped with more than one winding for accepting two or more currents. The output voltage can be measured directly by a millivoltmeter, displayed on an oscilloscope or amplified for measurement by a conventional voltmeter. The output voltage is equal to the product of the width of the semiconducting material, the density of the current in the material, the density of the magnetic field and a quantity known as the Hall coefficient. This quantity is a measure of the effectiveness with which the material generates output voltage when immersed in a magnetic field. The Hall coefficient of copper is quite low, which accounts for the low output Hall observed during his pioneering experiment. The output is highest when the direction of the magnetic field is perpendicular to the plane of the semiconducting material, and it decreases in proportion to the trigonometric sine of the angle when the direction of the field is inclined with respect to the plane.

"Perhaps the simplest application of the device is also the one of most interest to amateurs: its use as a fluxmeter for measuring the strength of magnetic fields. Few laymen normally have access to fluxmeters of the conventional type, in which a voltage is generated in a moving coil of wire by the field under measurement. The Hall device generates a continuous output that varies in direct proportion to the field strength as long as current is maintained through the sensitive element. It requires no moving parts and can measure fields of constant intensity as well as alternating fields, including those that oscillate at the radio frequencies. It senses only the portion of the field that passes through the semiconductor and it can therefore be used to explore variations in the pattern of the field and to chart the contour of nonuniform fields and of stray fields.



Circuitry for a gauss meter incorporating a Hall device



Schematic arrangement of a transistor amplifier

"A gauss meter for measuring the strength of magnetic fields can be built by amateurs. It makes use of a Hall device such as the Type BH200 manufactured by F. W. Bell, Inc. (1356 Norton Avenue, Columbus, Ohio 43212). The unit can be energized by a six-volt dry battery through a 33-ohm resistor. A 1,000-ohm potentiometer should be connected across the input terminals of the Hall device, with the sliding contact connected to one side of the meter circuit, for adjusting the meter indication to zero in the absence of a magnetic field. The output voltage is measured by connecting a millivoltmeter to the output terminals of the device. If the experimenter wishes, the calibration of the meter can be adjusted by inserting a resistor in series with one of its leads [*see upper illustration above*]. The value of this resistor will depend on the characteristics of the meter and the desired scale. It must be determined experimentally.

"A more sensitive instrument for measuring field strength (and one that is easily adaptable for other experiments) can be made. To do so one equips the Hall device with an amplifier that includes two transistors connected differentially so that the performance of one compensates for that of the other when the characteristics of the pair are altered by variations of temperature or other disturbances [*see lower illustration above*]. The unit will amplify both direct and alternating current. It includes a potentiometer for adjusting the meter to zero indication in the absence of a magnetic field. A vacuum-tube voltmeter, connected to the output of the amplifier, is used for measuring the amplified output voltage of the Hall device. The amplified output amounts to about two volts per 1,000 gauss of magnetic-field strength. Alternating fields of course generate an alternating output voltage, which must be measured by a voltmeter designed for use with alternating current.

"Accurate calibration of the fluxmeter requires the use of a magnetic field of known strength. The Hall device is

placed in the field and a note is made of the corresponding output voltage. The value of the voltage is divided by the known strength of the field in gauss. This computation expresses the intensity of the field in terms of voltage and constitutes the calibration. For example, a measurement of 1.2 volts when the Hall device is immersed in a magnetic field of 600 gauss indicates a sensitivity of 1.2/600, or .2 volt per 100 gauss. The scale of the meter can be graduated accordingly for direct reading. Alternatively a graph can be drawn as a reference for converting voltage indications into units of field strength. The amplifier can be used for measuring fields up to about 1,500 gauss; beyond that, in the range of 1,500 to 20,000 gauss with which an amateur is likely to be involved, it is not needed.

"A magnetic field of accurately known intensity for use in calibrating the instrument can be generated by constructing a 'long solenoid.' This is an insulated-wire coil at least 20 times longer than its width, wound of wire that is less than a tenth as thick as the width of the coil. The coil consists of a single closely wound layer of wire, each turn in contact with its neighbor. The intensity of the magnetic field developed by a direct current inside the center of the coil is equal to .4945 multiplied by the number of turns per inch of the coil length and by the current in amperes. ($B = .4945\ IN/l$, in which B is the magnetic flux in gauss, I the current in amperes, N the total number of turns of wire and l the length of the coil in inches.)

"A coil 15 inches long, for example, wound with 660 turns of 24-gauge enamel magnet wire on a phenolic tube 3/4 inch in diameter will develop at its center a magnetic-field strength of 21.7 gauss per ampere of current. If the wire is wound at constant pitch (accomplished most accurately by prethreading the phenolic tube), the computed value will not be in error by more than .5 percent. The coil can be used for setting up either direct-current or alternating-current fields. When the coil is energized by alternating current of relatively high frequency, losses in the winding and from resonance effects can cause the current in the coil to differ from that in the external circuit. Such losses do not produce significant inaccuracies at frequencies below 10,000 cycles per second. Above this frequency another source of error can appear in the form of voltages that may be induced in the leads attached to the Hall device. These four connections consti-

tute a pair of small loops that can act as single-turn pickup coils. To detect possible error from this source disconnect the current supply from the Hall device while maintaining the alternating field. The output voltage should then drop to zero. If it does not, stray voltages are being induced in the leads.

"It is not practical to leave a discontinuity in the coil winding for inserting the Hall device into the center of the coil, because the opening would alter the distribution of the magnetic field. The Hall device must be inserted through one end of the coil and carefully aligned so that the plane of the element is perpendicular to the axis of the coil. This can be done by attaching the unit to the end of a plastic rod.

"The current used to energize the calibration coil must not be so high that it heats the wire and thus raises the temperature of the Hall device appreciably above that of the room. The accuracy of the calibration will be no better than that of the ammeter used for measuring the current in the coil. Once a calibration point for the Hall device is found by use of the long solenoid the same constant can be applied to any field intensity within the limits of the Hall device. The unit recommended for this introductory experiment will generate output voltages that vary in direct linear proportion to the magnetic field within 3 percent to 10,000 gauss if the unit is supplied by a current of 100 milliamperes and the output leads are connected to a meter or an amplifier of 100 ohms. Other Hall devices, available at higher cost, are accurate to within .25 percent to 10,000 gauss.

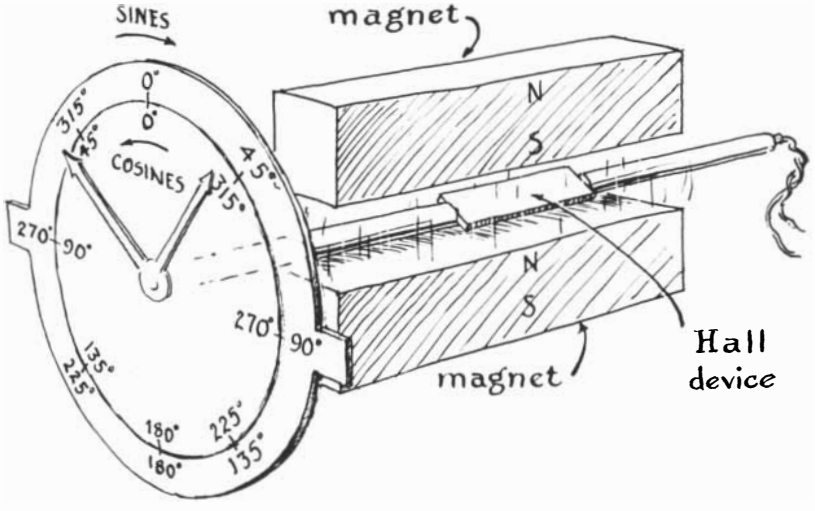
"The BH200 Hall device can be bought separately from F. W. Bell, Inc., or as part of an inexpensive kit that includes a transistor amplifier, a pair of bar magnets of approximately known field strength and an explanatory booklet. To calibrate the gauss meter with these magnets the experimenter simply places the Hall device against the surface of either magnet in a position that produces the highest meter reading. In addition to functioning as a gauss meter, the apparatus included in the kit can be used for demonstrating the pattern of magnetic-field lines in the space around a magnet and for investigating the directional effect of field lines with respect to the plane of the Hall device. The device also can act as a linear-displacement and an angular-displacement transducer and as a proximity detector and position sensor.

"To measure angles in terms of trigonometric functions, for example, the Hall device is mounted on a shaft that is free to rotate between the two bar magnets. One end of the shaft is equipped with a pair of pointers for indicating angular displacement [*see upper illustration below*]. When one of the pointers indicates zero degrees, the output voltage is proportional to the trigonometric sine of the angle indicated by the second pointer and to the cosine of the angle indicated by the first pointer. By mounting a pair of Hall devices on the shaft, displaced 90 degrees in relation to each other, the device will generate sine and cosine outputs simultaneously. It will also act as a rectangular-to-polar coordinate converter.

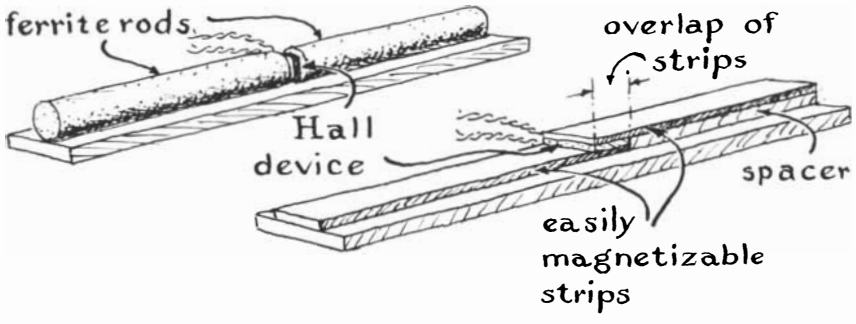
"The sensitivity of the Hall probe in large homogeneous fields can be increased by the use of flux collectors: two lengths of ferromagnetic material such as iron or ferrite bars. The device can be sandwiched between the butted ends of the bars or between the overlapped ends [*see lower illustration below*]. The lines of magnetic flux in the vicinity are attracted by the ferromagnetic material and are thereby concentrated in the Hall device. Long, slender rods are more effective than short ones. For maximum sensitivity the air gap between the rods occupied by the Hall device should be made as small as possible because an air gap has the effect of shortening the rods. The area of the rods in contact with the Hall device should equal that of the sensitive portion of the semiconducting material.

"An effective flux concentrator can be made from strips of laminated sheet iron removed from an old transformer. Several laminations are stacked to make a rod of rectangular cross section. If the laminations happen to be slightly magnetized, as may be the case with laminations removed from the transformer of an old radio set, they can be demagnetized before use by passing the strips through a coil of 100 turns or more that carries several amperes of 60-cycle alternating current. Flux concentrators can easily increase by a hundredfold the strength of the magnetic field in space that acts on the Hall device.

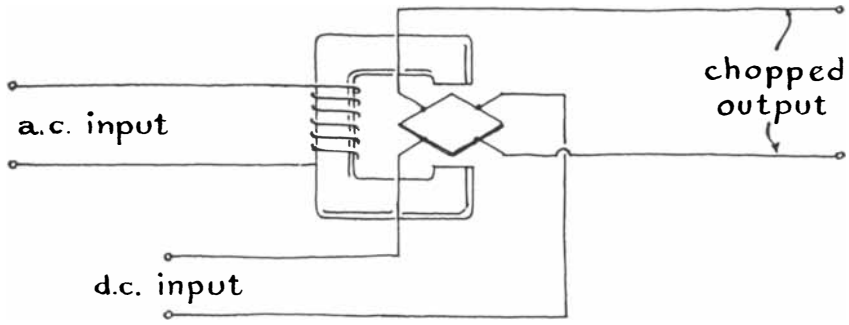
"Tests were made of a type BH700 Hall device, a more sensitive unit than the BH200. The BH700 device was equipped with a high-gain transistor amplifier and a flux concentrator consisting of a pair of nickel-alloy steel strips three inches long, 1/4 inch wide and .014 inch thick. The tests produced



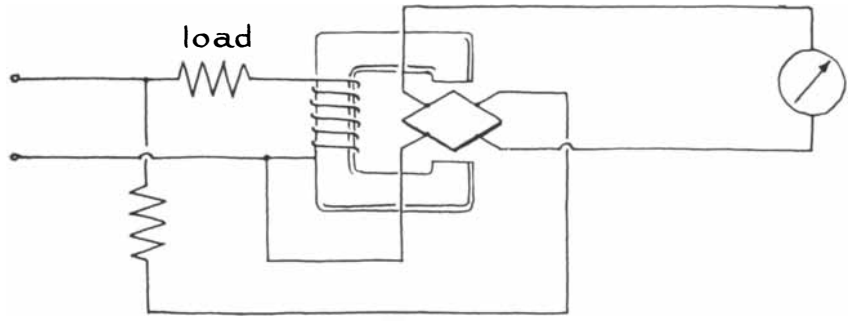
Arrangement of an apparatus for generating trigonometric functions



Details of devices for concentrating a magnetic flux



Circuitry of a "chopper" using a Hall device



A circuit using a Hall device in a wattmeter

a response of one volt in the earth's field when a current of 14 milliamperes was applied to the device. The probe was zeroed in the presence of the earth's field by adjusting the balancing potentiometer of the amplifier for peak readings as the probe was reversed in the earth's field and finally rotated to the intermediate or null position.

"An interesting sensitivity test was also made with the BH700 device. First the zero controls were adjusted so that the two peak responses due to the earth's field were equal in amplitude. The probe was rotated to the position of zero output, the null point, and securely fastened. The gain of the amplifier was then increased until the null potential measured five millivolts. The Hall device was energized by alternating current at 400 cycles per second for amplification by a high-gain alternating-current amplifier. The current in the Hall device was then increased to 200 milliamperes, its rated limit. At this point the zero controls became so sensitive that a 10-turn potentiometer had to be substituted for the single-turn 'coarse' potentiometer of the amplifier.

"At this critical adjustment the sensitivity of the instrument was computed to be 1.7×10^{-4} gauss, equivalent to 17 gammas. (The horizontal intensity of the earth's magnetic field in the Northern Hemisphere varies from about 8,000 to 30,000 gammas.) Operation at a sensitivity this extreme is not at all impossible. The stability of the instrument is quite good if care is taken to maintain a constant current through the Hall device and to protect it from strains and temperature changes. The probe must be fixed so that its plane is perpendicular to the earth's field. An angular displacement will generate a large change in the output voltage.

"Ferromagnetic rods for concentrating magnetic flux display their maximum response to fields that parallel the long axis of the assembly and their minimum response to transverse fields. Surfaces of equal response in the space near the rods appear approximately as spheres positioned around each end of the rod with centers on the axis of the rod. It is possible to distort or reshape the spatial response by adjusting the balancing potentiometer. This action reshapes the null plane of the concentrator into a conical surface around the axis of one pole of the concentrator, reduces the sensitive lobe inside the cone and increases the size of the opposite lobe.

"Continued adjustment will ultimately reduce the small lobe to zero and convert the null cone into a null line that coincides with the axis. With the apparatus adjusted for high sensitivity and the output zeroed with the probe in a fixed relation to the earth's field, any magnetic change in the environment, such as one caused by energizing an electromagnet, will be detected independently of the earth's field. Metal objects that are assumed to be demagnetized often show a residual field when they are brought near the probe of an instrument in critical adjustment. In fact, it is frequently difficult to determine whether the response of the instrument is due to residual magnetism in the object or to the distortion the object introduces in the earth's field.

"The Hall device can also be applied to the solution of many conventional circuit problems. For example, it makes an ideal 'chopper,' which is a device for converting weak direct current to alternating current for amplification by a high-gain alternating-current amplifier. Choppers are essential for converting weak bioelectric currents of low frequency to higher frequencies for amplification, and for converting the output of thermocouples to alternating current. Conventional choppers, particularly those of the mechanical variety that consist of motor-driven switches, tend to generate noise, or spurious currents. To chop with a Hall device, the signal to be converted is fed into the semiconductor and the electromagnet is energized with current at the frequency into which the experimenter wishes to chop the signal. The output voltage represents the arithmetical product of the direct-current signal and the alternating current supplied to the electromagnet [*see upper illustration on this page*].

"The Hall device also lends itself readily to the measurement of electric power. The energy consumed by a load in watts is equal to the product of the current in the load and the voltage across the load. The Hall device automatically multiplies its input quantities. To apply it as a wattmeter the current in the load (or a known portion of the load current) is used to energize the electromagnet. When the voltage across the load (or a known portion of the voltage) is applied to the semiconducting material, the output voltage varies directly with the power consumed by the load [*see lower illustration on this page*].

"Hall devices operate well over a wide range of frequencies. Hence they make ideal modulators. The carrier frequency of a radio transmitter can be modulated with voice signals simply by energizing the semiconductor with the high-frequency carrier current and modulating it by energizing the electromagnet with currents at voice frequency.

"Only a few of the known circuit applications have been described. Doubtless many others are possible. Their number appears to be limited only by the imagination of the experimenter."

Rio de Janeiro

New York City

Sampan City, Hong Kong

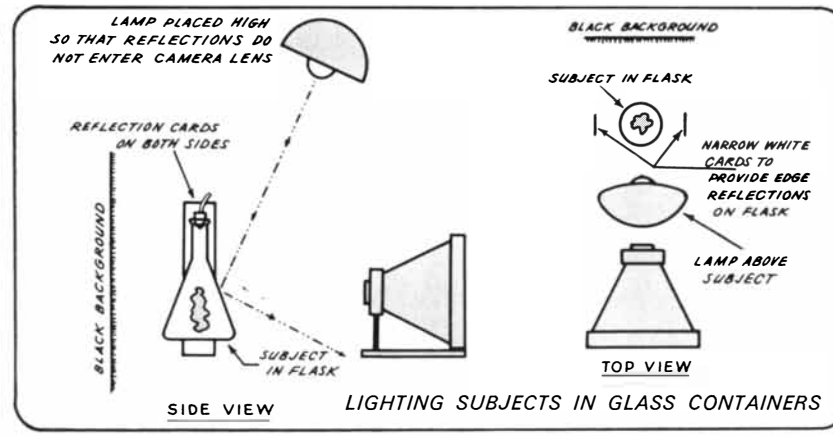
SCIENTIFIC AMERICAN

Announces a single-topic issue
concerned with the urbanization of the world population
under the title of

CITIES

To be published in September, 1965

FREEMAN BOOKS IN SCIENCE



LIGHTING SUBJECTS IN GLASS CONTAINERS

Photography for Scientific Publication, A HANDBOOK

ALFRED A. BLAKER, *Principal Photographer, Scientific Photographic Laboratory, University of California, Berkeley*

The first work of its kind, this handbook shows how to solve the many technical problems encountered in making photographs for scientific purposes. Based on experience gained in a laboratory devoted exclusively to this type of work, it provides instruction in photography for research illustration and serves as a convenient source of information on methods of photographing specific subjects.

Lighting, photography at magnifications, and rendition of monochromatic colorations are a few of the general photographic topics dealt with in the book, and dish cultures, insect materials, and liquids in tubes are typical of the many specific photographic subjects treated. Emphasis is on black-and-white still photography of objects falling between infinity focus and magnifications of up to about x60, using equipment that is normally available. The broad coverage includes biological and geological specimens, anthropological materials, and even art-object and portrait photography.

Throughout the book many line drawings are used to diagram lighting setups needed to achieve general effects or solve particular problems. At the end of the book a special collection of the author's photographs serves as a standard against which the reader may check the results of his own work.

A valuable sourcebook for researchers, teachers, and photographers in many fields, *Photography for Scientific Publication* is also a useful textbook.

1965, 158 pages, 55 illustrations (two plates in full color), $8.00

Handbook of Paleontological Techniques

Edited by **BERNHARD KUMMEL,** *Harvard University, and* **DAVID RAUP,** *The Johns Hopkins University*

1965

852 pages

102 illustrations

$18.00

Prepared under the auspices of the Paleontological Society, this volume fills the need of the paleontologist for an up-to-date and comprehensive sourcebook of field and laboratory techniques.

A collection of essays and two bibliographies make up the book. The essays, contributed by more than eighty specialists, describe both classic and modern techniques and range from methods of fossil-collecting to the preparation of illustrations for publication. Most of them are here published for the first time. The bibliographies provide a guide to the world literature on paleontological techniques.

There are five parts to the Handbook:

PART I. Procedures and techniques applicable to the major fossil groups.

PART II. Techniques having broad applicability—collecting techniques, mechanical methods of preparation, chemical methods of preparation, radiation and related techniques, casting and molding, and illustration.

PART III. Highly specialized methods used in the new and expanding field of palynology.

PART IV. Bibliography of additional literature on paleontological techniques.

PART V. Compilation of bibliographies of use to paleontologists and stratigraphers.

W. H. FREEMAN AND COMPANY
660 Market Street
San Francisco, California 94104

_____ Photography for Scientific Publication

_____ Handbook of Paleontological Techniques

I enclose payment of $_____ (California residents add 4% sales tax). Send me, postpaid, the books listed above. I understand that if I am not satisfied the books may be returned for full refund within 10 days after they are received.

NAME_____

ADDRESS_____

☐ Send me a complete catalogue of Freeman Books in Science.

W. H. FREEMAN AND COMPANY
660 Market St., San Francisco, Calif. 94104 / Warner House, 48 Upper Thames St., London, E.C. 4

BOOKS

The classes of stupidity

by Edwin G. Boring

MENTAL RETARDATION: A REVIEW OF RESEARCH, edited by Harvey A. Stevens and Rick Heber. The University of Chicago Press ($12.50).

Mental retardation is plainly not something new. The ancients had not only their famous wise men such as Socrates but also their idiots, whom they treated with aversion or derision and sometimes with persecution—even, as in Sparta, by casting them out to perish. Down through history the lot of the dim-witted, the insane and criminals was cruelty. They were kept in chains and abused, except that some of the idiots who failed to acquire speech came to be thought of as speaking only with God and were known as *les enfants du bon Dieu*. Thus in some communities the birth of an idiot was thought to be a mark of God's favor to the parents.

The movement for humane treatment did not come until the late 18th and early 19th centuries. It was in 1793 that Philippe Pinel released the insane from their chains at the great hospital at Bicêtre. Jean Jacques Rousseau, blaming man's ills on society, spoke of the "noble savage" and advocated a return to nature, but when in 1799 a wild boy was captured in the woods of Aveyron, it was hard to see nobility in his savagery. Pinel pronounced him an incurable idiot, but Jean Marie Itard took him up and sought to educate him, devoting patient years to his training, with only limited success. The boy never learned to talk or to walk instead of trot. Around 1850 asylums were established for the feebleminded and the insane, psychiatry was founded and the failure of education markedly to improve the dim-witted was taken to mean that this kind of mental deficiency was incurable. People came to realize that mental insufficiency exists in various degrees, with the idiots lowest in the scale and imbeciles a bit above them. Much later Henry H. Goddard added the class of morons above the inbeciles, and all three degrees were then known as feebleminded.

In 1912 Goddard published *The Kallikak Family*, the parallel histories of two lines of descent from the same paternal ancestor, one line from a normal wife and the other from an illegitimate union with a feebleminded girl. *Kallikak* is a name coined from the Greek roots for "good" and "bad," and the good line showed no feeblemindedness or delinquency, whereas the bad line was filled with records of stupidity, alcoholism, prostitution, epilepsy and other undesirable traits. This study seemed at the time to clinch the belief in the inheritance of feeblemindedness; actually, as was seen later, it did no such thing. Although it left open the possibility that mental retardation is inherited, eventually it also made clear that poverty, disease and low standards are products of family culture and are passed on from parents to their children by the character of the home and environment in which the children are raised.

Sir Francis Galton (in his study *Hereditary Genius*, published in 1869) had made a similar mistake at the other end of the scale of intelligence: he found that eminence in Britain tended to run in families and therefore he concluded that genius is heritable. He was confusing with biological inheritance the social inheritance of the opportunity for success that comes from affluence and a cultivated home, and he invented the term "eugenics" to popularize his belief in the need for selective human mating to increase the level of intelligence in Britain. The practice of sterilization of the feebleminded, which is legal in 30 states of the U.S., grew out of this misbelief.

The intelligence tests were not developed until the beginning of this century. First there was Alfred Binet and his scale of measurement, and then in 1916 Lewis M. Terman's revision of the scale and his adoption of the concept of the intelligence quotient (IQ). The scale measures mental age; the IQ is the ratio of mental to actual age. A child with intelligence appropriate for age 14 who is actually only 10 has an IQ of 140 (14/10 × 100), which is at the bottom edge of what Terman counted as genius. A child with a mental age of seven and an actual age of 10 has an IQ of 70, which Terman took to be at the top of the range of feeblemindedness. Terman firmly believed that the IQ is inherited and constant, and that it cannot be improved; he never quite gave up this view. Modern investigation shows that a great many social and educational influences can notably alter the IQ, although not change feeblemindedness to genius.

There is no doubt that public opinion can affect scientific belief. The rise of liberalism in the 1920's promoted a dubiousness about the immutability of mental retardation and the constancy of the IQ. Social psychologists wanted to believe in equal opportunity for all men. In the 1930's, however, opinion swung back toward an acceptance of the inheritance of behavior as well as of bodily structure. It was plain that a spider of one species always builds its particular style of web, adapting it to local circumstances, whereas the web of another species is characteristically different.

After the 1930's investigation of mental retardation became genuinely scientific, and this modern research is what the book reviewed here reports. The volume consists of 13 chapters by 14 authors, and it reports on some 1,200 researches. (The 1,741 studies cited include many duplications.) The book is really a handbook for the scientific in-group concerned with mental retardation; in undertaking to bring the review of research up to date it must frequently fail to come to any positive finding. Some researches are inconclusive, others contradict one another and often an author finishes his discussion simply by choosing the investigations he thinks ought to be undertaken next.

Altogether the book leaves one with the impression that mental retardation

is an enormously important problem that has not lacked for attention but that still remains unsolved. What would its solution mean? Presumably a solution would get at the causes and split mental retardation (or amentia, as it has often been called since the term "feeblemindedness" went out of style) into a variety of retardations due respectively to this or that other genetic "error," and to various diseases and accidents that insult the brain, and to the different social and cultural factors that hinder or prevent development. Knowing about intelligence as the tests test it is not enough; we still have a partly undefined problem. This book is aimed at specifying the varieties of retardation in order to define the problems and to map the areas of clear ignorance.

Can we assess the importance of the problem? This query could be met by a statement of the prevalence of mental retardation, but that in turn requires a definition of the term itself. Current practice and this book, unlike Goddard and Terman with their three grades, specify four levels of mental retardation:

I. *Profound.* IQ less than 20. Makes repetitive movements. No training is possible. Undoubtedly has brain damage or malformation. Usually supported by the state.

II. *Severe.* IQ 20–35. Total dependency throughout life but some training is possible. Usually brain damage.

III. *Moderate.* IQ 36–52. Can be trained and may fill a useful place in society.

IV. *Mild.* IQ 53–68. Slow in development and in learning but amenable to special training. If institutionalized, can be placed outside in unskilled or semiskilled useful work after training.

At all four levels there are about five million mental retardates in the U.S., and six million are predicted by 1970. Only about 200,000 of them are in state institutions, at a cost of $250 million a year for their special training. Mental retardation is about 10 times more frequent than diabetes, 20 times more frequent than tuberculosis, 25 times more frequent than muscular dystrophy. If those disabilities demand public concern, why not mental retardation? Perhaps because amentia is not specific; there are so many different ways it can be caused and so many different attacks on its problems are needed. Perhaps also because it cannot be cured or mitigated nearly so effectively as those other disabilities. Nevertheless, our society is awakening to the problem; we now have conferences on mental retardation, programs to cope with it and books such as this one.

The symptoms of mental retardation are behavioral maladaptation. The definition is not very satisfactory because such also are the symptoms of the psychoses. It is proper, however, to stress the social nature of mental retardation. One must deal, for instance, with such questions as why mental retardation is more common in cities than elsewhere (or is more likely to be recognized) and why it is that any census of retardates is so uncertain.

Chapter 2 of *Mental Retardation: A Review of Research* attempts to list the medical causes of mental retardation. It is a sorry list, because it has for an invariant only the fact that a normal brain is needed for normal mental development. What are the listed inhibitors to being bright? They are: (1) infection, (2) intoxication, (3) physical trauma, (4) metabolic dysfunction, (5) growths in the brain, (6) prenatal influence, which might be infection, bad metabolism or genetic error, including a genetic interference with metabolism.

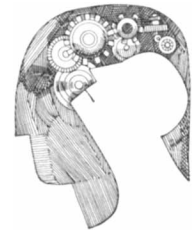
Chapter 3 tells about the education of mental retardates in special classes that have increased in attendance almost tenfold in 36 years, from 23,000 in 1922 to 197,000 in 1958. Belief in the value of special classes is obviously increasing, perhaps because belief grows more readily than fact. The author of the chapter complains that it is not at present clear whether or not the special classes really do increase the self-reliance, social adjustment and economic usefulness of the retardates. He thinks the research is poorly controlled and should be done again.

The investigations of learning, on the other hand, are more positive. The retardates of higher level learn slowly but their retention is good. They succeed well in rote learning, but their ceiling for complication is low; they fail at complex tasks. They have little capacity where abstraction is required. Training may not make them self-reliant in general but it can fit them for specific tasks that are not too complex. This rule applies in varying degrees for mild and moderate retardates, not, of course, for the severe and profound ones.

Do retardates have personalities? Of course they do, but personality is likely to be forgotten in the welter of negatives that describe them. They are subject to frustration, perhaps more so than the normal person, and like other people they react to frustration with aggression. The suggestion that they are more aggressive than other people, however, is based on observations made in institutions, where frustration is greater. Moreover, it is to institutions that the more aggressive retardates are likely to have been sent. Retardates like adult approval and resent disapproval. Their contact with caretaking adults works for their happiness when it yields praise and otherwise does not. They are not good self-appraisers; they tend to think they do better than is really the case, as indeed any of us might if we were frustrated and starved for praise. The late Kurt Lewin thought that retardates are more "rigid" than normal people, change less readily and find a variable environment less stimulating than others do; Rick Heber, the author of the chapter on personality and one of the editors of the book, thinks not. He notes experiments showing that this rigidity applies to change of motivation and not to change of learning and activity. The retardates do as well as normal subjects in relearning a suddenly reversed task.

In the opinion of this reviewer the best chapter in *Mental Retardation* is a discussion of environmental factors in relation to intellectual functioning by Boyd R. McCandless of Fort Hays Kansas State College. He restricts his discussion to "endogenous" mental retardation: to those who show no sign of brain injury and who might therefore be expected to profit from learning if any retardate can. How much is stolidity or alertness fixed by the pattern of social contact in which the child is involved from birth? McCandless gives somebody's list of stultifying early conditions: (1) stark poverty, (2) abandonment by parents, (3) social humiliation, rejection and defeat, (4) parental drunkenness, feeblemindedness and psychosis, (5) failure at school and rejection by peers, increasing with the child's age, (6) malnutrition, (7) cultural barrenness of the home, (8) parental indifference to the child's education and his social success, (9) poor health. He struggles for a definition of intelligence and decides that the ability to solve problems is the specification he likes best. Well, how much of the eager alertness to tackle the unexpected and to achieve the insight needed for solving a novel problem would you expect from a child handicapped by any of these listed conditions? No wonder retardates can have children brighter than themselves when foster parents give a child from an early age the stimulus that relative affluence and culture can provide. Adopted children rate in in-

WHERE CAN MAN GO...IN R & D?



To distant planets, to land-vehicles of the 1970's, to a region far beyond the grasp of man today—the ocean bottom. Lockheed's major R & D programs reach from deepest space to the ocean deep. Micro-particles, advanced concepts of re-entry, cryogenics, materials, lunar-crawling systems, automated hospitals, data compression systems, ocean mining and deep submergence—these are a few of Lockheed's R & D interests ranging across space, land and undersea environments. Scientists and engineers, particularly those with an R & D and systems bent, are invited to write Mr. K. R. Kiddoo, Professional Placement Manager, 576 Industrial Relations Building, Post Office Box 504, Sunnyvale, California. Lockheed is an equal opportunity employer.

LOCKHEED
MISSILES & SPACE COMPANY
A GROUP DIVISION OF LOCKHEED AIRCRAFT CORPORATION

telligence above the national average, because adoption nearly always involves a move upward in socioeconomic level. Adopted children of a retardate mother may have an IQ 20 to 30 points higher than that of the mother.

Language helps, because not only do intelligence tests tend to be verbal but also they usually demand abstraction, and language is the instrument of abstraction. Children who are neglected by their parents at home or who live in institutions lag in speech even more than in IQ. Twins who speak more with each other than with adults suffer in fluency and intelligence more than only children do. Bilingual children are at a disadvantage, because they command each language less well than they would one. Talk to your child when he is two months old, and start reading to him as soon as he is old enough to understand. He is likely then to demand more reading. Orphanage children who were read to for three hours every weekend for some six months gained 14 points in IQ, whereas a control group that was not read to gained only two points. The reading also increased the vocabulary of the group that was read to by 17 points, as against 10 points for the control group.

There is a great deal of evidence that the IQ of children depends to a significant degree on the intellectual and social status of the parents and thus quite often on their affluence. Herein lies a reason why racial equality in the U.S. cannot be achieved quickly. There may be a difference of 20 IQ points between children of parents of low and of high occupational status. Urban children test higher than rural, although that may be partly because the tests are prepared by urban people; no test is ever culture-free. Children in depressed rural areas, where the schools are poor, not only have IQ's below normal but also their IQ's diminish as they get older: their intelligence increases, but more slowly. The IQ of Negro children in Louisiana decreases with distance from New Orleans, the nearest urban center. The IQ of Negroes from the South increases the longer they remain in New York City. Here is a little table of average IQ's for the children of four kinds of parents:

Demanding parents	IQ 124
Overanxious parents	IQ 107
Normal parents	IQ 110
Unconcerned parents	IQ 97

One recalls what James Mill required of his son John Stuart Mill. The younger Mill was taught only by his father, was introduced to the study of Greek at the age of three, was required to teach himself Latin from a Greek-and-Latin dictionary at eight and eventually outshone his father with an estimated IQ of 190.

When it is argued that the effect of culture on the IQ is due to the verbal character of intelligence tests, one can turn to what isolation does to animals, who are clearly not dependent on words in tests of their intelligence. McCandless summarizes this literature too. Puppies petted and roughed up for a year from weaning do better than puppies kept in isolation when eventually both groups are given maze tests. Rats brought up in interesting cages with varied objects and inclined runways available excel in the maze compared with rats from plain cages. Rats that have been petted and gentled are brighter than rats left alone. Rats that have a full social life with other rats do better than rats that have been forced to live in isolation. A lively social life, with its challenge of constant problems, is good intellectually for rats and dogs as well as for men.

McCandless ends his chapter with a few telling questions. He asks: (1) Are the endogenous mentally retarded confined to the lower socioeconomic classes? Can improvement in social level become a preventive of mental retardation? (2) Should children of the endogenous mentally retarded be placed, when possible, at higher socioeconomic levels in order for them to be benefited? (3) Should high-grade nursery schools be established for children of two and over in blighted rural areas? Such questions obviously have considerable social significance.

Another chapter in *Mental Retardation* considers the role the mild and moderate retardates can play in the community. Most of them can be fitted in after job training in special classes, at least in good times. They are the first to lose their jobs in a depression. Automation works against them; the work they can do has to be unskilled or at best semiskilled. Many of them change jobs often, but in one survey a third of the retardates were found to have held their jobs for five years. Work in a good home is likely to be successful, but much depends on the home.

There is a chapter on epidemiology, which is unsatisfactory because its generalizations are only the obvious ones. There is great variability in the prevalence of mental retardation from community to community. An epidemic of rubella, or German measles, may be followed by an unusually large number of brain-injured births that result in an increase of the frequency of mental retardation. Retardation is associated not only with poverty but also with diseases such as epilepsy, cerebral palsy and cretinism, as well as with physical defects of vision, hearing, speech and locomotion. When one gets away from endogenous retardation, one has to begin to consider the incidence of many diseases.

The past 30 years have seen great strides in the study of the chemistry of the brain. No longer is knowledge limited to the effects of infection, hemorrhage and tumor or later to changes in the fine structure of the tissues. We have learned a great deal about the relation of carbohydrates, lipids, amino acids and proteins to cerebral function, and about the function of enzymes released by the liver and kidneys to produce substances that can sometimes be accepted by the brain and sometimes not (when they are so constituted as to be blocked by the "blood-brain barrier"). Biochemistry is a very active field and promises more knowledge about the causes of mental retardation. It would seem that these biochemical limitations are what place an upper limit on the training and social adjustment of the endogenous mental retardates.

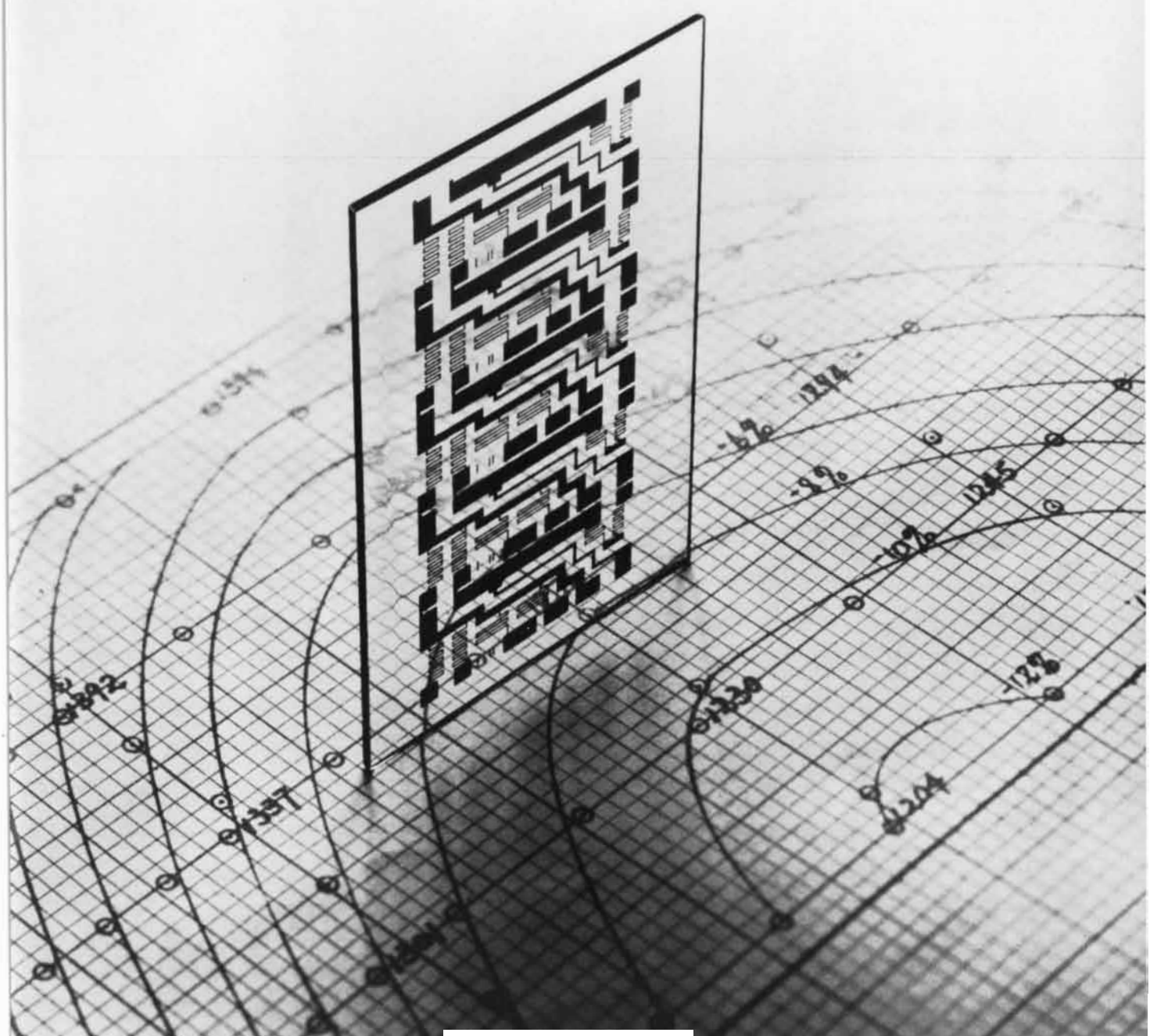
A chapter on genetics describes the great advances that have been made in this field too. Today the life history of chromosomes can be followed in tissue culture. Faulty genes, responsible for the creation of ineffective enzymes, can actually be identified. Behavior as well as structure is having its genetics determined. Large masses of genetic data can be handled by computers. There are long lists of disturbances in amino acid and hormone metabolism and in carbohydrate and lipid metabolism—disturbances that all lead to mental retardation. There are more than a score of chromosome anomalies that aggravate mental retardation. It seems to be an extra chromosome that is responsible for mongolism. The complexities of possible conditions of mental retardation are too great for any simple statement, but in general it is supposed that moderate retardation may have many causes but severe retardation a single primary cause.

Teratogenesis of the central nervous system is the genetic malformation of the brain with consequent psychologi-

PLOTTING THE UNMEASURABLE

In a tantalum thin film circuit where the film is 400 atom layers thick, a variation in thickness of as little as four atom layers can cause a significant difference in resistance. Since Western Electric is using an increasing number of thin film circuits in the sophisticated communications equipment we are now making for the Bell System, we needed a rapid but reliable method of measuring variations of that order. ☐ Originally Bell Telephone Laboratories engineers worked out a method of using a stylus measuring device to give measurements with a precision of about 15 atom layers. But this was not close enough. So Western Electric engineers went further. ☐ Using the statistical technique known as multiple regression analysis, they developed a computer program that enabled them to derive contour plots of the surface of the film which gave them results almost five times as accurate as their original measurements. This ability to analyze processes far more closely than straightforward measurement would allow is becoming an increasingly important part of Western Electric's job in helping the Bell telephone companies give America the most convenient communications on earth.

Western Electric
MANUFACTURING & SUPPLY UNIT OF THE BELL SYSTEM



Lockheed-Georgia Company

Eminent Engineers and Scientists

F. A. Cleveland

"I don't agree with you, but convince me," is the affable, interested reply made to a proposal spread out on the work table. The speaker is F. A. Cleveland, Assistant Chief Engineer and C-5A Deputy Program Manager at Lockheed-Georgia Company, and once again, a member of his staff feels the challenge and inspiration to work out a satisfactory solution to a baffling technical problem.

From the time he received his MA degree in Aeronautical Engineering in 1944 at Stanford University—following award of his AB degree in Mechanical Engineering in 1943—F. A. Cleveland has applied his inherent abilities as an organizer and a humanitarian to aid his brilliant technical knowledge in getting the job done.

In July 1944, upon graduation from Stanford, he went to work at the NACA's Lewis Laboratory, working mainly in wind tunnel testing of power plants, both subsonic and supersonic. It was here that he assisted in the development and design of the first supersonic wind tunnel of reasonable size, and where he designed the first supersonic model ramjet for test and also conducted the test program.

Cleveland joined the Lockheed Aircraft Corporation in 1946 as an aerodynamics engineer. He became Preliminary Design Group Engineer in charge of the Company's nuclear powered aircraft project in 1951 and was advanced to Department Manager, Bomber Design, in charge of the nuclear powered Weapon System 125A program in 1954.

In July 1956, Cleveland joined Lockheed's Georgia Division as Preliminary Design Division Engineer. In June 1958, he was promoted to Chief Advanced Design Engineer. In April 1961, he became Assistant Chief Engineer and Engineering Program Manager for the C-141 StarLifter.

Cleveland is a prolific writer of technical articles and briefings, and has achieved wide-spread recognition for his ability in this area. He has written several basic memoranda on the preparation of technical reports which have served as partial bases for a company manual on "How to Prepare Reports."

A member of many professional organizations and societies, Cleveland was the general chairman of the first annual national meeting of the American Institute of Aeronautics and Astronautics, held in Washington, D.C. in June of 1964.

Engineers and Scientists who are interested in becoming associated with the group at Lockheed-Georgia Company are invited to address inquiries to: Mr. F. A. Cleveland, Assistant Chief Engineer, Lockheed-Georgia Company, 834 West Peachtree Street, Atlanta, Georgia 30308, Dept. AA-80.

Lockheed-Georgia Company
A Division of Lockheed Aircraft Corporation
An Equal Opportunity Employer

cal disturbances. (The Greek word *teras* means monster.) The morphogenetic processes by which tissues become specialized in the development of the embryo are not well understood, and therefore aberrant development is also mysterious. The genetic cause of error is sometimes apparent, however; there is even an experimental teratology. Teratogens that act to distort development may be effective at one stage of the genetic process and not at another, depending on the kind of differentiation a given tissue is ready to undergo. Thus the effects of external agents are likely to be unpredictable. Radiation may affect the ovum or the embryo. Rubella may cause cataract or deafness if it occurs in the first four weeks of pregnancy. It may be that Asian influenza results sometimes in anencephalic, or brainless, babies. The effects of thalidomide on some embryos in early human pregnancy are now well known and have been confirmed with rabbits.

The final chapter of the book asks whether or not mental retardation is associated with behavioral disturbance and whether or not the relation is causal. Any retardate is capable of any psychosis or neurosis that unretarded people may have and for the same reasons. Neurosis is not common among retardates, however, whereas schizophrenia is. That could be because the retardates have a particularly hard time of it. The psychoses emerge more clearly in the older retardates. Next to mental retardation itself, psychoses are the most common reason for the institutionalization of retardates. Retardates are quite as capable as others of delinquency, but retardation cannot be considered a principal cause of delinquency. Of course intellectual deficiency does make it easier to be led into crime, even inadvertently, but this would seem to be a negative rather than a positive cause.

Our understanding of mental retardation has come a long way since Galton applied the normal, or "bell-shaped," curve of distribution to genius and stupidity, and the most rapid and varied scientific progress has been made in the past two or three decades. Research continues, and the main purpose of this book is to promote more such investigation. The relative importance of heredity and environment in human development is a central issue in human affairs, as is the related question of the degree to which man can safely intervene in shaping his heredity and environment. If progress in the study of mental retardation continues at its present pace, we can expect a considerable decrease in the confusion that surrounds this fundamental problem.

Short Reviews

THE WORLD OF PROBABILITY, by Solomon Diamond. Basic Books, Inc., Publishers ($4.95). STATISTICALLY SPEAKING, by Warren K. Garlington and Helen E. Shimota. Charles C Thomas, Publisher ($5.50). HOW TO TAKE A CHANCE, by Darrell Huff. W. W. Norton & Company, Inc. ($1.25). Winston Churchill was once asked, "If you had it all to do over, would you change anything?" "Yes," he replied, "I wish I had played the black instead of the red at Cannes and Monte Carlo." To all men probability and statistics are lifelong companions, whether or not they are recognized or understood, whether or not heeded. These three books are attempts to give the average reader at least a rudimentary comprehension of the mathematics of chance and errors (although it is doubtful that they would have helped Sir Winston, who had an ingrained hostility to such subjects).

The best of the group is Diamond's account of the fundamentals of statistics and their use in science. In fewer than 200 pages, using nothing more than arithmetic and elementary algebra, he explains, among other things, deviations, correlations, the chi square, Brownian movement, factor analysis, the Gaussian curve, Fisher's exact test, Pascal's triangle, Poisson's distribution and inverse probability. He writes clearly and patiently and covers a wide ground. Where his book is difficult, in spite of the minimum mathematical demands it makes, it is because the ideas themselves are difficult and not because of any shortcomings of his as an expositor.

Garlington and Shimota, whose little volume is a somewhat revised version of a series of articles published in a medical journal, skip through the subject superficially. The authors try relentlessly to be funny and offer nothing more than brief glimpses of half a dozen or so topics. About 100 pages long, and illustrated by diagrams and cartoons, the volume is overpriced.

The Huff is a paperback reprint of a rambling essay on poker and dice, game theory, the uses of probability in studies of extrasensory perception, the famous problem of the drunkard's walk, the ancient game of Morra, the normal-distribution curve, the laws of chance and weather prediction, and related matters.

Not too strenuous, consistently lively and moderately enlightening. Illustrations by Irving Geis.

DIRECT USE OF THE SUN'S ENERGY, by Farrington Daniels. Yale University Press ($7.50). With the huge and still accelerating consumption—not to say waste—of the earth's fossil fuels it becomes increasingly necessary to study the direct utilization of the sun's energy. In half a century, or perhaps less, chemical fuels will run short and atomic energy, which apparently can be produced only in large, expensive units and accordingly entails a high cost of distribution through power lines, will not meet many needs. In a single desert area of northern Chile some 28,000 square miles annually receive an amount of solar heat greater than all the heat produced in the world during the same period by the burning of coal, oil, gas and wood. Many difficulties have to be overcome before solar energy can be tapped on a large scale. Among the disadvantages of its use are the intermittency of sunlight in most regions, the cost of collecting the energy and of the collecting devices, and the problem of energy storage. This last problem involves such steps as the conversion of heat into electricity, the construction of electric storage cells, the raising of water to higher levels so that it can be used to drive turbines and the manufacture of chemicals that can subsequently serve as fuels. At the present time solar energy has greater promise in small undeveloped regions than in large industrial areas, where electricity is available at much lower cost. Research in this field has a fairly long history but has progressed only by fits and starts. Intensive work has been done by the author of this book at the University of Wisconsin. Here he sketches the evolution of solar-energy studies, explains the nature of solar radiation and describes in detail the various types of collectors, solar cookers and water heaters, methods of agricultural and industrial drying, space-heating systems, the distillation of water, solar furnaces and solar engines, thermoelectric and thermionic conversion, photochemical conversion, storage and transportation of power. This is an ably written book of practical as well as theoretical value. Illustrations.

PSYCHOANALYSIS AND CURRENT BIOLOGICAL THOUGHT, edited by Norman S. Greenfield and William C. Lewis. The University of Wisconsin Press ($8). For many years physiologists, psychologists and philosophers (but not all theologians) have been sure that the events of consciousness do not occur independently of neural activity. Freud, who trained with the physiologist Ernst Wilhelm von Brücke, the great Johannes Müller's pupil and thus Hermann von Helmholtz' confederate, sought to attack the problems of the mind neurologically and, failing, turned to the mentalistic procedures that evolved into psychoanalysis. Freud held to the end of his life, however, that ultimately psychology must become biological. This book presents 18 sophisticated "interdisciplinarians" in 16 well-documented essays (there are almost 1,000 references) that show how far psychoanalysis has moved toward biology since the time of Freud. It has moved far but still has much further to go. For example, we now know that the brain is perpetually active and that stimulation does not impress itself on the brain in kind but acts by modulating activity already in existence. The philosophy of science has also helped psychoanalysis to become biological; this volume, except for its last chapter, escapes from the constraints of mind-body dualism and deals with psychological events as constructs—quite as real as electrons—inferred from neural, behavioral and "conscious" data as scientific advantage indicates. One begins to wonder if psychoanalysis is not close to becoming a biological science, but then one suddenly recalls that it is a science of man and not of life; as yet there is no animal psychoanalysis. The old Cartesian limitation is still on us: psychoanalysis is limited to organisms with souls.

On what kind of psychobiological relations or identities does the argument of this book depend? On all psychosomatic phenomena. On the negative correlation between ego strength and somatic adaptation (Norman S. Greenfield and A. A. Alexander). On the way that eczema and asthma in babies depend on the relation of the parents to the child (Sheldon T. Selesnick and Zanwill Sperber). On the way in which the paradox of instinct as being both complex and simple is explained by the continuous activity of the brain (William C. Lewis). On the way in which change of attitude or instruction can alter autonomic response (Richard A. Sternbach), and the manner in which stress and ego defense change skin conductance, an autonomic response (Joseph C. Speisman). There is also the discovery of periodic occurrence and cessation of rapid eye movements dur-

ing sleep, movements that are correlated with dreaming as well as with the electroencephalogram (Frederick Snyder). Of especial interest is George S. Klein's report of his discovery that a person speaks with fewer inhibitions when he cannot hear what he is saying, that is, when the sound of his own voice is masked with "white" noise.

ON ANCIENT CENTRAL-ASIAN TRACKS, by Sir Aurel Stein. Pantheon Books ($5.95). A reprint of the late Sir Aurel Stein's account of archaeological explorations he carried out in the early part of the century in Chinese Turkestan and adjacent areas of innermost Asia. Stein's journeys in central Asia were conducted over a period of about 45 years with great energy and skill and with an almost joyous indifference to the hardships they entailed. This Hungarian-born and -educated "little gnome of a man with apple cheeks and the eyes of a youth"—as he was described in his 70's—had an unrivaled fluency in Oriental languages and a natural bent for looking and finding. He produced a series of classic archaeological and geographical reports, all of which are now out of print. This volume, addressed to a wider public than his lavishly illustrated quarto-sized records, is a graceful and illuminating summary of his labors. The book has an introduction by Jeannette Mirsky, who, according to the title page, also edited the work, but exactly what this editing consisted in is a well-hidden mystery. Illustrations.

THE WORKS OF SIR THOMAS BROWNE, edited by Geoffrey Keynes. The University of Chicago Press ($37.50). A four-volume revision of Sir Geoffrey Keynes's admirable edition of the works of Sir Thomas Browne, which was published in Britain during the years 1928–1931. Although this has long been a much coveted publication, when it originally appeared it was not very successful with the reading public; a special edition of 250 copies on handmade paper even had to be remaindered. In 1941, as Keynes recounts in his preface, a large stock of the regular edition was destroyed in a fire during the German bombing of London. Since then sets have become so scarce that the publisher was unable even after months of advertising to locate a set and the firm's file set had to be sacrificed to provide "printer's copy." In addition to correcting earlier errors the present version contains substantial addenda to Browne's correspondence. This is a delectable

piece of bookmaking, a pleasure to hold and to read and worthy of Browne's incomparable prose.

SOVIET EDUCATORS ON SOVIET EDUCATION, edited and translated by Helen B. Redl, with a foreword by Fritz Redl. The Free Press of Glencoe ($6.95). In this volume Mrs. Redl has presented selections from the writings of contemporary Soviet educators on such topics as heredity and upbringing, sex education, the family and Soviet society, schools and children's literature. The material is enlightening as a sample of Soviet theory and practice; a good deal of it, however, is either in the most banal educationist's vein or simply absurd. For example, an article by T. S. Atarov, a blushing savant in sex matters, tells us that stimuli "such as reading, cynical conversations, observing sex life of animals, inactivity, isolation from the collective life, can... produce desire for a sexual outlet through masturbation"; that "at all times children must sleep on hard beds, never lie on the stomach or keep their hands under the blankets"; that "children of early marriages tend to be listless and their life expectancy is shorter. These early marriages result in an early loss of husbands' potency and their premature aging." Fritz Redl has written an excellent foreword that explains how difficult it is for the visitor to the U.S.S.R. to judge the true character of Soviet behavior, even if he is a skilled observer. Mrs. Redl deserves credit for taking such pains with what must in many respects have been a painful task.

HITLER MOVES EAST: 1941–1943, by Paul Carell. Little, Brown and Company ($10). A history of "Operation Barbarossa," translated from the German, that unashamedly glorifies every feature and phase of the Nazi invasion of the U.S.S.R. and unabashedly bemoans the "tragic" failure of the campaign. It will be a long time—if ever—before a relatively disinterested account of this part of the war is written; this book is little better than a caricature.

THE AMBIDEXTROUS UNIVERSE, by Martin Gardner. Basic Books, Inc., Publishers ($5.95). An essay on symmetry and antisymmetry in nature and the constructs of science. Amusing and intriguing things are to be found in this book on such diverse matters as the properties of mirrors, right-handed mollusk shells, left-handed crystals, the bathtub vortex, the fourth dimension, interplanetary communication and the concept of parity. It is worth mentioning that Gardner's explanation of parity is more skillful than most of the other articles on the subject that have appeared.

THE MIND OF PRIMITIVE MAN, by Franz Boas. The Free Press ($1.95). A soft-cover edition, with a foreword by Melville J. Herskovits, of Franz Boas' major study of the relation between race and culture. *The Mind of Primitive Man* was one of the first books to win the distinction of being burned by the Nazis. It is a volume whose profound insights, although they were framed more than half a century ago, continue to hold living force and validity.

RAILROADS OF THE BLACK HILLS, by Mildred Fielder. Superior Publishing Company ($12.50). A lavishly illustrated, loving account of the little railroads of the Black Hills, which flourished during the last 20 years of the 19th century and provided, in addition to transportation, a full share of exciting, narrow-gauge, steep-grade yarns about engines and men. Indispensable for lovers of iron horseflesh.

CLASSIC NEW YORK: GEORGIAN GENTILITY TO GREEK ELEGANCE, by Ada Louise Huxtable. Anchor Books ($1.95). A profusely illustrated, well-written guide to the architectural remains of the buildings in the Georgian and Federal tradition found in Manhattan. This attractive paperback is the first of a six-volume series and should provide many pleasant hours for walkers of the New York streets.

Notes

BORES, BREAKERS, WAVES AND WAKES, by R. A. R. Tricker. American Elsevier Publishing Company, Inc. ($6.50). A handsomely illustrated, ably written, mostly nonmathematical introduction to the engrossing study of waves on water.

SUICIDE AND ATTEMPTED SUICIDE, by Erwin Stengel. Penguin Books (95 cents). A quiet, useful examination of the ubiquitous problem of suicide that draws on the author's own research and many other studies in different countries of incidence, methods, motivation and steps toward prevention.

ESSAYS 1958–1962 ON ATOMIC PHYSICS AND HUMAN KNOWLEDGE, by Niels Bohr. John Wiley & Sons, Inc. ($5). This collection of articles written by Bohr during the last five years of his

life considers such topics as causality and complementarity, the unity of human knowledge, reminiscences of Lord Rutherford and the genesis of quantum mechanics—all dear to Bohr's heart.

THE SCIENCE OF IONIZING RADIATION: MODES OF APPLICATION, compiled and edited by Lewis E. Etter. Charles C Thomas, Publisher ($26.50). A comprehensive, cooperative survey of the science of ionizing radiation, ranging from the history of the subject through such topics as radiation physics and radiation chemistry, radiobiology, medical uses of ionizing radiation and applications in industry, crystallography, archaeology, anthropology, agriculture and the graphic arts.

THE WORLD OF LEARNING, 1964–1965. Europa Publications Limited ($23.50). The 15th edition of this guide to educational, scientific and cultural institutions throughout the world.

THE EXPLORERS OF NORTH AMERICA: 1492–1806, by John Bartlet Brebner. The World Publishing Company ($2.45). A reprint of the revised (1955) edition of Brebner's valued account of American exploration from Columbus to Lewis and Clark. Paperback.

A COLLECTION OF PROBLEMS OF MECHANICS, by I. V. Meshcherskii. Pergamon Press ($12). An English translation of the 26th Soviet edition (1960) of a standard collection of some 1,300 problems of mechanics at the college level.

FLORA EUROPAEA: VOLUME I, LYCOPODIACEAE TO PLATANACEAE, edited by T. G. Tutin, V. H. Heywood and others. Cambridge University Press ($16). This handsome volume on the flora of Europe is the first installment of a work intended to provide a synthesis of the existing literature.

HANDBOOK OF BIBLICAL CHRONOLOGY, by Jack Finegan. Princeton University Press ($8.50). A leading student of New Testament history and archaeology examines the methods of time-reckoning in the ancient Biblical world and attempts to lay down the full sequence of Biblical dates from 605 B.C. on.

INTRODUCTION TO LOGIC AND TO THE METHODOLOGY OF DEDUCTIVE SCIENCES, by Alfred Tarski. Oxford University Press ($1.85). Tarski's skillful and much used elementary guide to mathematical logic is presented in soft covers.

INDEX OF ADVERTISERS

JULY 1965

BIBLIOGRAPHY

Readers interested in further reading on the subjects covered by articles in this issue may find the lists below helpful.

THE SUPPORT OF SCIENCE IN THE U.S.

BASIC RESEARCH AND NATIONAL GOALS: A REPORT TO THE HOUSE OF REPRESENTATIVES COMMITTEE ON SCIENCE AND ASTRONAUTICS FROM THE NATIONAL ACADEMY OF SCIENCES' PANEL ON BASIC RESEARCH AND NATIONAL GOALS. Government Printing Office, 1965.

REVIEWS OF DATA ON SCIENCE RESOURCES, Vol. 1, No. 4. National Science Foundation, 1965.

HYDROXYL RADICALS IN SPACE

DETECTION OF THE INTERSTELLAR OH LINES AT 1612 AND 1720 MC/SEC. F. F. Gardner, B. J. Robinson, J. G. Bolton and K. J. van Damme in *Physical Review Letters,* Vol. 13, No. 1, pages 3–7; July 6, 1964.

DISTRIBUTION AND MOTIONS OF OH NEAR THE GALACTIC CENTRE. J. G. Bolton, F. F. Gardner, R. X. McGee and B. J. Robinson in *Nature,* Vol. 204, No. 4953, pages 30–31; October 3, 1964.

OH ABSORPTION SPECTRA IN SAGITTARIUS. S. J. Goldstein, Jr., E. J. Gundermann, A. A. Penzias and A. E. Lilley in *Nature,* Vol. 203, No. 4940, pages 65–66; July 4, 1964.

RADIO OBSERVATIONS OF OH IN THE INTERSTELLAR MEDIUM. S. Weinreb, A. H. Barrett, M. L. Meeks and J. C. Henry in *Nature,* Vol. 200, No. 4909, pages 829–831; November 30, 1963.

THE SECONDARY RECOVERY OF PETROLEUM

A CALCULATION METHOD FOR CARBONATED WATER FLOODING. Noel de Nevers in *Society of Petroleum Engineers Journal,* Vol. 4, No. 1, pages 9–18; March, 1964.

FUNDAMENTALS OF RESERVOIR ENGINEERING. John C. Calhoun, Jr. University of Oklahoma Press, 1953.

THE LINEAR DISPLACEMENT OF OIL FROM POROUS MEDIA BY ENRICHED GAS. H. J. Welge, E. F. Johnson, S. P. Ewing, Jr., and F. H. Brinkman in *Journal of Petroleum Technology,* Vol. 13, pages 787–796; August, 1961.

THERMAL, SOLVENT AND IMPROVED GAS-DRIVE OIL-RECOVERY METHODS. L. E. Elkins in *History of Petroleum Engineering.* American Petroleum Institute, 1961.

THE PINEAL GLAND

CONTROL OF HYDROXYINDOLE O-METHYLTRANSFERASE ACTIVITY IN THE RAT PINEAL GLAND BY ENVIRONMENTAL LIGHTING. Julius Axelrod, Richard J. Wurtman and Solomon H. Snyder in *The Journal of Biological Chemistry,* Vol. 240, No. 2, pages 949–954; February, 1965.

MELATONIN SYNTHESIS IN THE PINEAL GLAND: EFFECT OF LIGHT MEDIATED BY THE SYMPATHETIC NERVOUS SYSTEM. Richard J. Wurtman, Julius Axelrod and Josef E. Fischer in *Science,* Vol. 143, No. 3612, pages 1328–1329; March 20, 1964.

THE PINEAL GLAND: A REVIEW OF THE PHYSIOLOGIC LITERATURE. Julian I. Kitay and Mark D. Altschule. Harvard University Press, 1954.

SOME CLINICAL, BIOCHEMICAL, AND PHYSIOLOGICAL ACTIONS OF THE PINEAL GLAND. Robert A. Cohen, Richard J. Wurtman, Julius Axelrod and Solomon H. Snyder in *Annals of Internal Medicine,* Vol. 61, No. 6, pages 1144–1161; December, 1964.

STRUCTURE AND FUNCTION OF THE EPIPHYSIS CEREBRI. Edited by J. Ariëns Kappers and J. P. Schadé in *Progress in Brain Research,* Vol. X. Elsevier Publishing Company, 1965.

ULTRASTRONG MAGNETIC FIELDS

HIGH MAGNETIC FIELDS. Edited by Henry Kolm, Benjamin Lax, Francis Bitter and Robert Mills. The M.I.T. Press and John Wiley & Sons, Inc., 1962.

PRODUCTION OF VERY HIGH MAGNETIC FIELDS BY IMPLOSION. C. M. Fowler, W. B. Garn and R. S. Caird in *Journal of Applied Physics,* Vol. 31, No. 3, pages 588–594; March, 1960.

SOLIDS UNDER PRESSURE. Edited by William Paul and Douglas M. Warschauer. McGraw-Hill Book Company, Inc., 1963.

THE ROLE OF CHLOROPHYLL IN PHOTOSYNTHESIS

ACTION SPECTRUM OF THE "SECOND EMERSON EFFECT." Govindjee and E. Rabinowitch in *Biophysical Journal,* Vol. 1, No. 1, pages 1–14; September, 1960.

FLUORESCENCE CHANGES IN PORPHYRIDIUM EXPOSED TO GREEN LIGHT OF DIFFERENT INTENSITY: A NEW EMISSION BAND AT 693 Mμ AND ITS SIGNIFICANCE TO PHOTOSYNTHESIS. A. Krey and Govindjee in *Proceedings of the National Academy of Sciences,* Vol. 52, No. 6, pages 1568–1572; December, 1964.

PRIMARY PROCESSES IN PHOTOSYNTHESIS. Martin D. Kamen. Academic Press, 1963.

ANCIENT JERUSALEM

BRITISH ARCHAEOLOGY ABROAD, 1964: REPORT NO. 3. Kathleen Kenyon in *Antiquity: A Quarterly Review of Archaeology,* Vol. 39, No. 153, pages 36–37; March, 1965.

EXCAVATIONS AT JERUSALEM, 1961. Kathleen M. Kenyon in *Antiquity: A Quarterly Review of Archaeology,* Vol. 36, No. 141, pages 93–96; March, 1962.

FLIES AND DISEASE

EXPERIMENTAL TRANSMISSION OF SALMONELLA TYPHIMURIUM BY HOUSEFLIES TO MAN. Bernard Greenberg in *The American Journal of Hygiene,* Vol. 80, No. 2, pages 149–156; September, 1964.

FLIES IN RELATION TO DISEASE: NON-BLOODSUCKING FLIES. G. S. Graham-Smith. Cambridge University Press, 1914.

THE HOUSEFLY: ITS NATURAL HISTORY, MEDICAL IMPORTANCE, AND CONTROL. Luther S. West. Comstock Publishing Company, 1951.

MATHEMATICAL GAMES

ART THAT PULSES, QUIVERS AND FASCINATES. John Canaday in *The New York Times Magazine,* Vol. 114, No. 39,110, pages 12–59; February 21, 1965.

THE AMATEUR SCIENTIST

THE HALL EFFECT AND RELATED PHENOMENA. E. H. Putney. Butterworth & Co., Limited, 1960.

ON A NEW ACTION OF THE MAGNET ON ELECTRIC CURRENTS. E. H. Hall in *American Journal of Mathematics,* Vol. 2, pages 287–292; 1879.

PRINCIPLES AND APPLICATIONS OF HALL-EFFECT DEVICES. M. Epstein, L. J. Greenstein and H. M. Sachs in *Proceedings of the National Electronics Conference, Vol. XV.* National Electronics Conference, Inc., 1960.

Yankee doodle dandy

This is our new aluminum electrolytic capacitor. It's used to keep AC from mingling with DC in radio and TV sets and other home instruments. We make it right here in the U.S.A. But its modest price tag would make you think it just came off the boat. And its quality is exceptionally high.

What's the secret? A large measure of home-grown ingenuity, for one thing. Mallory capacitor specialists worked out a way to mold the capacitor's case around its inside, all in one shot . . . sealing out moisture and assuring long, stable life. Ultra-modern manufacturing facilities, for another . . . the newest, brightest capacitor plant in the business.

All Mallory components—electronic, electrical and metallurgical—offer unusual value. And they bring extra value to products ranging from household appliances to supersonic aircraft. They help miniaturize. They help raise performance levels. They hold production costs in line. They add sales appeal. Value-building is a Mallory specialty. P. R. Mallory & Co. Inc., Indianapolis, Indiana 46206.

It's good business to do business with Mallory

MALLORY